Internet Topology Discovery

Ramesh Govindan

USC/Information Sciences Institute

Joint work with Hongsuda Tangmunarunkit

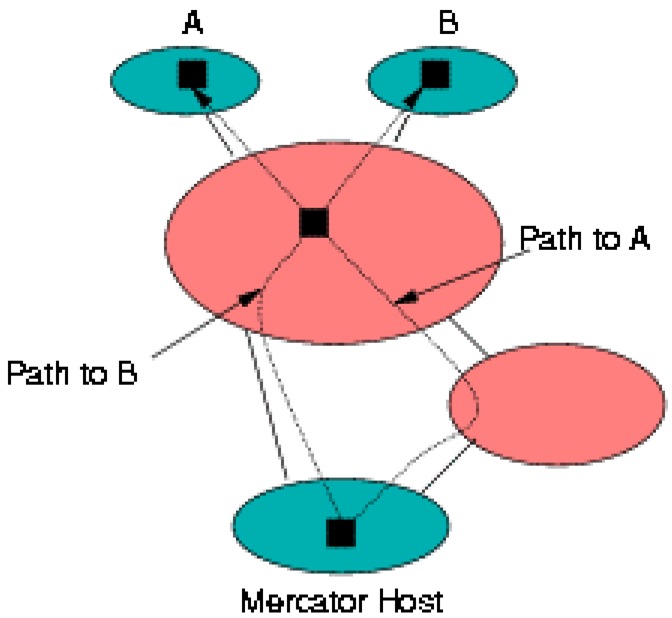
Mapping the Internet

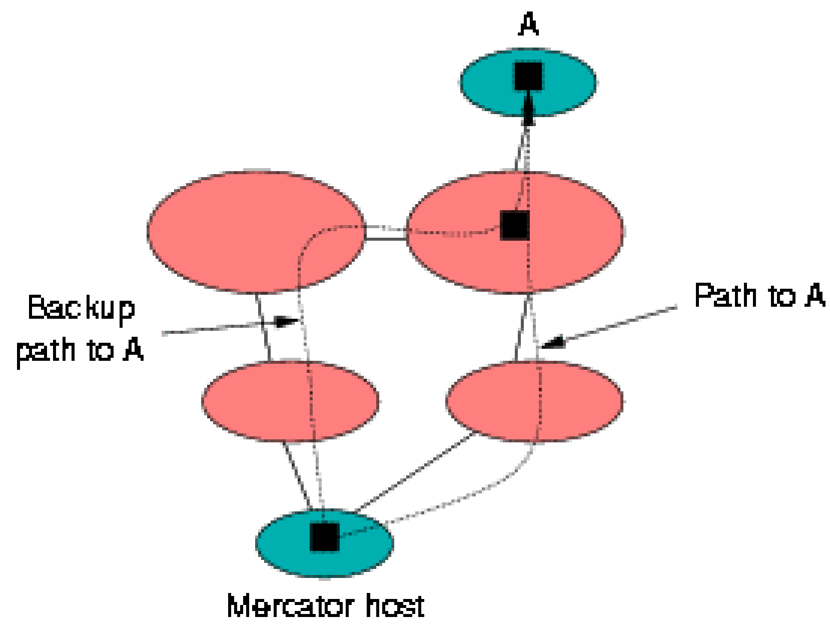
- Motivation
  - context for fault isolation
  - obtain realistic topologies for simulation
  - understand network structure
- Features of Mercator
  - Requires no input, gradually discovers topology
  - Careful probing minimizes overhead
  - Uses only hop-limited UDP probes
- Heuristics
  - Informed Random Probing
    - learn which address to probe from results of probe
  - Source-routed path probing
    - discover some cross-links
  - Alias resolution
    - infer which interfaces belong to same router

Informed Random Probing

- In the absence of external databases
  - guess which portions of the space contain addressable nodes
- Guessing addressable prefixes
  - If an address A is seen in a path probe, assume that some prefix P of A contains addressable nodes
    - hence, *informed* by path probes
  - If a prefix P contains addressable nodes, then some neighboring prefix also contains addressable nodes
- What prefix lengths to use?
  - 8 for class A, 16 for class B and 19 for class C
  - mirrors registry allocation policies

