Applied Networking Research Workshop 2020

On the Accuracy of Country-Level IP Geolocation

Ioana Livadariu, Thomas Dreibholz, Anas Saeed Al-Selwi
Haakon Bryhni, Olav Lysne, Steinar Bjørnstad, Ahmed Elmokashfi

simulamet
Simula Metropolitan Center for Digital Engineering AS

IP geolocation is an open research area

Geolocating IP addresses:

- Edge vs core of the Internet
- User-centric vs research oriented

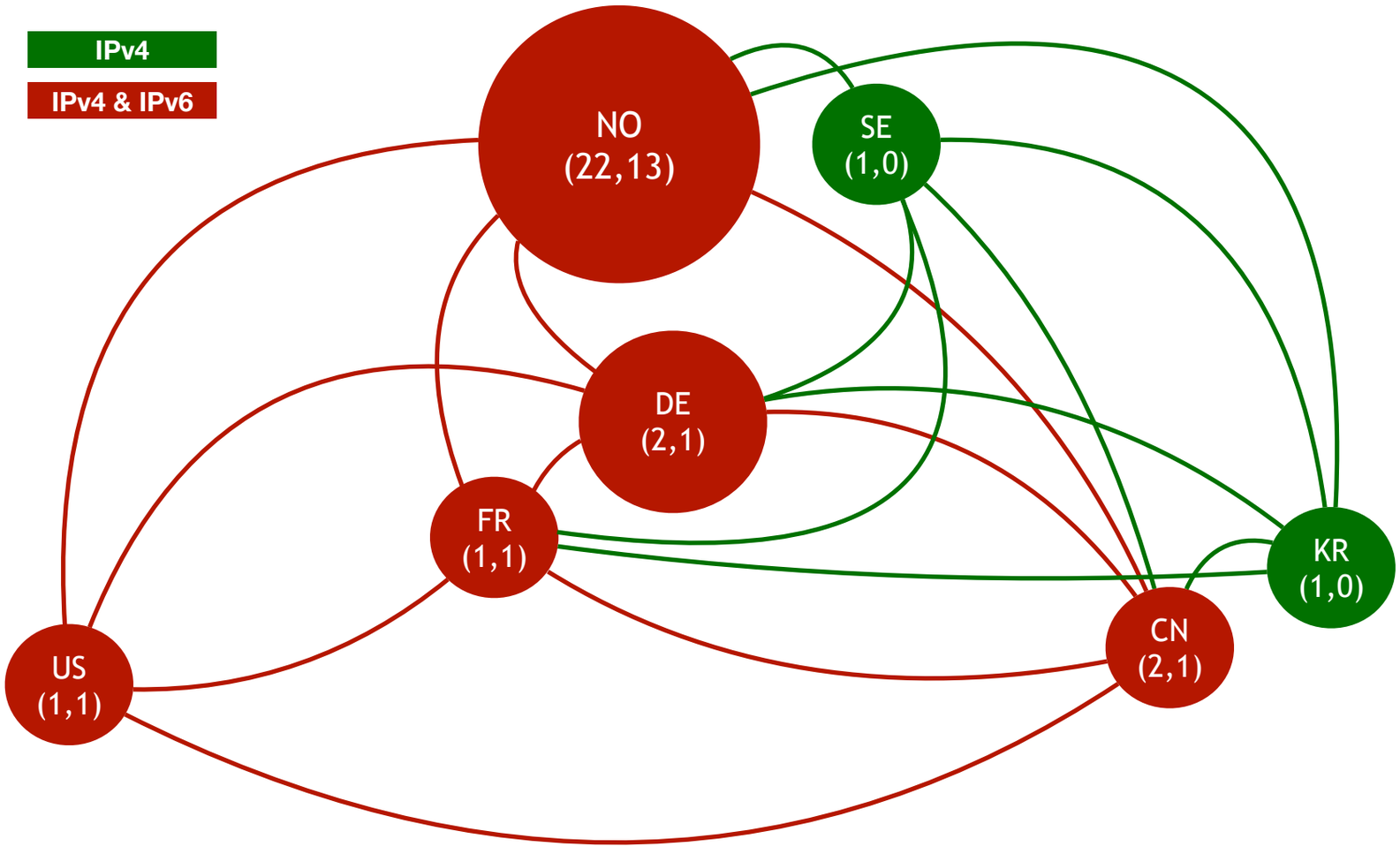
Geolocating approaches:

- Commercial Geolocation Databases (e.g. MaxMind*, IP2Location**,NetAcuity***)
- Measurement-based approaches (latency, geo-hints in DNS names)
- Evaluate the IP geolocating datasets.

Evaluate IP geolocation by studying country-level end-to-end path geo-mappings.

*MaxMind, https://www.maxmind.com/en/home
**IP2Location Lite, https://lite.ip2location.com/
***NetAcuity, https://www.digitalenvoy.com/

Measurement Setup and Collected Data



Geolocation datasets: overview

MaxMind and IP2Location: Dedicated IP geolocation datasets (commercial and *free* version)

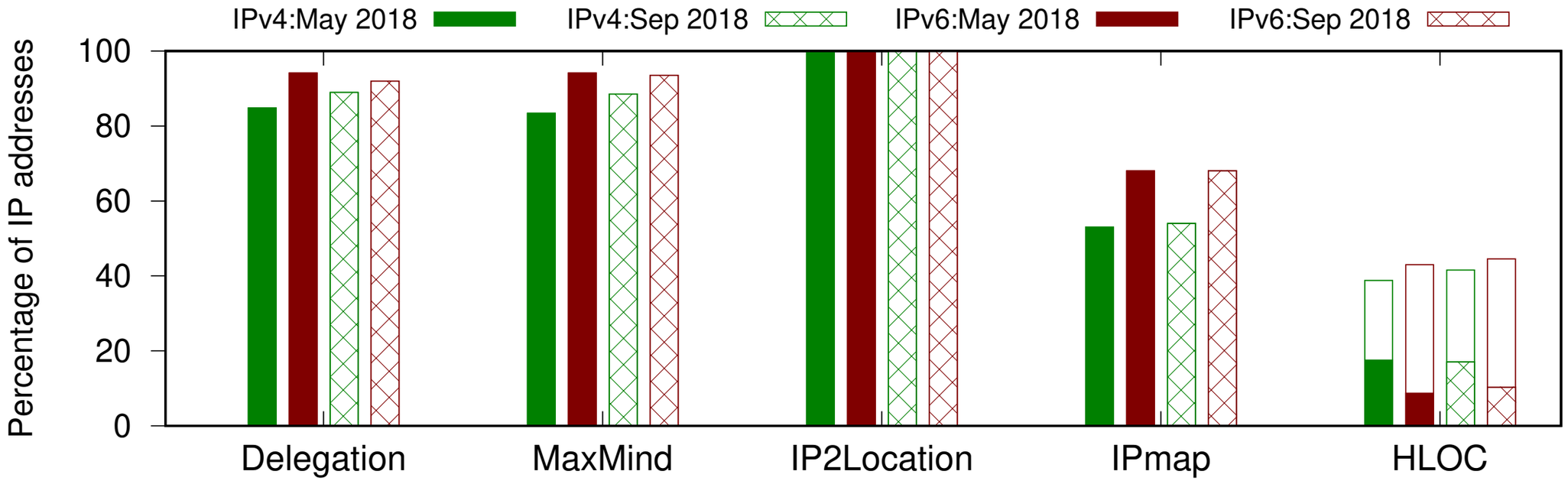
RIR Delegation Files: Daily published by the Regional Internet Registry. Contains registration information regarding Internet resources (IP addresses)

