AS Assignment for Routers

Bradley Huffaker
bradley@caida.org

Amogh Dhamdhere, Marina Fomenkov, kc claffy

CAIDA
University of California
at San Diego, La Jolla, CA

AIMS Workshop -- February 2010

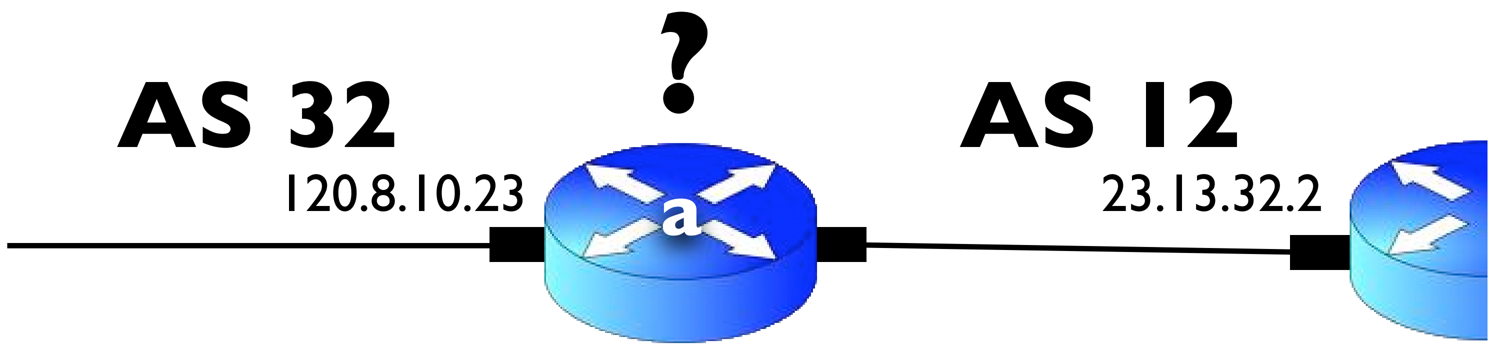
Overview

- motivation
- methodology
- analysis
- conclusions

AS Assignment Problem

motivation

Which AS, 32 or 12, owns/controls the router **a**?



IP address	120.8.10.23	23.13.32.2
prefix	120.8.10.0/24	23.13.0.0/16
AS	32	12
router	?	

BGP longest matching prefix — BGP origin AS — As assignment

23.13.32.2 → 23.13.0.0/16 → 12

120.8.10.23 → 120.8.10.0/24 → 32

12, 32 → ?

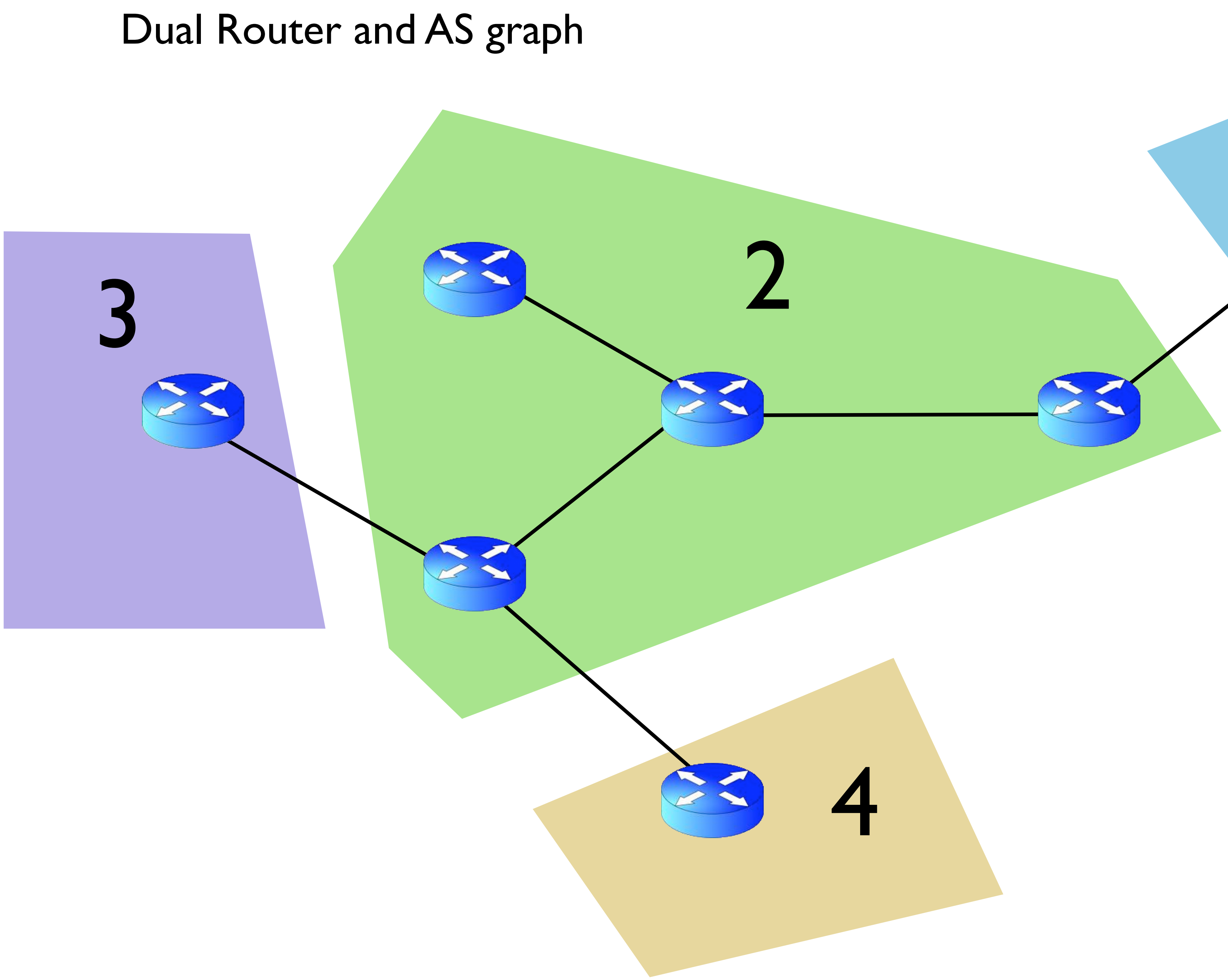
Motivation

- Dual graph
  - a combined router and AS graph
- Dual graph analysis
  - Relationship between AS degree and the AS's number of routers.
  - how does heuristic assignment affect the inferred number of routers in an AS
- More accurate AS traceroute
  - resolving AS loops

