On Third-party Addresses in Traceroute Paths

Young Hyun, Andre Broido, kc claffy
CAIDA, San Diego Supercomputer Center
University of California, San Diego
{**youngh, broido, kc**}**@caida.org**

PAM2003, La Jolla, April 2003

Motivation

Background

- data collection
- broader question: BGP and traceroute incongruity
- 3rd-party addresses

Methodology

Analysis

- frequency of 3rd-party addresses
- distribution relative to path beginning
- distribution relative to path end
- multihoming

Conclusions

Motivation

- AS-level Internet topology is very useful
  - for studying growth, performance, resiliency, convergence times
  - for supporting design of routing protocols
- complete, up-to-date topology not available
- sources of partial topology
  - ask ISPs about their peering relationships
  - routing registries (e.g., RADB)
  - BGP tables at RouteViews and RIPE
  - AS paths derived from traceroute paths

Traceroute advantages

- does not require ISP involvement
- can see stub networks

Traceroute disadvantages

- limited vantage point $\Rightarrow$ miss lateral peering
- miss backup links
- IP paths sometimes inaccurate
- IP-to-AS mapping sometimes inaccurate

Two kinds of inaccuracies

1. actual vs. intended routing
   - study misconfigurations, circuitous paths, etc.
2. *actual vs. observed routing*
   - study precision, completeness, and truthfulness of observation tools/methods
   - traceroute is an observation tool
   - the kind of inaccuracy covered in this talk

Actual vs. observed inaccuracies in traceroute paths

- gaps in measurement—some hops unresponsive
  - packet filtering by firewalls
  - ICMP rate limiting (ICMP processing is on slow path)
- routers with private addresses: RFC1918, multicast, loop-back
- non-atomic snapshot of path
  - caused by normally occurring routing changes during the several seconds needed to trace full path
  - caused by load balancing
    - probe packets take alternate paths
    - resulting path shows interleaving of alternate paths
- not see forward path—3rd-party addresses

Background

Source of traceroute paths

- CAIDA's skitter monitors
  - around two dozen monitors deployed worldwide
  - technique based on TTL, like `traceroute` (but use ICMP `ECHO_REQUEST`)
  - predetermined set of addresses ('destination list')
  - one pass through destination list = one cycle
  - IP path + RTT for one destination = one trace

skitter monitors used

monitor	location	network
a-root	Herndon, VA	Verisign
k-peer	Amsterdam	RIPE, near AMS-IX
m-root	Tokyo	WIDE, near NSPIXP
champagne	Urbana, IL	U. of Illinois at Urbana-Champaign
lhr	London	MFN/AboveNet
sjc	San Jose, CA	MFN/AboveNet

Destination lists used

- DNS (200K responding): clients of DNS root servers
  - `a-root`, `k-peer`, `m-root`
- IPv4 (80K responding): broad cross-section of Internet hosts
  - web servers, backbone routers, business desktops, consumer dial-up/broadband desktops
  - `champagne`, `lhr`, `sjc`
- DNS and IPv4 lists have 8K responding dests in common

Stable paths

- for meaningful analysis, we must ensure each traceroute path reflects a single path
- therefore, only use *complete traces*—destination and all intermediate hops responded
- furthermore, only use *stable paths*—remained the same in 3 consecutive cycles
- all numbers about paths in this talk are in terms of stable paths
- 3rd-party addresses are not routing anomalies $\Rightarrow$ not excluded by use of stable paths

Data collected

- 3 consecutive cycles on Jan 10–13, 2003
- avg. 79% paths complete
- avg. 52% paths stable
- avg. # stable paths: DNS = 108K, IPv4 = 40K
- avg. coverage of BGP prefixes (127K) by stable paths: DNS = 17%, IPv4 = 20%

Broader question: BGP and traceroute incongruity

- simplistically: BGP AS path = specified, traceroute AS path = actual
- how do they differ?
- studied in prior paper “Traceroute and BGP AS Path Incongruities”
- causes of incongruity:
  - insertion of exchange point (IX) ASes
  - insertion of ASes under the same ownership
  - unidentified causes (e.g., 3rd-party addresses)
- next few slides summarize these results

Summary: exchange point ASes

- prefixes with origin AS owned by exchange point
  - AS6695 DE-CIX, AS5459 LINX, AS1200 AMS-IX, etc.
- many BGP and traceroute paths differ only by IX ASes

cause of incongruity	sjc		k-peer		m-root	
only IX ASes	3,749	(33%)	30,163	(82%)	20,601	(54%)
only non-IX ASes	6,818	(60%)	4	(0%)	6,759	(18%)
both IX & non-IX ASes	712	(6%)	6,721	(18%)	11,100	(29%)
total incongruent paths	11,279	(100%)	36,888	(100%)	38,460	(100%)

Summary: ASes under same ownership

- many organizations have more than one AS
  - after a merger or acquisition
  - for convenience in implementing routing policy, such as segregating . . .
    - academic vs. commercial traffic
    - transit vs. customer traffic
    - regional vs. national vs. international traffic
- some closely related organizations
  - MCI/WorldCom/UUNET/AlterNet/ANS/Bertelsmanns
  - SBC/Pacific Bell/Nevada Bell/Southwestern Bell
  - C&W/Exodus/PSI
  - Qwest/US West/SuperNet/Touch America
  - Cogent/PSINet/NetRail
- rework analysis questions: e.g., not “peering between ASes” but “peering between *organizations*”