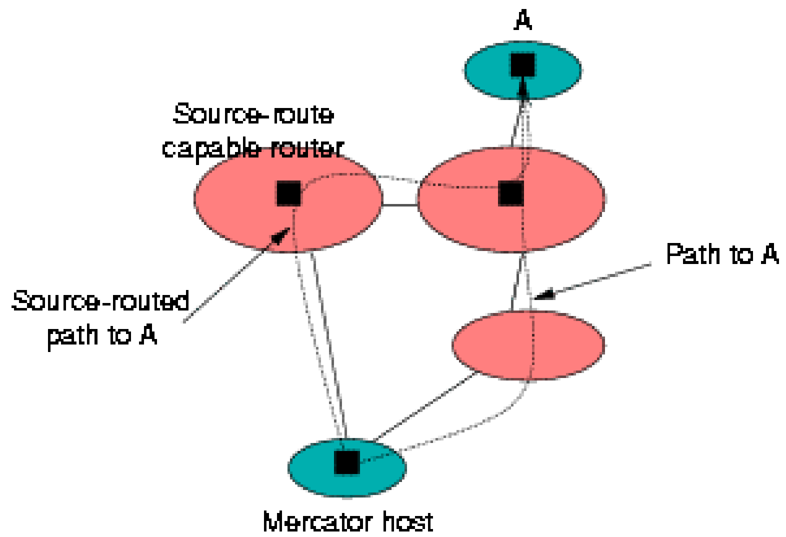
Traceroutes





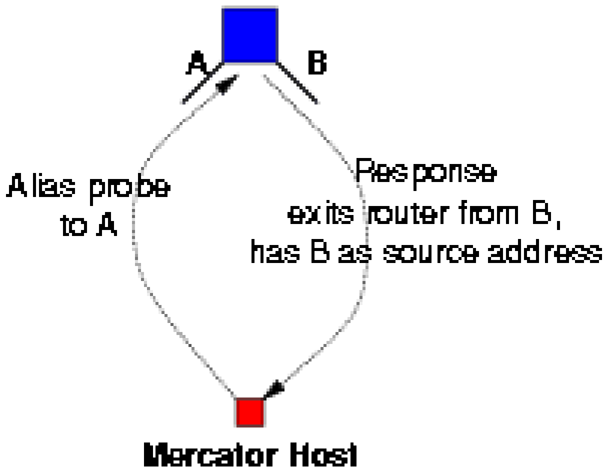
… can provide some richness in topology

Source-routed Traceroutes



Alias Resolution

- For ICMP unreachables
  - most routers use as source address
  - … the address of the outgoing interface towards host
- So, a UDP packet sent to A
  - … can reveal A and B to be aliases for the same router
- Need to do this repeatedly!
  - … because routes can change
  - only a small number of such probes outstanding at any instant



What the Map is Not

- Not a topology map
  - does not discover all interfaces of a router
  - cannot distinguish between shared media and point-to-point media
- Not an instantaneous map of the Internet
  - a time-averaged picture of the *routed* topology
- Not a complete map
  - does not capture detail in stubs
  - perhaps captures a fair bit of the transit portions of the network

Experiences

- Misconfiguration
  - class D address as source
  - traceroutes to net 10
- Oddity
  - router with more than 1800 neighbors!
- 10% of routers are source-route capable

- About 1 in 6 links are discovered by source routing
- Source routing not entirely bug free
  - missing hops in source-routed traceroutes
  - at least one router which ignores the source route option and processes packet
  - older versions of router code may not decrement TTLs

Experiences

- About 15% of interfaces were not reachable
- Reasons
  - numbered out of private space
  - routes not propagated beyond ISP border
  - routes not widely propagated
    - source-routed alias resolution works, but is slow
- About 100 abuse complaints over a six month period
- Mostly from firewall operators with logging enabled
  - financial institutions
  - the DoD
- International representation
  - few from Europe, one each from South Africa, Australia
- Only one major ISP
  - and only because the probes reached their corporate offices…
- Mostly supportive, for others we use exclusion lists

Social Experiences

Admin: Stop probing my network

Me: My program randomly probes the address space.
Can you tell me your address prefixes?

Admin (sounding officious): This is **classified** information

... Some more similar exchanges...

... Tiring of these exchanges, it takes me all of 2 minutes to find out the address prefixes from the RADB

Impact of Policy on Internet Paths

Joint work with Hongsuda Tangmunarunkit, Scott Shenker and Deborah Estrin

Does Policy Skew Paths?

- Observation
  - Can approximately overlay AS structure on router-level map
  - With a simplified policy model, can compare shortest router-level path with policy path
- Approximate because
  - … we don’t have all interfaces of router
  - … of inconsistent interface labeling
  - … of incompletely populated DNS
  - … simplistic policy assumptions

AS Overlay Methodology

- Internet map
  - Collected between March 26 – April 10
  - 102,639 nodes, 142,303 links
  - 61,485 traceroute paths
- AS overlay
  - Use BGP routing table and RADB data to determine domain/AS# for each node
  - 3210 private addresses → assigned unused AS#
  - 497 unresolved addresses → assigned unused AS#

AS Overlay Methodology

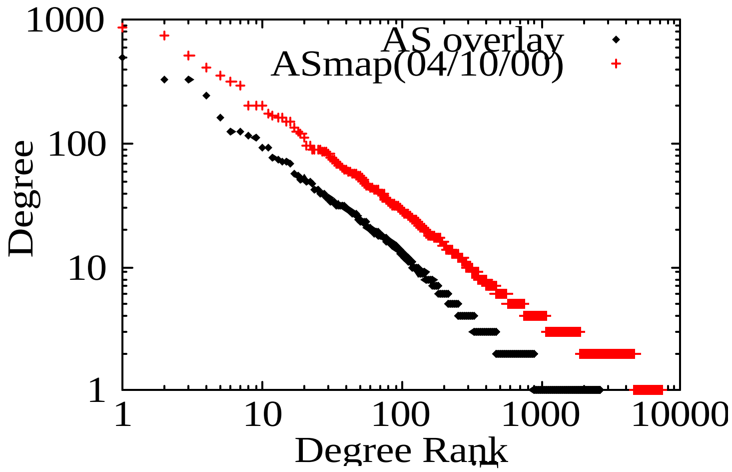
- Collapse nodes into domains
  - Recursively marks neighbors with the same AS# to the same AS domain
  - Due to incomplete Internet map and AS# assignment, there are many disjoint clusters of nodes belonging to the same AS domain
  - Most AS's have one big component and many small clusters
  - Identify the biggest component for each domain
  - Reassign the rest (20,000 nodes) to the nearest AS