Here is What We Want

motivation

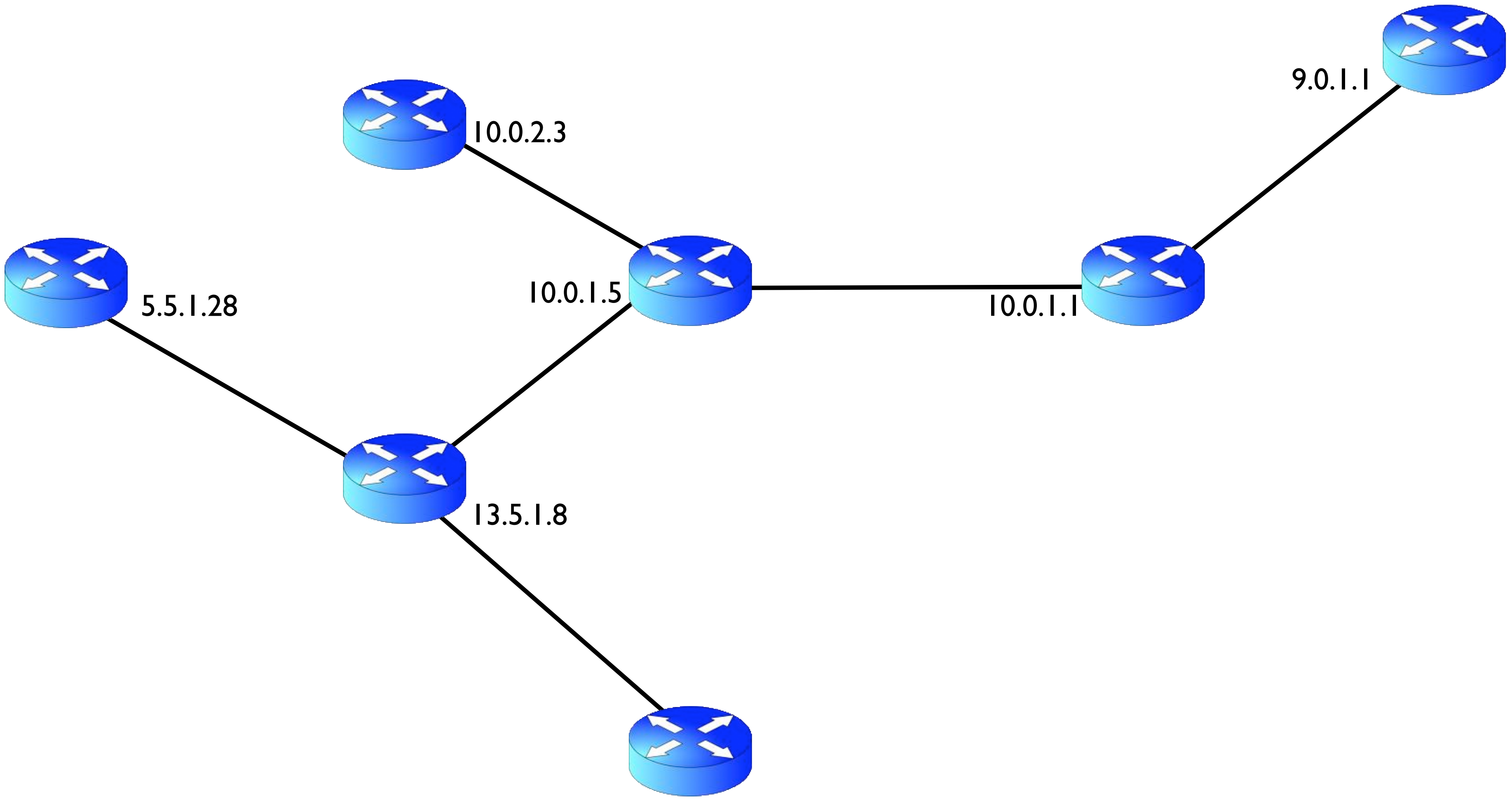
Dual Router and AS graph



This Is What We Have

motivation

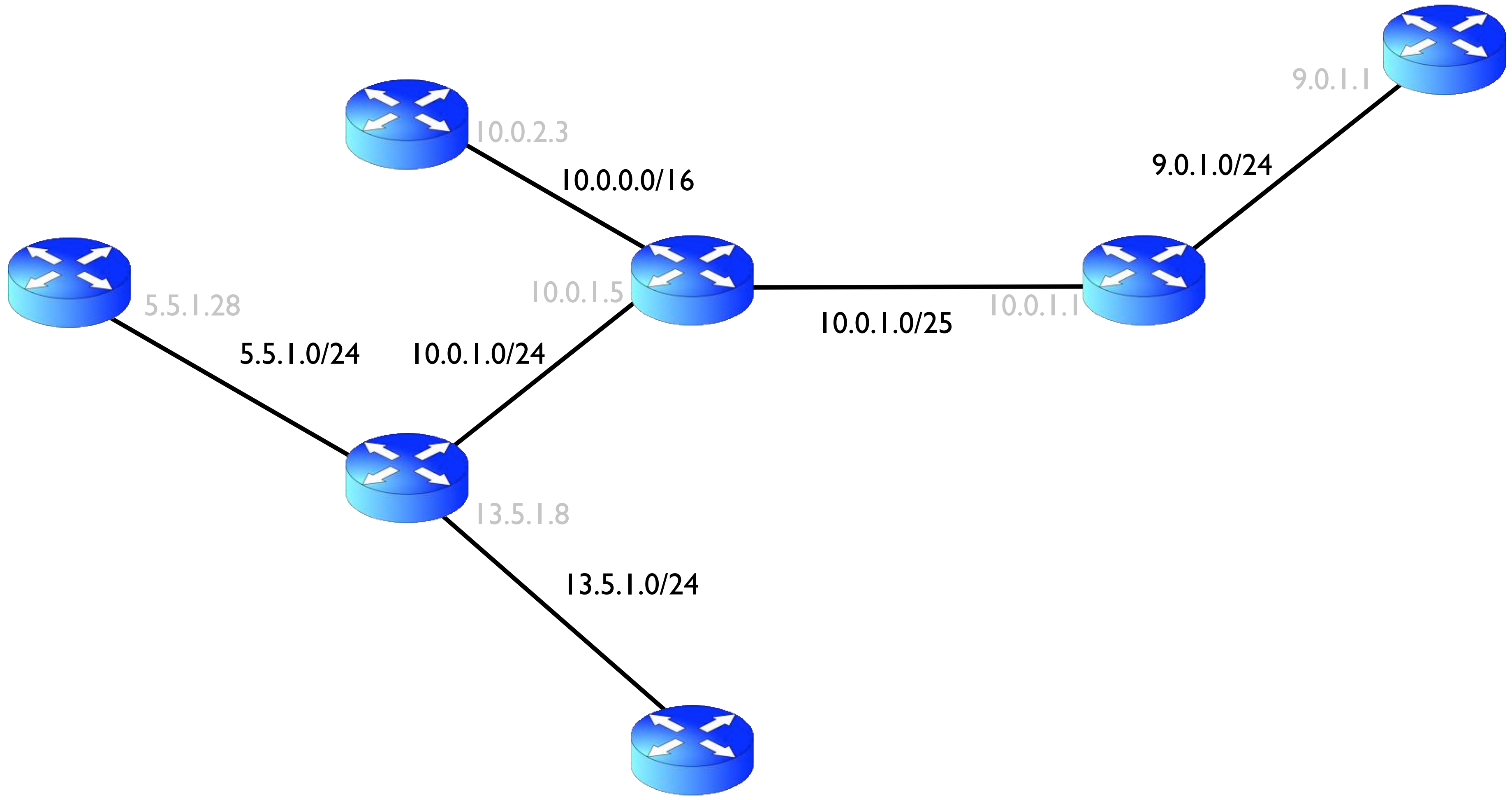
Router graph with interfaces.