Summary: Unidentified causes of incongruity

- performed textual comparison of BGP and traceroute AS paths in terms of editing operations
- for example: (a) delete 11422, (b) insert 1

```
BGP           207.99.128.0/17    6461 209 11422 2151   2920
Traceroute    207.99.161.1       6461 209       2151 1 2920
```

- most traceroute AS paths longer than corresponding BGP AS paths

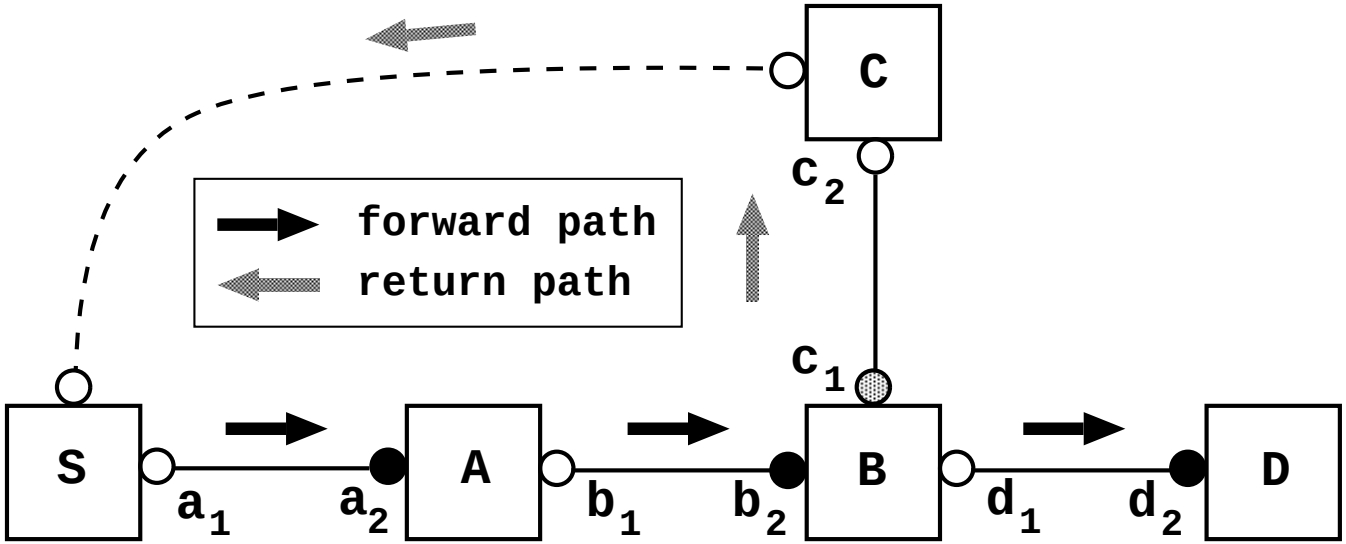
sjc			k-peer			m-root		
+	3,125	(65%)	+	3,673	(70%)	+	15,765	(93%)
−	1,220	(25%)	−	103	(2%)	−	36	(0%)
total	4,819		total	5,216		total	16,927	

- usually insertions of ASes in traceroute paths

Operation	sjc		k-peer		m-root	
insertions only	2,788	(58%)	2,764	(53%)	13,661	(81%)
deletions only	1,132	(23%)	1	(0%)	0	(0%)
substitutions only	813	(17%)	1,813	(34%)	2,648	(16%)
mixture	86	(2%)	683	(13%)	618	(4%)
total paths	4,819	(100%)	5,216	(100%)	16,927	(100%)

What are 3rd-party addresses?

- possible cause of some incongruities between BGP and traceroute paths
- addresses in *return* path, not forward path
- RFC1812: Set source address of ICMP response packet to address of *outgoing* interface.



expect IP path expect AS path	$a_2\ b_2\ d_2$ $A\ B\ D$
get IP path get AS path	$a_2\ c_1\ d_2$ $A\ C\ D$

Methodology

Definitions

intermediate address: address other than the source or destination in IP path

adjacent address: address appearing before or after a given address in IP path

candidate 3rd-party address ('candidate address'): intermediate address that maps to a different AS than adjacent addresses

- candidate addresses are locations where derived AS path may be wrong
- only candidate addresses (subset of all 3rd-party addresses) studied

Procedure

1. perform traceroutes from multiple locations
2. reduce to stable paths
3. derive AS paths
4. identify initial candidate addresses
5. refine set of candidate addresses

Analysis

Frequency of candidate addresses

1. by # unique candidate addresses—suggests # locations in network with phenomenon
2. by # paths having candidate addresses—suggests impact of phenomenon

