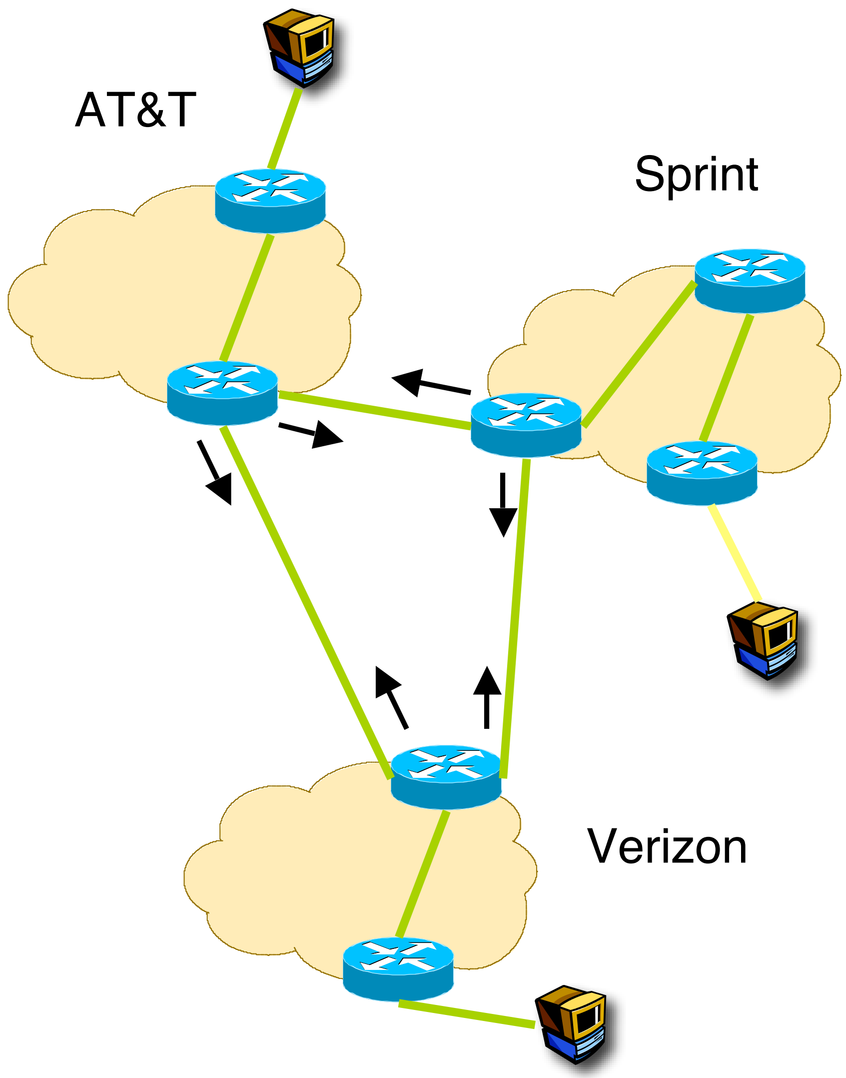
In Search of the Elusive Ground Truth: The Internet's AS-level Connectivity Structure

Ricardo Oliveira, UCLA
Dan Pei, AT&T
Walter Willinger, AT&T
Beichuan Zhang, U. Arizona
Lixia Zhang, UCLA

