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

Robert Beverly, Arthur Berger*, Geoffrey Xie

Naval Postgraduate School
*MIT/Akamai

November 2, 2010

ACM Internet Measurement Conference

Internet Topology

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)

$\Rightarrow$ Poorly understood topology (interface, router, or AS level)

Internet Topology

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)

$\Rightarrow$ 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 Work

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

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 $\sim$ 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):

	Ark	iPlane
Traces	263K	150K
Probes	4.4M	2.5M
Prefixes	55K	30K

Path-pair Distance Metric

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`

Path-pair Distance Metric

Q1: *How similar are traceroutes to the same destination BGP prefix?*

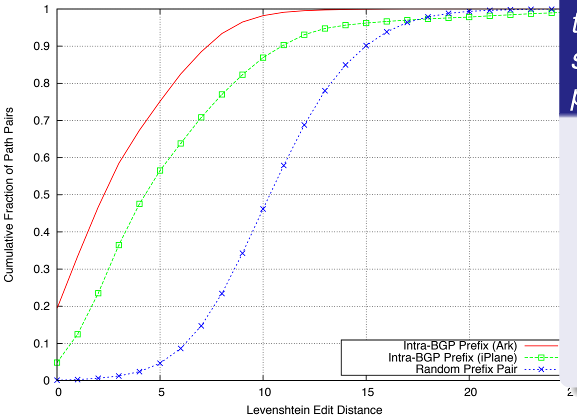
- Use Levenshtein "edit" distance DP algorithm
- Determine the minimum number of edits (insert, delete, substitute) to transform one string into another
- e.g. "robert" $\rightarrow$ "robber" = 2