Frequency of candidate addresses

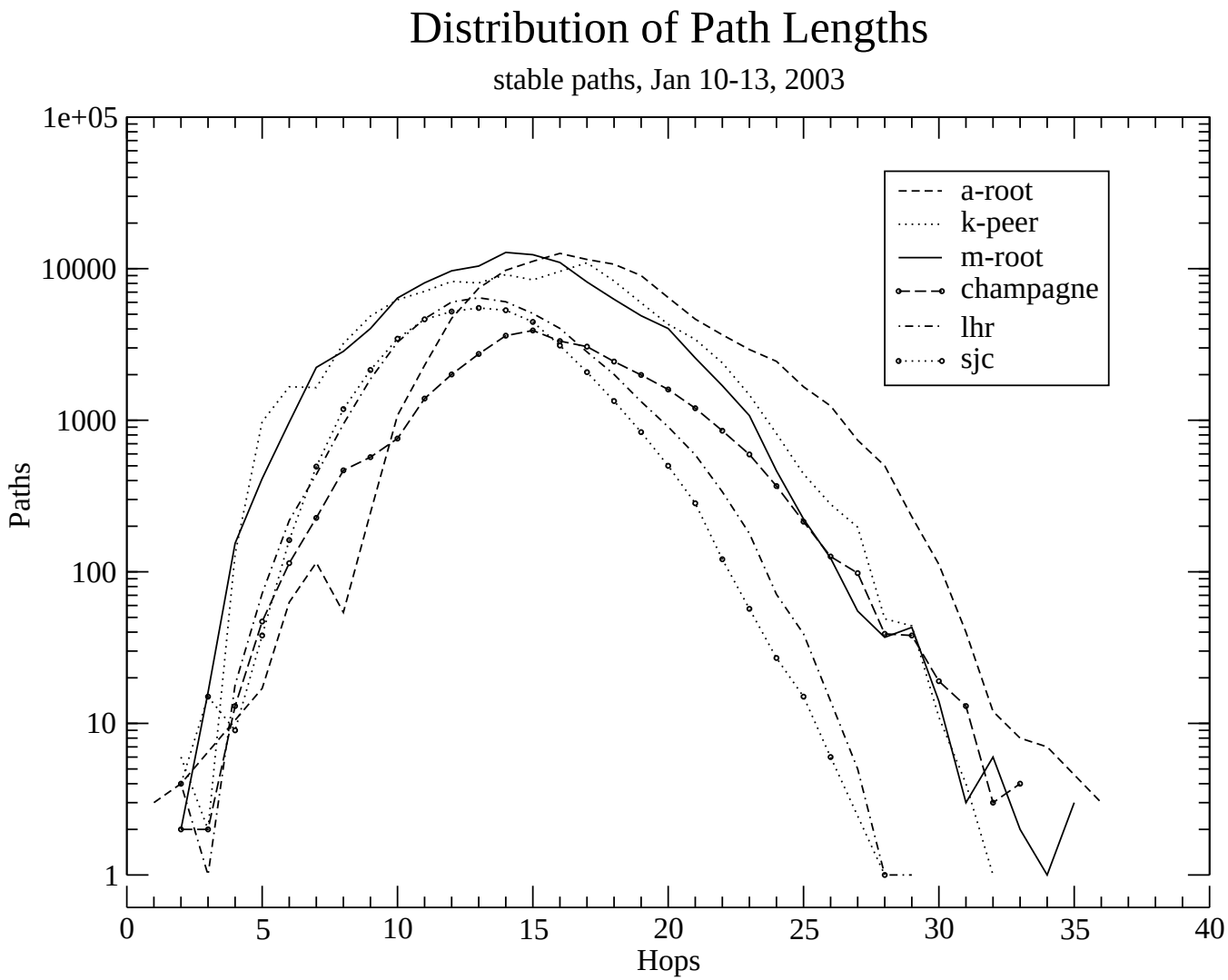
	candidate addrs	% of intermediates	paths with candidates	% of stable paths
a-root	1,617	1.4%	8,266	7.8%
k-peer	1,407	1.3%	6,253	5.8%
m-root	1,482	1.3%	39,479	35.6%
champagne	1,145	2.9%	3,337	10.5%
lhr	1,414	2.6%	3,800	8.0%
sjc	1,202	2.4%	3,222	7.9%
m-root (-top3)			8,233	7.4%

- frequency not negligible but generally low
- typically a small # of candidate addresses responsible for bulk of appearances
- m-root path counts inflated by 3 candidate addresses occurring within 2-3 hops of monitor
- m-root path counts without top 3 candidate addresses more in line with others

