What we know and what we don't about the Internet

Dmitri Krioukov
CAIDA
dima@caida.org
X. Dimitropoulos, **P. Mahadevan**,
M. Fomenkov, B. Huffaker,
Y. Hyun, and kc claffy.

Aveiro, May 15th, 2008

What the Internet does

- The Internet was designed for and exists to transfer information packets from *A* to *B*, where *A* and *B* are any two Internet-Protocol- (IP-)talking devices

IP packet format

+	Bits 0–3	4–7	8–15	16–18	19–31
0	Version	Header length	Type of Service (now DiffServ and ECN)	Total Length	
32	Identification			Flags	Fragment Offset
64	Time to Live		Protocol	Header Checksum	
96	Source Address				
128	Destination Address				
160	Options				
160 or 192+	Data				

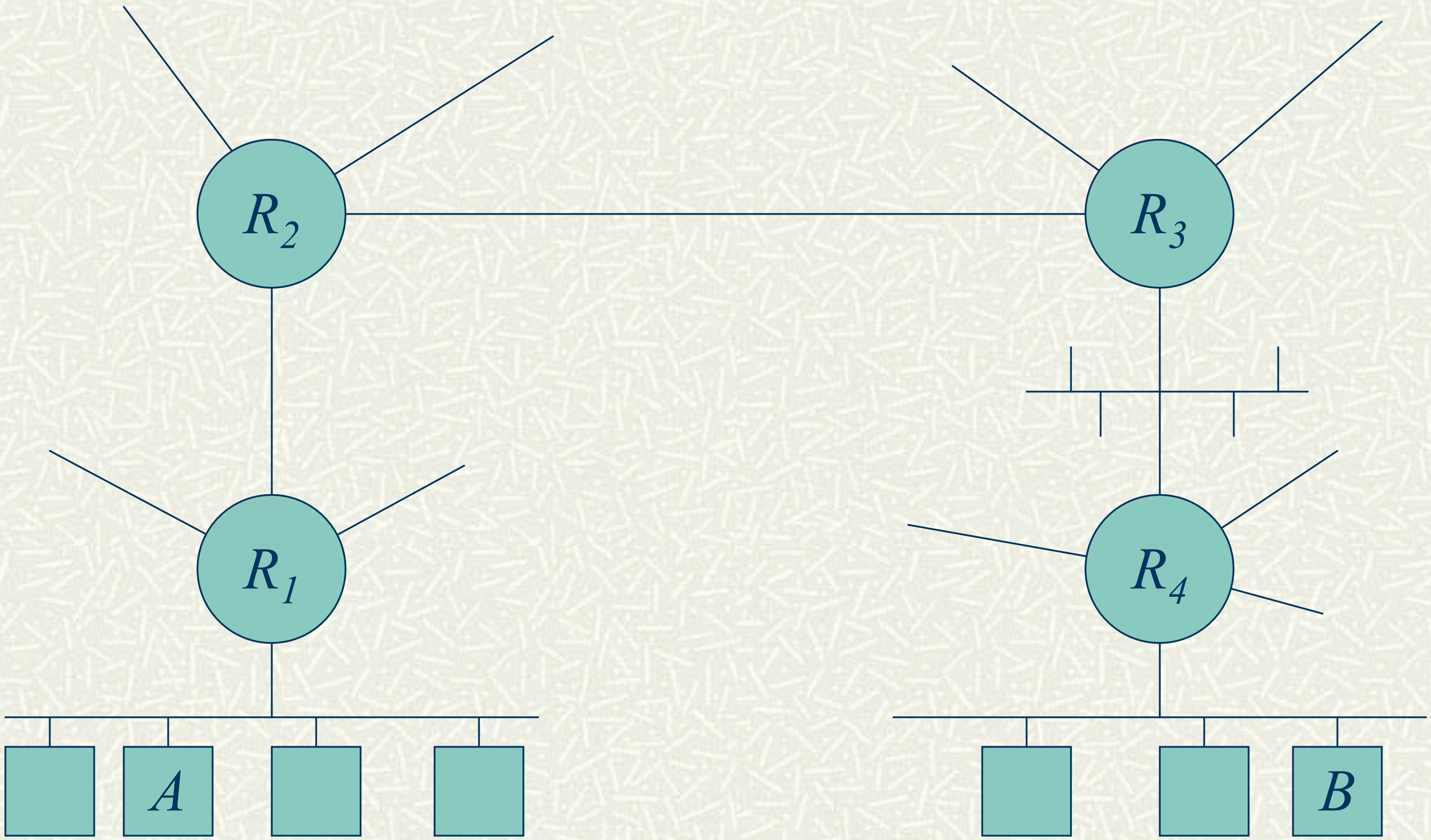
IP addresses

- A = 193.137.168.155
- B = 192.172.226.78

IP routes

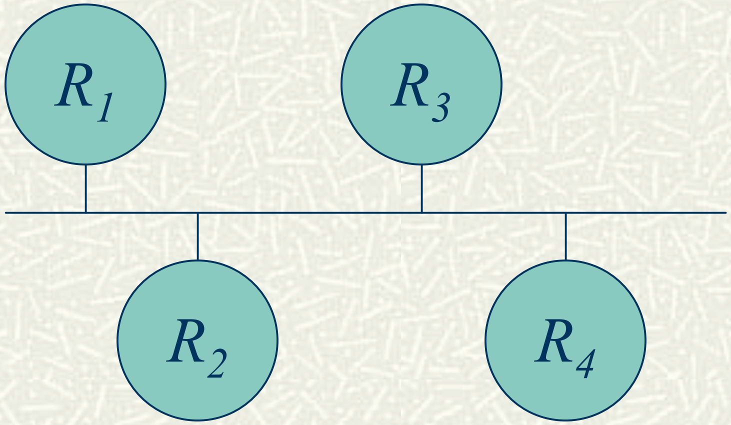
- traceroute 192.172.226.78