Mapping to Prefix

motivation

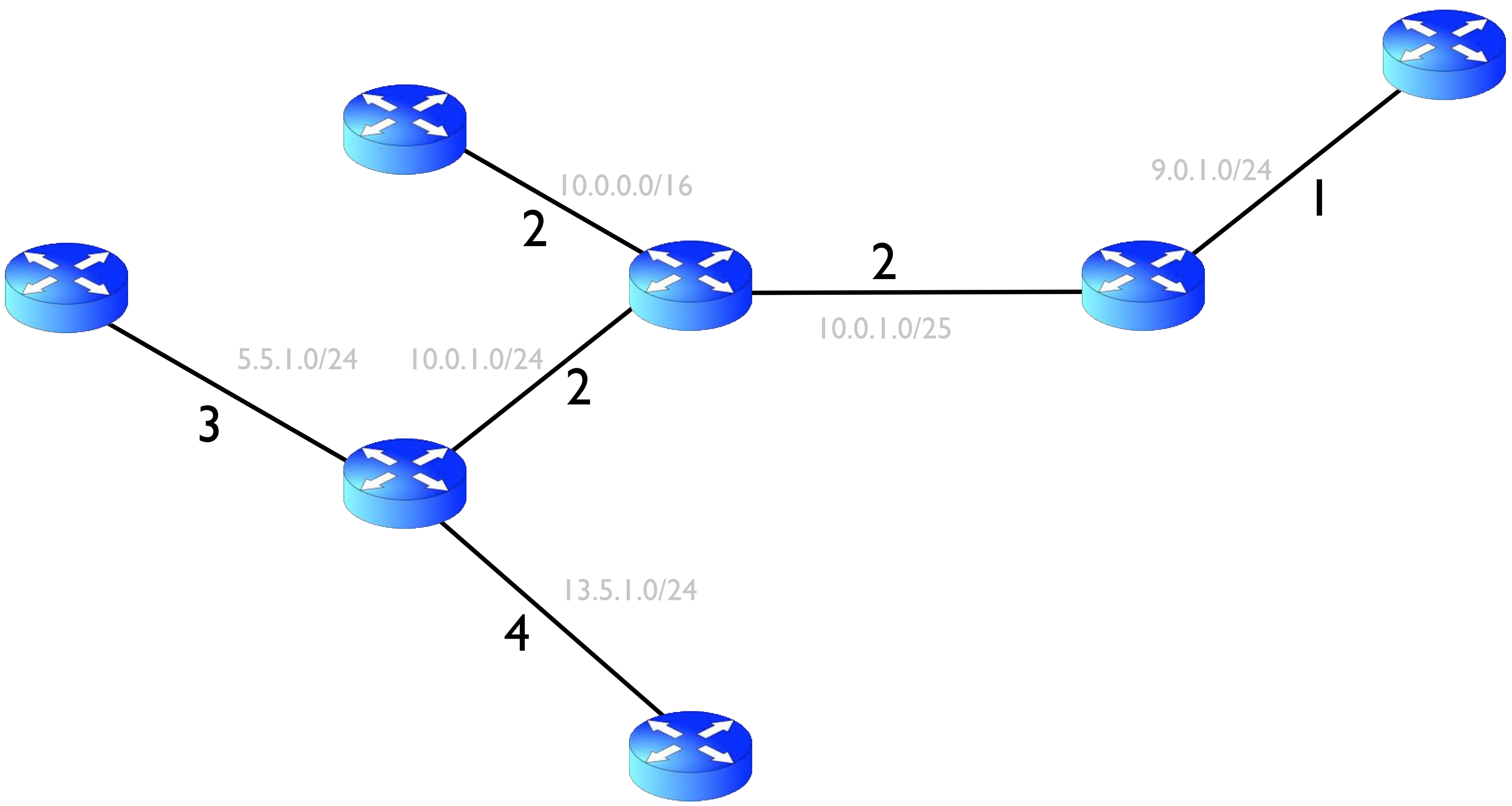
Router graph with prefixes assigned to links.



Mapping to ASes

motivation

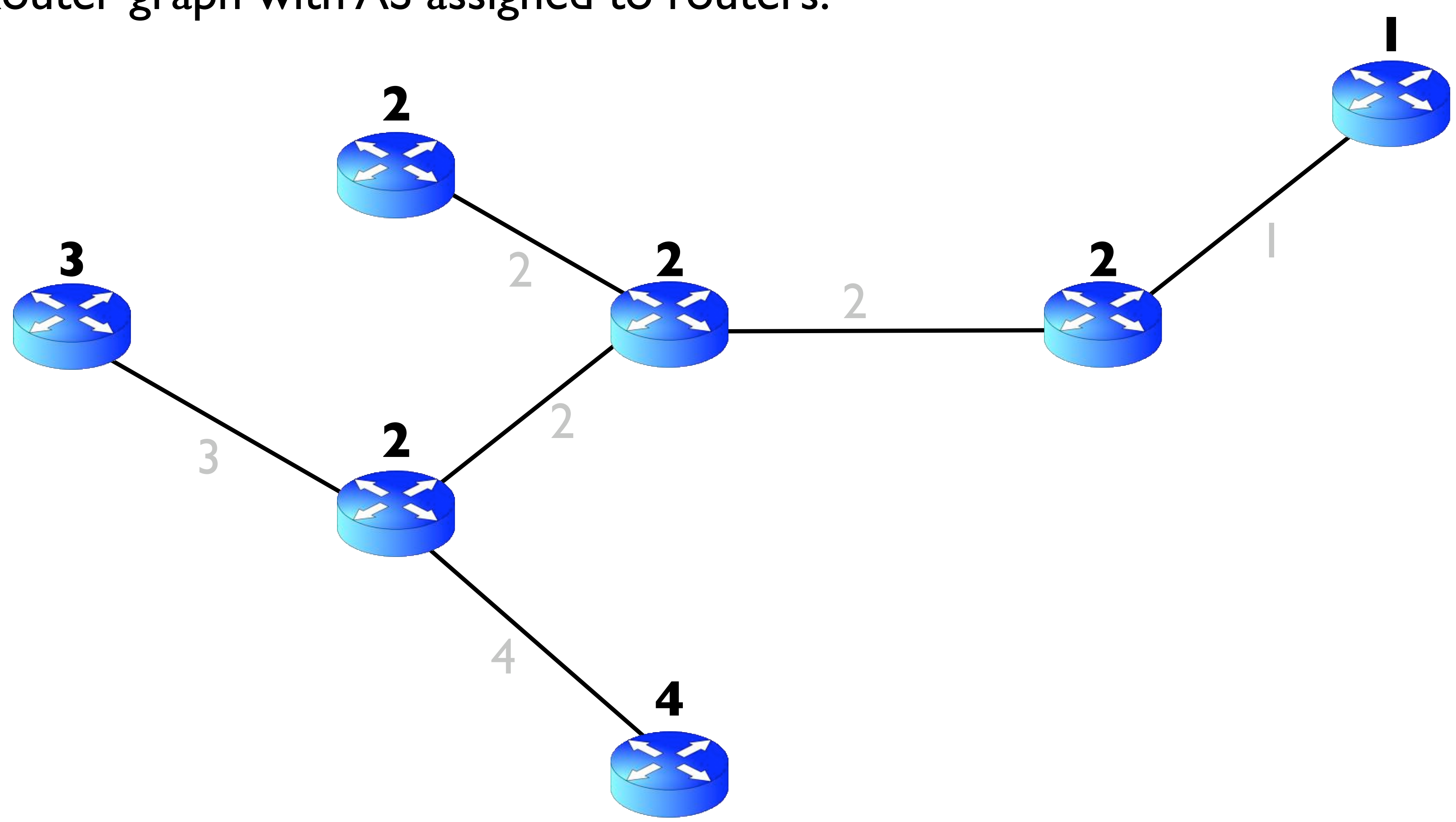
Router graph with AS assigned to links.