Distribution of candidate addresses by hop distance

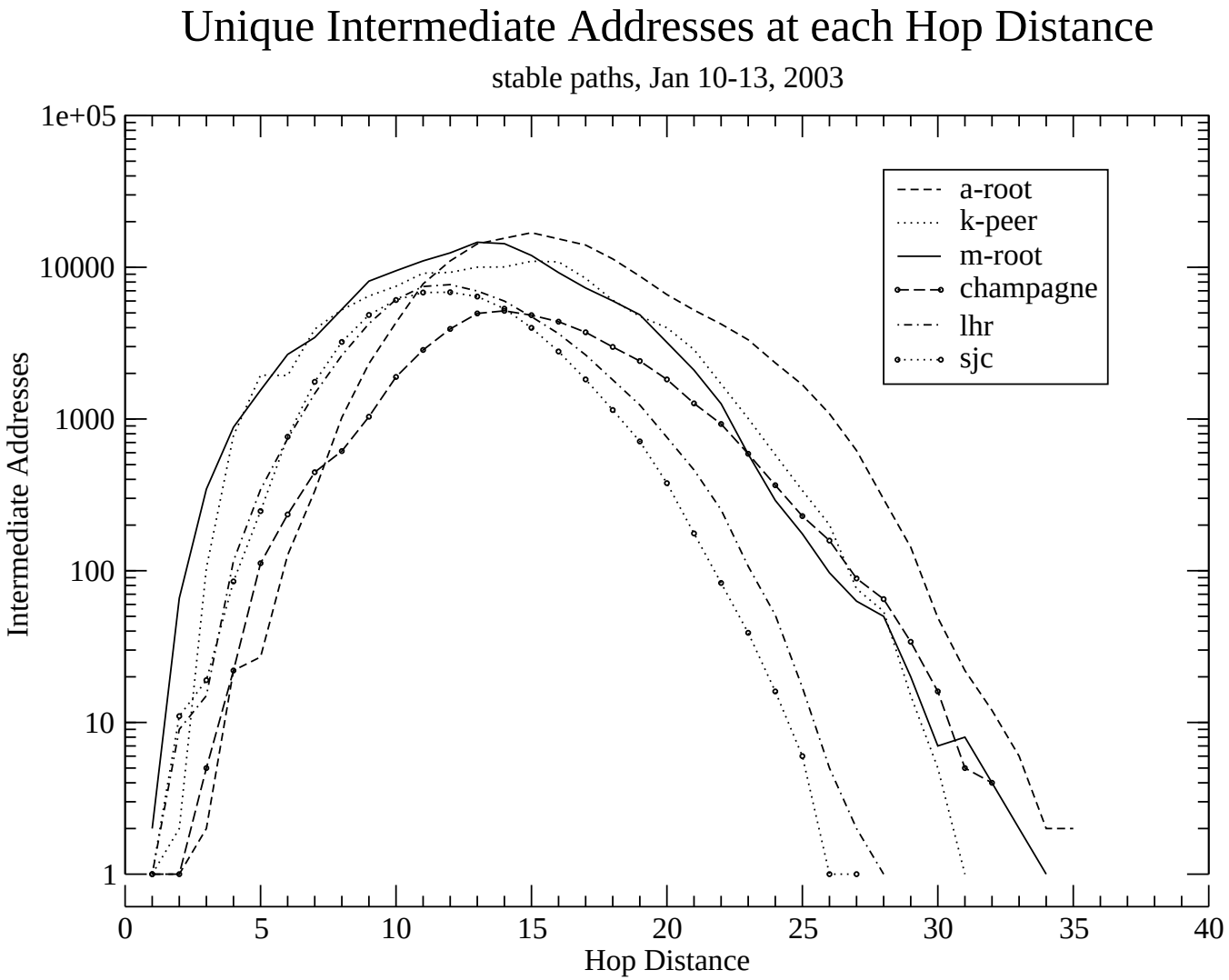
- hop distance relative to (1) beginning and (2) end of paths
  - in path $S\ I_1\ I_2\ I_3\ I_4\ D$: I_2 is 2nd hop from beginning, 3rd from end
- # unique candidate addresses vs. total # intermediate addresses at each hop distance
- try to answer: what is contribution of candidate addresses to total variation in intermediate addresses at each hop?

Relative to path beginning



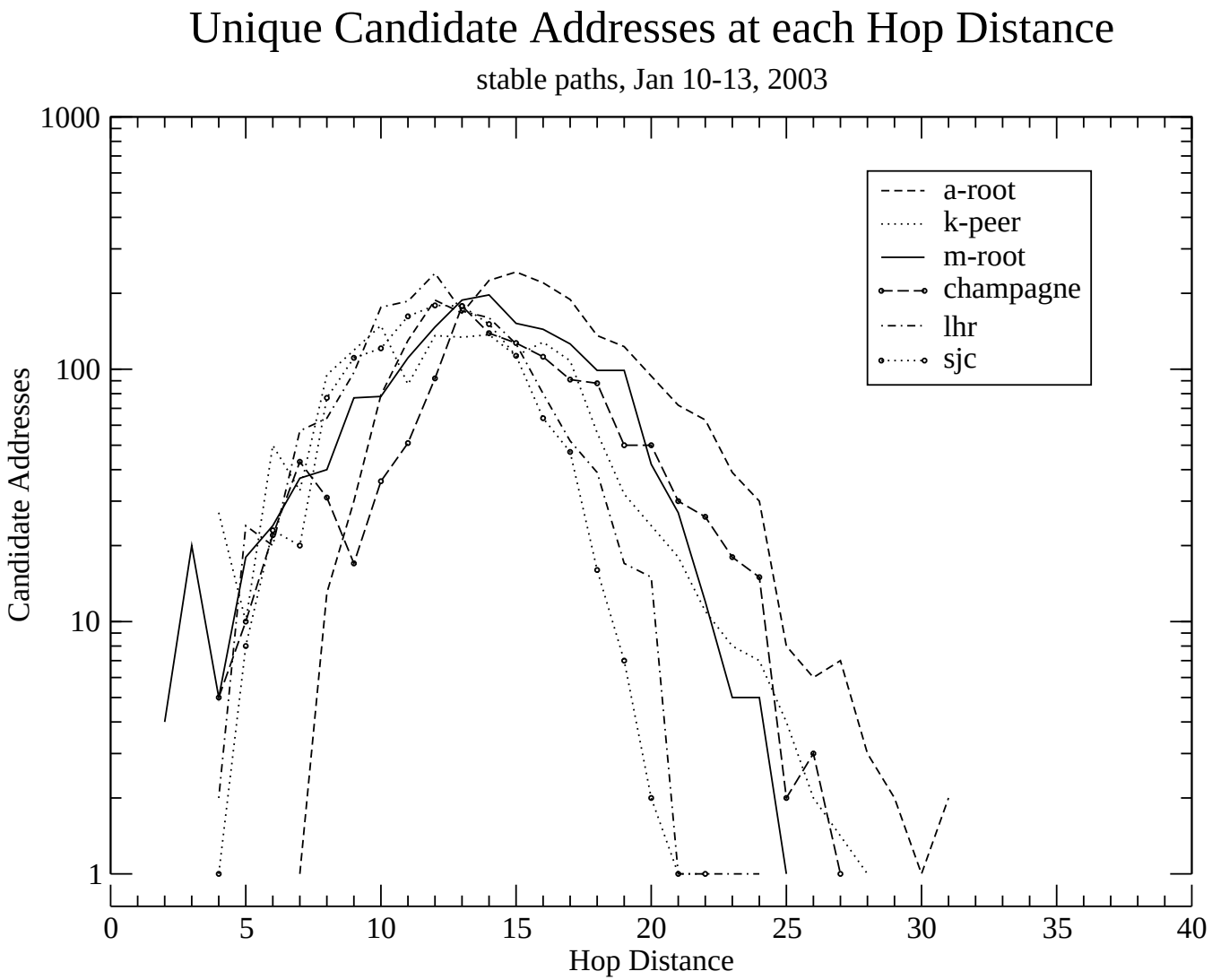
- typical distribution of Internet path lengths
- bell-curve shape centered around 15–16