AS Overlay Validation

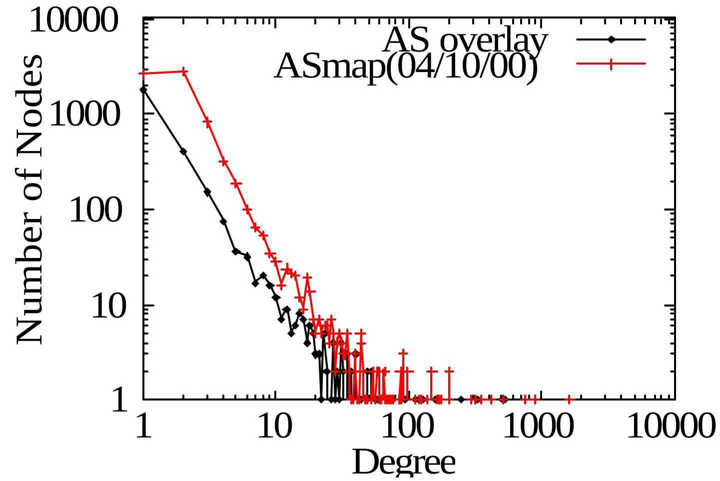
- Comparison of macroscopic properties
  - Degree Rank Distribution
  - Degree Distribution
  - Average Hop-pair Distribution
- … between AS overlay and actual AS map