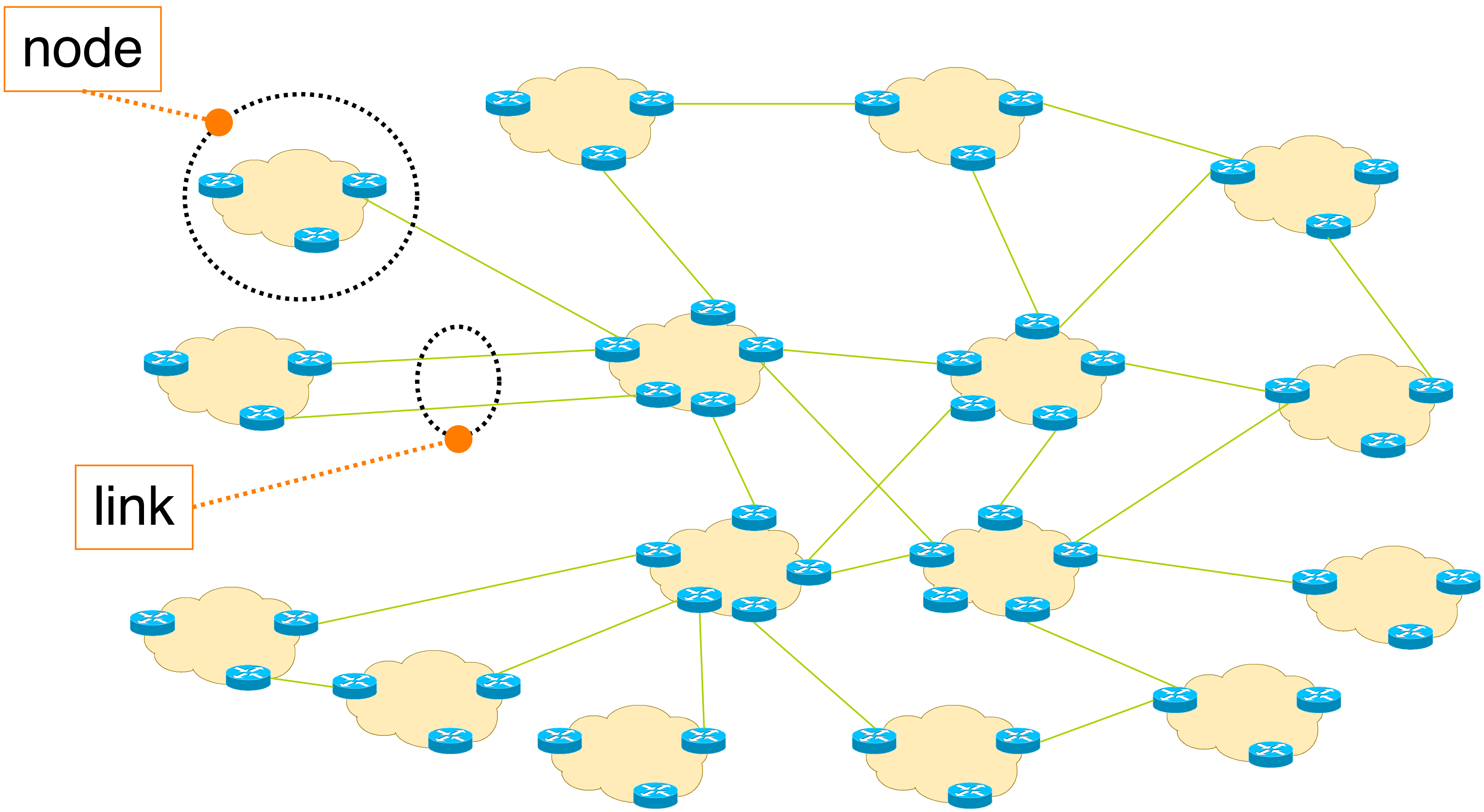
Internet and Autonomous Systems



- **Autonomous System:** a set of routers or networks under the same administration

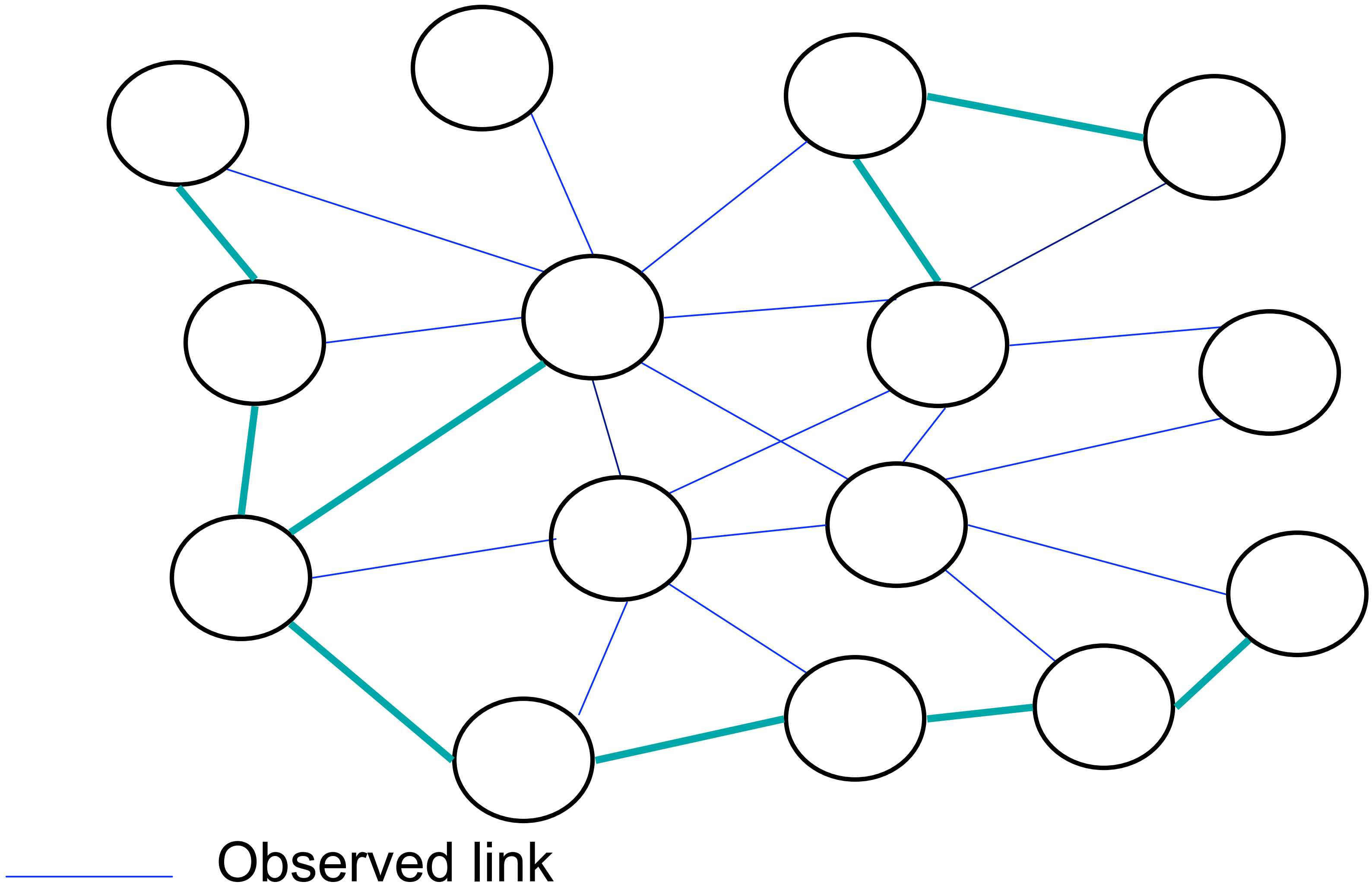
- Border routers exchange routing updates via the **Border Gateway Protocol (BGP)**

