Importance of IP Alias Resolution in Sampling Internet Topologies

Mehmet H. Gunes and Kamil Sarac

Global Internet Symposium 2007, Anchorage, AK

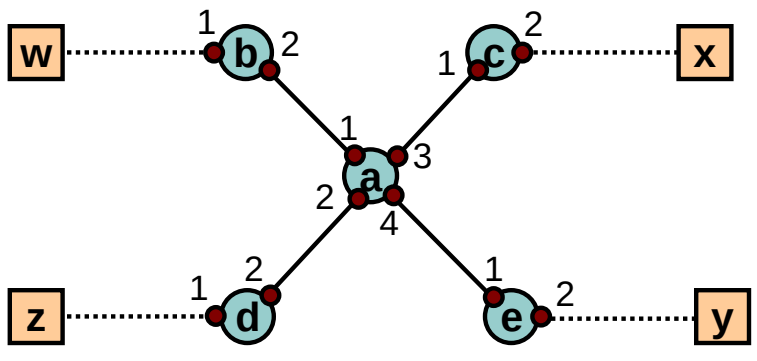
Introduction:
Internet Mapping

- Topology measurement studies require representative Internet maps.
  - AS-level, POP-level, Router-level, etc.
- Such maps are not publicly available
  - ISPs consider their network topology information as confidential.
- *Topology sampling* (at router-level)
  - Probe the Internet (using *traceroute) to collect* path traces.
- *Map construction*
  - Verify the accuracy of the collected path traces,
  - Resolve anonymous routers,
  - **Resolve IP aliases.**

Overview

- What:
  - Impact of alias resolution on *traceroute*-based sample network topologies.
- Why:
  - Whether the alias resolution distorts the results significantly.
- How:
  - Use synthetic and genuine topologies to analyze topological changes of resulting graphs with varying alias resolution success.

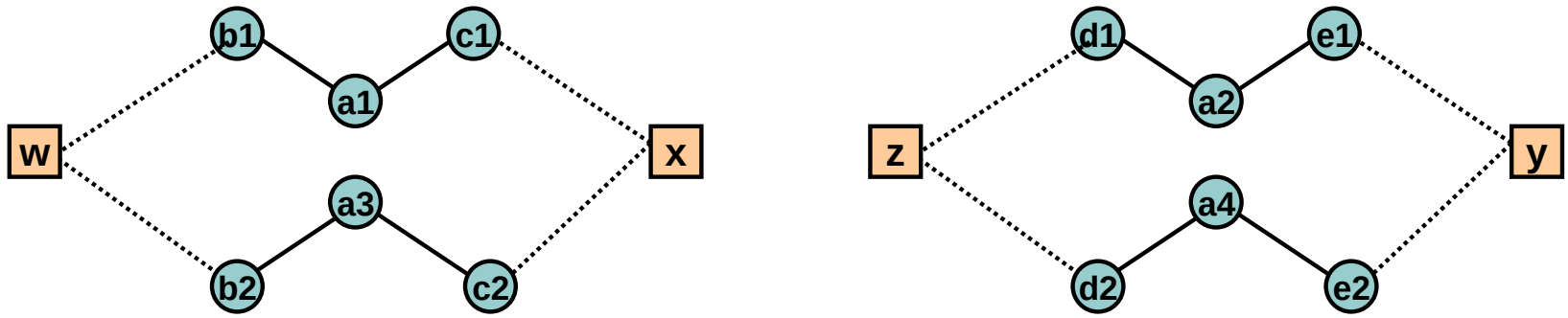
Alias Resolution:
Problem



a sub-graph

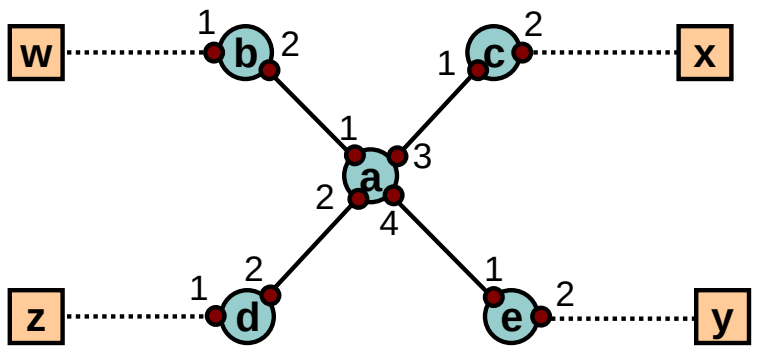
- A set of collected traces
  - w, …,b1, a1, c1, …, x
  - z, …,d1, a2, e1, …, y
  - x, …,c2, a3, b2, …, w
  - y, …,e2, a4, d2, …, z

