Inferring Autonomous System Relationships in the Internet

Lixin Gao
Dept. of Electrical and Computer Engineering
University of Massachusetts, Amherst
http://www-unix.ecs.umass.edu/~lgao

Outline

• Internet Architecture and Routing
• AS Relationships
• Heuristic Algorithms
• Experimental Results
AS Commercial Relationships

• Provider-customer:
  – customer pays its provider for transit services
• Peer-peer:
  – exchange traffic between customers
  – no money exchange
• Sibling-sibling:
  – have mutual transit agreement
  – merging ISPs, Internet connection backup

Route Propagation Policy

• Constrained by contractual commercial agreements between administrative domains

e.g., An AS does not provide transit services between its providers
Why Infer AS Relationships?

• Crucial part of Internet structure
  – Connectivity does not imply reachability
  – Connectivity alone can not fully characterize structural properties of Internet

• No registry of AS relationships
  – Many ISPs are not willing to reveal their relationships to others in IRR
  – Relationships are evolving; hard to be up-to-date

Applications of AS Relationships

• Construct distance map
• Place proxy or mirror site servers
• Potentially avoid route divergence

• Help ISPs or domain administrators to achieve load balancing and congestion avoidance
• Help ISPs or companies to plan for future contractual agreements
• Help ISPs to reduce effect of misconfiguration and to debug router configuration files
Route Propagation Rule

- An AS or a set of ASes with sibling relationship does not provide transit services between any two of its providers and peers

- BGP routing table entries have certain patterns
Routing Table Entry

<table>
<thead>
<tr>
<th>Network</th>
<th>Next hop</th>
<th>AS Path</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.2.24.0/21</td>
<td>134.24.127.3</td>
<td>1740 i</td>
</tr>
<tr>
<td></td>
<td>194.68.130.254</td>
<td>5459 5413 i</td>
</tr>
<tr>
<td></td>
<td>158.43.133.48</td>
<td>1849 704 702 701 i</td>
</tr>
<tr>
<td></td>
<td>193.0.0.242</td>
<td>3333 286 i</td>
</tr>
<tr>
<td></td>
<td>144.228.240.93</td>
<td>1239 i</td>
</tr>
</tbody>
</table>

Routing Table Entry Patterns

uphill top provider

downhill top provider
Heuristic Algorithms

• Infer provider-customer and sibling-sibling
  – basic
  – refined
• Infer peer-peer
  – final

Basic Algorithms

• Heuristics:
  – Top provider has largest degree
• Based on patterns on BGP routing table entries
  – Consecutive AS pairs on the left of top provider are customer-to-provider or sibling-sibling edges
  – Consecutive AS pairs on the right of top provider are provider-to-customer or sibling-sibling edges
Initialize Consecutive AS Pair Relationship

Maximum degree AS
Sibling-sibling[$u_i,u_2$] = 1

**Refined Algorithm**

- Bogus Routing Entries
- Each routing table entry votes on AS relationships
- Ignore sibling-to-sibling relationship concluded by only one entry
Inferring Peer-Peer Relationships

- Peer-peer edge is between top provider and one of its neighbors only
- Heuristics:
  - peer-to-peer edge is between top provider and its higher degree neighbor
  - degrees of two peers do not differ significantly
    - $< R$ times

Final Algorithm

\[
\begin{align*}
\text{degree}[u_{j-1}] & \leq \text{degree}[u_{j+1}] \\
\end{align*}
\]
Final Algorithm

Experimental Verification

• Routing table from Route Views
  – Connected to 22 ISPs at 24 locations
  – Daily routing table dump
• Routing table from 3 days
  – ~1 million routing entries
Inference Results

<table>
<thead>
<tr>
<th>Date</th>
<th>TOTAL ROUTING ENTRIES</th>
<th>TOTAL EDGES</th>
<th>SIBLING - SIBLING EDGES INFERRED BY BASIC (PERCENTAGE)</th>
<th>SIBLING - SIBLING EDGES INFERRED BY REFINED (IGNORED ENTRIES)</th>
<th>PEER-PEER EDGES INFERRED BY FINAL [R=INFINITY] (PERCENTAGE)</th>
<th>PEER-PEER EDGES INFERRED BY FINAL [R=60] (PERCENTAGE)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1999/9/27</td>
<td>968674</td>
<td>11288</td>
<td>149 (1.3%)</td>
<td>124 (25)</td>
<td>884 (7.8%)</td>
<td>733 (6.5%)</td>
</tr>
<tr>
<td>2000/1/2</td>
<td>936058</td>
<td>12571</td>
<td>186 (1.47%)</td>
<td>135 (51)</td>
<td>838 (6.7%)</td>
<td>668 (5.3%)</td>
</tr>
<tr>
<td>2000/3/9</td>
<td>1227596</td>
<td>13800</td>
<td>203 (1.47%)</td>
<td>157 (46)</td>
<td>857 (6.2%)</td>
<td>713 (5.7%)</td>
</tr>
</tbody>
</table>