Internet AS topology (the abstraction)



• About **26k+ ASes** in the Internet, at least **80k links** interconnecting them

Completeness of the AS graph



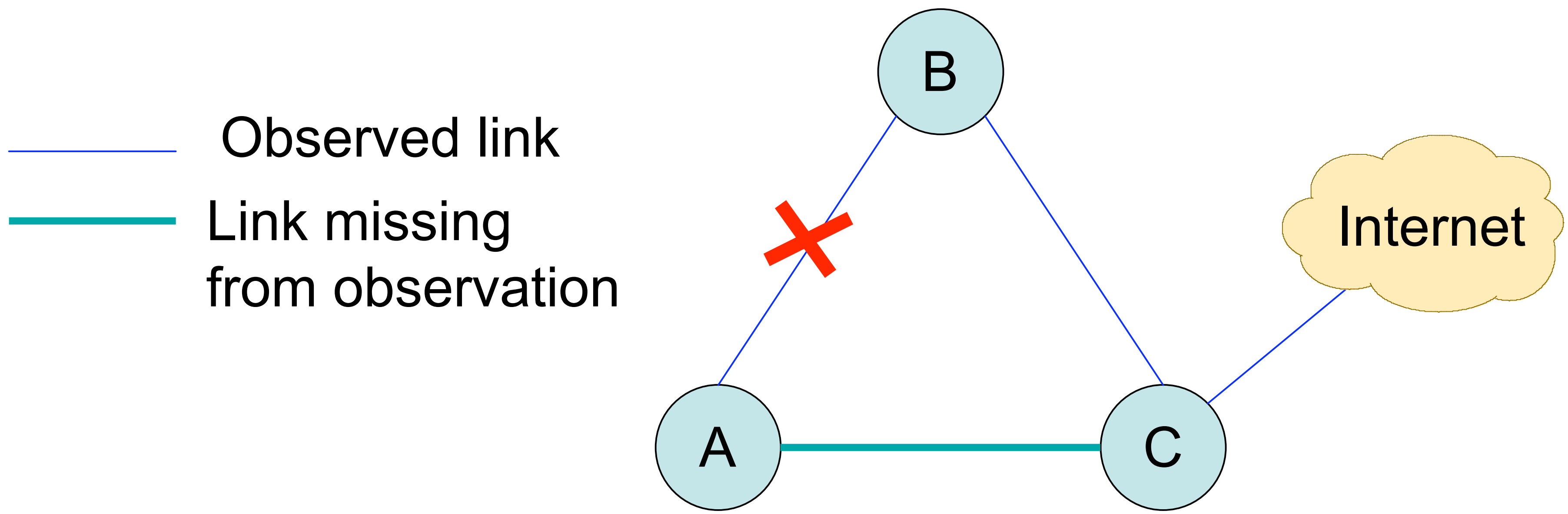
Observed link

Link missing from observation

The completeness problem

- The AS-level connectivity is inferred from BGP tables+updates collected from hundreds of vantage points - the **public view** (PV)
- Previous work tried to make the observed AS topology “more complete”, **w/out knowing “how complete”…**
- Many Internet-scale studies depend on the completeness problem…

A simple example: resilience of topology to failures

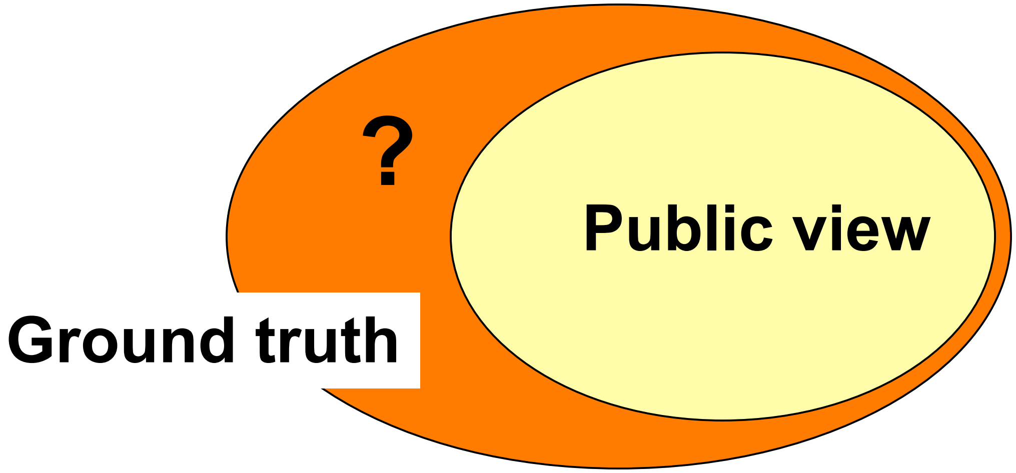


A is estimated to be disconnected, while in reality it's still connected to **C**