Relative to path beginning, cont'd



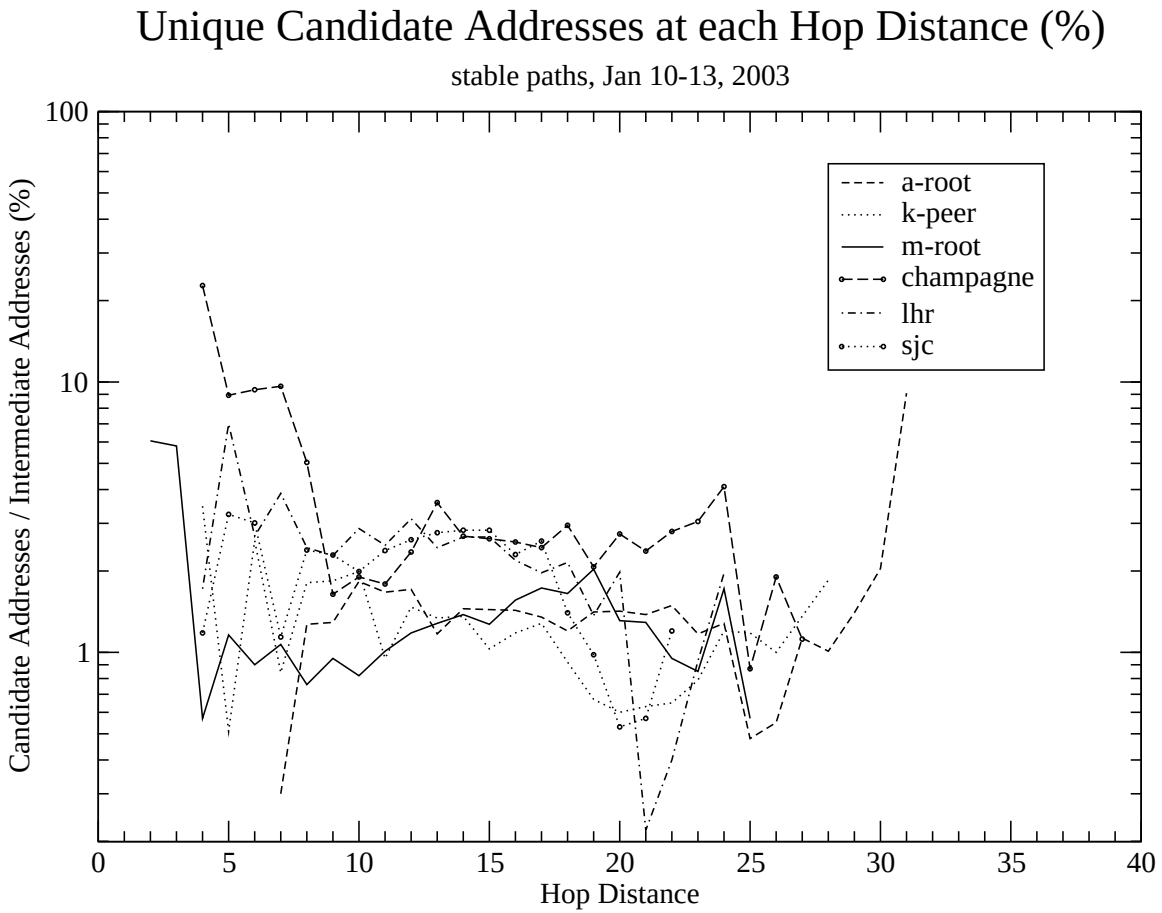
- nearly identical to path-length distribution