The sample map from the collected path traces



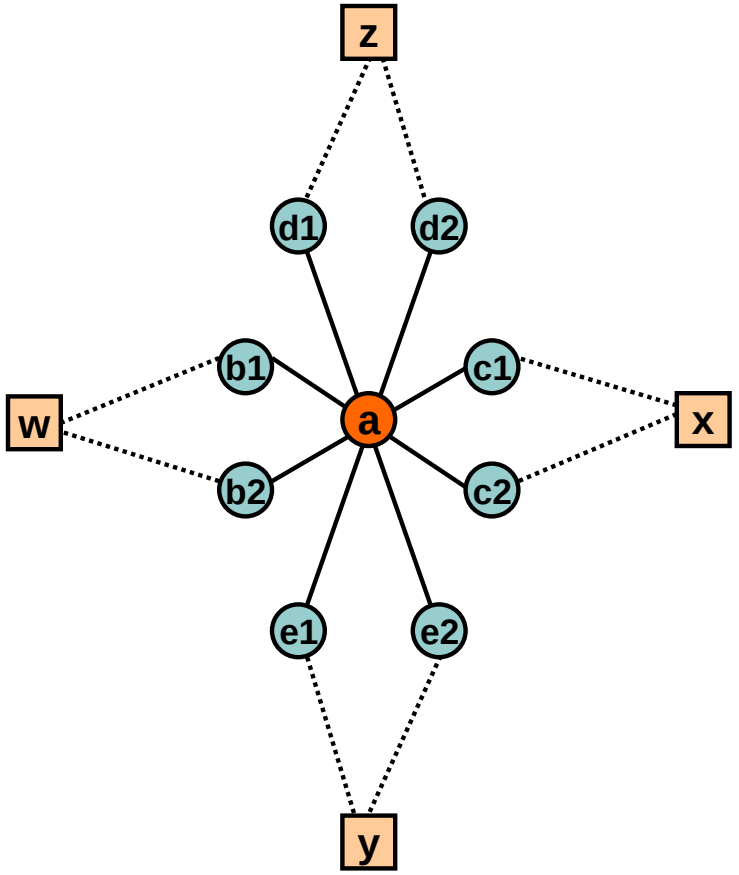
Alias Resolution is the process of identifying the IP addresses of a router.

Alias Resolution:
Problem



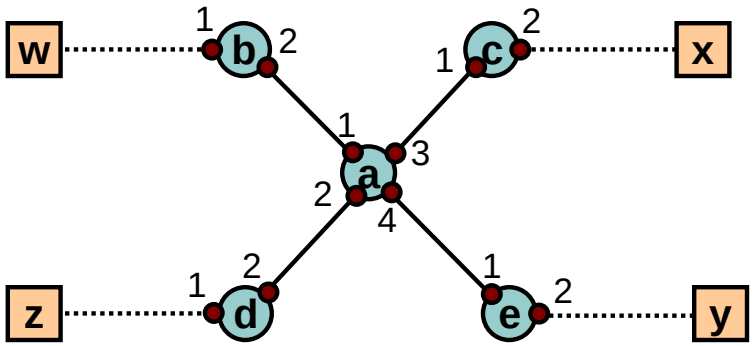
a sub-graph

- A set of collected traces
  - w, …,b1, a1, c1, …, x
  - z, …,d1, a2, e1, …, y
  - x, …,c2, a3, b2, …, w
  - y, …,e2, a4, d2, …, z



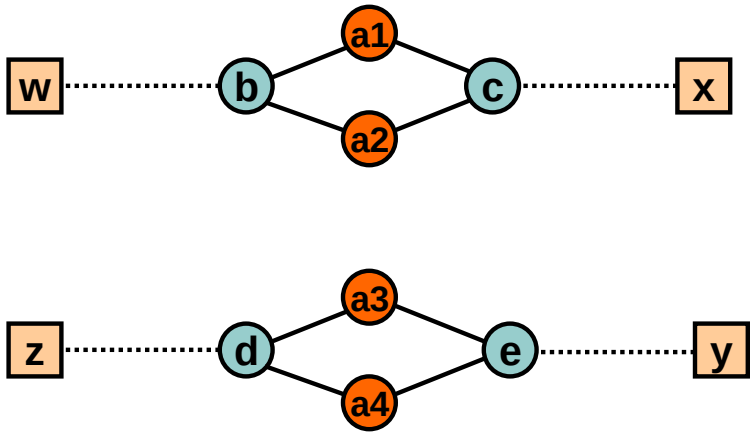
partial alias resolution
(only router **a** is resolved)