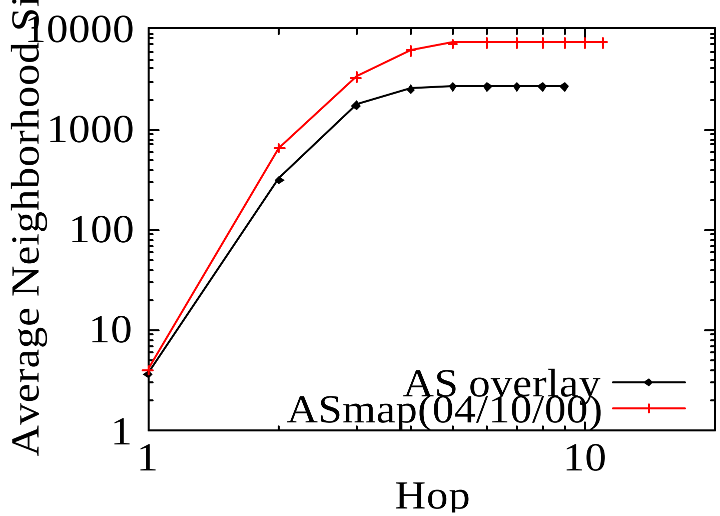
Qualitatively Similar Properties

Degree Rank Distribution



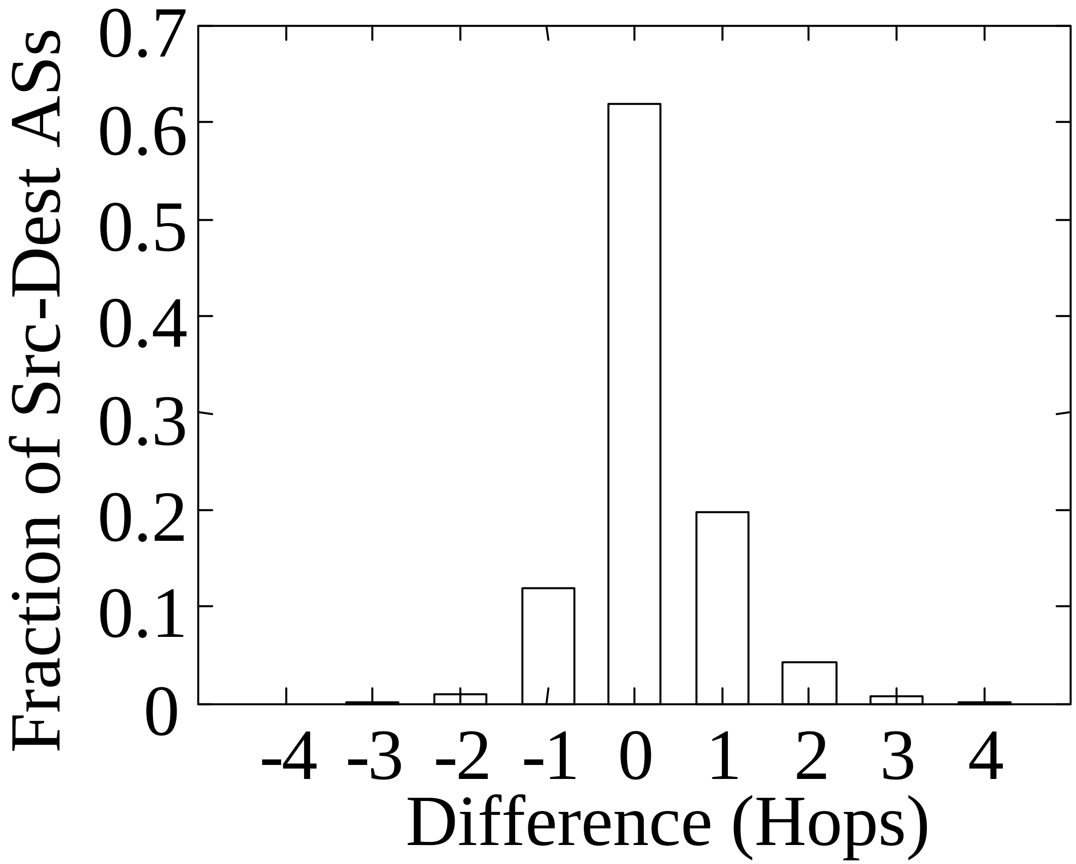
Degree Distribution





Avg Hop-pair Distribution

Validation: Comparing AS Path Lengths



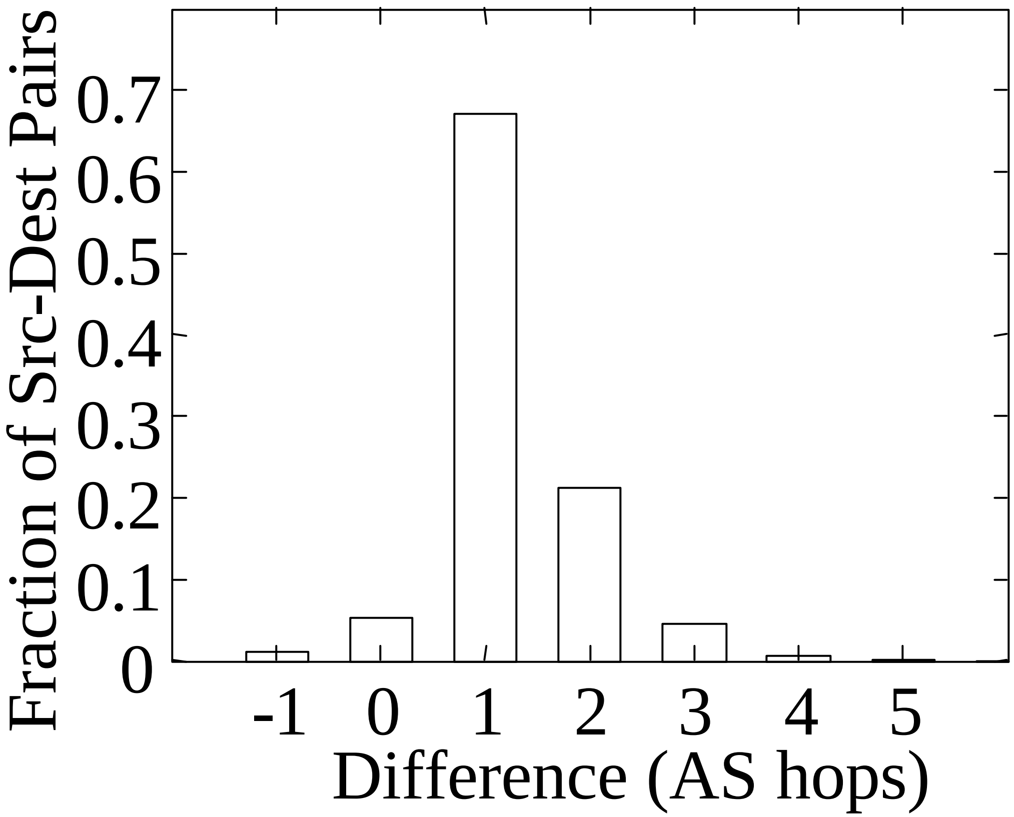
Distribution of path length difference

- Compare AS path length in table with path length in AS overlay
- 60% of node pairs have the same length
- 93% of node pairs are within one AS hop of the corresponding path in the BGP table

Simple Policy Model

- Routing policy model: Shortest AS Path
- Actual routing policy
  - Select customer routes only
  - Use shortest AS path otherwise
- Validating the model
  - Use our traceroutes to
    - Compare differences between AS path corresponding to traceroute, and shortest AS path on overlay