Relative to path beginning, cont'd



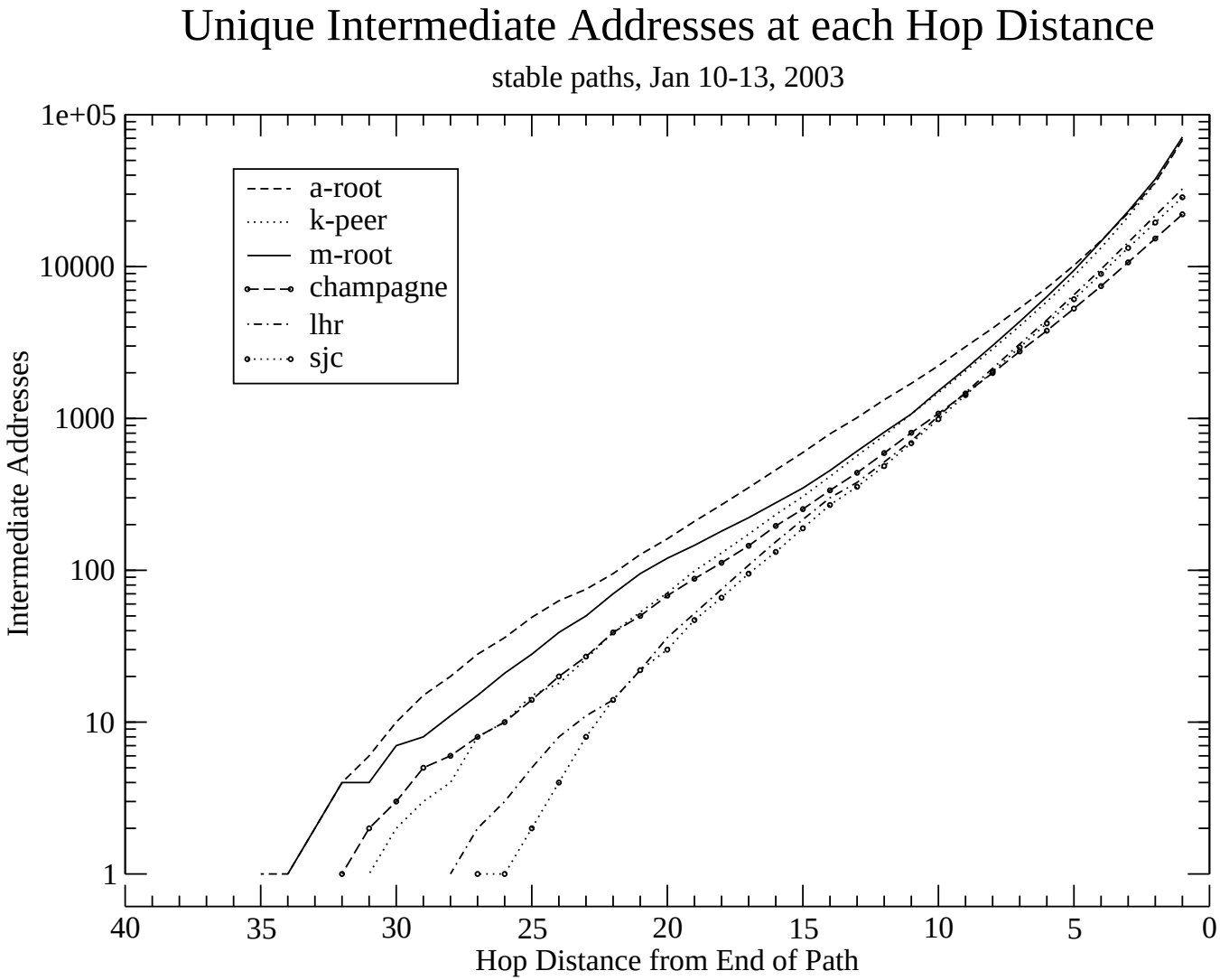
- similar shape as distribution of all intermediate addresses, only smaller in scale