Alias Resolution:
Problem



a sub-graph

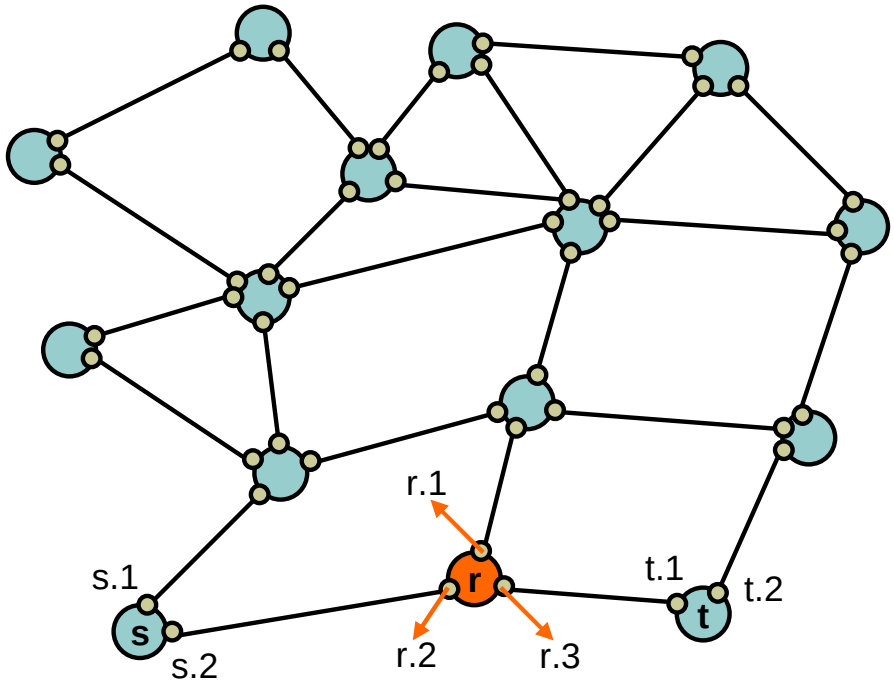
- A set of collected traces
  - w, …,b1, a1, c1, …, x
  - z, …,d1, a2, e1, …, y
  - x, …,c2, a3, b2, …, w
  - y, …,e2, a4, d2, …, z



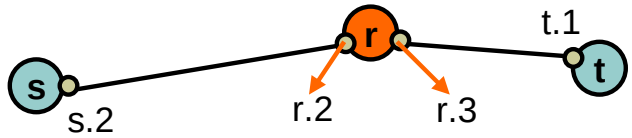
partial alias resolution
(only router **a** is *not* resolved)

Alias Resolution:
Experimental Procedure

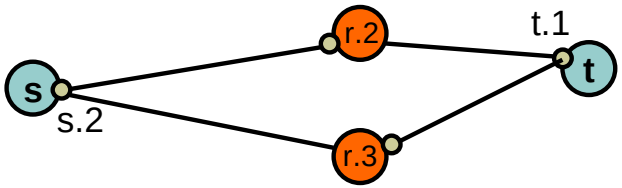
- First, generate a synthetic network topology,
- Next, annotate it to add interface addresses,
- Then, emulate *traceroute* to collect path traces,



- Consider an example
  - A path from s to t: s.2 – r.2 – t.1
  - A path from t to s: t.1 – r.3 – s.2
- Case 1: resolve aliases @ r



- Case 2: do not resolve aliases @ r



- Finally, build sample topologies
  - with different *alias resolution* success rate

Alias Resolution:
Experimental Procedure

Apply *alias resolution* with different success rate
- 0%, 20%, 40%, 60%, 80%, and 100% success rates.

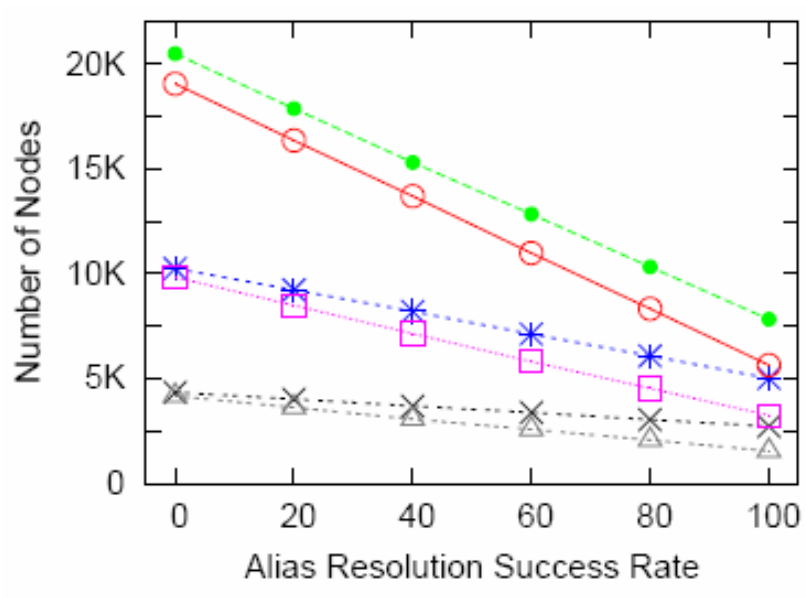
- Generate various synthetic graphs to represent the Internet
  - Random : *Waxman* (WA),
  - Power-law : *Barabasi-Albert* (BA) and *Inet*,
  - Hierarchical : *Transit-Stub* (TS)

