



Towards Capturing Representative AS-Level Internet Topologies

Hyunseok Chang
U. Of Michigan

Ramesh Govindan
USC

Sugih Jamin
U. Of Michigan

Scott J. Shenker
ICIR

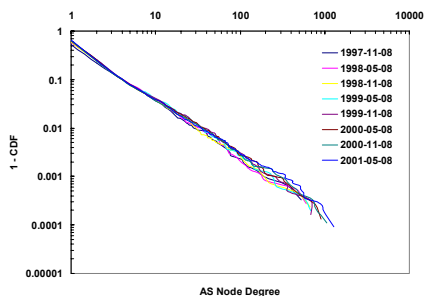
Walter Willinger
AT&T

Objective

- How **complete** & how **representative** is the BGP-derived AS graph?

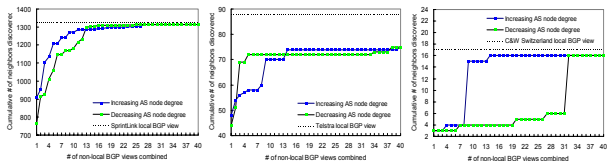
Background

- BGP-inferred AS graph: **Highly-variable** node degree distribution, characterized by $f_d \propto d^{-\alpha}$ ($d=1,2,\dots$; α : const)



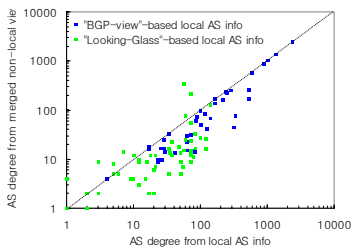
Local view vs. non-local view

- Question: How well is **my** connectivity observed by **others**?
- Dataset:
 - full BGP dumps of 41 different ASs,
 - Looking Glass data from ~70 different ASs.



Observations:

- The connectivity of tier-1 ASs is more easily observed than that of non tier-1 ASs.
- The connectivity of an AS is more easily observed by its **customer** than by its **peer**.



Visualize the **incompleteness** of non-local BGP views.

Augmenting AS topology data: IRR

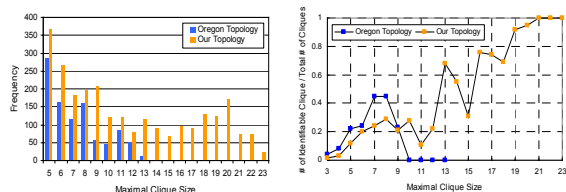
- Internet Routing Registry: coordinate/facilitate inter-AS policy routing.
- Obtaining AS peering relationship info from IRR:
 - Data sanity check: identify (1) **void**, (2) **obsolete**, and (3) **incomplete** records.

Validity analysis of IRR RIPE database

	# of records (ASs)
All records	4,203
- void	3,917
- void - obsolete	1,582
- void - obsolete - complete	1,026

Validating Our Topology

- Do the newly found edges reflect **physical** inter-AS connectivity?
- Methodology: Identify those ASs that are physically co-located at existing **public exchange points (EPs)**.
- Retrieved AS co-location info of 16 European EPs by LG query.



- More than 50% of newly found edges occur at one of the 16 EPs.
- A majority of large maximal cliques in "our topology" consist of physically co-located ASs.

More Representative AS graph?

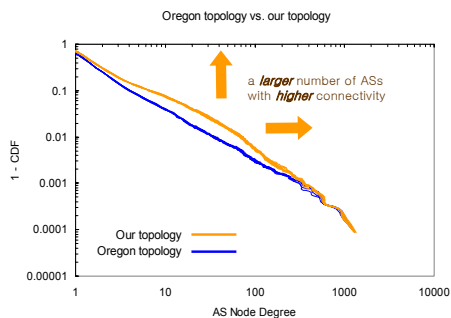
- A more complete "our topology" is much more **denser** than Oregon topology.

AS graph statistics

Source	# of node (%inc)	# of edges (%inc)
Oregon route-views	11,174	23,409
+ RSs	11,268 (0.84%)	26,324 (12.5%)
+ RSs + LG	11,320 (1.3%)	27,899 (19.2%)
+ RSs + LG + RIPE	11,456 (2.5%)	32,759 (40.0%)

"Our topology"

- The degree distribution of "our topology" is still heavy-tailed or highly-variable, but deviates significantly from strict power-law.



- Predict more representative AS Internet topology, based on:

- We know how many ASs exist globally.
- We know the connectivity of tier-1 ASs.
- We know the connectivity of end-customer ASs.
- There are many public EPs worldwide, where ~10 to ~100 ASs are co-located.

- Potentially, the degree distribution will be curved **more outwardly**.

Conclusion

- The completeness issue of BGP-inferred AS Internet topologies.
- Important lesson for network-related measurement studies in general.
- <http://topology.eecs.umich.edu>.