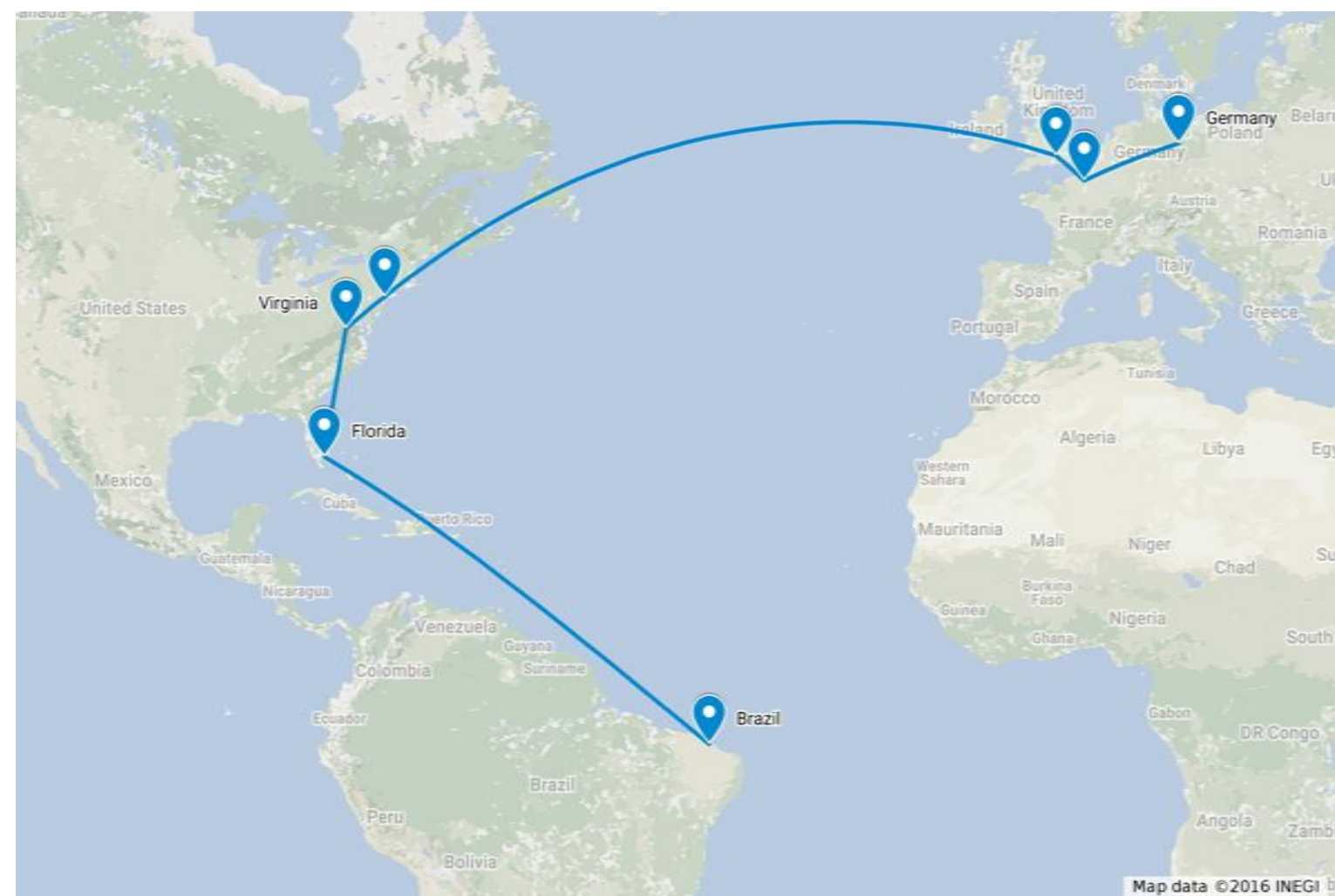


## Classifier-Based Country-Level Router Geolocation

Muzammil Abdul Rehman, David Choffnes, Northeastern University, Sharon Goldberg, Boston University

### Motivation

- Internet paths traversing different countries have performance and privacy implications
- Routing does not respect international borders
- Internet does not reveal path geography



- Key questions:
  - Is it possible to accurately find router location on a country-level granularity?
  - What are the implications of geopolitical paths?