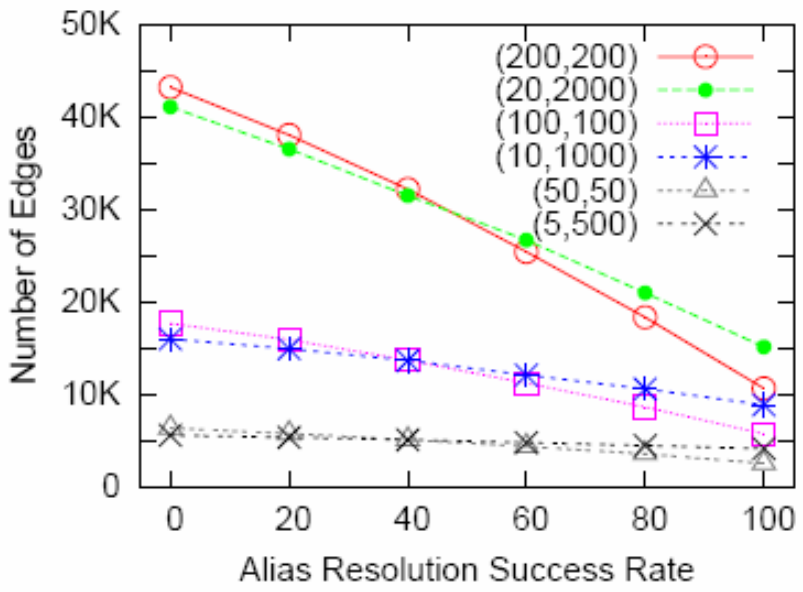
- Analyze changes in topological characteristics
  - Topology Size,
  - Node Degree,
  - Degree Distribution,
  - Joint Degree Distribution,
  - Characteristic Path Length,
  - Hop Distribution,
  - Betweenness,
  - Clustering.

- Analyze a genuine Internet sample
  - Utilize state-of-the art alias resolution tools.

Effects on Topological Characteristics :
Topology Size

- Number of nodes and edges reduces by 57% and 62%, on average, as alias resolution improves from 0% to 100%.

