Orbis: Rescaling Degree Correlations To Generate Annotated Internet Topologies

Priya Mahadevan

UC San Diego

Calvin Hubble (UC San Diego), Dmitri Krioukov (CAIDA), Bradley Huffaker (CAIDA) and Amin Vahdat (UC San Diego)
Motivation

• Network topology critical input to a range of important research areas
  • Performance of routing protocols, robustness of network, traffic engineering, application performance

• Two options:
  • Real measurements (skitter, Rocketfuel, RouteViews)
    • Fixed size
    • Do not allow sensitivity analysis
  • Topology generators (GT-ITM, Inet, BRITE)
    • Do not reproduce important metrics: betweenness (load), distance distribution, spectrum, assortativity coefficient
Orbis Topology Generator

• Generate structurally similar graphs  [SIGCOMM 2006]
  • Reproduce metrics of real observed networks

• Generate random graphs of a range of sizes  [This paper]
  • Scale up to understand protocol scalability
  • Scale down for tractable simulations
  • Sensitivity to random variations in graph properties

• Support domain-specific annotations
  • Internet graphs are more than just nodes and edges
  • Application behavior depends on node and edge properties
    • Annotating nodes in a router topology with AS membership
**dK-distributions**

*dK-distributions*: *degree correlations within non-isomorphic simple connected subgraphs of size d; d = 0, 1,… n* (number of nodes)

- **0K** (average degree)
- **1K** (degree distribution)
- **2K**
  - Probability of an edge between two degrees
- **3K**
  - Probability of $\Delta$ and $\Lambda$ among three degrees
- …
- **nK**
  - Given graph itself
Generating $dK$-Graphs

As $d$ increases, accuracy increases, so does complexity
Orbis Topology Generator

• Generate structurally similar graphs
  • Reproduce metrics of real observed networks
• Generate random graphs of a range of sizes
  • Scale graph sizes both up and down
• Support annotations
  • Node annotations - AS membership
Rescaling Challenges

Goal: To rescale given $dK$-distributions

What should the new degrees be?
How does the maximum degree scale?
Historical Skitter AS Graph Summary

- Analyzed skitter AS graphs from years 2000 to 2005
  - Graph grew from 3000 to 10,000 nodes
- Graph properties remain largely unchanged
  - Average degree ~ 6
  - Maximum degree scales almost linearly with graph size
  - Other graph metrics do not vary significantly
Degree Distribution in Skitter AS Graphs

![Graph showing degree distribution with PDF on the y-axis and node degree on the x-axis for different years: Year 2000, Year 2001, Year 2002, Year 2003, Year 2004, Year 2005. The graph indicates a power-law distribution.]

SIGCOMM 2007                                   Priya Mahadevan
1K Rescaling Technique

- Rescale the degree distribution vector
- Attempt to preserve the shape of the PDF of the degree distribution
  - Keep proportion of low degree nodes in rescaled graph same as original
  - Linearly scale with size (of rescaled graph) high degrees
  - Keep number of high degree nodes in rescaled graph same as original
1K Rescaling – Degree Distribution in Generated AS-level Topologies

![Graph showing the degree distribution of AS-level topologies with varying node counts. The x-axis represents the node degree, while the y-axis represents the probability density function (PDF). The graph includes lines for different node counts: Original (9200 nodes), 9200 nodes, 6000 nodes, 12000 nodes, 15000 nodes, and 30000 nodes. Each line is color-coded and differentiated by markers. The graph demonstrates the distribution of node degrees across the generated topologies.]
### Scalar Metrics for 1K Rescaled AS Graphs

<table>
<thead>
<tr>
<th>Metric</th>
<th>Skitter (9200 nodes)</th>
<th>6000 nodes</th>
<th>12000 nodes</th>
<th>15000 nodes</th>
<th>30000 nodes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Avg. degree</td>
<td>6.29</td>
<td>6.17</td>
<td>6.38</td>
<td>6.35</td>
<td>6.44</td>
</tr>
<tr>
<td>Assortativity</td>
<td>-0.24</td>
<td>-0.23</td>
<td>-0.23</td>
<td>-0.22</td>
<td>-0.21</td>
</tr>
<tr>
<td>Mean Clustering</td>
<td>0.46</td>
<td>0.23</td>
<td>0.25</td>
<td>0.27</td>
<td>0.27</td>
</tr>
<tr>
<td>Avg. distance</td>
<td>3.12</td>
<td>3.18</td>
<td>3.15</td>
<td>3.13</td>
<td>3.1</td>
</tr>
<tr>
<td>Smallest eigenvalue</td>
<td>0.1</td>
<td>0.1</td>
<td>0.09</td>
<td>0.09</td>
<td>0.08</td>
</tr>
<tr>
<td>Largest eigenvalue</td>
<td>1.9</td>
<td>1.89</td>
<td>1.9</td>
<td>1.9</td>
<td>1.93</td>
</tr>
</tbody>
</table>

1K reproduces most metrics for AS graphs
Predictive Power

- Skitter AS graph for year 2000 has ~ 3,000 nodes
- From 1K distribution of this graph, we generated a 10,000 node topology
- Skitter AS graph for year 2005 has ~ 10,000 nodes
- Compared our predicted graph with the actual graph
  - All metric values (except clustering) of both graphs match
  - By definition, clustering will be reproduced by 3K
- We cannot predict evolution, but graph properties stay the same between 2000 and 2005
1K Rescaling for HOT Graph