Assigning AS to Routers

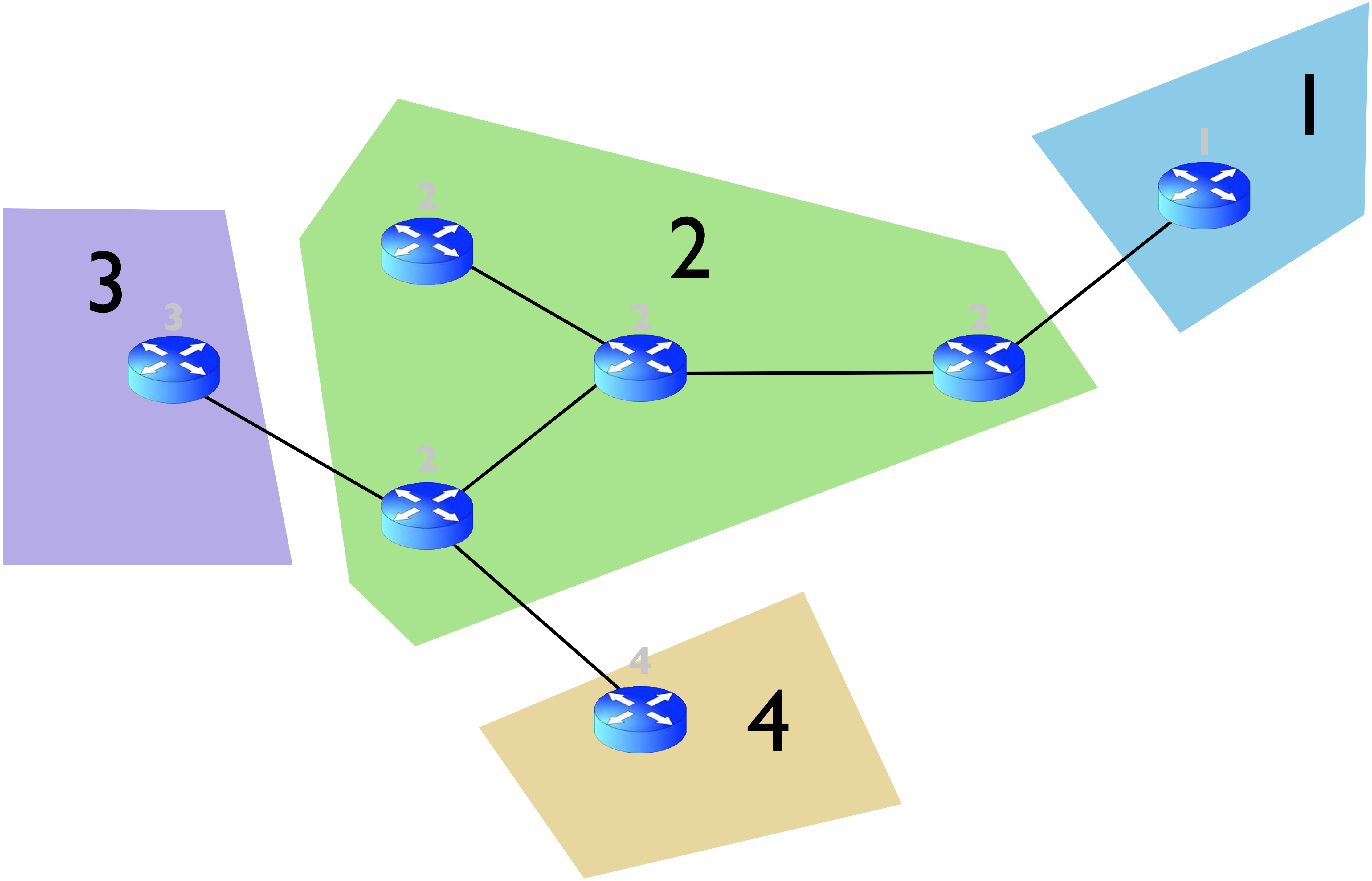
motivation

Router graph with AS assigned to routers.



Dual Graph

motivation



Methodology

We compared the success rates of four different AS assignment heuristics against our ground truth data sets.

Ground Truth

methodology

- **ISPs (i)**
  - Tier 1, Tier 2, and five research networks
- **interface sets**
  - $\mathbf{I}_i$ interfaces in the address space of ISP_i, on routers that do belong to ISP_i
  - $\overline{\mathbf{I}}_i$ interfaces in the address space of ISP_i on routers that do **not** belong to ISP_i
- **router sets**
  - $\mathbf{R}_i$ is the set of routers with interfaces in the address space of ISP_i that do belong to ISP_i
  - $\overline{\mathbf{R}}_i$ is the set of routers with interfaces in the address space of ISP_i that do **not** belong to ISP_i
- **AS sets**
  - $\mathbf{A}_i$ is the set of ASes that do belong to ISP_i
  - $\overline{\mathbf{A}}_i$ is the set of ASes that do **not** belong to ISP_i

Ground Truth

methodology

	R routers owned	$\overline{\mathbf{R}}$ routers not owned
Tier 1[f,h]	3,405	2,254
Tier 2[h]	241	86
GEANT[f]	37	0
I-Light[f]	32	0
Internet 2[f]	17	0
National LambdaRail[f]	16	0
CANET[f]	8	0

[f] Organization provided **f**ull interface list
[h] Organization provided naming **h**euristic that allowed for inference of $\overline{\mathbf{R}}$

Data sources

methodology

- Router Graph (MAARS[1])
  - Sept. - Oct. 2009
  - 268 million traceroute paths
  - 22 million nodes[2] / 22 million links[3]

- BGP Data
  - Oct. 2009
  - 311,230 prefixes

- AS relationships
  - Oct. 2009
  - BGP data
  - 148,565 AS relationship pairs

[1] router alias resolver
[2] node = set of IPs on same router
[3] link can connect > 2 nodes

Data Topology

methodology

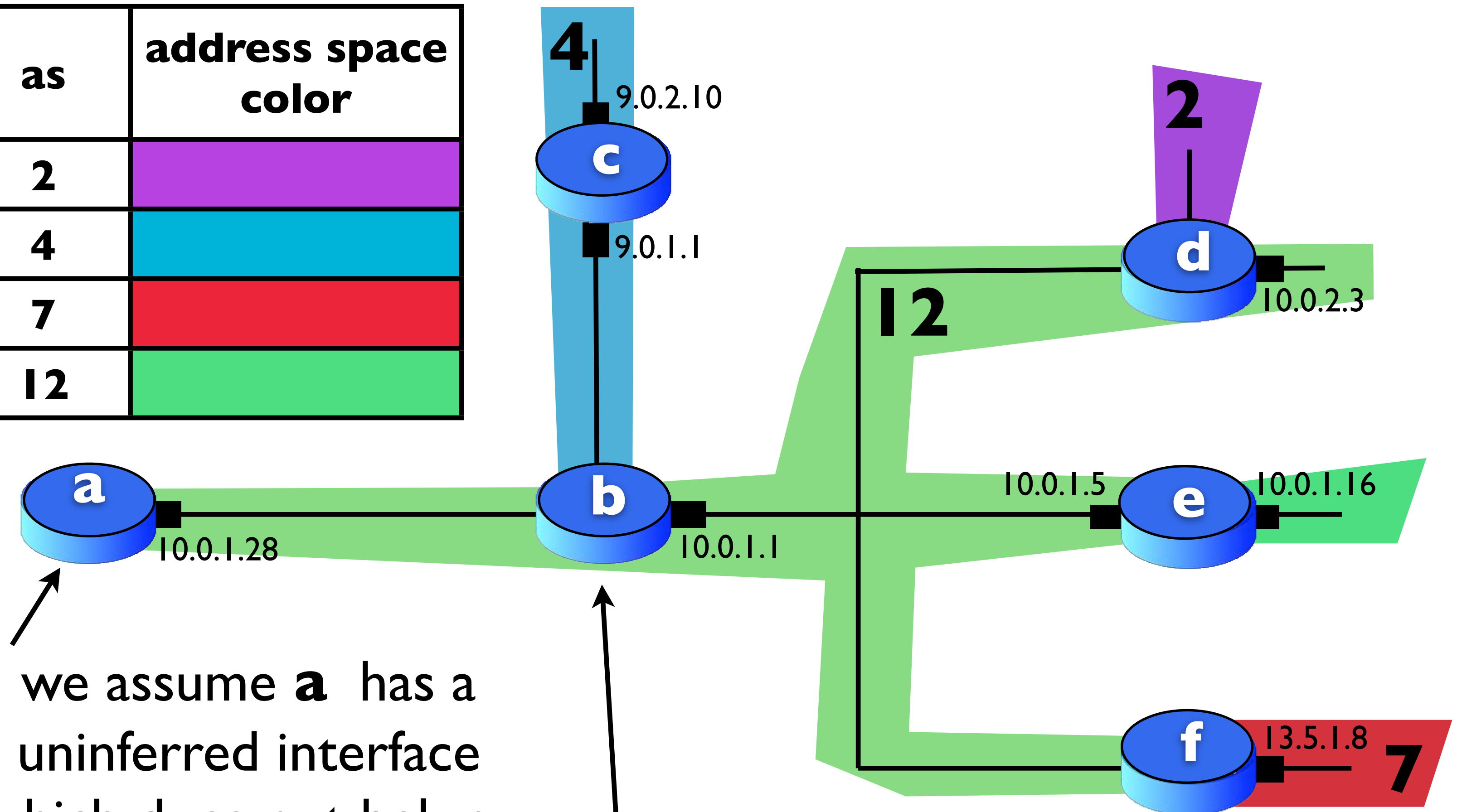
—	Interface sets
I_{12}	10.0.1.1, 10.0.2.3, 10.0.1.6
I_{12}	10.0.1.28