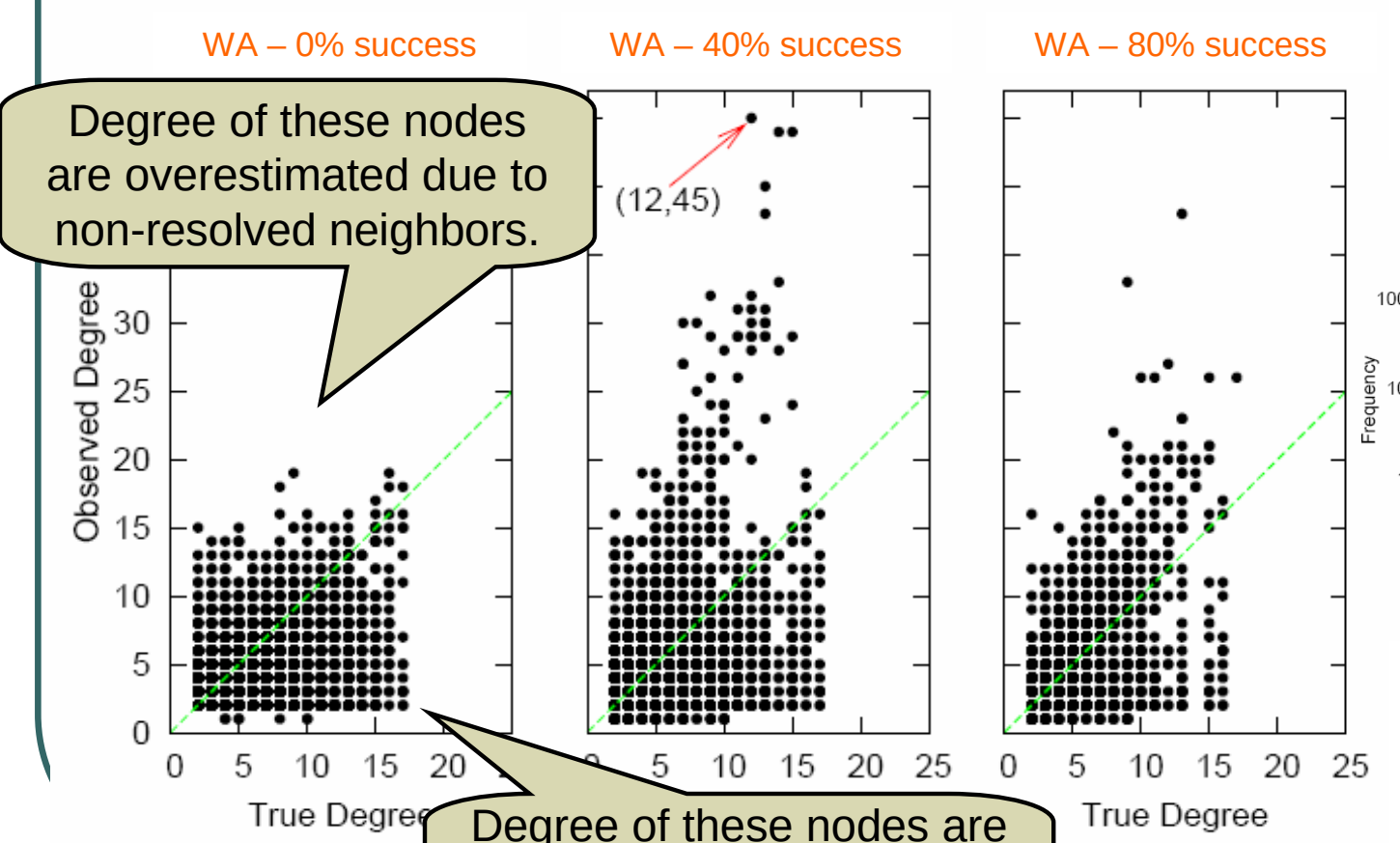


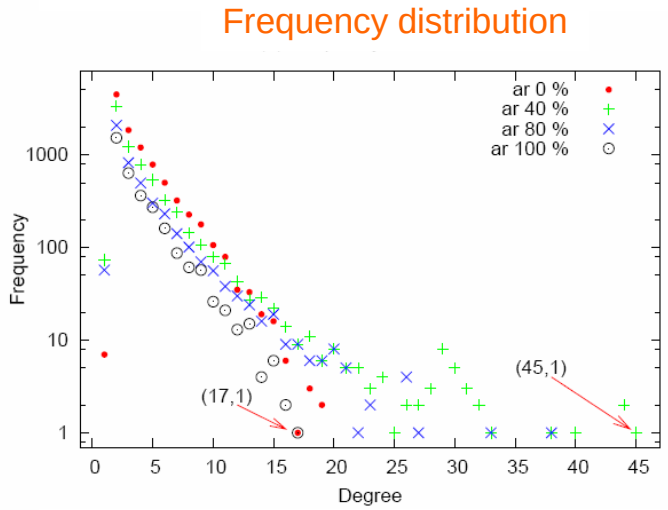


- The impact of imperfect alias resolution increases as the size of the sample topology increases.
- (n,n) samples are affected more by imperfect alias resolution.

Effects on Topological Characteristics :
Node Degree

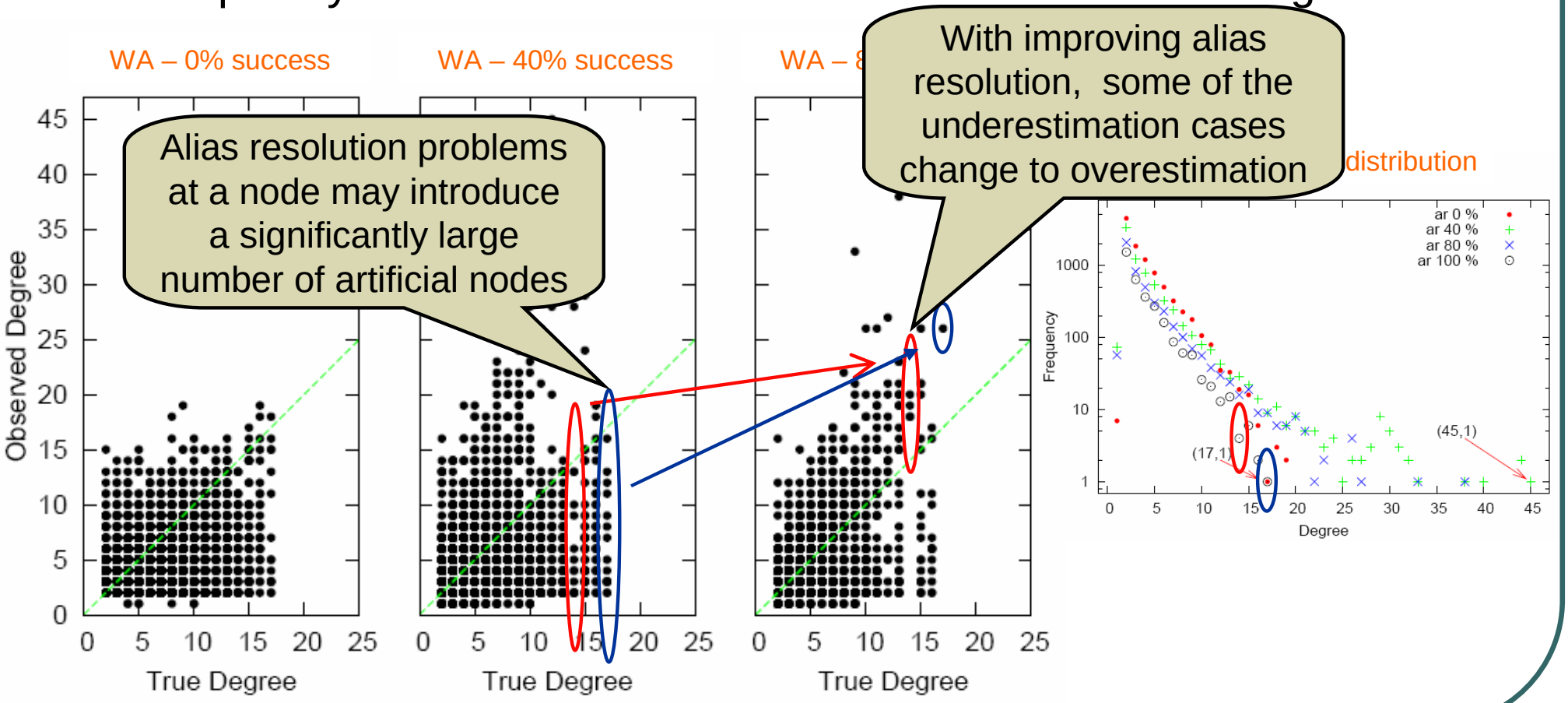
- *Observed degree*: degrees with imperfect alias resolution
- *True degree:* degrees with perfect alias resolution.
- Frequency distribution: number of nodes at each node degree