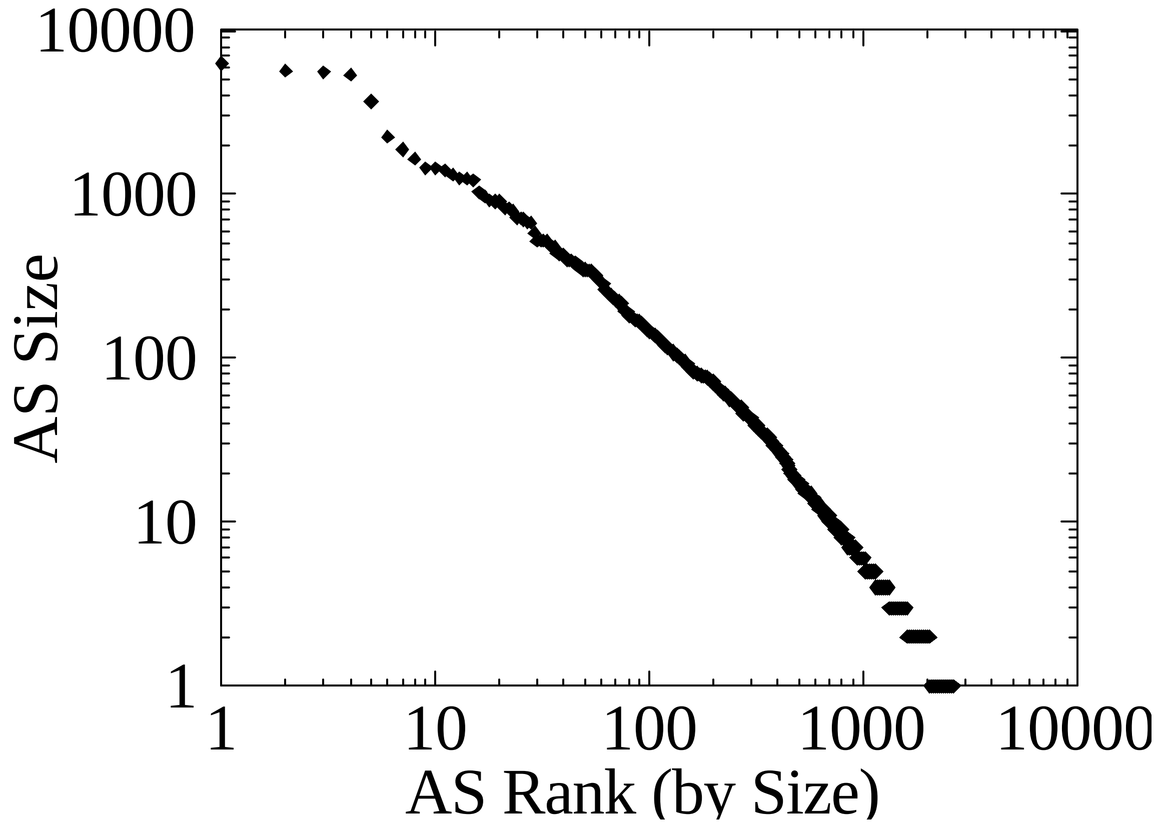
Policy Model Validation



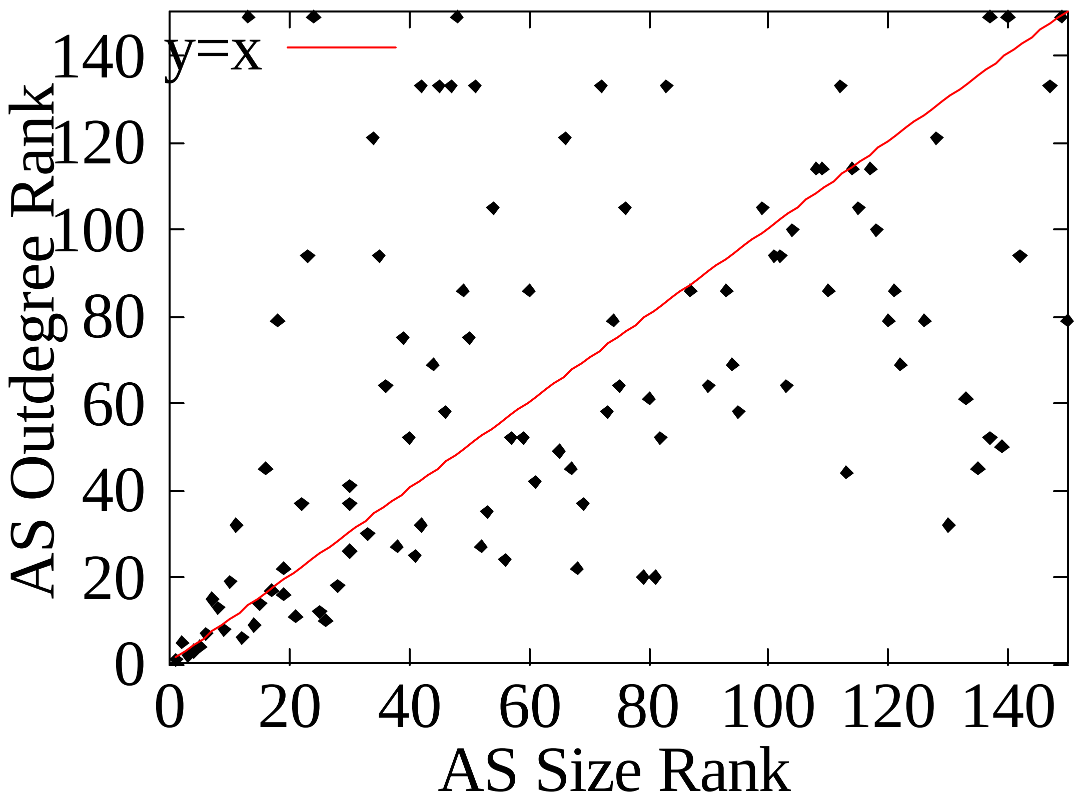
AS path length difference distribution

- Policy model under-estimate traceroute AS path by
  - <= 1 AS hop for 70%
  - <= 2 AS hops for 95%
- Many traceroute AS paths are relatively small

Power Laws: ISP Size Distribution



Degree-Size Correlation?