- 1 <1 ms 2 ms 2 ms 193.137.81.254
- 2 <1 ms <1 ms <1 ms 192.168.255.253
- 3 1 ms 1 ms 1 ms 193.137.173.254
- 4 1 ms 1 ms 1 ms 193.136.4.26
- 5 5 ms 5 ms 5 ms 193.136.1.221
- 6 5 ms 6 ms 6 ms 193.137.0.30
- 7 6 ms 6 ms 6 ms 62.40.124.185
- 8 32 ms 33 ms 32 ms 62.40.112.146
- 9 41 ms 40 ms 40 ms 62.40.112.137
- 10 123 ms 124 ms 124 ms 62.40.112.134
- 11 130 ms 130 ms 129 ms 216.24.184.85
- 12 134 ms 131 ms 130 ms 216.24.186.23
- 13 143 ms 144 ms 143 ms 216.24.186.20
- 14 167 ms 167 ms 167 ms 216.24.186.8
- 15 199 ms 199 ms 198 ms 216.24.186.30
- 16 197 ms 197 ms 197 ms 137.164.26.130
- 17 203 ms 203 ms 203 ms 137.164.25.5
- 18 204 ms 203 ms 203 ms 137.164.27.50
- 19 204 ms 205 ms 204 ms 198.17.46.56
- 20 203 ms 204 ms 204 ms 192.172.226.78

IP routes

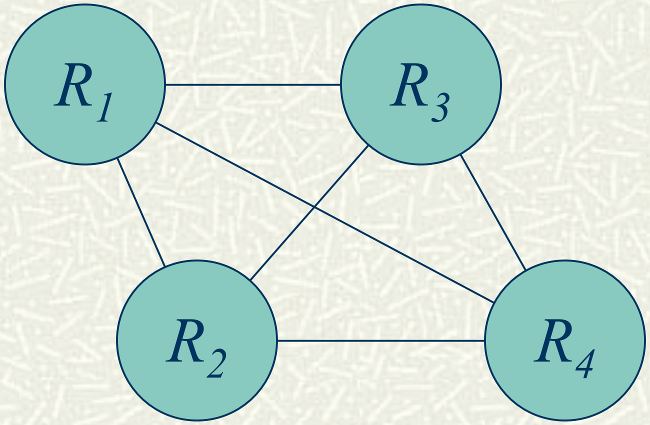


Broadcast media (e.g., ethernet)