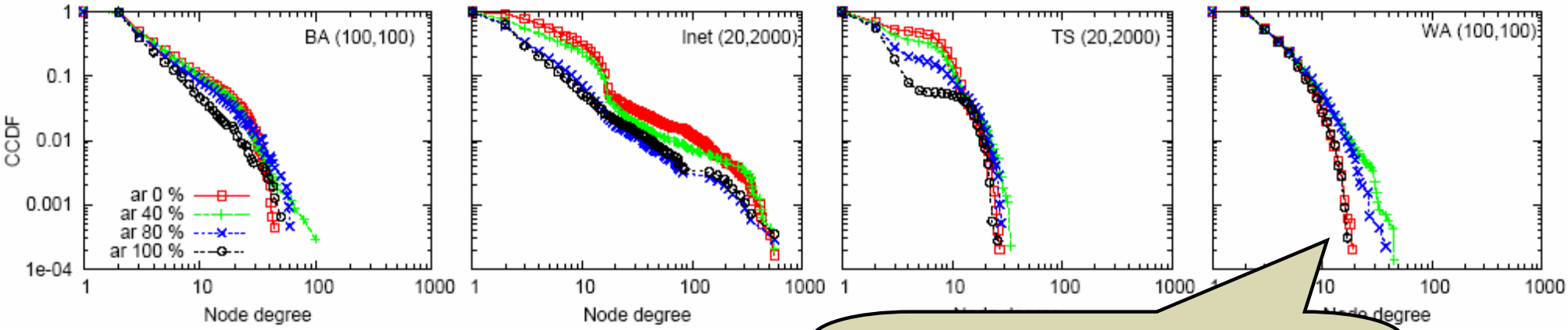
Effects on Topological Characteristics :
Node Degree

- *Observed degree*: degrees with imperfect alias resolution
- *True degree:* degrees with perfect alias resolution.
- Frequency distribution: number of nodes at each node degree



Effects on Topological Characteristics :
Degree Distribution

- The probability *P(k)* that a randomly chosen node has degree *k*.



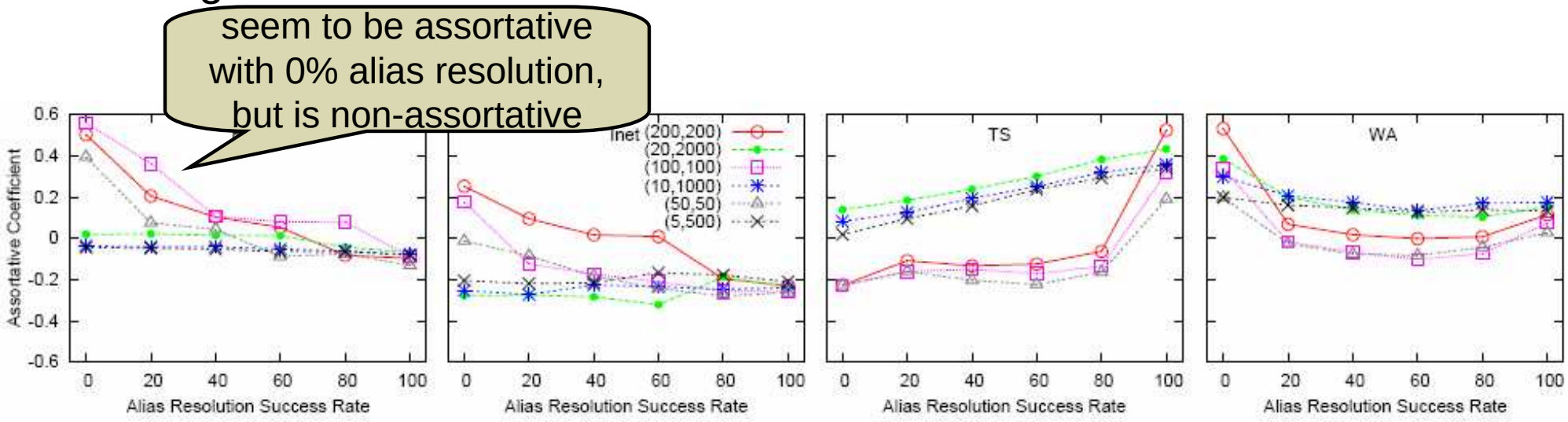
Degree-related characteristics do not always improve with an increasing success rate

- Imperfect alias resolution, especially,
  - distorts the power-law characteristic of *BA- and Inet-based samples,*
  - impacts especially low degree ranges (3-13) of *TS-based samples,*
  - impacts especially high degree ranges (20-up) of *WA-based samples.*

Effects on Topological Characteristics :

Joint Degree Distribution

- The probability *P*(*k1, k2*) that a node of degree *k1* and a node of degree *k2* are connected.