—	router sets
R_{12}	**b, d, f**
R_{12}	**a**

—	AS sets
A_{12}	**12**
A_{12}	**4, 2, 7**

route	AS	type
a	**12**	single-AS
b	**4, 12**	multi-AS
c	**4**	single-AS
d	**2, 12**	multi-AS
e	**12**	single-AS
f	**12, 7**	multi-AS

as	address space color
2	
4	
7	
12	



we assume **a** has a uninferred interface which does not belong to **12**

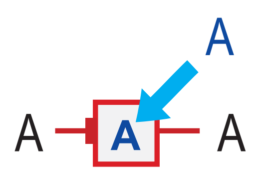
b gets candidate AS from its interface 10.0.1.1 and the link it shares with **c**.

f has no interface in $\mathbf{I_{12}}$ and $\mathbf{\bar{I}_{12}}$, so has no known ownership

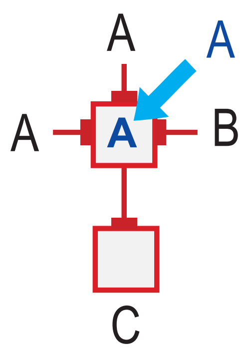
AS assignment methods

methodology

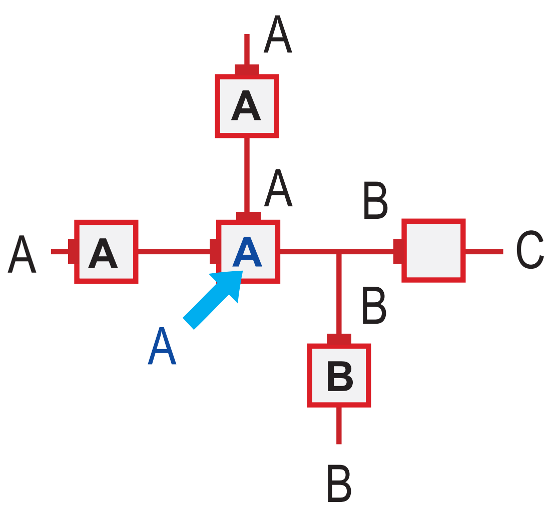
Single

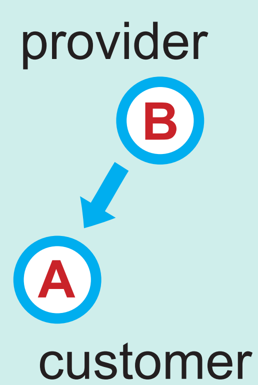


Election

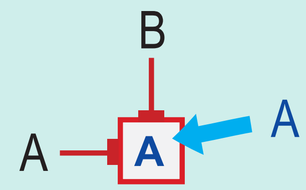


Neighbor



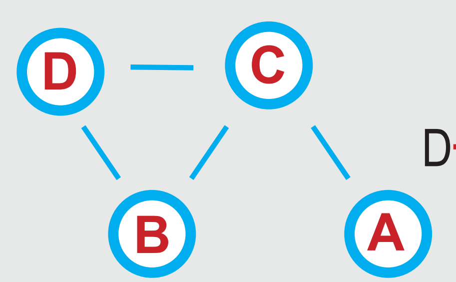


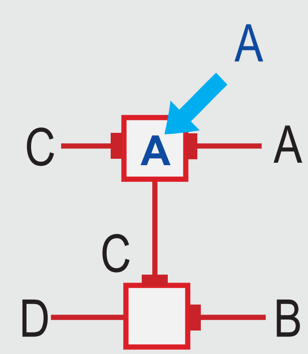
Customer



Degree

AS DEGREE	
A	1
B	2
C	3





Single: only one choice

Election: most interfaces
- more links into router's ISP's address space

Neighbor: most single AS neighbors
- connected to more routers owned by the router's ISP

Customer: customer AS
- customer's router uses provider's address space for the interconnect

Degree: smallest degree AS
- proxy for Customer, large degree AS typically is provider of small degree AS

Methodology

- primary method
  - assignment is used if it is not ambiguous
- tie-breaker method
  - method with highest success rate on routers for which primary method yields ambiguous results

	ambiguous
election	no majority AS among links
neighbor	no majority AS among neighbors
customer	no unambiguous customer relationship among ASes
degree	tie between smallest degree ASes

counting success?

successful assignment:

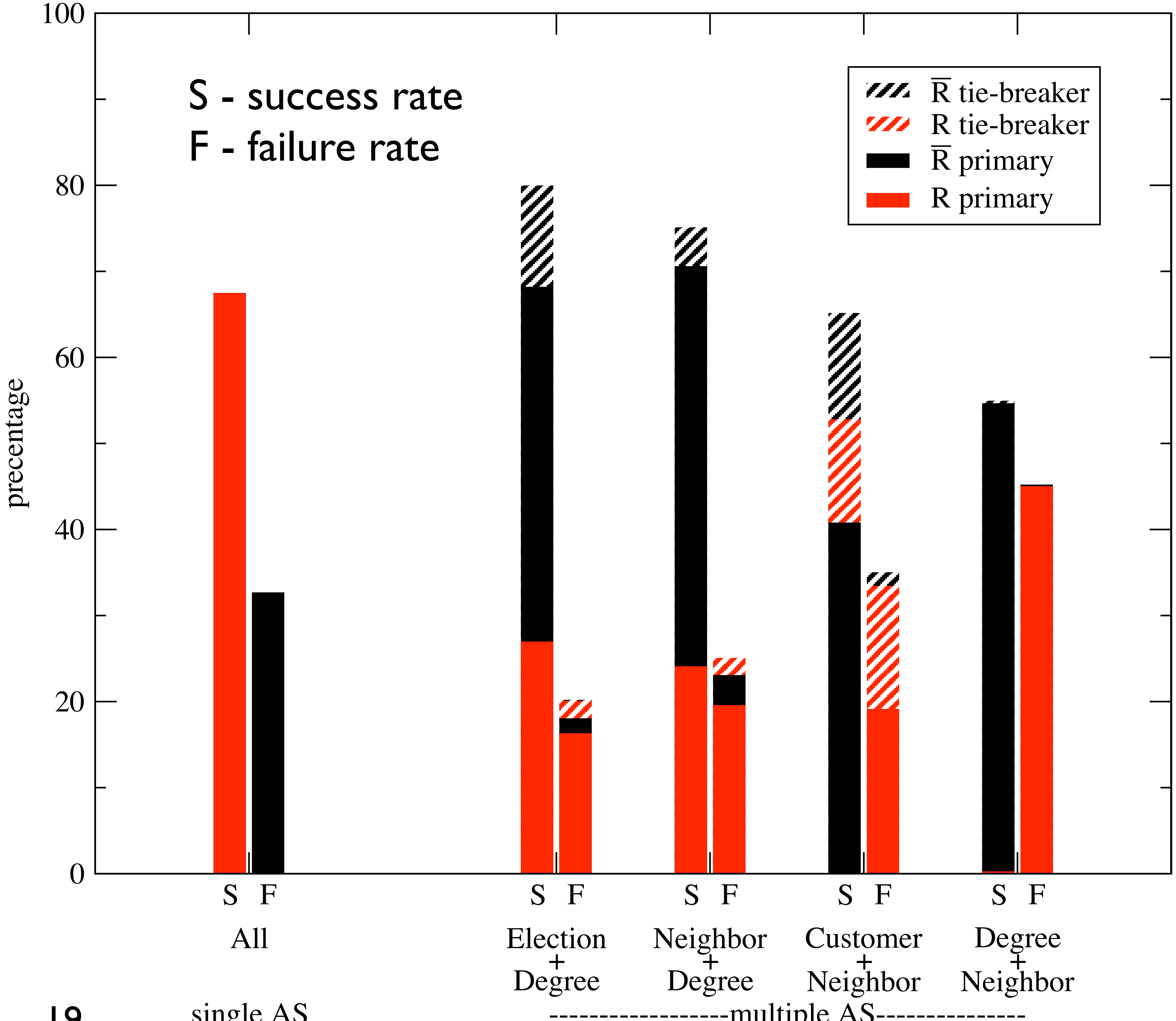
If router r is known to be owned by ISP_i and method H(r) selects an AS owned by ISP_i,

or

if r is known to not be owned by an ISP_i and method H(r) selects an AS not owned by ISP_i.