IPmap: IP geolocation approach that uses crowdsourcing and active measurements

HLOC: IP geolocation active-based approaches that use geo-hints and active measurements to geolocate IP addresses

Massimo Candela,RIPE IPmap - What's Under the Hood?, RIPE Labs, 2019
Scheitle et al., "HLOC: Hints-based geolocation leveraging multiple measurement frameworks", TMA 2017
Gharaibeh et al., "A look at Router Geolocation in Public and Commercial Databases", IMC 2017

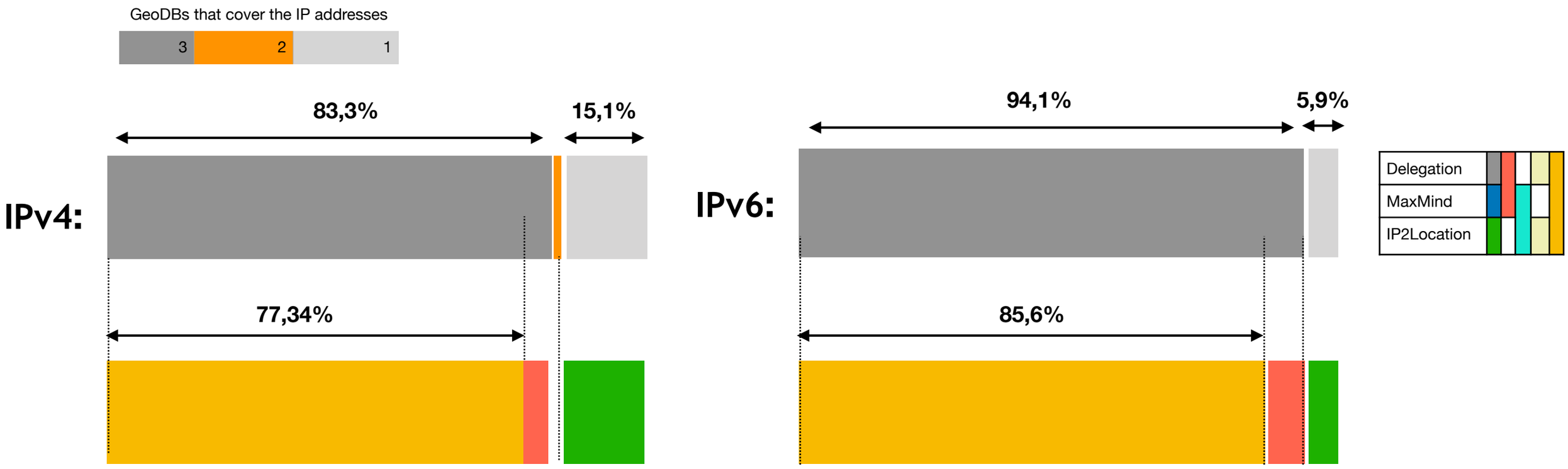
Geolocation dataset IP coverage



Delegation, **MaxMind** and **IP2Location** cover more at least 80% of our collected IP addresses.

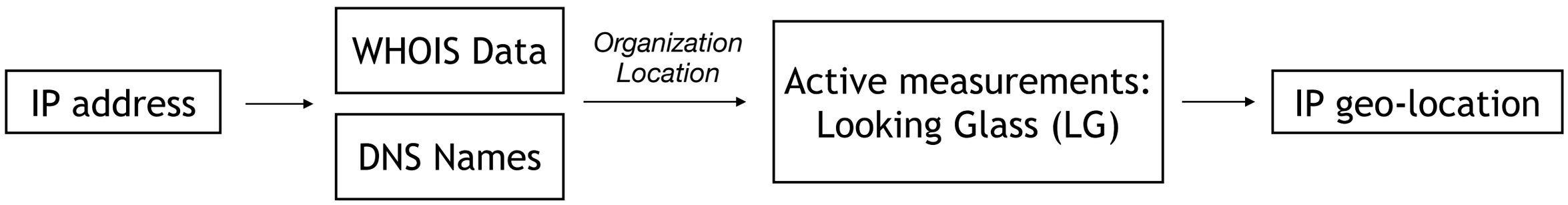
IPmap and **HLOC** have limited coverage of the IP addresses.

How many IP addresses are mapped to the same location?

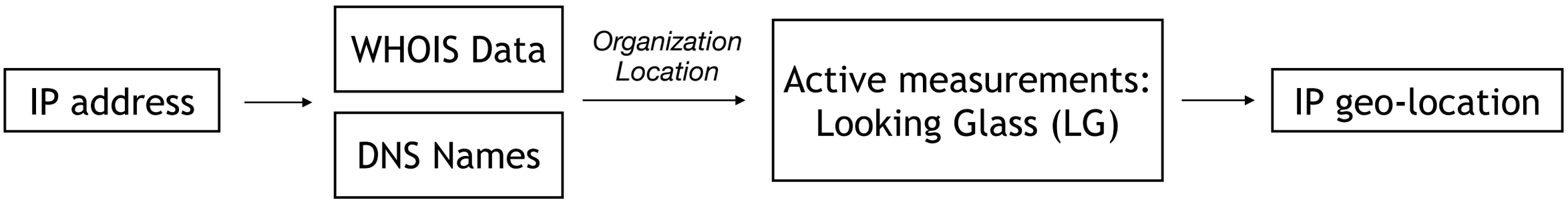


- IP addresses geolocated by the three geo-location datasets are most likely mapped to the same country.
- Found both partial and complete disagreements between the geo-location datasets.