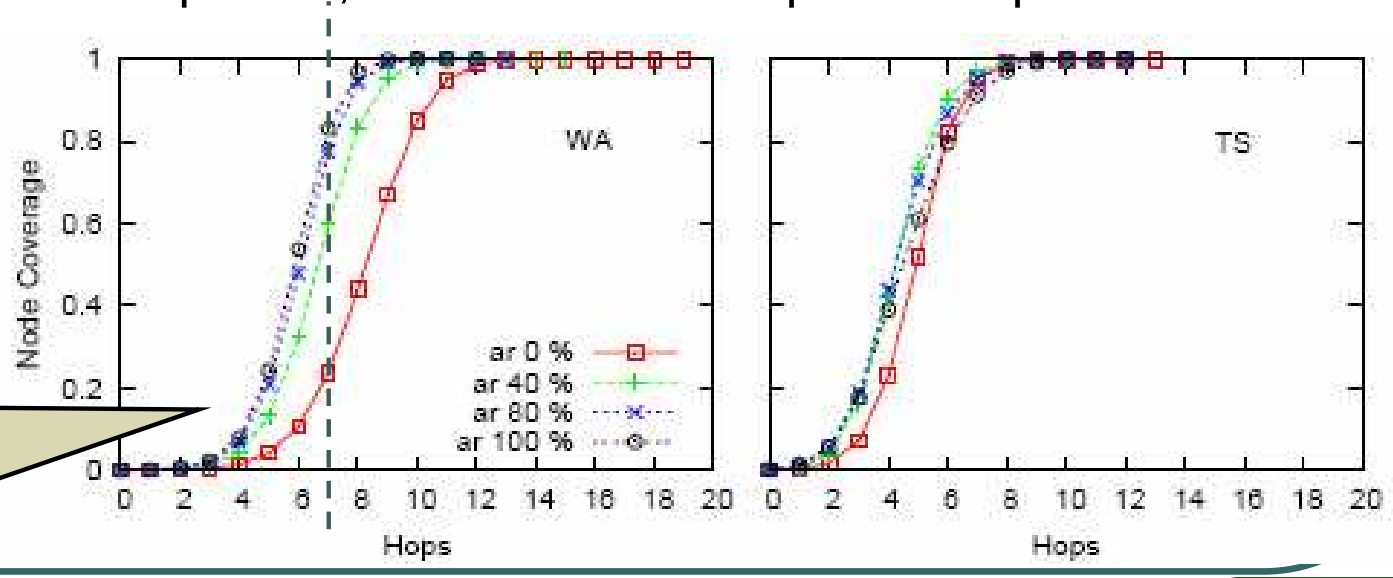
- Assortativity coefficient: The tendency of a network to connect nodes of the same or different degrees.
  - Positive values indicate assortativity
    - most of the links are between similar degree nodes.
  - Negative values indicate disassortativity
    - most of the links are between dissimilar degree nodes.
  - 0 indicates non-assortativity.

Effects on Topological Characteristics :
Characteristic Path Length & Hop Distribution

- Characteristic Path Length
  - The average of the shortest path lengths between all node pairs.
  - Reduces with the increasing alias resolution success rate.
    - On average 30% for BA, Inet and WA-based sample topologies.

- Hop Distribution
  - The average percentage of the nodes reached at each hop
  - As alias resolution improves, less number of hops are required to reach others.

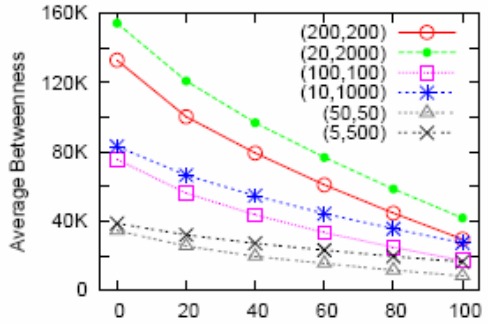
24%, 60%, 78%, and 83% of the nodes are reachable within 7 hops with 0%, 40%, 80% and 100% alias resolution, respectively

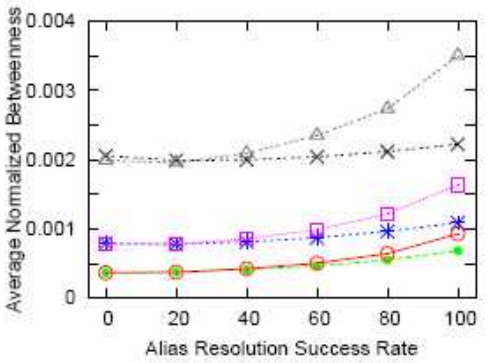


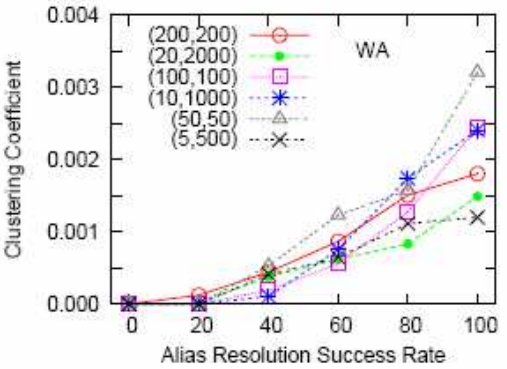
Effects on Topological Characteristics :
Betweenness & Clustering