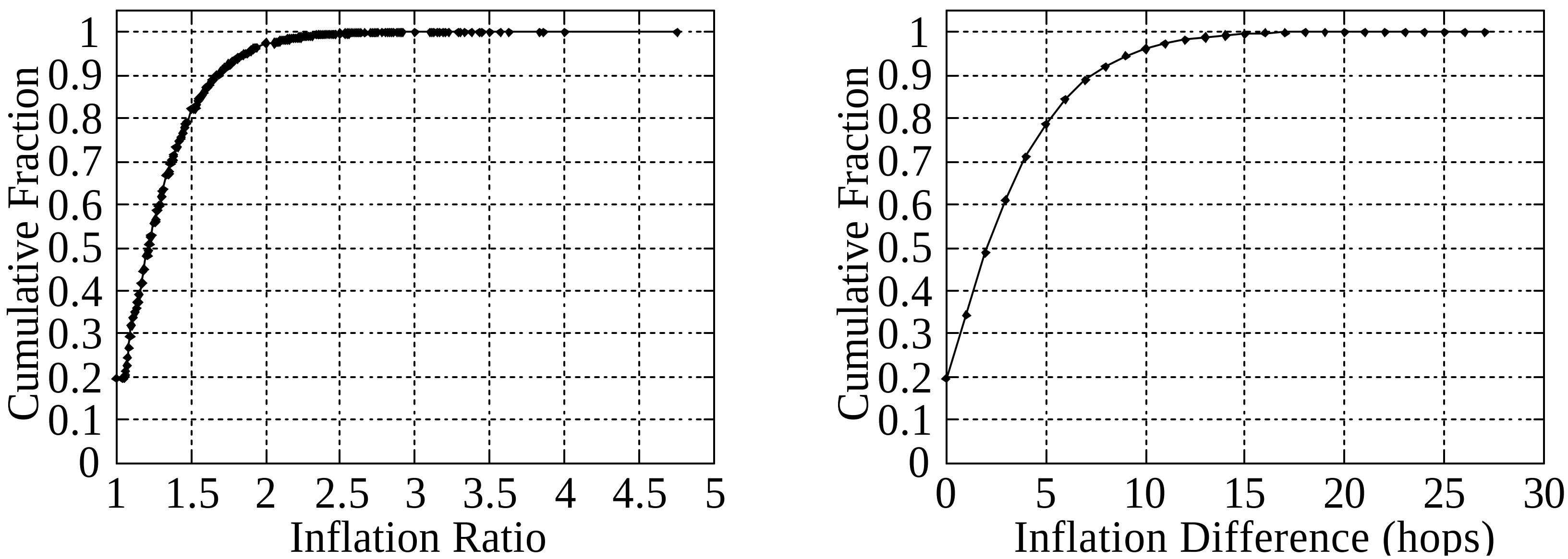
Impact of Policy

- Unicast path inflation
- Detour paths
- Path concentration
- Policy impact on mcast

Path Length Inflation

- How different are policy paths from shortest paths on the graph?
- Metrics
  - Inflation Ratio
  - Inflation Difference

Path Inflation



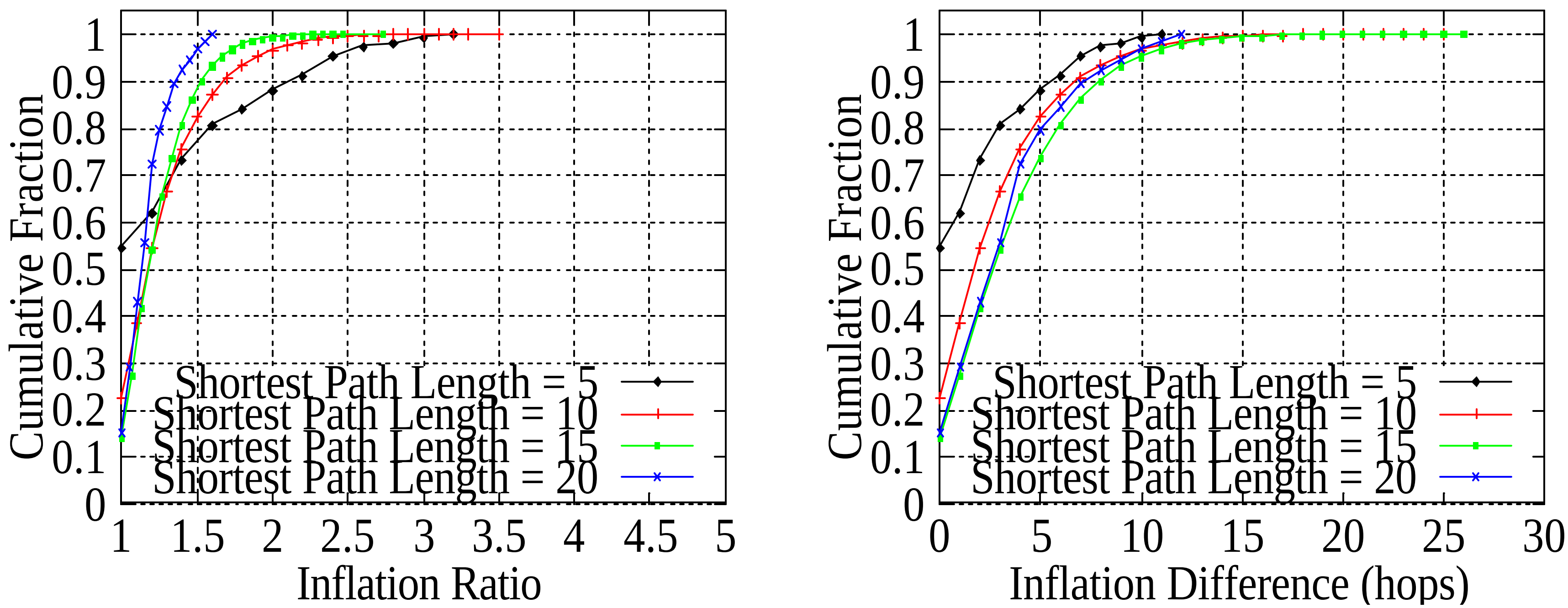
Cumulative distribution of inflation ratio and inflation difference