Verification of Inferred Relationships by AT&T

<table>
<thead>
<tr>
<th>OUR INFEERENCE</th>
<th>AT&amp;T INFORMATION</th>
<th>PERCENTAGE OF AS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Customer</td>
<td>Customer</td>
<td>99.8%</td>
</tr>
<tr>
<td></td>
<td>Peer</td>
<td>0.2%</td>
</tr>
<tr>
<td>Peer</td>
<td>Peer</td>
<td>76.5%</td>
</tr>
<tr>
<td></td>
<td>Customer</td>
<td>23.5%</td>
</tr>
<tr>
<td>Sibling</td>
<td>Sibling</td>
<td>20%</td>
</tr>
<tr>
<td></td>
<td>Peer</td>
<td>60%</td>
</tr>
<tr>
<td></td>
<td>Customer</td>
<td>20%</td>
</tr>
<tr>
<td>Nonexistent</td>
<td>Customer</td>
<td>95.6%</td>
</tr>
<tr>
<td></td>
<td>Peer</td>
<td>4.4%</td>
</tr>
</tbody>
</table>

Comparing inference results from Basic and Final(R=∞) with AT&T internal information
### Verification of Inferred Relationships by AT&T

<table>
<thead>
<tr>
<th>OUR INFEERENCE</th>
<th>AT&amp;T INFORMATION</th>
<th>PERCENTAGE OF AS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Customer</td>
<td>Customer</td>
<td>99.5%</td>
</tr>
<tr>
<td>Peer</td>
<td>Peer</td>
<td>0.5%</td>
</tr>
<tr>
<td>Peer</td>
<td>Peer</td>
<td>76.5%</td>
</tr>
<tr>
<td></td>
<td>Customer</td>
<td>23.5%</td>
</tr>
<tr>
<td>Sibling</td>
<td>Sibling</td>
<td>25%</td>
</tr>
<tr>
<td>Peer</td>
<td>Peer</td>
<td>50%</td>
</tr>
<tr>
<td></td>
<td>Customer</td>
<td>25%</td>
</tr>
<tr>
<td>Nonexistent</td>
<td>Customer</td>
<td>95.6%</td>
</tr>
<tr>
<td></td>
<td>Peer</td>
<td>4.4%</td>
</tr>
</tbody>
</table>

Comparing inference results from Refined and Final (R=∞) with AT&T internal information.

### Verification of Inferred Relationships by AT&T

<table>
<thead>
<tr>
<th>OUR INFEERENCE</th>
<th>AT&amp;T INFORMATION</th>
<th>PERCENTAGE OF AS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Customer</td>
<td>Customer</td>
<td>99.8%</td>
</tr>
<tr>
<td>Peer</td>
<td>Peer</td>
<td>0.2%</td>
</tr>
<tr>
<td>Peer</td>
<td>Peer</td>
<td>100%</td>
</tr>
<tr>
<td>Sibling</td>
<td>Sibling</td>
<td>20%</td>
</tr>
<tr>
<td></td>
<td>Peer</td>
<td>60%</td>
</tr>
<tr>
<td>Nonexistent</td>
<td>Customer</td>
<td>20%</td>
</tr>
<tr>
<td></td>
<td>Peer</td>
<td>4.4%</td>
</tr>
</tbody>
</table>

Comparing inference results from Basic and Final (R=60) with AT&T internal information.
WHOIS Lookup Service

- Supplies name and address of company that owns an AS
- AS pair might have sibling-sibling relation if
  - belong to the same company or two merging companies
  - belong to two small companies located closeby

Verification by WHOIS lookup Service

- Confirm 101 of 186 inferred sibling-sibling relationships (> 50%)
- Some unconfirmed sibling-sibling relationships might be due to
  - WHOIS service is not up to date
  - Not enough information
- Bogus Routes:
  - Router configuration typo: 7018 3561 7057 7075 7057
  - Misconfiguration of small ISPs:1239 11116 701 7018
  - ...
Conclusions and Further Work

• AS relationships are inherent aspect of Internet architecture
• Our heuristic algorithm is based on routing entry pattern derived from policy rules
• Verification:
  – AT&T (99%)
  – Whois services (>50%)
• Further Work:
  – Policy effect on AS path length
  – AS relationship evolution