- Betweenness
  - The total number of shortest paths that pass through a node.
  - As the alias resolution success rate increases
    - The average betweenness reduces
    - The normalized betweenness increases

- Clustering
  - Characterizes the density of the connections in the neighborhood of a node.
  - All samples yield a clustering coefficient of 0 with 0% alias resolution success rate
  - It almost always increases with the improving alias resolution.







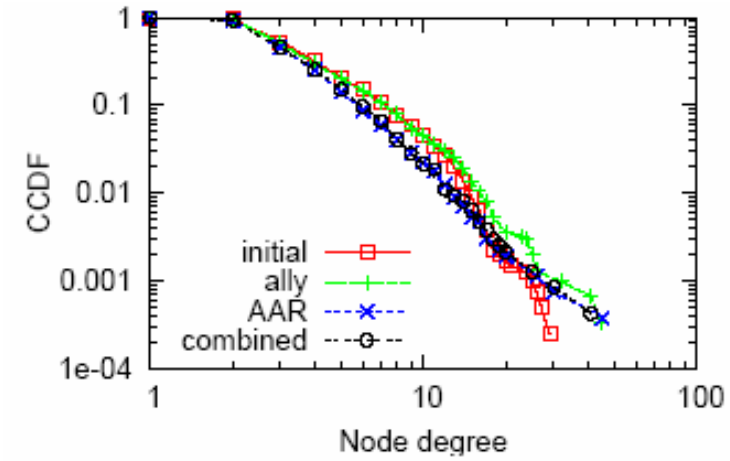
Effects on Topological Characteristics :
Impact on Genuine Topologies

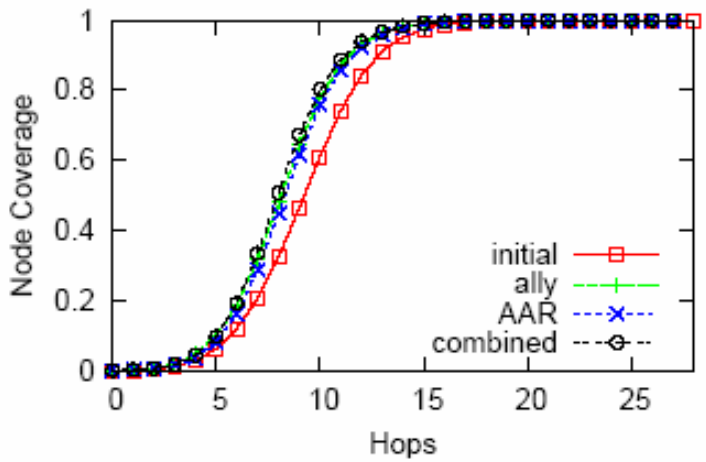
- We analyze the effect of alias resolution in building a topology map from a set of collected path traces.
  - *Ally* is the current state-of-the-art probe based approach.
  - APAR is a recent analytical approach.

- We study the changes in observed topological characteristics of the networks with respect to the initial topology where none of the tools is used.

	Initial	Ally	APAR	Ally & APAR
Number of Nodes	4085	3080	2659	2376
Number of Edges	7313	5502	4132	3727
Average Degree	3.58	3.57	3.11	3.14
Assortativity Coefficient	0.39	0.14	0.08	0.09
Characteristic Path Length	9.77	8.67	8.90	8.57
Normalized Betweenness	0.0021	0.0025	0.0030	0.0032
Clustering Coefficient	0.0061	0.0275	0.074	0.0566

Effects on Topological Characteristics :
Impact on Genuine Topologies





- .s resolu
  path trac
  ne-art pr
  approac
- served topol
  the initial to

Degree related characteristics are mostly similar to that of BA samples

Path length related characteristics are closer to that of TS samples.

	Initial	Ally	APAR	Ally & APAR
Number of Nodes	4085	3080	2659	2376
Number of Edges	7313	5502	4132	3727
Average Degree	3.58	3.57	3.11	3.14
Assortativity Coefficient	0.39	0.14	0.08	0.09
Characteristic Path Length	9.77	8.67	8.90	8.57
Normalized Betweenness	0.0021	0.0025	0.0030	0.0032
Clustering Coefficient	0.0061	0.0275	0.074	0.0566

Conclusion

- We analyzed the impact of alias resolution on *traceroute*-based sample network topologies.
- We observed that the accuracy of the alias resolution process may significantly distort, almost all, topological characteristics that we consider in this study.
- Internet measurement studies should employ all the means possible to increase the accuracy/ completeness of the alias resolution process.
  - The combination of *ally* and APAR achieves better results.