Primitives for Active Internet Topology Mapping: Toward High-Frequency Characterization

Robert Beverly, Arthur Berger*, Geoffrey Xie

Naval Postgraduate School
*MIT/Akamai

February 9, 2011

CAIDA Workshop on Active Internet Measurements
Long-standing question: *What is the topology of the Internet?*

Difficult to answer – Internet is:

- A large, complex distributed system (organism)
- Non-stationary (in time)
- Difficult to observe, multi-party (information hiding)
- Poorly instrumented (not part of original design)

⇒ Poorly understood topology (interface, router, or AS level)
Long-standing question: *What is the topology of the Internet?*

Difficult to answer – Internet is:

- A large, complex distributed system (organism)
- Non-stationary (in time)
- Difficult to observe, multi-party (information hiding)
- Poorly instrumented (not part of original design)

⇒ Poorly understood topology (interface, router, or AS level)
What is the topology of the Internet?

Why care?

- Network Robustness: to failure, to attacks, and how to best improve. (antithesis – how to mount attacks)
- Impact on Research: network modeling, routing protocol validation, new architectures, Internet evolution, etc.
- Easy to get wrong (see e.g. “What are our standards for validation of measurement-based networking research?” [KW08])

These challenges and opportunities are well-known. We bring some novel insights to bear on the problem.
Our focus:

- Active probing from a fixed set of vantage points
- High-frequency, high-fidelity continuous characterization
- Use external knowledge and adaptive sampling to solve:
  - Which destinations to probe
  - How/where to perform the probe

This Talk:

1. Characterize production topology mapping systems
2. Develop/analyze new primitives for active topology discovery
The Problem
Measurement Techniques

Archipelago/Skitter/iPlane

Production Topology Measurement
- Ark/Skitter (CAIDA), iPlane (UW)
- Multiple days and significant resources for complete cycle

Ark probing strategy:
- IPv4 space divided into /24’s; partitioned across ~ 41 monitors
- From each /24, select a single address at random to probe
- Probe == Scamper [L10]; record router interfaces on forward path
- A “cycle” == probes to all routed /24’s

Investigate one vantage point (Jan, 2010):

<table>
<thead>
<tr>
<th></th>
<th>Ark</th>
<th>iPlane</th>
</tr>
</thead>
<tbody>
<tr>
<td>Traces</td>
<td>263K</td>
<td>150K</td>
</tr>
<tr>
<td>Probes</td>
<td>4.4M</td>
<td>2.5M</td>
</tr>
<tr>
<td>Prefixes</td>
<td>55K</td>
<td>30K</td>
</tr>
</tbody>
</table>

R. Beverly, A. Berger, G. Xie (NPS)
Q1: How similar are traceroutes to the same destination BGP prefix?

- Use Levenshtein “edit” distance DP algorithm
- Determine the minimum number of edits (insert, delete, substitute) to transform one string into another
- e.g. “robert” → “robber” = 2

- We use: \( \Sigma = \{0, 1, \ldots, 2^{32} - 1\} \)
- Each unsigned 32-bit IP address along traceroute paths \( \in \Sigma \)
Q1: How similar are traceroutes to the same destination BGP prefix?

- Use Levenshtein “edit” distance DP algorithm
- Determine the minimum number of edits (insert, delete, substitute) to transform one string into another
- e.g. “robert” → “robber” = 2

We use: $\Sigma = \{0, 1, \ldots, 2^{32} - 1\}$
- Each unsigned 32-bit IP address along traceroute paths $\in \Sigma$

ED=2

129.186.6.251 129.186.254.131 192.245.179.52 4.53.34.13
129.186.6.251 192.245.179.52 4.69.145.12
Q1: How similar are traceroutes to the same destination BGP prefix?

- ~60% of traces to destinations in same BGP prefix have $ED \leq 3$
- Fewer than 50% of random traces have $ED \leq 10$
Q1: How similar are traceroutes to the same destination BGP prefix?

- ≈60% of traces to destinations in same BGP prefix have $ED \leq 3$
- Fewer than 50% of random traces have $ED \leq 10$

Confirms our intuition
Q2: *How much path variance is due to the last-hop AS?*

- Intuitively, number of potential paths exponential in the depth
- More information gain at the end of the traceroute?
The Problem

Measurement Techniques

Edit Distance

Q2: Variance due to the last-hop AS?

- Lob off last AS
- Answer: lots!
- For \( \sim 70\% \) of probes to same prefix, we get no additional information beyond leaf AS
Q2: Variance due to the last-hop AS?

- Lob off last AS
- Answer: lots!
- For \(~70\%\) of probes to same prefix, we get no additional information beyond leaf AS

Significant packet savings possible (DoubleTree)
Meta-Conclusion: adaptive probing a useful strategy

We develop three primitives:

1. Subnet Centric Probing
2. Vantage Point Spreading
3. Interface Set Cover

These primitives leverage adaptive sampling, external knowledge (e.g., common subnetting structure, BGP, etc), and data from prior cycles to maximize efficiency and information gain of each probe.
We develop three primitives:

1. Subnet Centric Probing
2. Vantage Point Spreading
3. Interface Set Cover

Best explained by understanding sources of path diversity:
### Methodology

**Subnet Centric Probing**

#### Granularity vs. Scaling

- $\sim 2^{32-1}$ possible destinations (2.9B from Jan 2010 routeviews)

#### Subnet Centric Probing

- From a single vantage point, no path diversity into the AS
- Path diversity due to AS-internal structure

![Diagram of Subnet Centric Probing](image-url)
**Goal:** adapt granularity, discover internal structure

- Leverage BGP as coarse structure
- Follow *least common prefix*: iteratively pick destinations within prefix that are maximally distant (in subnetting sense)
- Address “distance” is misleading: e.g. 18.255.255.100 vs. 19.0.0.4 vs. 18.0.0.5
- Stopping criterion: $ED(t_i, t_{i+1}) \leq \tau; \tau = 3$
Inferred degree distribution well-approximates ground-truth

Captures ≥ 90% of the vertex and edge fidelity
Captures $\geq 90\%$ of the vertex and edge fidelity

Using $\sim 60\%$ of ground-truth load
Discover AS ingress points and paths to the AS via multiple vantage points

Random assignment of destinations to vantage points is wasteful

E.g. empirically, the 16 /24’s in a /20 prefix are hit on average by 12 unique VPs
Using BGP knowledge, maximize the number of distinct VPs per-prefix

Note, this is complimentary to SCP
Diminishing return of vantage point influence

Vertices in resulting graph as compared to random: ~6% increase “for free.”
As shown in preceding analysis, full traces very inefficient
Perform greedy minimum set cover approximation (NP-complete)
Select subset of prior round probe packets for current round
Interface Set Cover

Generalizes DoubleTree [DRFC05] without parametrization
Efficient
Inherently multi-round
Additional probing for validation mis-matches (e.g. load balancing, new paths)
20K random IP destinations each day over a two-week period, fraction of missing interface using ISC

Uses $\leq 20\%$ of the full probing load ($\sim 30\%$ of full trace set cover)
Summary

Take-Aways:

- Deconstructed Ark/iPlane topology tracing as case study
- Developed primitives for faster, more efficient probing:
  - Subnet Centric Probing, Interface Set Cover, Vantage Point Spreading
  - Significant load savings without sacrificing fidelity

Future

- Combining our primitives on production system
- Refine ISC “change-driven” logic
- Build a better Internet scope to detect small-scale dynamics

Thanks!

Questions?

R. Beverly, A. Berger, G. Xie (NPS)