Method Success Rates

analysis



72%

28%

Election + Degree performs best with **80% success** rate.

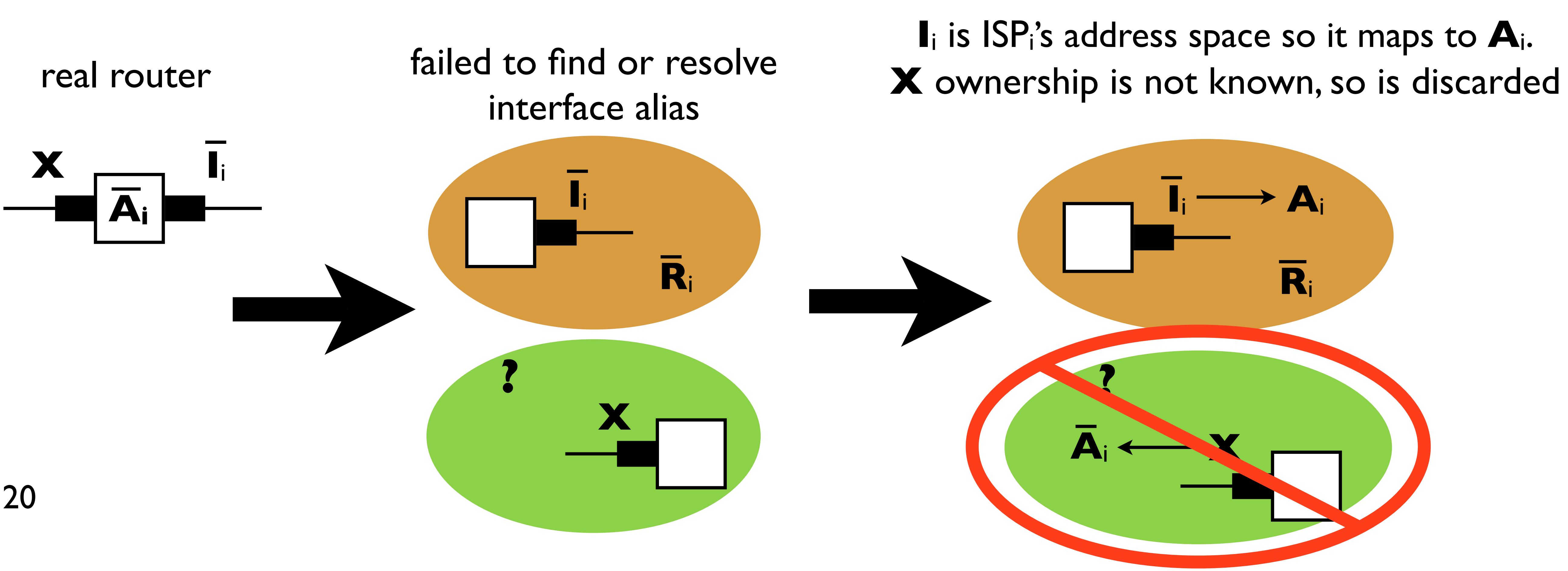
Tier 1 bias in ground truth reduces accuracy of customer and degree heuristics

Tie-breaker ambiguous assignments not counted

Success Rates

- single AS routers
  - all methods successful for R (67% of single AS routers)
  - all methods fail for $\bar{R}$ (33% of single AS routers)

    routers in $\bar{R}_i$ must have an interface in A_i, therefore single AS routers only have an AS in A_i, making it impossible for any method to select an AS in $\bar{A}_i$.