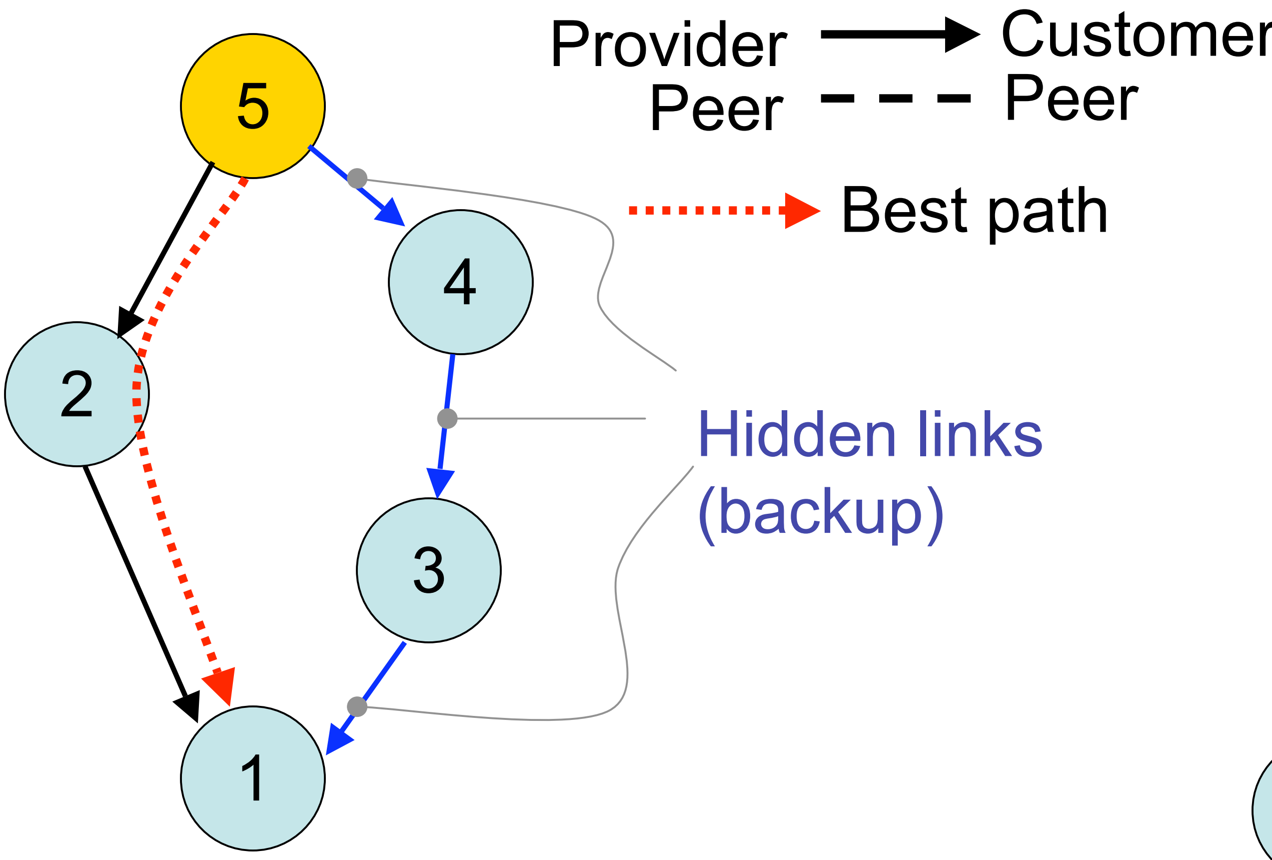
The completeness problem

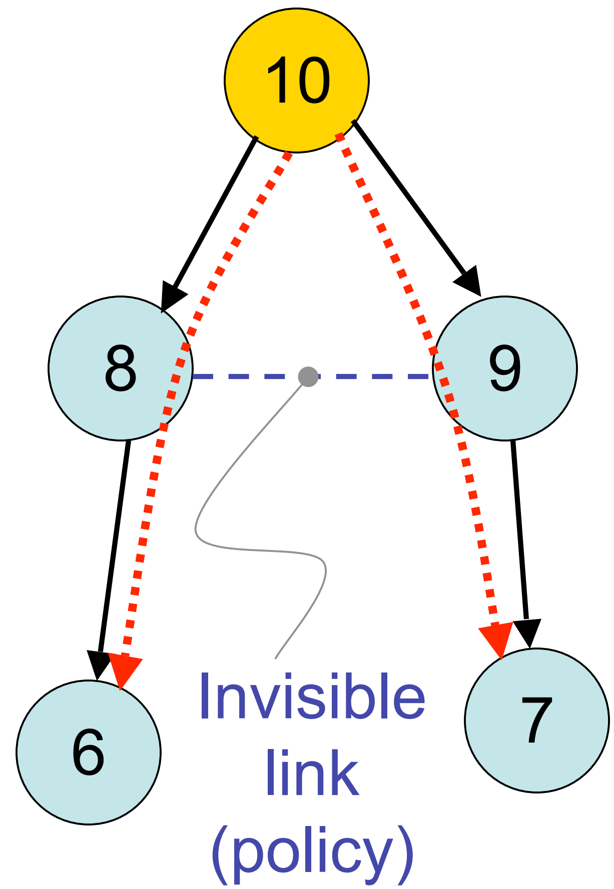
- To date, no quantitative evaluation on how (in)complete is the connectivity inferred from the PV …

- **This work is the first that uses ground truth info to evaluate the (in)completeness of the AS topology**



Observing AS level links





- **5** prefers shortest routes
- Hidden links are used if primary path fails; **need observation over time to capture them**