Original HOT graph – 939 nodes

1K-random graph – 2000 nodes
## Scalar Metrics for 1K Rescaled Router Graphs

<table>
<thead>
<tr>
<th>Metric</th>
<th>HOT (939)</th>
<th>250 nodes</th>
<th>1000 nodes</th>
<th>5000 nodes</th>
<th>8000 nodes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Avg. degree</td>
<td>2.1</td>
<td>2.09</td>
<td>2.5</td>
<td>2.28</td>
<td>2.22</td>
</tr>
<tr>
<td>Assortativity</td>
<td>-0.22</td>
<td>-0.32</td>
<td>-0.14</td>
<td>-0.1</td>
<td>-0.1</td>
</tr>
<tr>
<td>Mean Clustering</td>
<td>0</td>
<td>0.002</td>
<td>0.009</td>
<td>0.003</td>
<td>0.001</td>
</tr>
<tr>
<td>Avg. distance</td>
<td>6.81</td>
<td>5.4</td>
<td>4.4</td>
<td>4.94</td>
<td>5.4</td>
</tr>
<tr>
<td>Smallest eigenvalue</td>
<td>0.004</td>
<td>0.006</td>
<td>0.034</td>
<td>0.015</td>
<td>0.007</td>
</tr>
<tr>
<td>Largest eigenvalue</td>
<td>1.997</td>
<td>1.994</td>
<td>1.967</td>
<td>1.985</td>
<td>1.994</td>
</tr>
</tbody>
</table>

1K is insufficient for router graphs
2K Rescaling

2K-distribution: Probability of edge between nodes of any two degrees

- In 1K, we preserve the shape of the line
- In 2K, we preserve the shape of the surface
Rescaling for HOT Graph

Original HOT graph – 939 nodes

1K-random graph – 2000 nodes
Rescaling for HOT Graph

Original HOT graph – 939 nodes

2K-random graph – 2000 nodes
### Scalar Metrics for 2K Rescaled Router Graphs

<table>
<thead>
<tr>
<th>Metric</th>
<th>HOT (939)</th>
<th>250 nodes</th>
<th>1000 nodes</th>
<th>5000 nodes</th>
<th>8000 nodes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Avg. degree</td>
<td>2.1</td>
<td>2.09</td>
<td>2.18</td>
<td>2.2</td>
<td>2.21</td>
</tr>
<tr>
<td>Assortativity</td>
<td>-0.22</td>
<td>-0.36</td>
<td>-0.23</td>
<td>-0.18</td>
<td>-0.18</td>
</tr>
<tr>
<td>Mean Clustering</td>
<td>0</td>
<td>0.0005</td>
<td>0.0004</td>
<td>0.0005</td>
<td>0.0001</td>
</tr>
<tr>
<td>Avg. distance</td>
<td>6.81</td>
<td>5.4</td>
<td>6.32</td>
<td>6.6</td>
<td>6.92</td>
</tr>
<tr>
<td>Smallest eigenvalue</td>
<td>0.004</td>
<td>0.01</td>
<td>0.005</td>
<td>0.004</td>
<td>0.005</td>
</tr>
<tr>
<td>Largest eigenvalue</td>
<td>1.997</td>
<td>1.986</td>
<td>1.996</td>
<td>1.996</td>
<td>1.996</td>
</tr>
</tbody>
</table>

2K reproduces most metrics for router graphs.
Orbis Topology Generator

- Generate structurally similar graphs
  - Reproduce metrics of real observed networks
- Generate random graphs of a range of sizes
  - Scale graph sizes both up and down
- Support annotations
  - Node annotations - AS membership
**Annotating routers with AS-membership**

- **Goal:** Generate a specified size router-level topology with AS-membership for each router
- **Based on skitter traces**
  - Extract both router- and AS-level topology from traceroute data

For each router:

- **AS-membership**
- **Is router peering or internal**
Generating Router Topologies Annotated with AS-membership

Generate AS topology of required size using 2K distribution of AS graph from skitter data
Generating Router Topologies Annotated with AS-membership

Generate router topology for each AS based on the AS’s degree
Generating Router Topologies Annotated with AS-membership

For each AS, assign peering routers in the router-level topology based on degree distributions of peering routers in original skitter data.
Generating Router Topologies Annotated with AS-membership

Connect peering routers for every connected AS pair to obtain final topology
Practical Issues

• Scaling factor?
  • Scaling up and down by a factor of 10
  • Higher scaling will likely require understanding fundamental network evolution principles

• Measurement accuracy
  • Generated topologies depend on the accuracy of measurements
Contributions and Conclusions

• Topology is critical for a range of networking studies
• Orbis: A topology generator suite based on $dK$-distributions
  • Reproduces structure of measured networks
  • Generates graphs of a range of sizes (factor of 10)
  • Supports annotations (AS-membership for routers in router-level topology)
• Source code will be available for download soon
Thank You

pmahadevan@cs.ucsd.edu
http://sysnet.ucsd.edu/~pmahadevan/