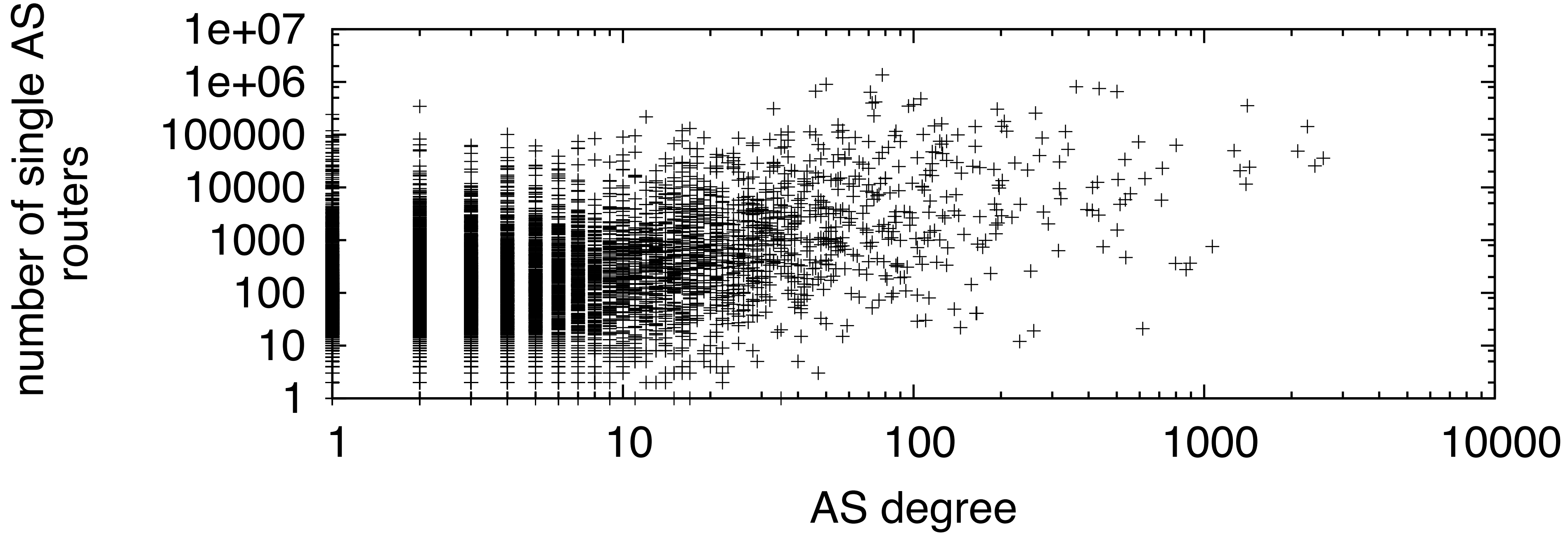


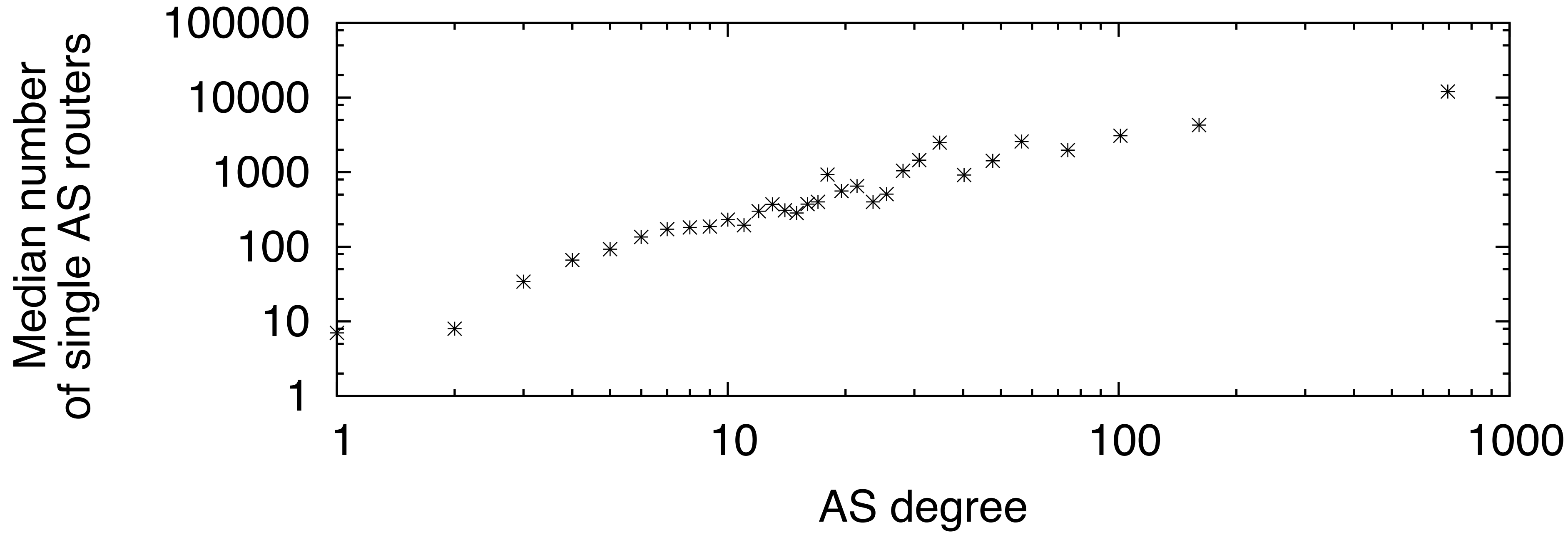
Success Rates

analysis

- multiple AS routers (28%)
  - **Election + Degree** best with **80% success** rate.
- single AS routers (72%)
  - all methods successful for R (67% of single AS routers)
  - all methods fail for $\overline{R}$ (33% of single AS routers)
- overall
  - **Election + Degree** best with **70% success** rate.

Analysis of Dual Topology



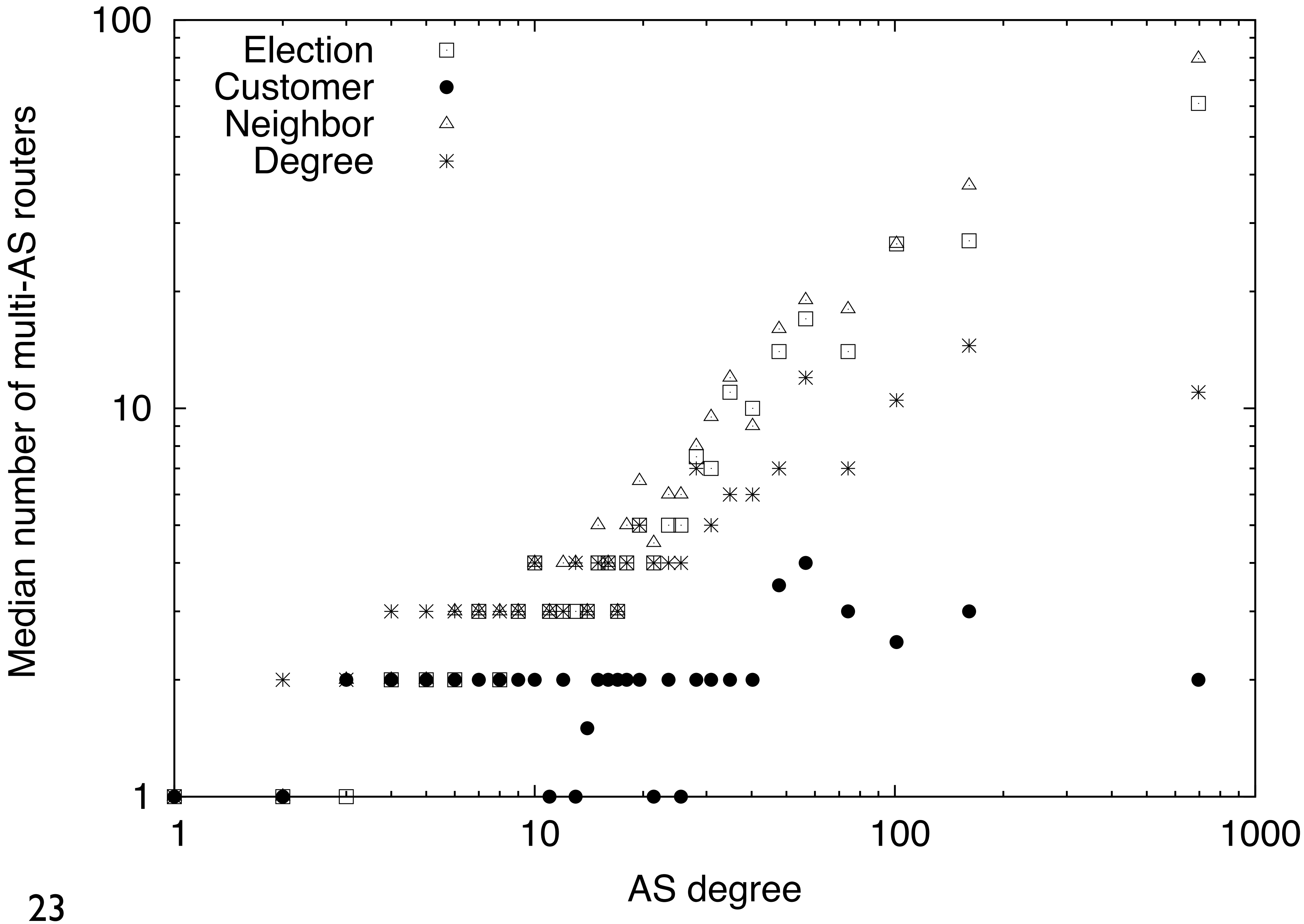


statistical correlation that we can use for topology scaling and generation

Heuristic Effect on AS Router Count

analysis

how do different heuristics affect number of inferred routers per AS



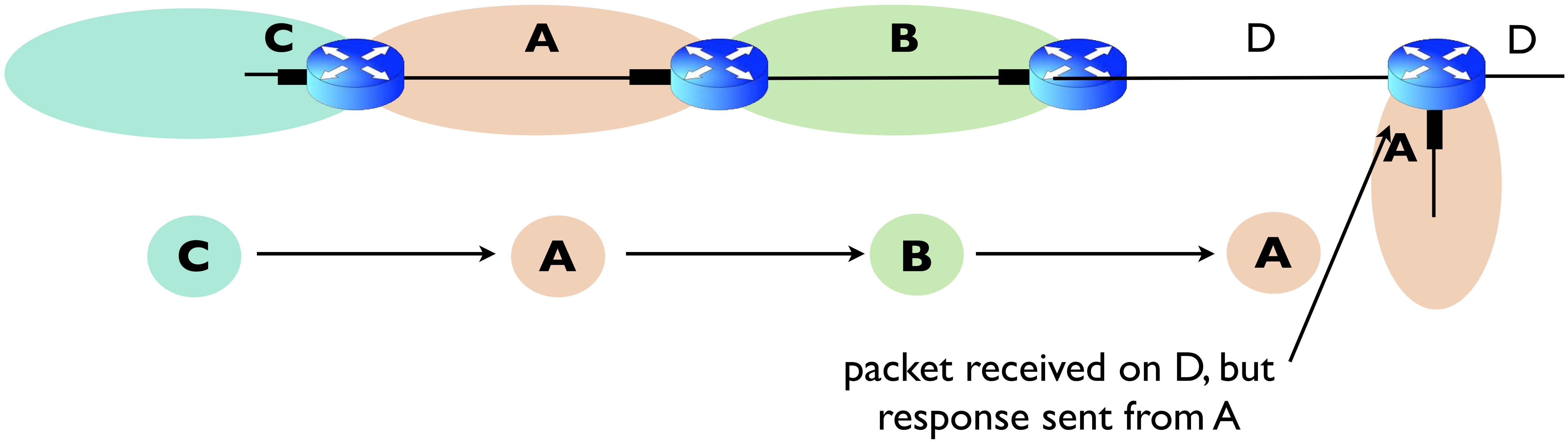
Neighbor assigns more nodes to large degree ASes