Relative to path beginning, cont'd



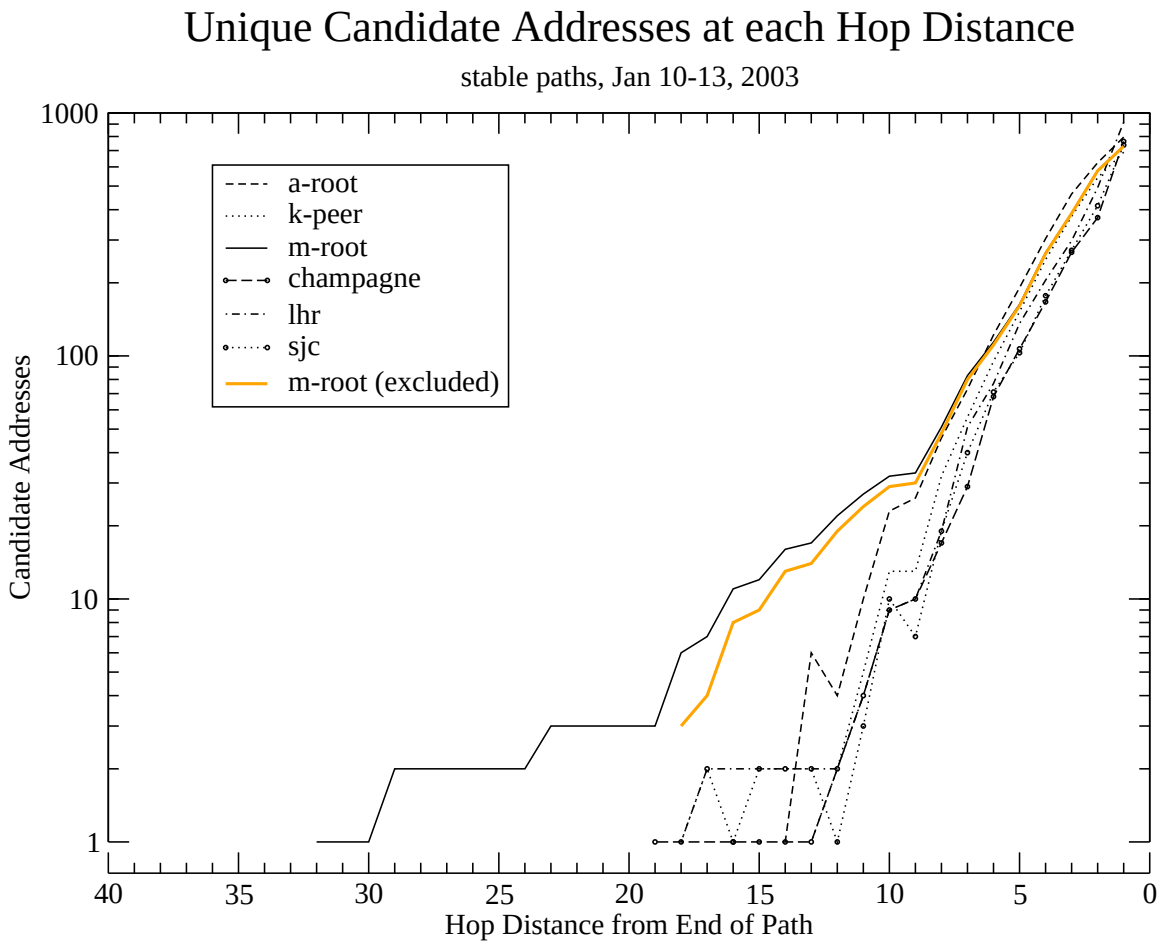
- distribution generally flat (between 1–3%)
- noise at lower and upper hops attributable to small # of intermediate addresses (e.g., 22 at 31st hop for a-root)
- no dependence between frequency of candidate addresses and hop distance from source

Relative to path end



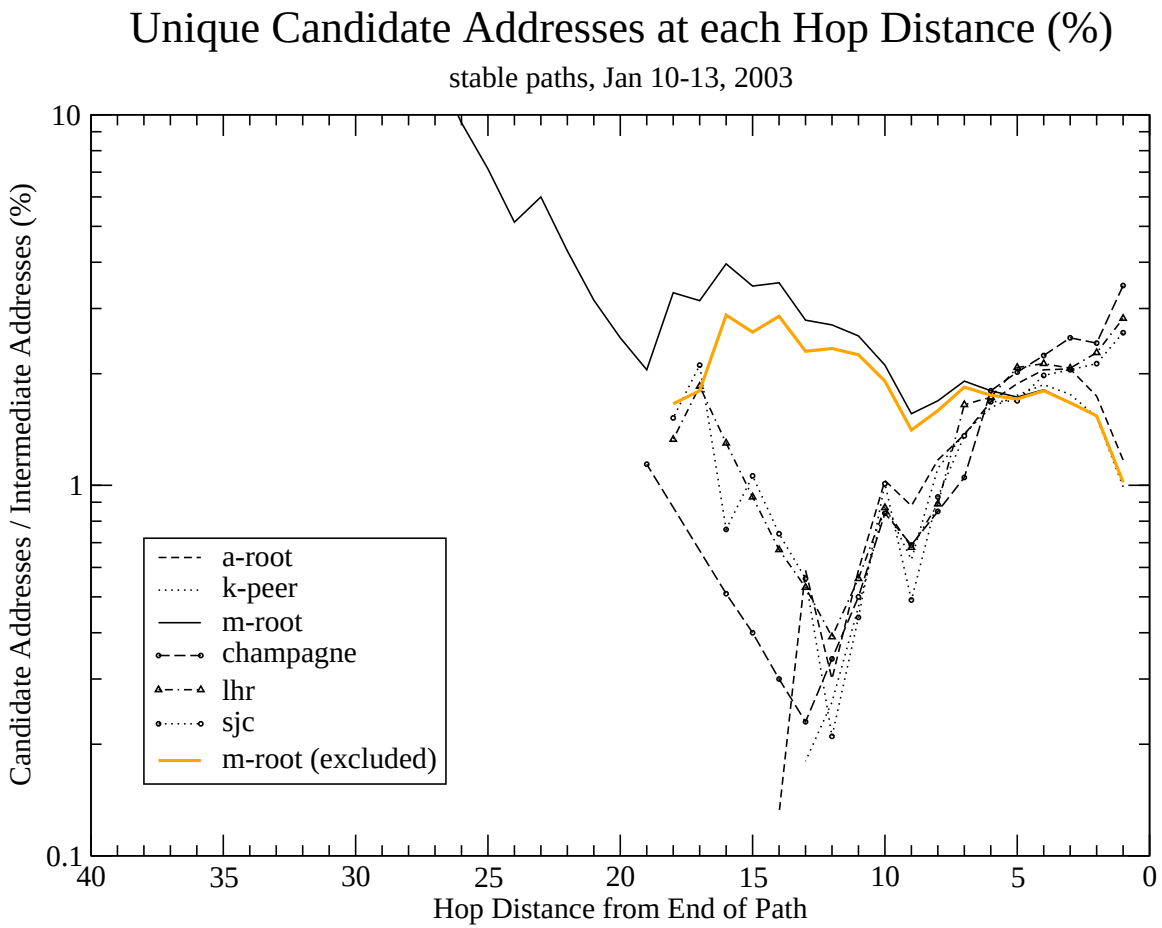
- exponential shape implies our set of stable paths is approximately a tree