- 20% of policy paths are inflated by >= 50%
- 20% of policy paths are inflated by >= 5 hops

Inflation by Path Length

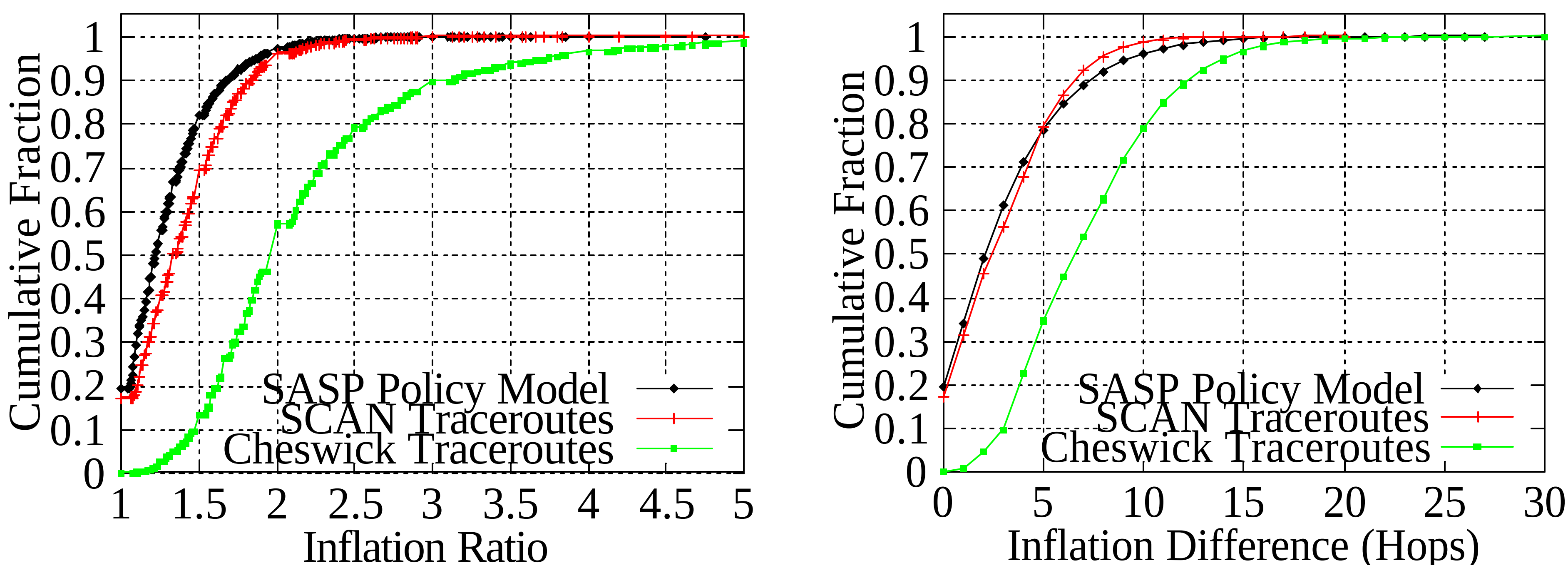
- Widely different path lengths
  - Short: 5 hops
  - Medium : 10 and 15 hops
  - Long: 20 hops
- Compared to shorter paths, longer paths are
  - Less inflated in proportion to their lengths
  - More inflated in the absolute path length