- Peer links are **NOT** advertised upstream to node **10**

Tackling the completeness problem

- We examine several ASes of different categories, for which we know the connectivity ground truth:
  - Tier-1
  - Tier-2
  - Simple stubs
  - Large content provider
- And quantify how much of this connectivity is revealed in the PV and what types of links are missing

Tier-1 network

- Extracted BGP connectivity information from **routers' configuration files**

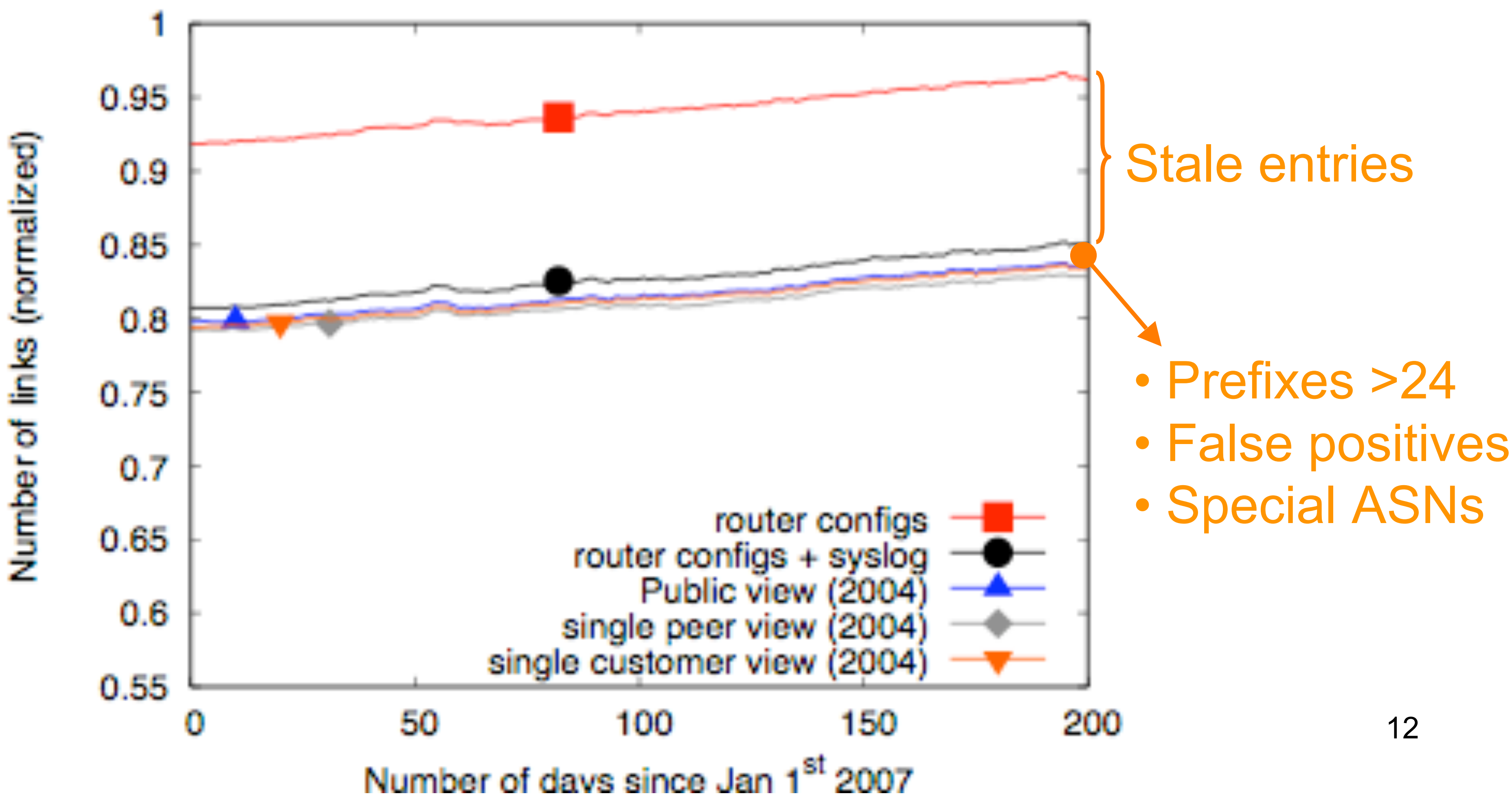
- Two types of BGP connections:
  - **Single-hop**: "ip address" or "ip route"
  - **Multi-hop**: "neighbor <*ip-address*> ebgp-multihop <*#hops*>"