- 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
```

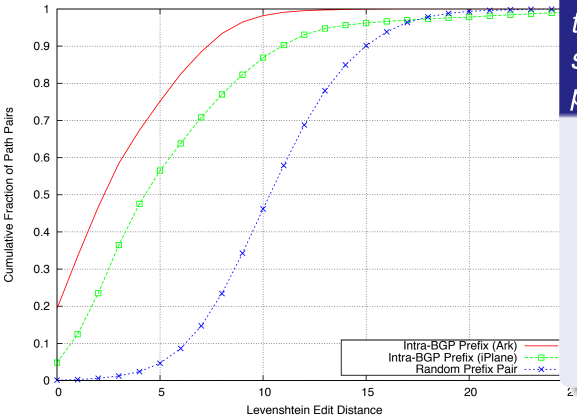
Path-pair Distance Metric



Q1: ***How similar are traceroutes to the same destination BGP prefix?***

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

Path-pair Distance Metric



Q1: *How similar are traceroutes to the same destination BGP prefix?*

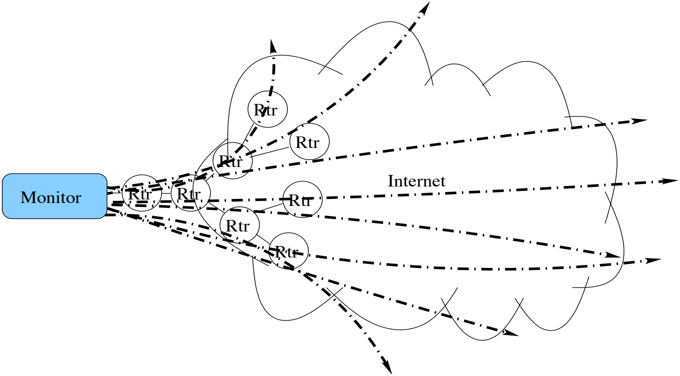
- ∼60% of traces to destinations in same BGP prefix have $ED \leq 3$
- Fewer than 50% of random traces have $ED \leq 10$

Confirms our intuition

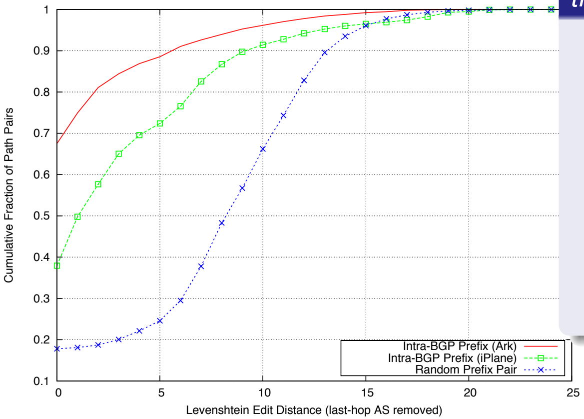
Edit Distance

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?



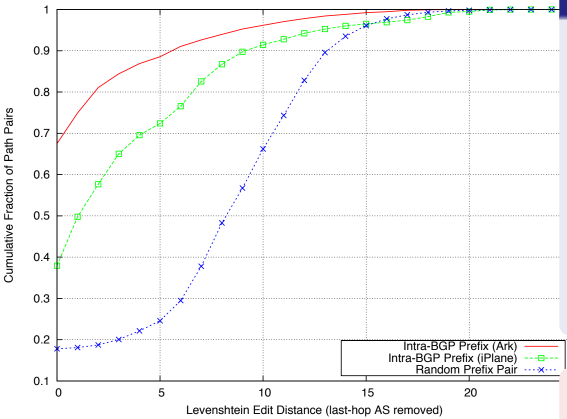
Edit Distance



Q2: *Variance due to the last-hop AS?*

- Lob off last AS
- Answer: lots!
- For $\sim$ 70% of probes to <u>same</u> prefix, we get <u>no</u> additional information beyond leaf AS

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

Significant *packet* savings possible (DoubleTree)

Adaptive Probing Methodology

Meta-Conclusion: adaptive probing a useful strategy

We develop three primitives:

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

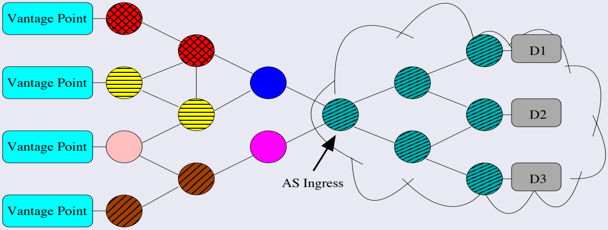
These primitives leverage adaptive sampling, external knowledge (e.g., common subnetting structure, BGP, etc), and data from prior cycles to *maximize efficiency and information gain of each probe*.