Improving IP geo-location accuracy



Improving IP geo-location accuracy

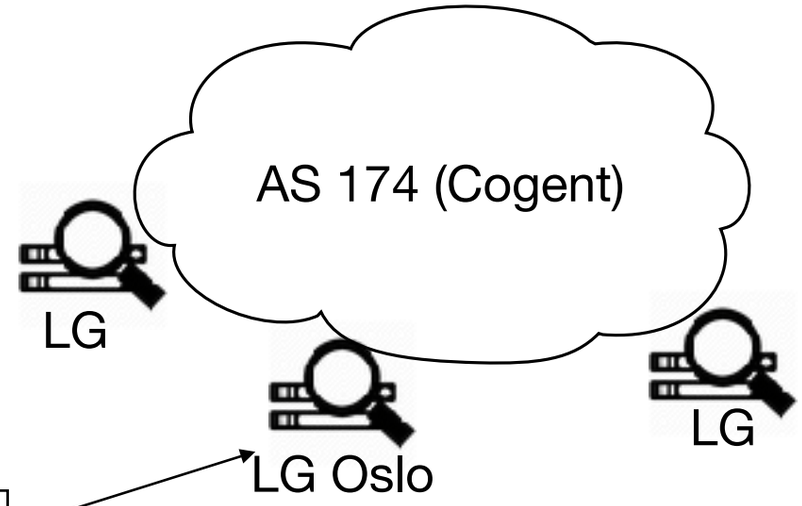


IP address = 154.25.4.213

name=be3561.rcr21.**osl**01.atlas.**cogentco**.com.

```
NetRange:       154.25.0.0 - 154.25.255.255
CIDR:           154.25.0.0/16
NetName:        COGENT-154-25-16
NetHandle:      NET-154-25-0-0-1
Parent:         NET154 (NET-154-0-0-0-0)
NetType:        Direct Allocation
OriginAS:       AS174
Organization:   PSINet, Inc. (PSI-2)
RegDate:        1992-02-05
Updated:        2017-10-30
```

LG Location = Oslo, NO



LG Query Results:

```
traceroute to 154.25.4.213 (154.25.4.213), 30 hops max, 60 byte packets
 1  gi0-6-1-19.201.rcr21.osl01.atlas.cogentco.com (130.117.254.161)  0.944 ms *
```

Sources of IP address geo-location disagreements

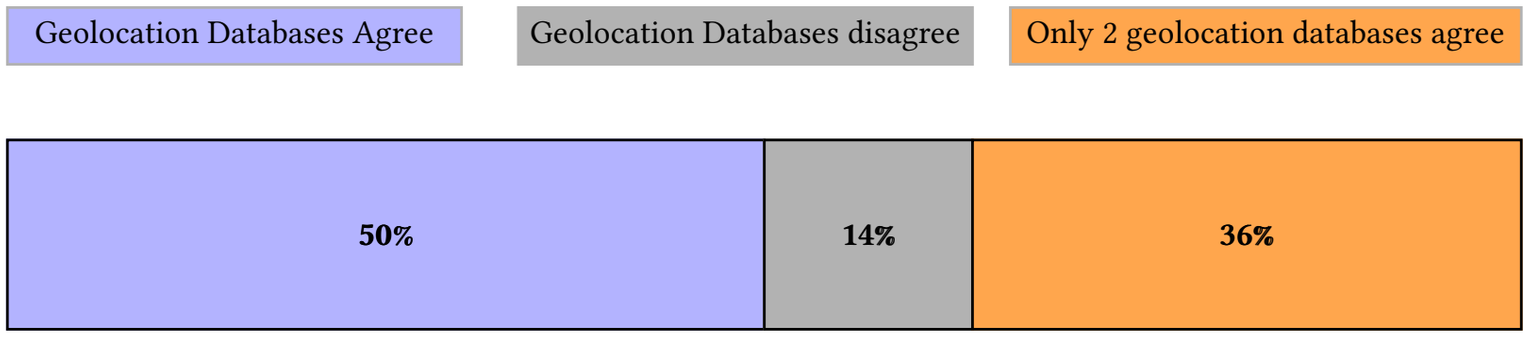
- IP addresses owned by *global organizations*:

IP address	Delegation	MaxMind	IP2Location	IPmap	HLOC	**Accurate location**
109.105.97.10	SE	SE	GB	NaN	NaN	**DK**

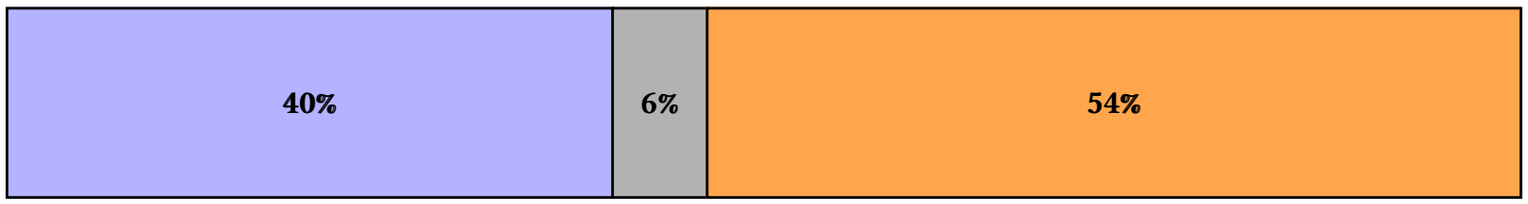
- IP addresses acquired by organizations through *merges & acquisitions*:

IP address	Delegation	MaxMind	IP2Location	IPmap	HLOC	**Accurate location**
149.6.154.202	US	IT	CA	NaN	NaN	**FR**

How many IP paths are geolocated similarly?