Filtering stale entries

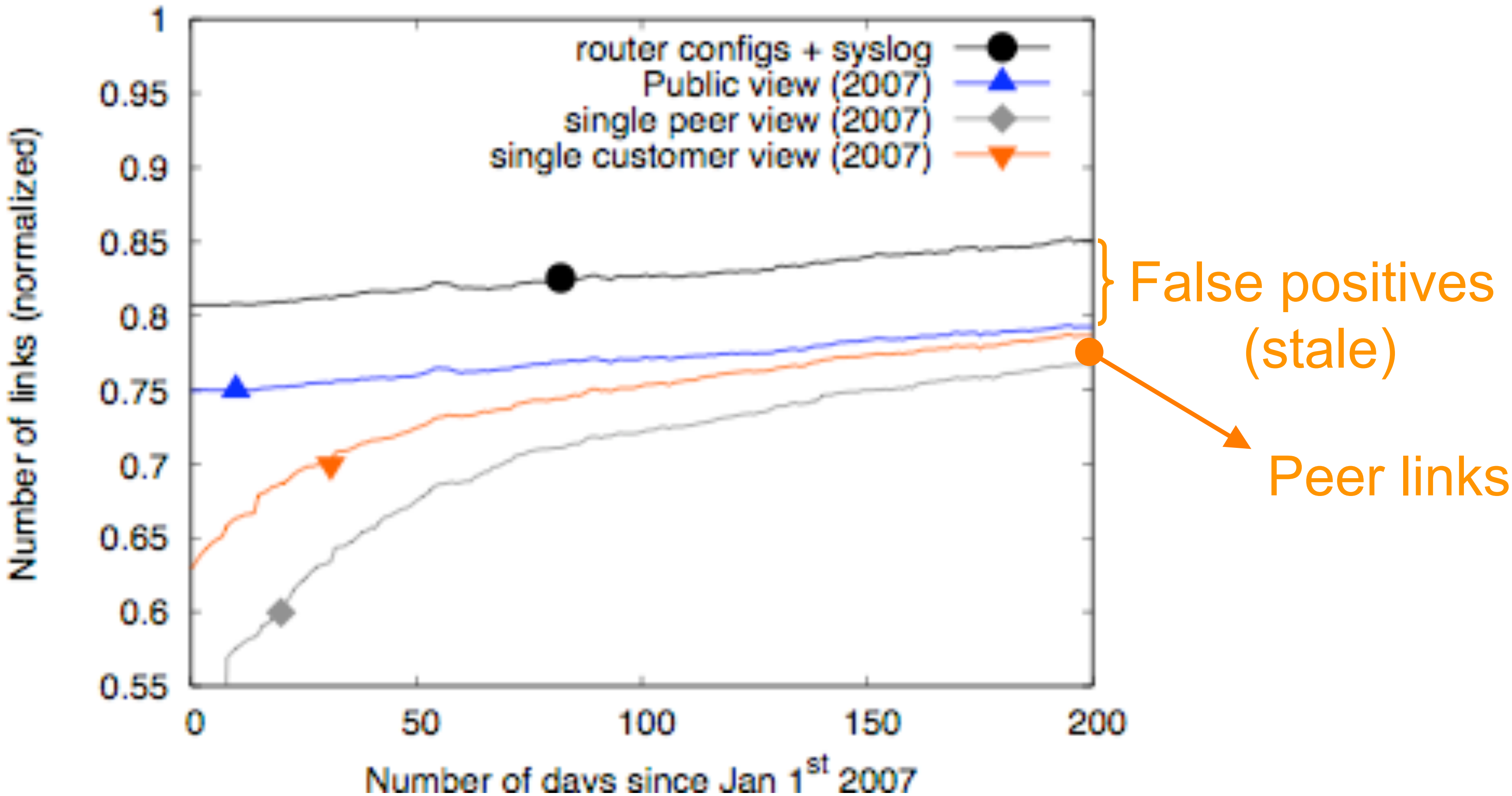
- Neighbor ASes w/ minimal configuration in routers' configs but with no established session => need to filter stale entries
- Use **syslog information** to nail down stale entries:
  - Filter cases which last syslog msg is "session down"
  - Filter cases where last syslog msg is older than 1 year
- Lesson learned: even w/ router config information, obtaining an accurate connectivity info in large networks is still a challenge

Connectivity of a Tier-1(2004)

- Need observation over time to reveal **hidden links**



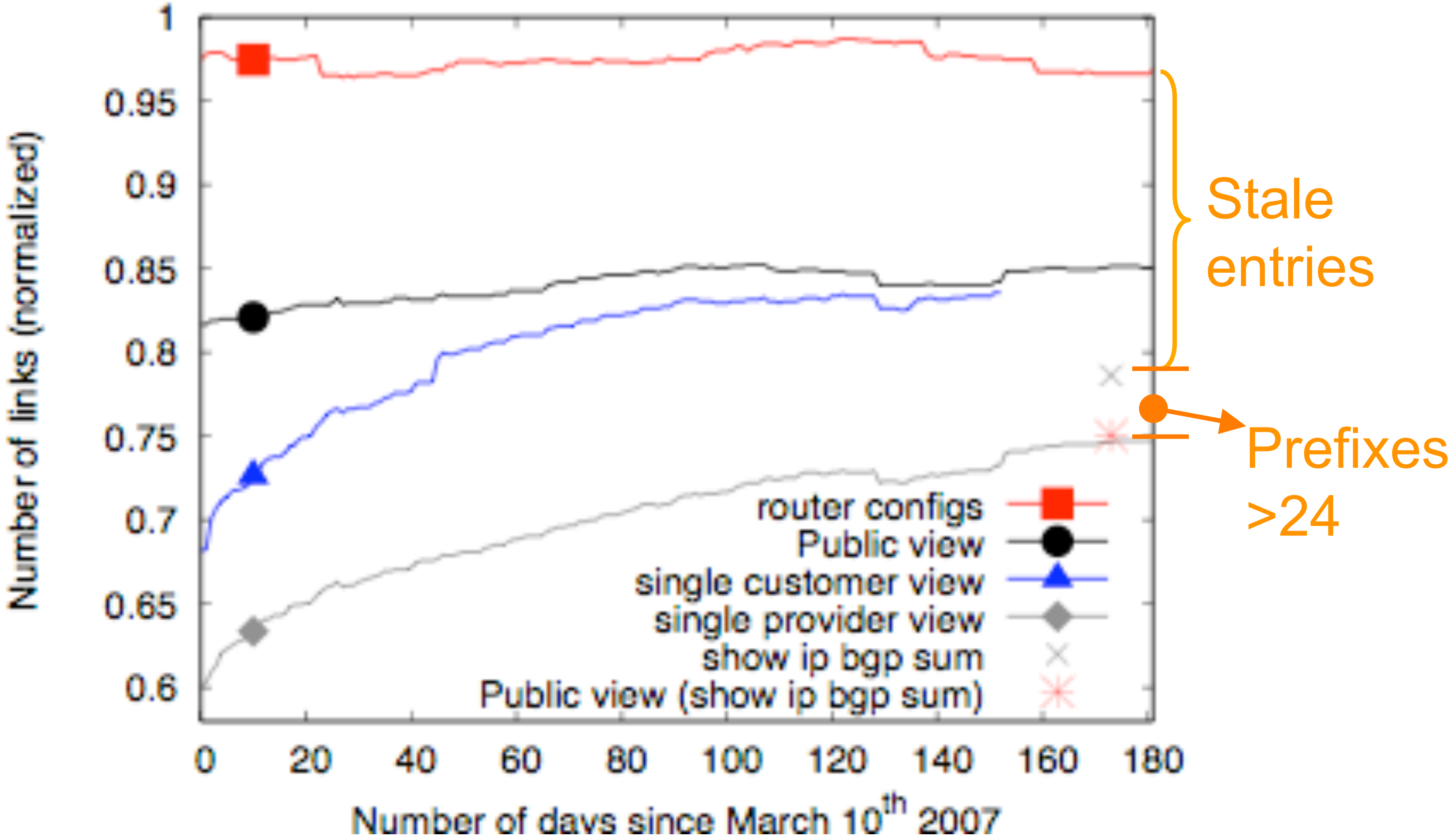
Connectivity of a Tier-1 (2007)