Adaptive Probing Methodology

We develop three primitives:

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

Best explained by understanding sources of path diversity:

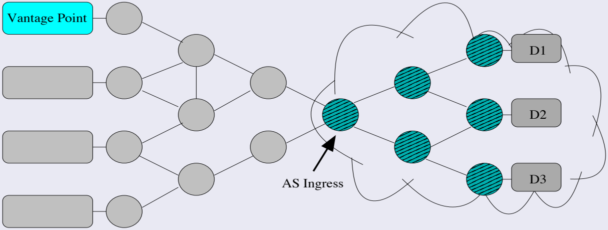


Subnet Centric Probing

Granularity vs. Scaling

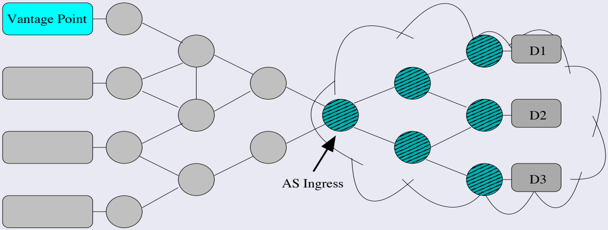
- $\sim 2^{32-1}$ possible destinations (2.9B from Jan 2010 routeviews)
- What granularity? /24's? Prefixes? AS's?

Subnet Centric Probing



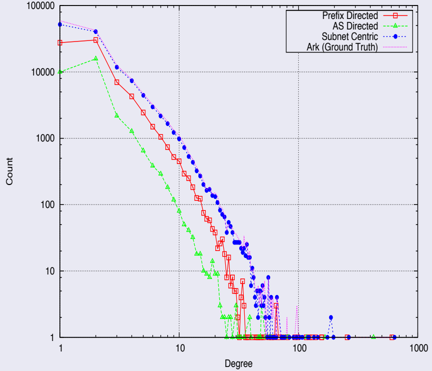
- From a single vantage point, no path diversity into the AS
- Path diversity due to AS-internal structure

Subnet Centric Probing

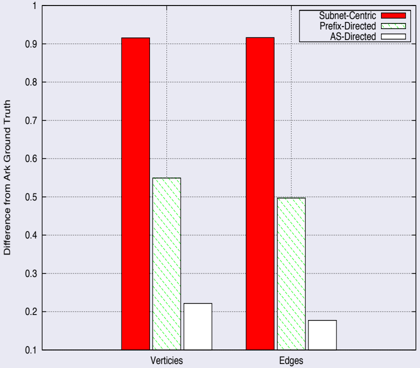


- **Goal:** adapt granularity, discover internal structure
- Leverage BGP as coarse structure
- Follow *least common prefix:* iteratively pick destinations within prefix that are maximally distant (in subnetting sense)
- Address “distance” is misleading: e.g. `18.255.255.100` vs. `19.0.0.4` vs. `18.0.0.5`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

Subnet Centric Probing

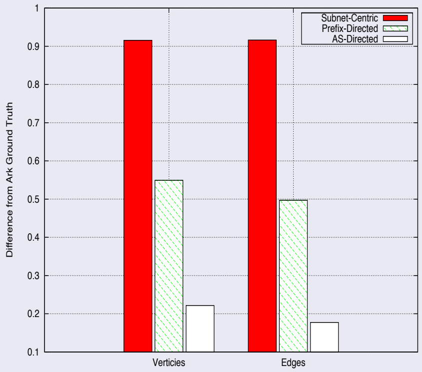


Inferred degree distribution well-approximates ground-truth

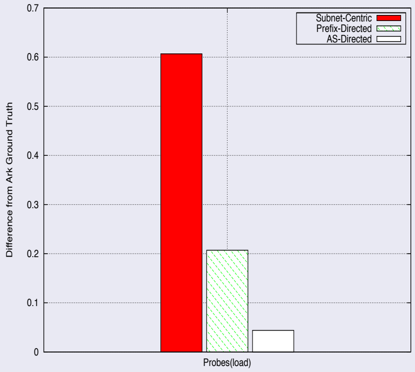


Captures $\geq$ 90% of the vertex and edge fidelity

Subnet Centric Probing



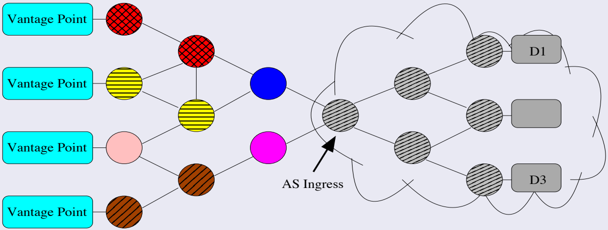