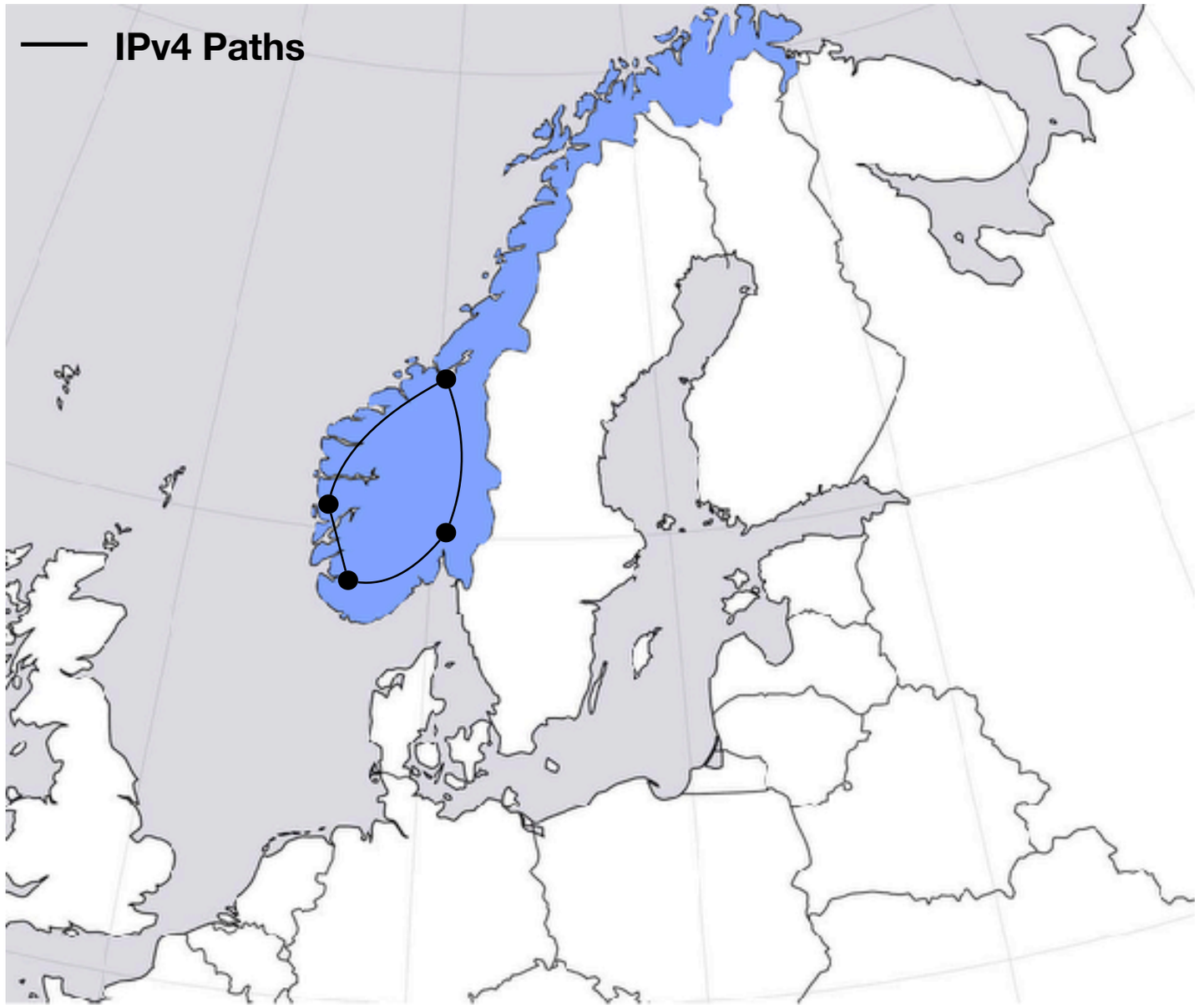
(a) IPv4-level paths



(b) IPv6-level paths

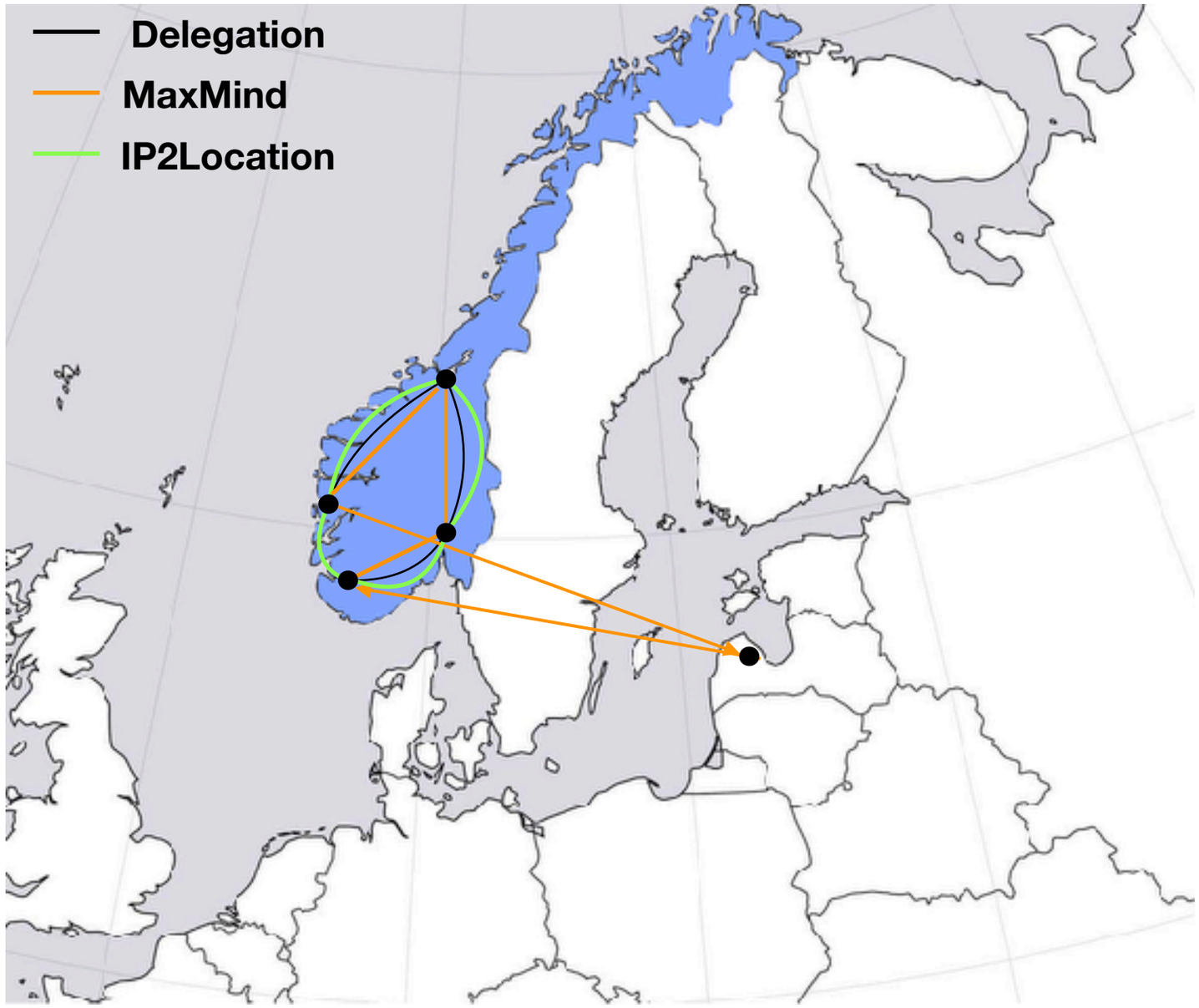
- At best, half of the IP paths are geo-mapped similarly by the three datasets. Most of the agreements occur between Delegation and MaxMind
- IP-to-country geolocation disagreements appear along the IP path