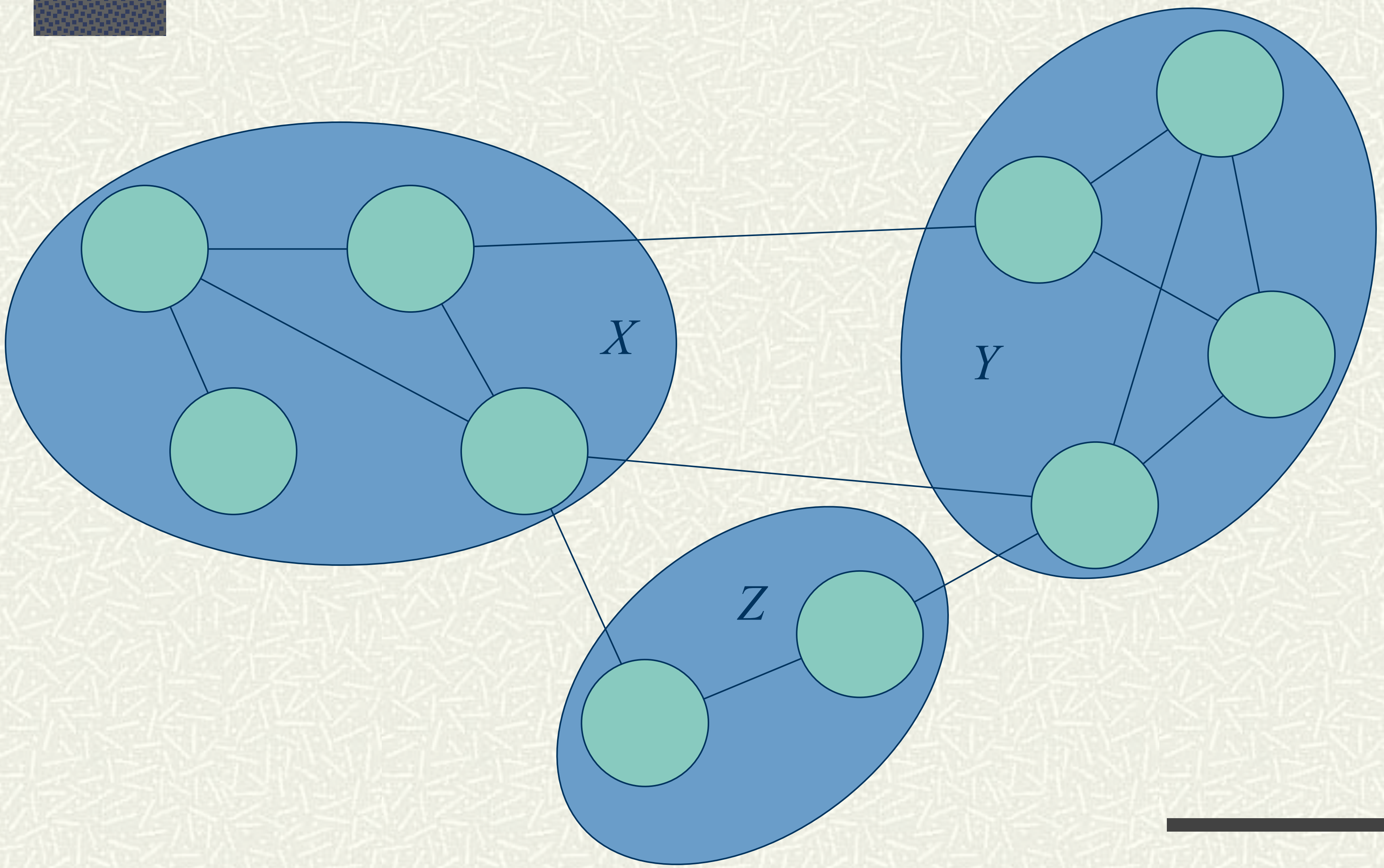
Reality



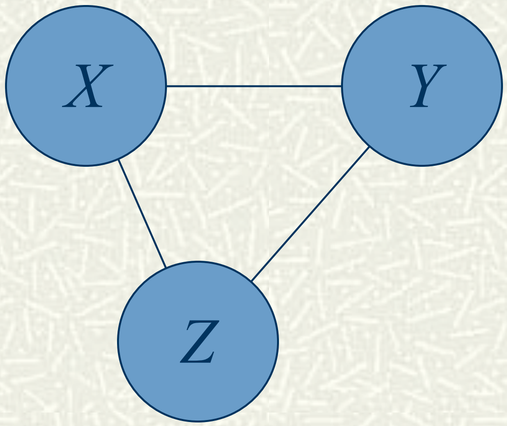
Perception



Autonomous Systems



AS topology



IP routing

- Intradomain (Interior Gateway Protocols (IGPs))
  - routing within an Autonomous System (AS)
  - protocols:
    - Open Shortest Path First (OSPF)
    - Intermediate System to Intermediate System (ISIS)
  - Links State (LS) routing protocols
- Interdomain (Exterior Gateway Protocols (EGPs))
  - routing between Autonomous Systems (ASs)
  - protocols:
    - Border Gateway Protocol (BGP)
  - Path Vector (PV) routing protocol

BGP

- Each AS advertises IP addresses that it has
  - AS 1930 (U. Aveiro) advertises:
    193.137.168.0 - 193.137.175.255 (193.136.0.0/15)
- All neighboring ASs receiving such advertisement re-advertise them to their neighbors after pre-pending their AS numbers
- The result is that each AS has a routing entry for (193.136.0.0/15) which looks like:
  193.136.0.0/15: AS X_1 , AS X_2 , …, AS 1930

The two main sources of the Internet topology data

- Traceroute data
  - gives a glimpse of the router topology
    - too many vagaries in IP-to-router resolution
  - gives a view of the AS topology
    - many vagaries in IP-to-AS resolution
- BGP data
  - gives another view of the AS topology
    - but there are still some missing links due to sampling biases

Router vs. AS topology

- We do not know the router topology
- We know the AS topology much better

AS relationships and BGP policies