Captures $\geq$ 90% of the vertex and edge fidelity



Using $\sim$ 60% of ground-truth load

Vantage Point Spreading

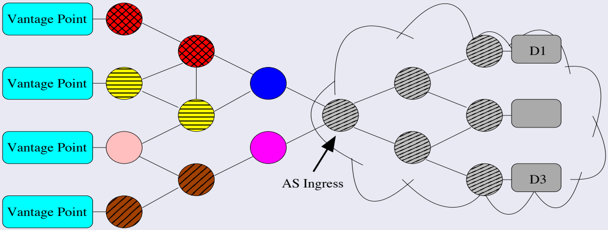
Vantage Point Spreading



- 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

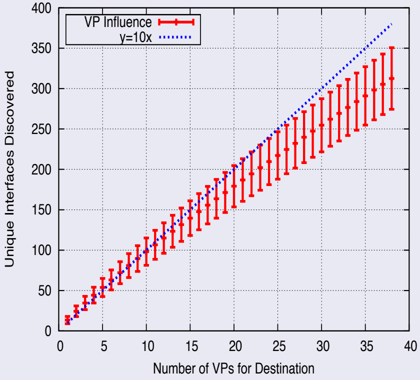
Vantage Point Spreading

Vantage Point Spreading

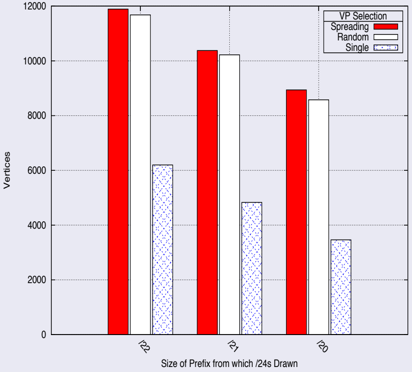


- Using BGP knowledge, maximize the number of distinct VPs per-prefix
- Note, this is complimentary to SCP

Vantage Point Spreading



Diminishing return of vantage point influence

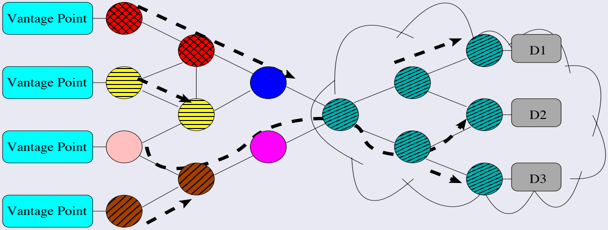


Vertices in resulting graph as compared to random: ∼ 6% increase "for free."

Interface Set Cover

Interface Set Cover

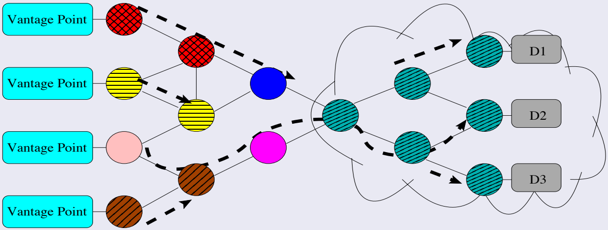
- As shown in preceding analysis, full traces very inefficient
- Perform greedy minimum set cover approximation (NP-complete)
- Select subset of *prior* round probe packets for *current* round



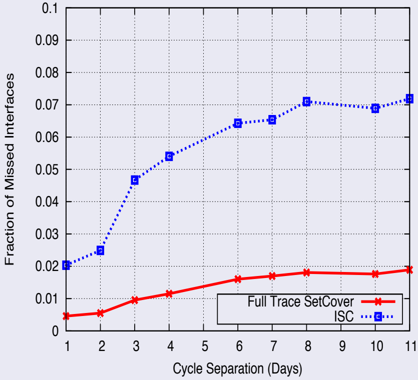
Interface Set Cover

Interface Set Cover

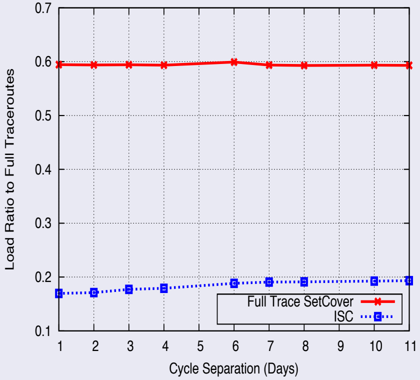
- Generalizes DoubleTree [DRFC05] without parametrization
- Efficient
- Inherently multi-round
- Additional probing for validation mis-matches (e.g. load balancing, new paths)



Interface Set Cover



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?