Observations and Implication: path tromboning



- 30% IPv4 and 26% IPv6 paths start and end in Norway
- No occurrence of path tromboning for IPv4 paths

Observations and Implication: path tromboning

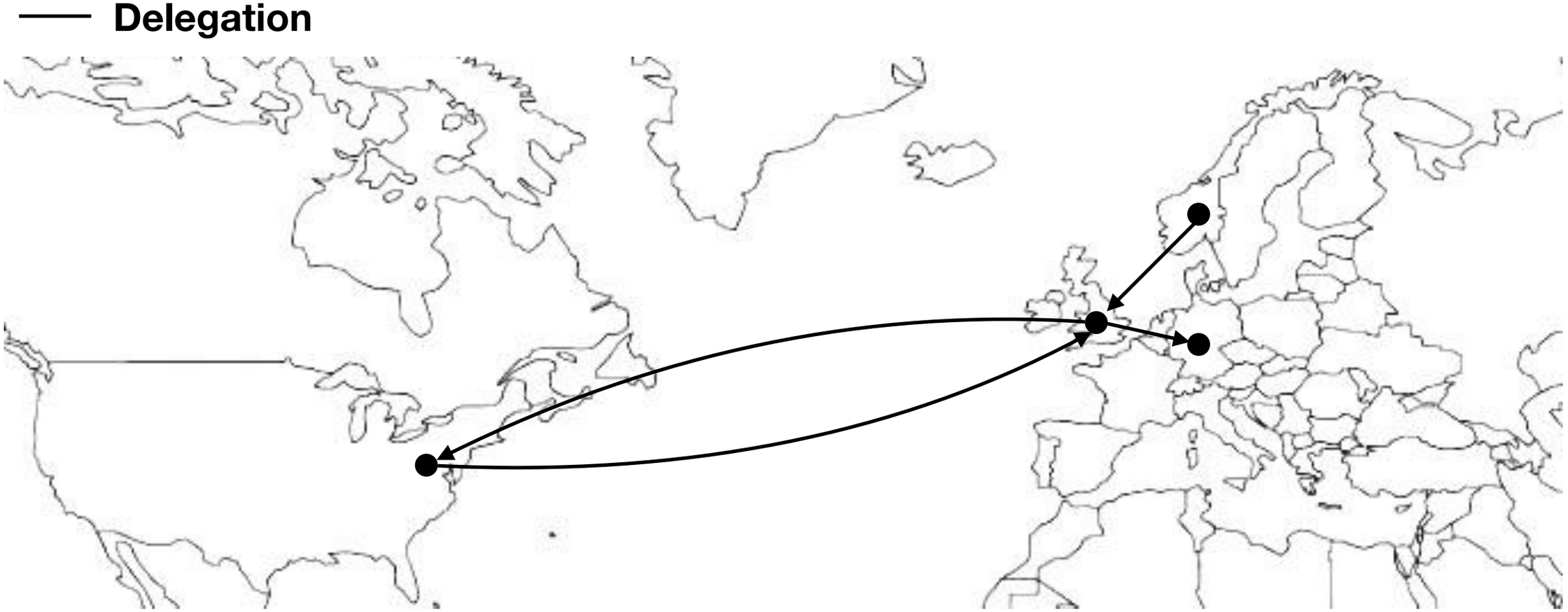


- 30% IPv4 and 26% IPv6 paths start and end in Norway
- No evidence of path tromboning for IPv4 paths
- Inaccurate MaxMind IPv6 geo-mappings cause path tromboning.