Inflation by Path Length



Cumulative distribution of inflation ratio/difference by path length

Policy Model v.s. Actual Paths

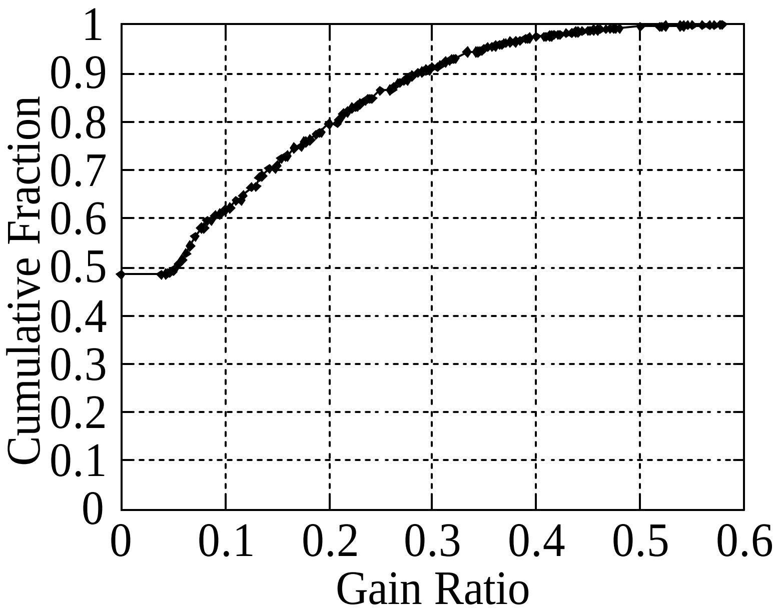


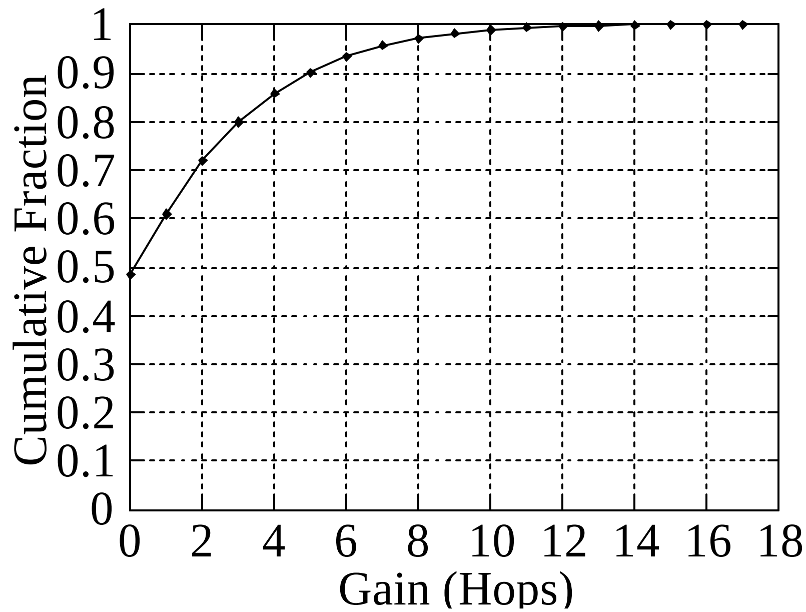
- Our routing policy model is conservative

Detour Paths

- Between nodes A and B, a detour I exists if there is an intermediate node I such that
  - I lies in a different AS than A and B
  - The AS path via I is longer than the AS path between A and B
  - The policy path via I is shorter than the policy path between A and B
- If many detours exist, pick shortest detour
- Metrics:
  - Gain: difference between policy path and path via detour
  - Gain ratio

Detour Gain/Gain Ratio



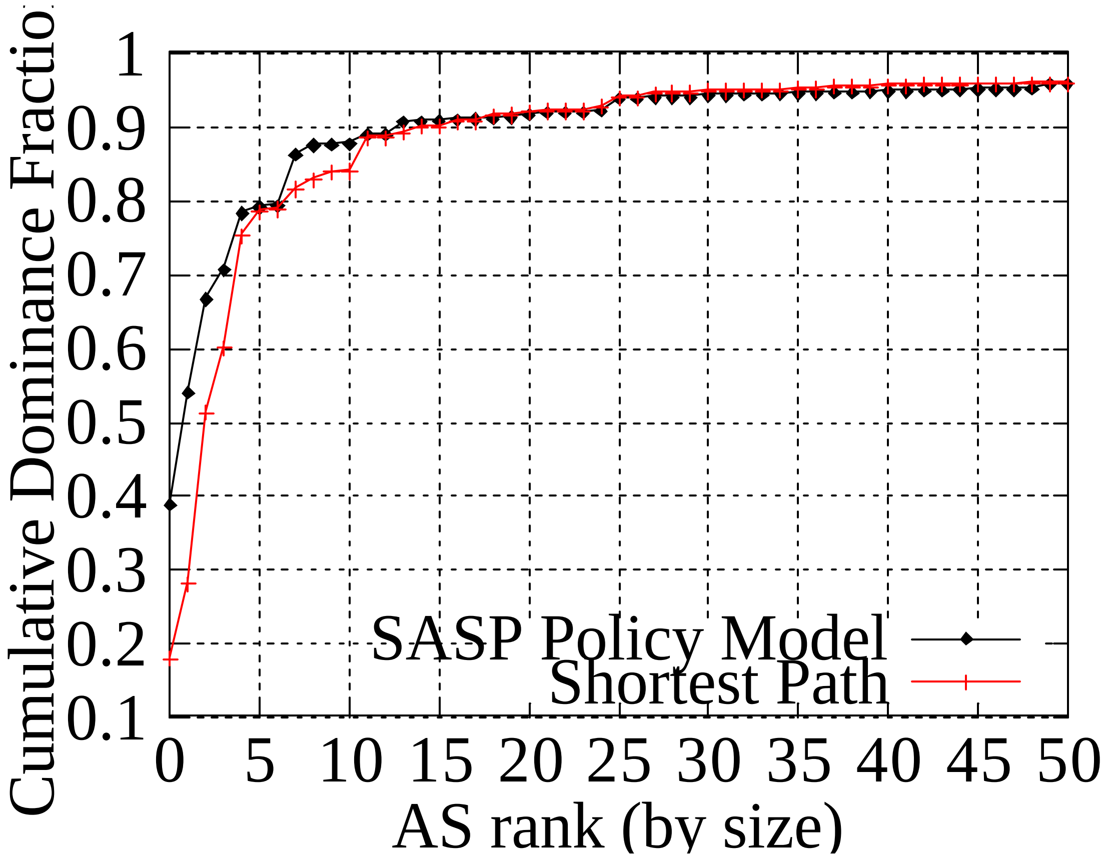


About 50% of paths have a detour

Path Concentration

- Does routing policy force Internet paths through large AS's?
- More general Question: Is the Internet connectivity rich enough that logical connectivity imposed by policy skews the paths?
- Dominant AS for (A,B): the largest **transit** AS encountered in the path between A and B
  - Ignore all nodes pairs within a single AS, or in adjacent AS's
- Dominance Fraction of AS X: fraction of node pairs for which X is the dominant AS

Path Concentration



http://www.isi.edu/div7/scan