Tier-1 summary

- PV covers almost all links of the Tier-1 network; the small % of invisible links have customers w/ prefixes >24
- Links of the Tier-1 network are covered by a single customer view
- The Tier-1 announces all its customer links to peers

Connectivity of a Tier-2 network



Tier-2 summary

- Tier-2's links are covered by a direct customer (except a small % of >24)

- A provider misses the Tier-2's peer links because of policy output filters in route advertisements

Simple stubs

- Simple stubs are end networks that do not have customers, e.g. academic networks
- Examined the connectivity of 4 stub ASes
- PV captures all customer-provider links
  - 3 false positive links, which should be eliminated after tweaking the death threshold period

IXPs and content providers

• IXP=**I**nternet e**X**change **P**oint=a shared L-2 switching fabric interconnecting different networks in a same location

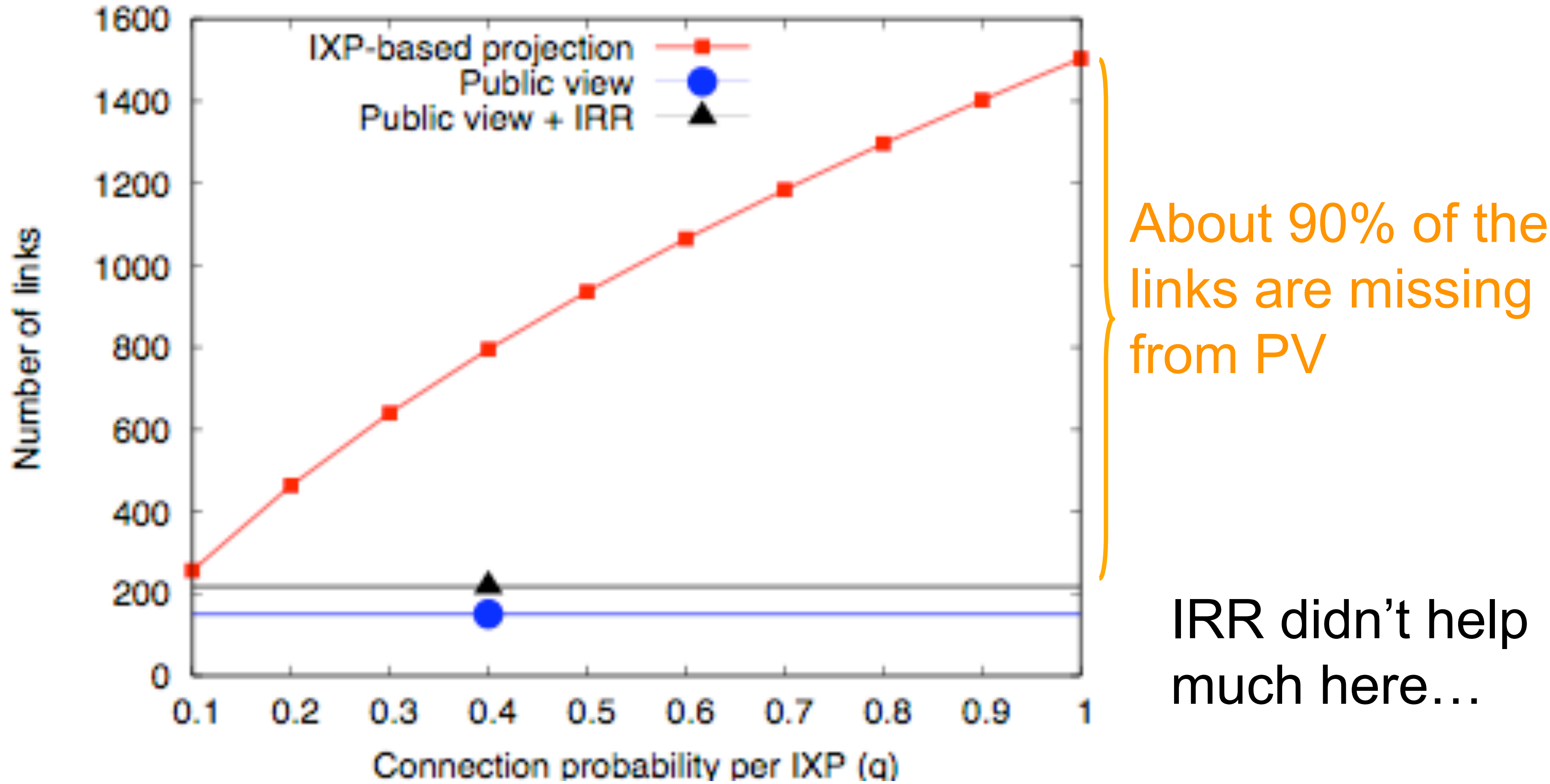
• Networks that participate in IXPs usually have **settlement-free** or **peer** agreements with each other