Observations and Implication: path detours

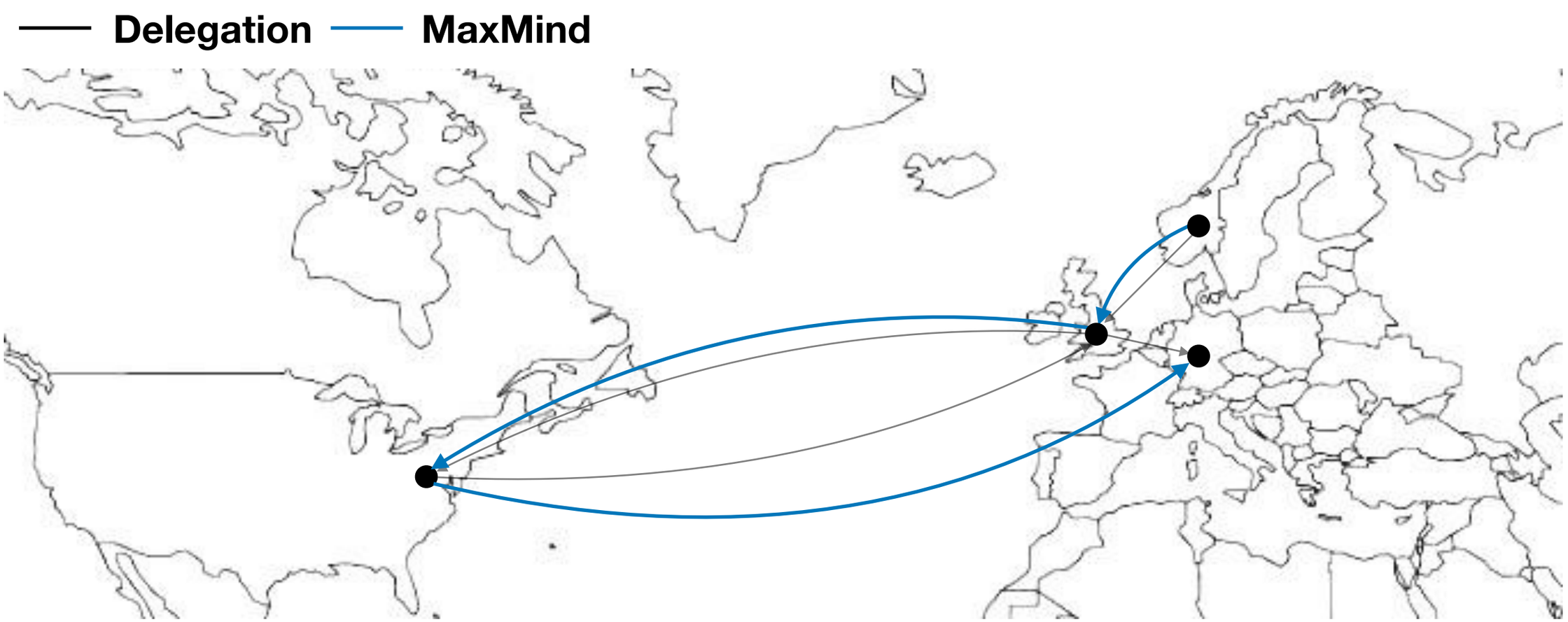
Assumption: IP hops on paths that starts and end in the same geographic region should be mapped within the same region.

Observations and Implication: path detours



Delegation: NO->**GB**->**US**->**GB**->DE

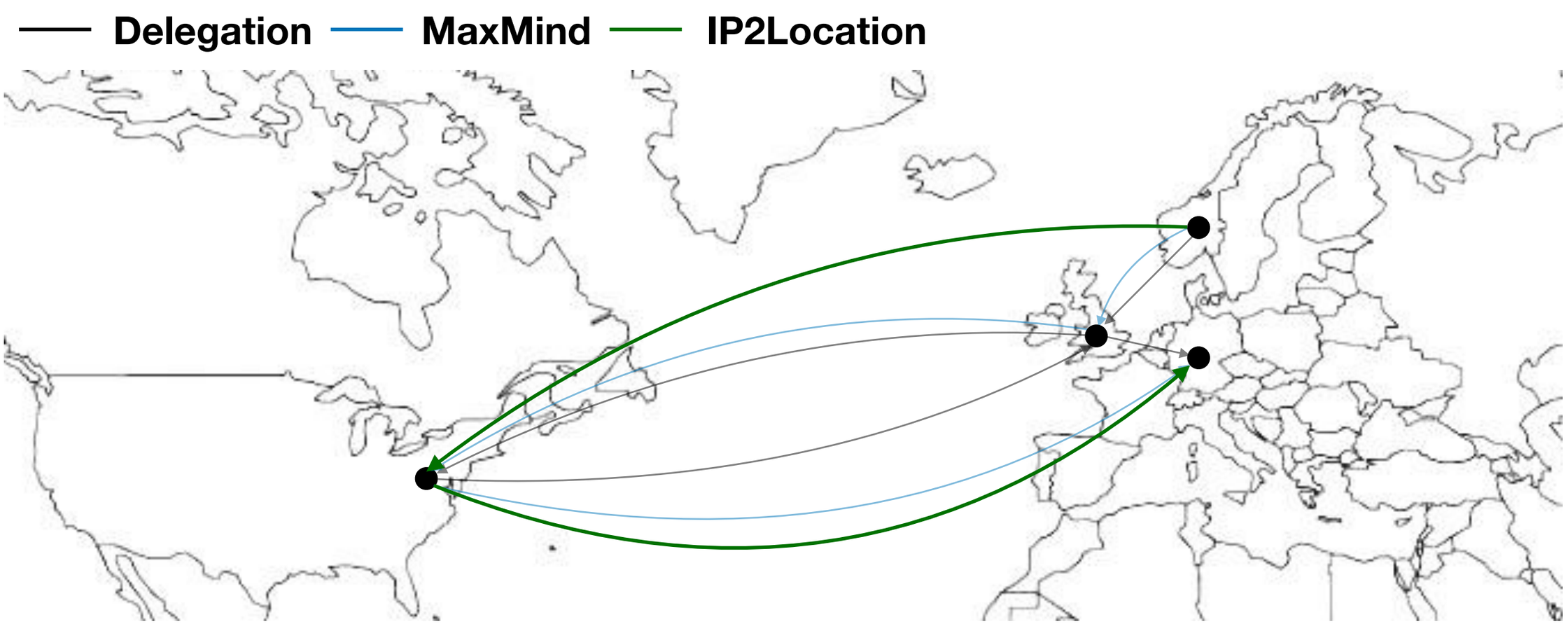
Observations and Implication: path detours