Relative to path end, cont'd



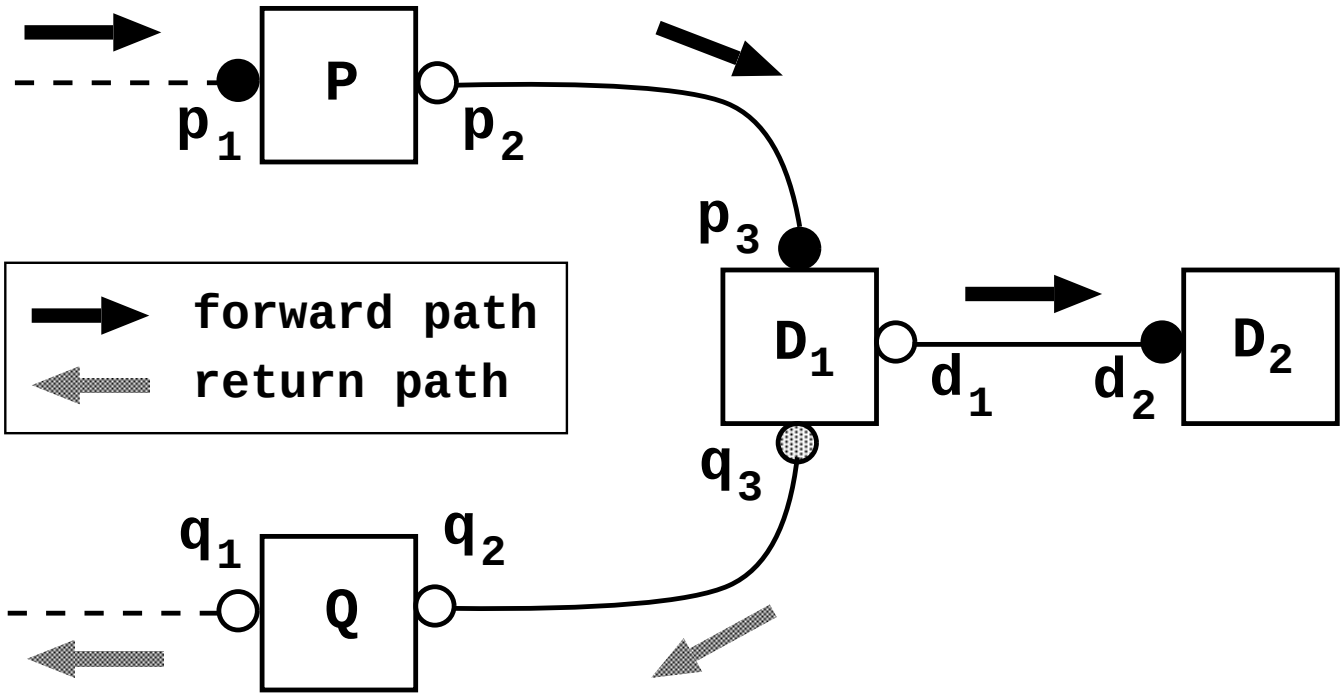
- exponential in last dozen hops
- $\frac{3}{4}$ of candidate addresses occur within last 3-4 hops
- extended tail of m-root caused by frequently occurring candidate addresses near path beginning

Relative to path end, cont'd



- percentage increases by order of magnitude (0.3 to 3%) over last 14 hops
- implies higher probability of occurrence near end of path

Multihoming



	possible path 1	possible path 2
expect IP path expect AS path	$p_1\ p_3\ d_2$ $P\ D$	$q_1\ q_3\ d_2$ $Q\ D$
get IP path get AS path	$p_1\ q_3\ d_2$ $P\ Q\ D$	$q_1\ p_3\ d_2$ $Q\ P\ D$

- ASes of *both* providers appear in single path!

Conclusions

- 3rd-party addresses that cause incorrect AS paths relatively uncommon
- tend to occur near the destination
- multihoming can lead to 3rd-party addresses
- impact on AS-level analysis probably small

Resources

- “Traceroute and BGP AS Path Incongruities” <www.caida.org/outreach/papers/2003/ASP/>