- Each AS link is the relationship (i.e., business, contractual agreement) between the two ASs
- There are roughly three types of such relationships
  - customer-provider (c2p)
  - peer-peer (p2p)
  - sibling-sibling (s2s)
- They stem from combinations of the following two BGP route re-advertisement policies
  - re-advertising to provider or peer, an AS advertises only its own IP addresses and IP routes learnt from its customers
  - re-advertising to customer or sibling, an AS advertises everything
- BGP advertisement policy combinations vs. AS relationships
  - asymmetric combination: c2p
  - symmetric combinations: p2p and s2s

Valid paths

- uphill: zero or more links from customer to provider
- pass: zero or one link from peer to peer
- downhill: zero or more links from provider to customer
- any number of sibling links anywhere in the path

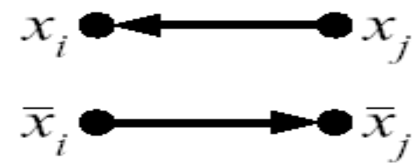
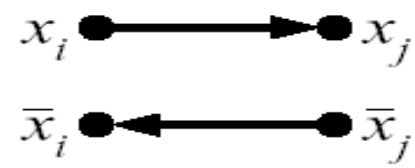
Type of Relationship (ToR) problem formulations

- Given a set of BGP paths *P*,
- Extract the undirected AS-level graph *G*.
  - Every edge in *G* is a link between pair of ASs.
- Assuming edge direction is from customer to provider,
- Direct all edges in *G* (2^m combinations),
- Inducing direction of edges in *P*,
- Such that the number of invalid paths in *P* is minimized.
  - Invalid path is a path containing a provider-to-customer link followed by customer-to-provider link

ToR and MAX2SAT

- Split all paths in *P* into pairs of adjacent links (involving triplets of nodes)
- Perform mapping…