Delegation: NO->GB->US->GB->DE

MaxMind: NO->**GB**->**US**->DE

Observations and Implication: path detours



Delegation: NO->GB->US->GB->DE

MaxMind: NO->GB->US->DE

IP2Location: NO->**US**->DE

Observations and Implication: path detours

Delegation: NO->GB->US->GB->DE

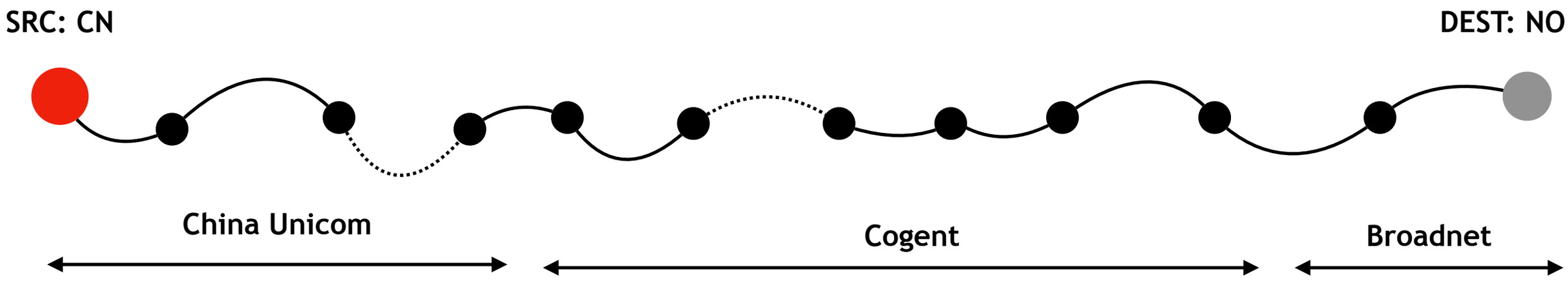
MaxMind: NO->GB->US->DE

IP2Location: NO->US->DE

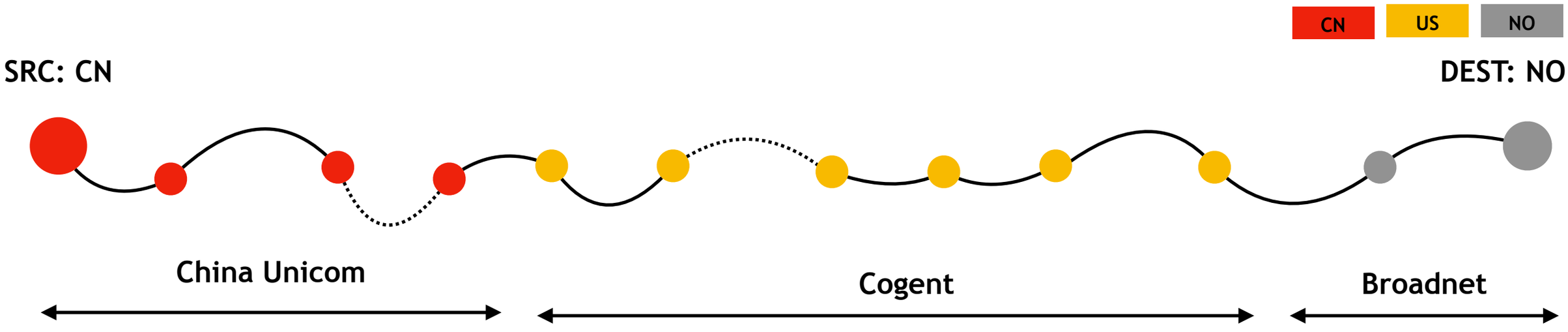
Country-level path: **NO->DE**

Path detours caused by Level3 IP addresses inaccurately mapped to US and GB.

High percentage of IP paths appear to miss countries

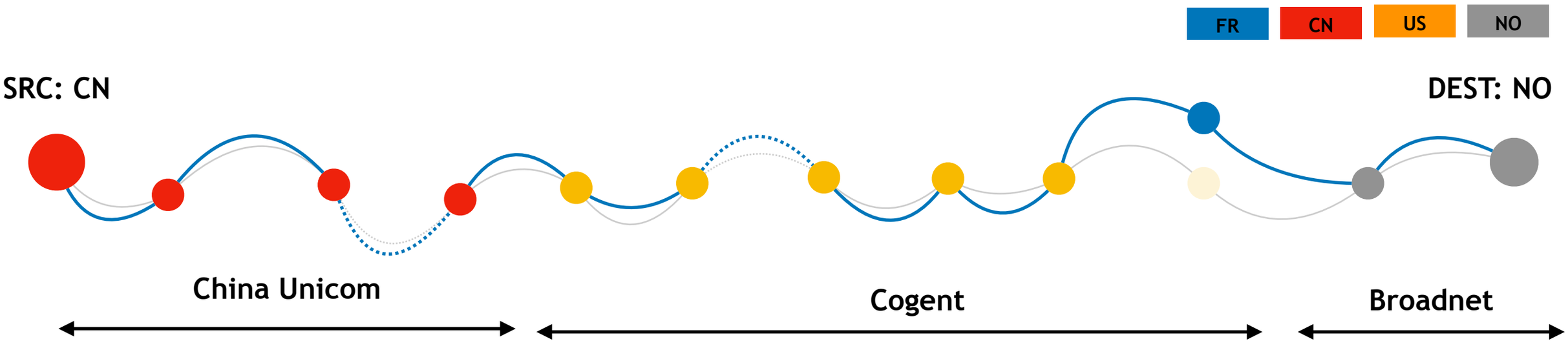


High percentage of IP paths appear to miss countries



—— **Delegation:** CN->US->NO

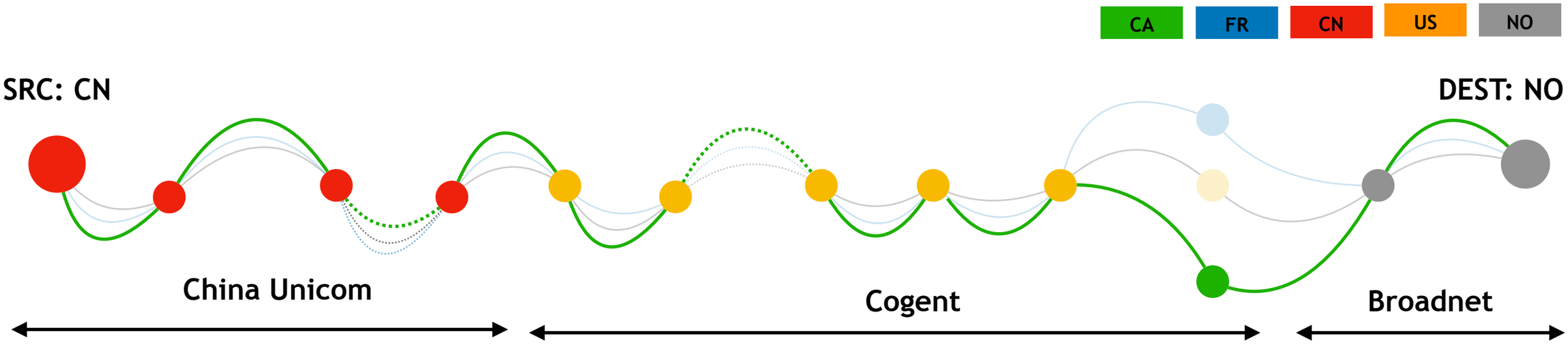
High percentage of IP paths appear to miss countries