Customer assigns more nodes to small degree ASes

Resolving AS Loops

analysis

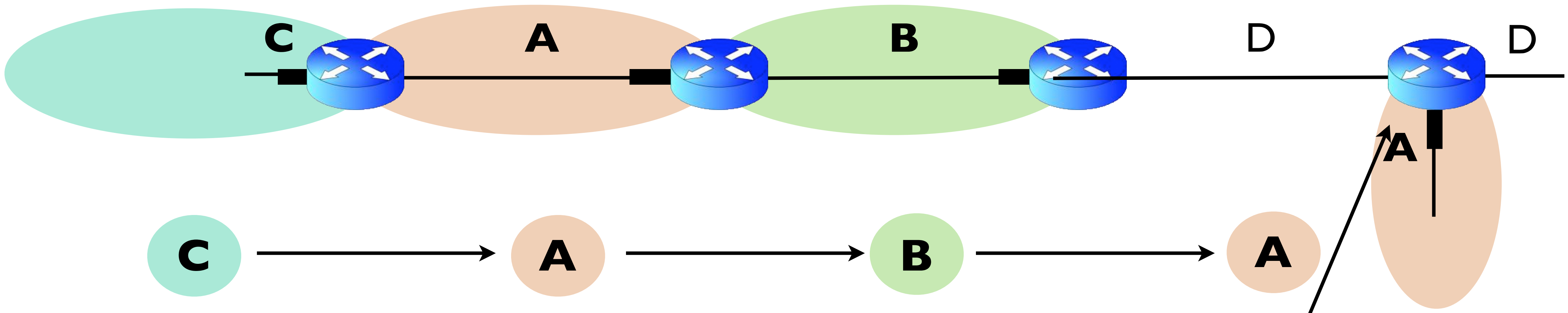
interface/link path



Resolving AS Loops

analysis

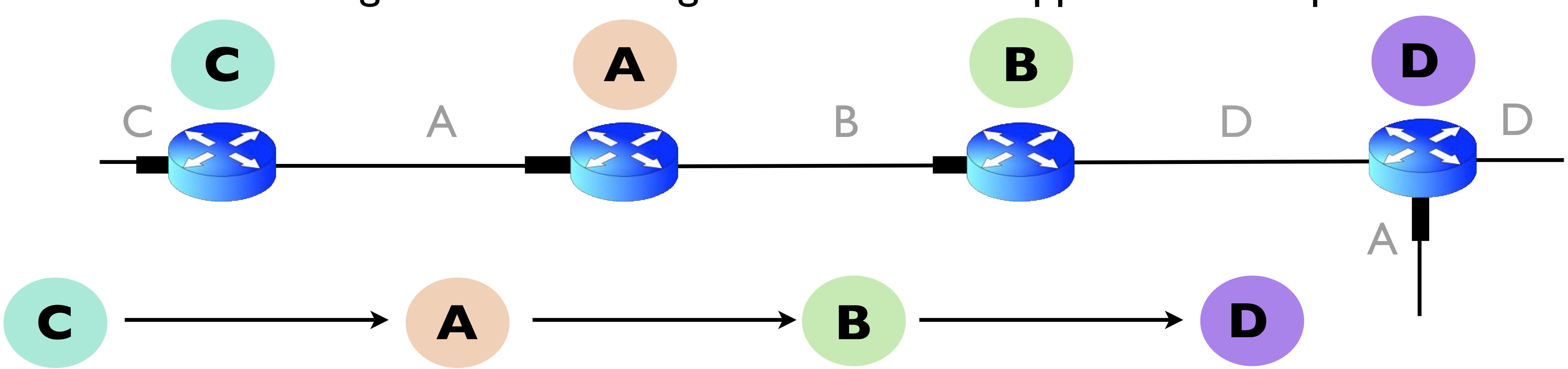
interface/link path



packet received on D, but
response sent from A

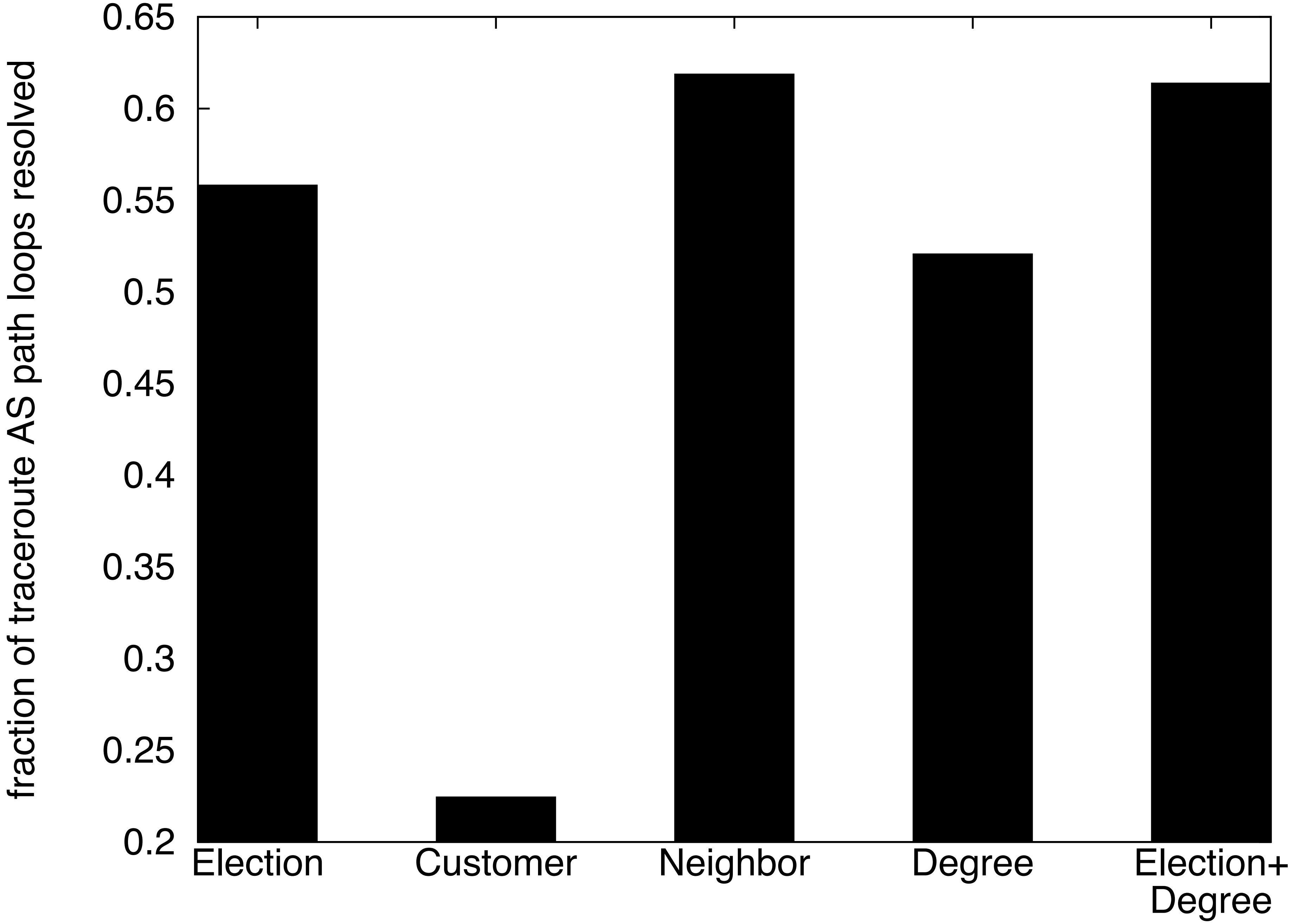
router path

Using inferred AS assignments resolves apparent AS loop.



Resolved AS Loops

analysis



Neighbor resolved the most loops with 63%.

Election+Degree (the combination with the greatest success rate) **resolves 62%** of AS loops

1~5% of paths contain AS loops, depending on the monitor.

Conclusion

- multiple AS routers
  - **Election + Degree** best with **80% success** rate.
- all routers
  - **Election + Degree** best with **70% success** rate.
- AS loop resolution
  - **Election+Degree resolves 62%** or AS loops

Future Work/What we need

- More ground truth
- alternative AS assignment heuristics

Bradley Huffaker bradley@caida.org

http://www.caida.org/publications/papers/2010/as_assignment/