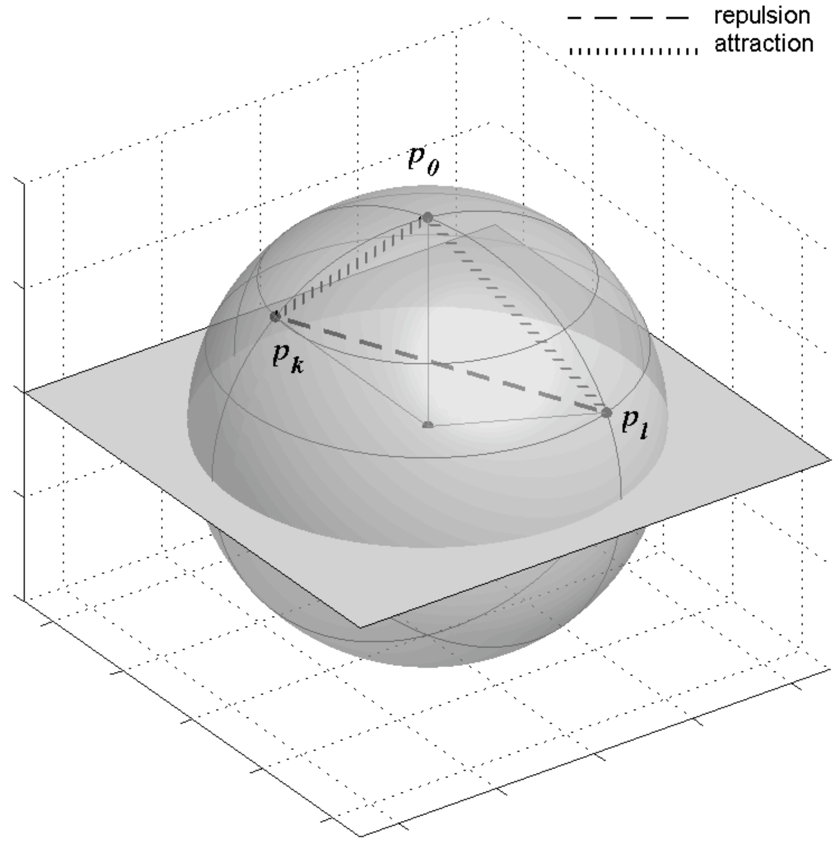
Mapping to MAX2SAT

Edges in P	2SAT clause	Edges in G_{2SAT}
	$x_i \vee x_j$	
	$x_i \vee \bar{x}_j$	
	$\bar{x}_i \vee x_j$	
	$\bar{x}_i \vee \bar{x}_j$	

SDP relaxation to MAX2SAT

$$
\begin{array}{ll}
\max & \displaystyle\frac{1}{4}\sum_{k,l=1}^{2m_1} w_{kl}(3 + v_0 \cdot v_k + v_0 \cdot v_l - v_k \cdot v_l) \\
\text{s.t.} & v_0 \cdot v_0 = v_k \cdot v_k = 1,\ v_i \cdot v_{m_1+i} = -1, \\
 & k = 1 \ldots 2m_1,\ i = 1 \ldots m_1.
\end{array}
$$

Physical interpretation



Infer c2p links using multiobjective optimization

- Maximize number of invalid paths:
  - 2-link clauses $w_{kl}(x_k \vee x_l)$
- Direct along the node degree gradient:
  - 1-link clauses $w_{kk}(x_k \vee x_k)$