Delegation: CN->US->NO

MaxMind: CN->US->FR->NO

High percentage of IP paths appear to miss countries

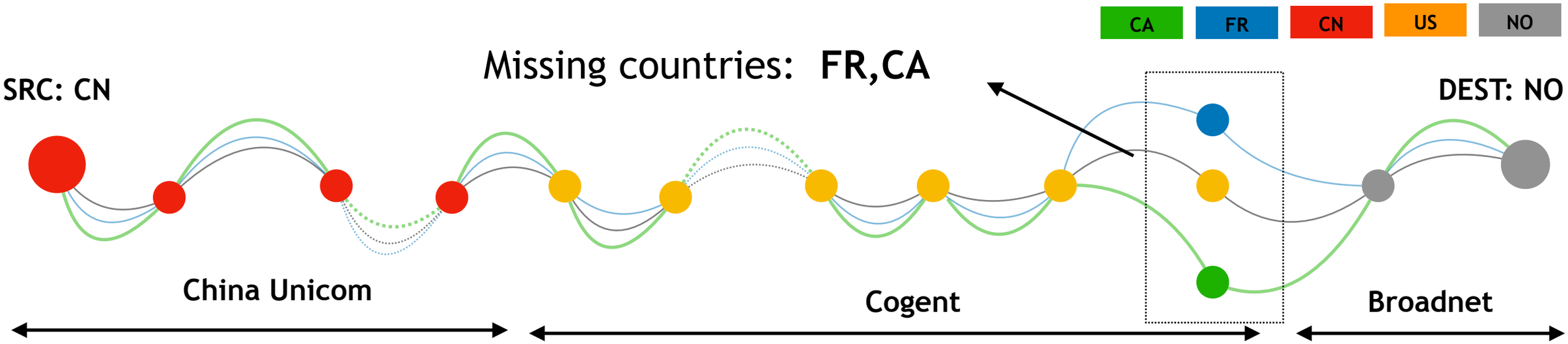


Delegation: CN->US->NO

MaxMind: CN->US->FR->NO

IP2Location: CN->US->CA->NO

High percentage of IP paths appear to miss countries.

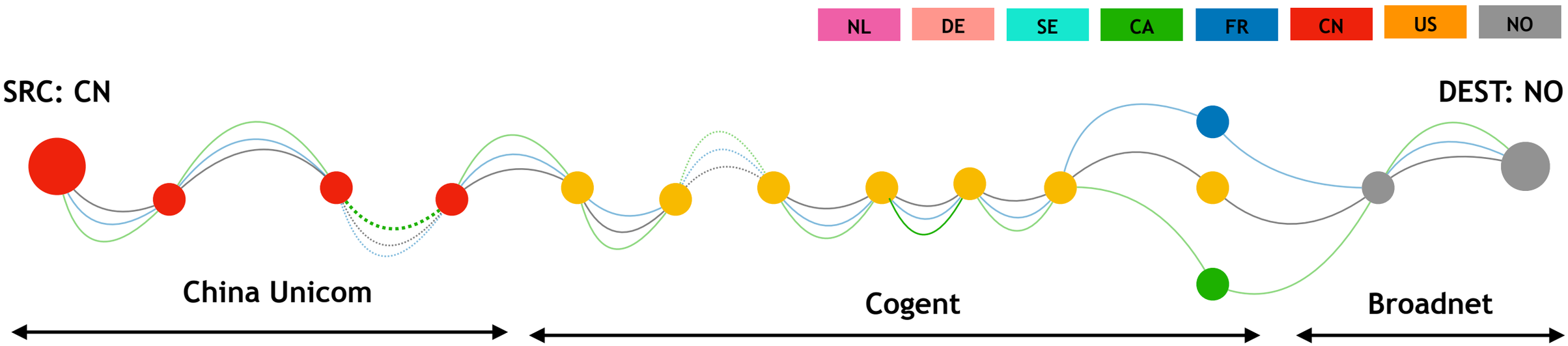


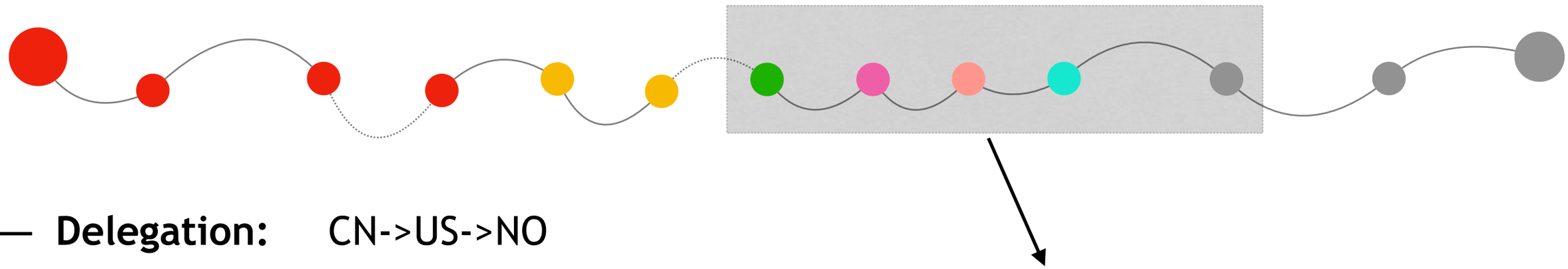
Delegation: CN->US->NO

MaxMind: CN->US->FR->NO

IP2Location: CN->US->CA->NO

High percentage of IP paths appear to miss countries





False negatives: DE, NL, SE

Delegation: CN->US->NO

MaxMind: CN->US->FR->NO

IP2Location: CN->US->CA->NO

Country-level path: CN->US->CA->NL->DE->SE->NO

Conclusions

- High level of agreement among the geolocation datasets hints that IP2Location and Maxmind use RIR information
- M&A activity causes IP geolocation inaccuracies
- Geolocation inaccuracies can cause misleading path geo-mappings — add or miss countries on the country-level paths
- Geolocating one week of RIPE traceroute data validates our observations
- Approach for improving IP geolocation IP