• Content providers tend to connect in IXPs to **maximize number of peers** (save $$ sent upstream)

• We pick a large content provider X with close to a **thousand BGP sessions**

• X peers heavily in ~30 **IXPs**

Large content provider X

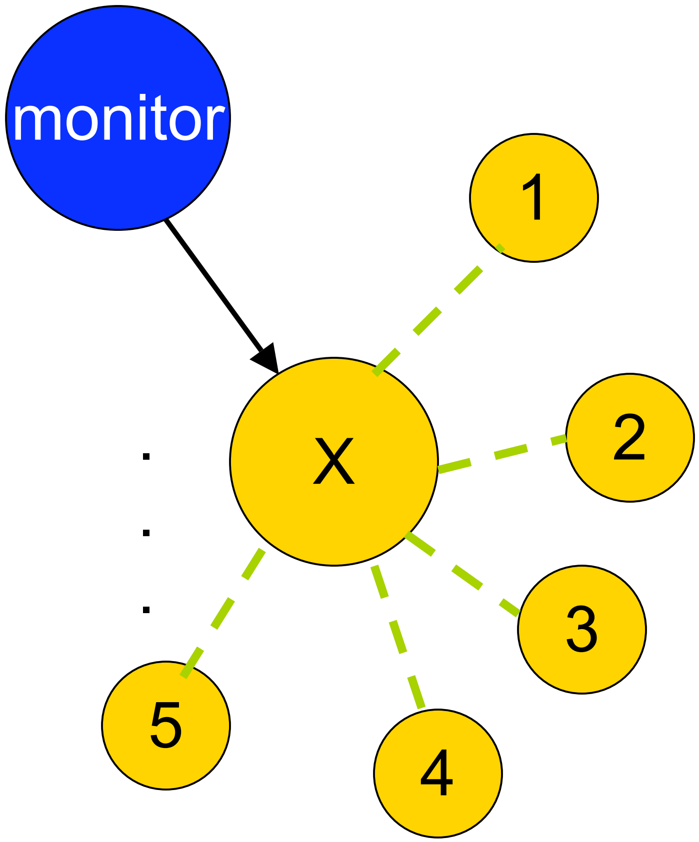


IRR didn't help much here…

Projection=1 - (1 - q)$^{n(i)}$

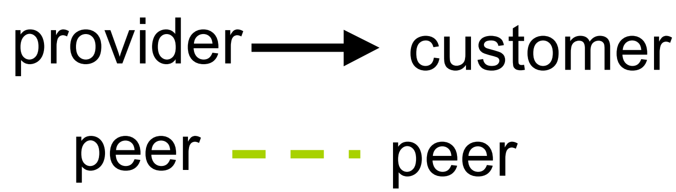
- **q** is connection probability per IXP, q~80-95%
- **n(i)** is the number of common IXPs net i has in common w/ X

Content nets



- Peer links 1-5 are not propagated upstream…

- These links are **invisible**



Observation about PV: What are captured, what missed?

- Coverage of **Tier-1 links**: OK

- Coverage of **customer-provider links**: OK, need observation over time

- Coverage of **peer links**: NOT OK; PV misses a great % of peer links (up to 90% according to our estimates); mainly between stubs

Using PV in practice

- Stub AS growth rates and net diameter: OK
- General graph-theoretic metrics: NOT OK
- Impact of prefix hijack: depends on who originates the prefix
- Resilience studies to failures: NOT OK
- Evaluation of new inter-domain protocols: be careful w/ assumptions made

Limitations of the PV

- On the absence of having **a BGP feed in each stub AS**, there will inevitably be **gaps** in the inferred AS topology map …

Moving beyond the PV

Q: How to fill the gaps of missing peer links in AS topology maps w/ only relying in measurements from vantage points?

A: Hard problem... Start by exploring IXP colocation info where content nets peer heavily. Assume peer nets connect to 80-95% of net there. Private peering inference still open problem.

Summary

- Confirmed that PV (if cumulated over long enough time) can **capture all the customer-provider links**
- **PV misses most peer links** between lower tier ISPs, content providers, and stub nets
- **New inference techniques are needed** to capture or estimate peer links that do not rely solely on measurements from small set of vantage points

Questions?

Contact: rveloso@cs.ucla.edu

Thanks!