Final form of the generalized problem formulation

$$
\begin{array}{ll}
\max & \displaystyle \frac{1}{4}\sum_{k,l=1}^{2m_1} w_{kl}(3 + v_0 \cdot v_k + v_0 \cdot v_l - v_k \cdot v_l) \\
\text{s.t.} & v_0 \cdot v_0 = v_k \cdot v_k = 1,\ v_i \cdot v_{m_1+i} = -1, \\
 & k = 1 \ldots 2m_1,\ i = 1 \ldots m_1.
\end{array}
$$

$$
w_{kl}(\alpha) = \begin{cases} c_2\alpha & \text{if} \quad \{kl\} \in P, \\ c_1(1-\alpha) f(d_k^-, d_k^+) & \text{if} \quad k = l \leqslant m_1, \\ 0 & \text{otherwise.} \end{cases}
$$

$$
f(d_i^-, d_i^+) = \frac{d_i^+ - d_i^-}{d_i^+ + d_i^-} \log(d_i^+ + d_i^-).
$$

AS relationship results

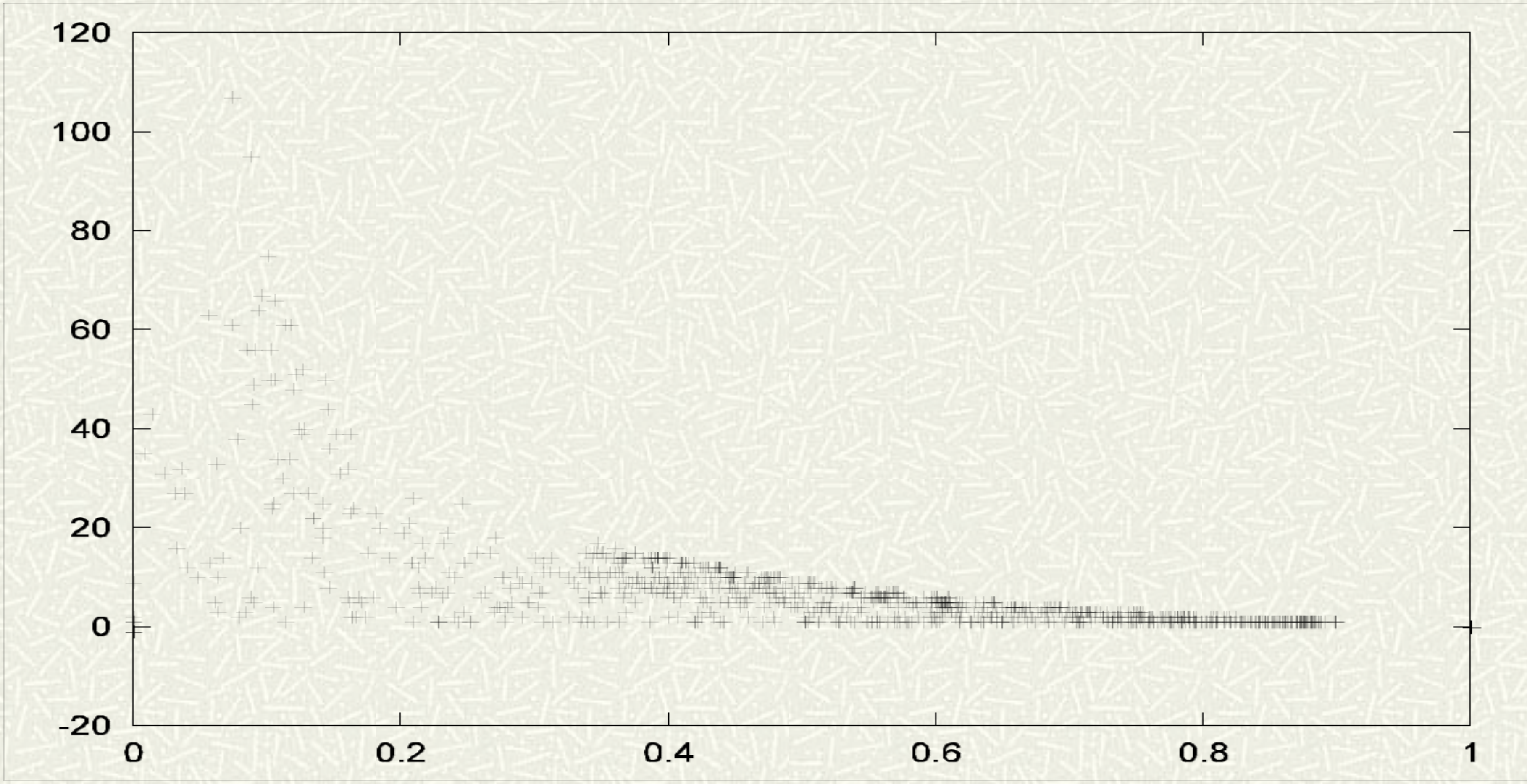
- Input: RouteViews, 8-hour interval snapshots between 03/01/05 and 03/05/05
- Output:

	Total $\|E\|$	c2p links $\|E \setminus F \setminus S\|$	p2p links $\|F \setminus S\|$	s2s links $\|S\|$
number of links	38,282	34,552	3,553	177
percentage	100%	90.26%	9.28%	0.46%

AS hierarchy

				$\alpha = 0.00$		$\alpha = 0.01$		$\alpha = 0.05$		$\alpha = 0.10$		$\alpha = 0.50$		$\alpha = 1.00$	
Percentage of invalid paths															
				12.75%		1.79%		0.69%		0.46%		0.36%		0.33%	
Top of reachability based hierarchy															
	AS #	name	degree	dep.	wid.	dep.	wid.	dep.	wid.	dep.	wid.	dep.	wid.	dep.	wid.
$\alpha = 0$	701	UUNET	2334	0	1	1	1	0	105	0	120	2	201	11	319
	7018	AT&T	1911	1	1	2	1	0	105	0	120	2	201	11	319
	1239	Sprint	1703	2	1	0	1	0	105	0	120	2	201	11	319
	3356	Level 3	1228	3	1	3	1	0	105	0	120	2	201	11	319
	209	Qwest	1105	4	1	4	1	0	105	0	120	2	201	11	319
$\alpha = 1$	14551	UUNET	35	128	1	137	2	138	1	151	1	260	2	0	1
	13987	IBASIS Inc.	3	1792	955	1802	963	1830	976	1847	971	1885	966	1	2
	8631	Routing Arbiter	48	108	1	123	1	122	2	0	120	0	1	1	2
	23649	Hong Kong Teleport	4	1792	955	1802	963	899	121	916	121	967	119	3	8
	4474	Village Communications	2	2747	16136	2765	16118	2806	16077	2818	16065	2	201	3	8

Phase transition in mean field approximation



Validation

- Previous validation efforts
  - Gao: AT&T
  - SARK: Gao
  - Subsequent: SARK/Gao
- Our validation
  - 38 ASs (5 Tier-1 ISPs, 13 smaller ISPs, 19 universities, and 1 content provider)
  - 3,724 links (9,7% of the total)
  - 94.2% overall accuracy

	links	inferred c2p links	inferred p2p links	inferred s2s links
total number of	3,724	3,070	623	31
number of correct	3,508	2,964	516	28
percentage of correct	94.2%	96.5%	82.8%	90.3%

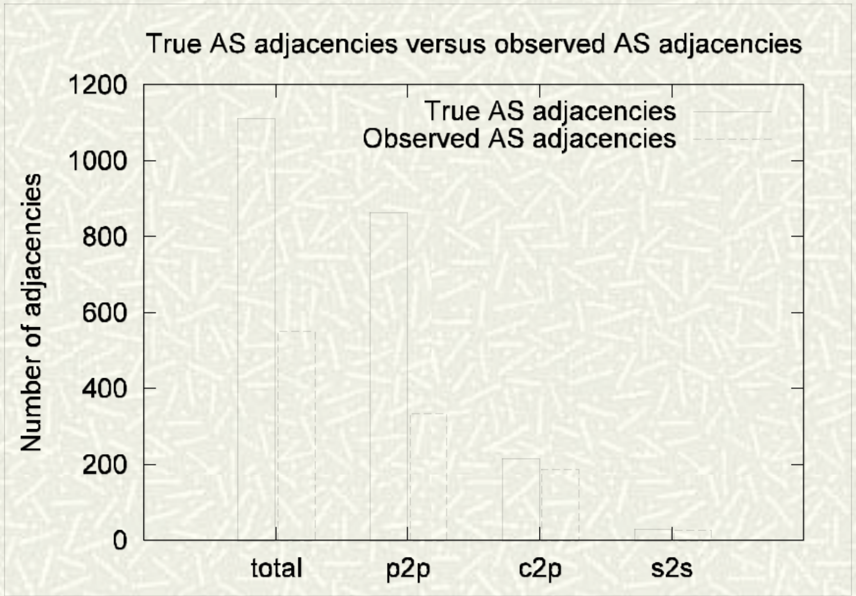
Questions in the questionnaire

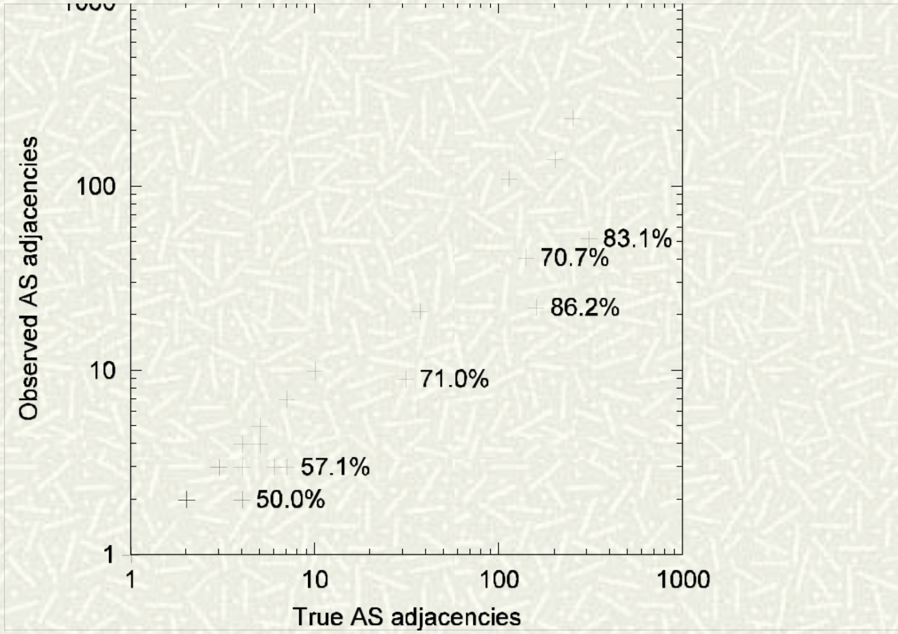
- For the listed inferred AS relationships, specify how many are incorrect, and what are the correct types of the relationships that we mis-inferred?
- What fraction of the total number of your AS neighbors is included in our list?
- Can you describe any AS relationships, more complex than c2p, p2p, or s2s, that are used in your networks?

Missing links

- 27 (3 tier-1 ISPs) out of 38 answered the second question, too, and provided us with their full AS relationship data: 1,114 links
- Among these, we see only 552 (49.6%):
  - 38.7% out of the 865 (77.6%) p2p links
  - 86.7% out of the 218 (19.6%) c2p links
  - 93.3% out of the 30 (2.7%) s2s links
- Maximum percentage of missing links per node is 86.2% (50% of ASs miss >70% links)

Missing links visualized





More complex policies

- Space
- Time
- Prefix

AS taxonomy

- Assign the following six attributes to every AS
  - organization description (IRR data, stop words are filtered out and the rest of words are stemmed)
  - number of customers
  - number of providers
  - number of peers
  - number of advertised IP prefixed
  - size of the advertised IP address space
- Feed this data into a machine learning algorithm (AdaBoost) with a training set of 1200 ASs
- Classify all ASs into the following six categories
  - Large ISPs
  - Small ISPs
  - Customer ASs
  - Universities
  - IXPs
  - NICs

AS taxonomy results

Classified 95.3% of ASs (non-abstained) with expected accuracy of 78.1%

	Large ISPs	Small ISPs	Customer ASes	Universities	IXPs	NICs
ASes	44	5,599	11,729	877	33	332
%	0.2	30.1	63.0	4.7	0.2	1.8

AS rank

That’s not all we now about the Internet but it’s pretty much all we know about the Internet AS topology ☺

Thank you!