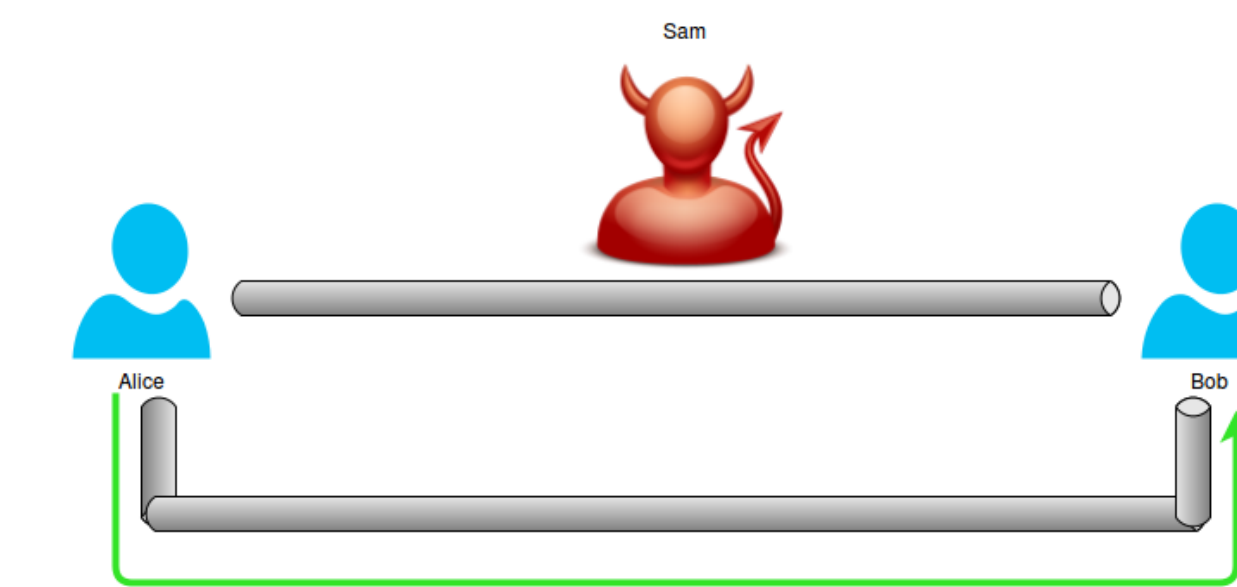
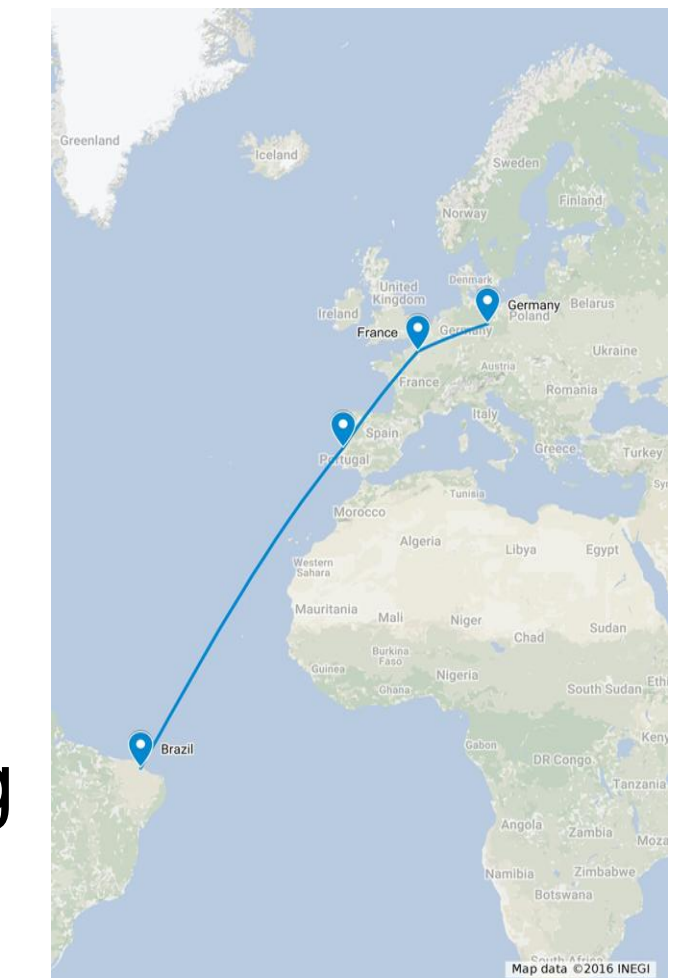
### How do we geolocate routers?

**Approach:** Use existing, imperfect data to reliably predict router locations

- Collect measurements (traceroutes)
- Get locations from partially reliable geolocation sources (GeoIP databases, hints in hostname, etc)
- Use **machine learning** to *infer* which sources are accuracy for a given router
- Use speed-of-light constraints to improve predictions

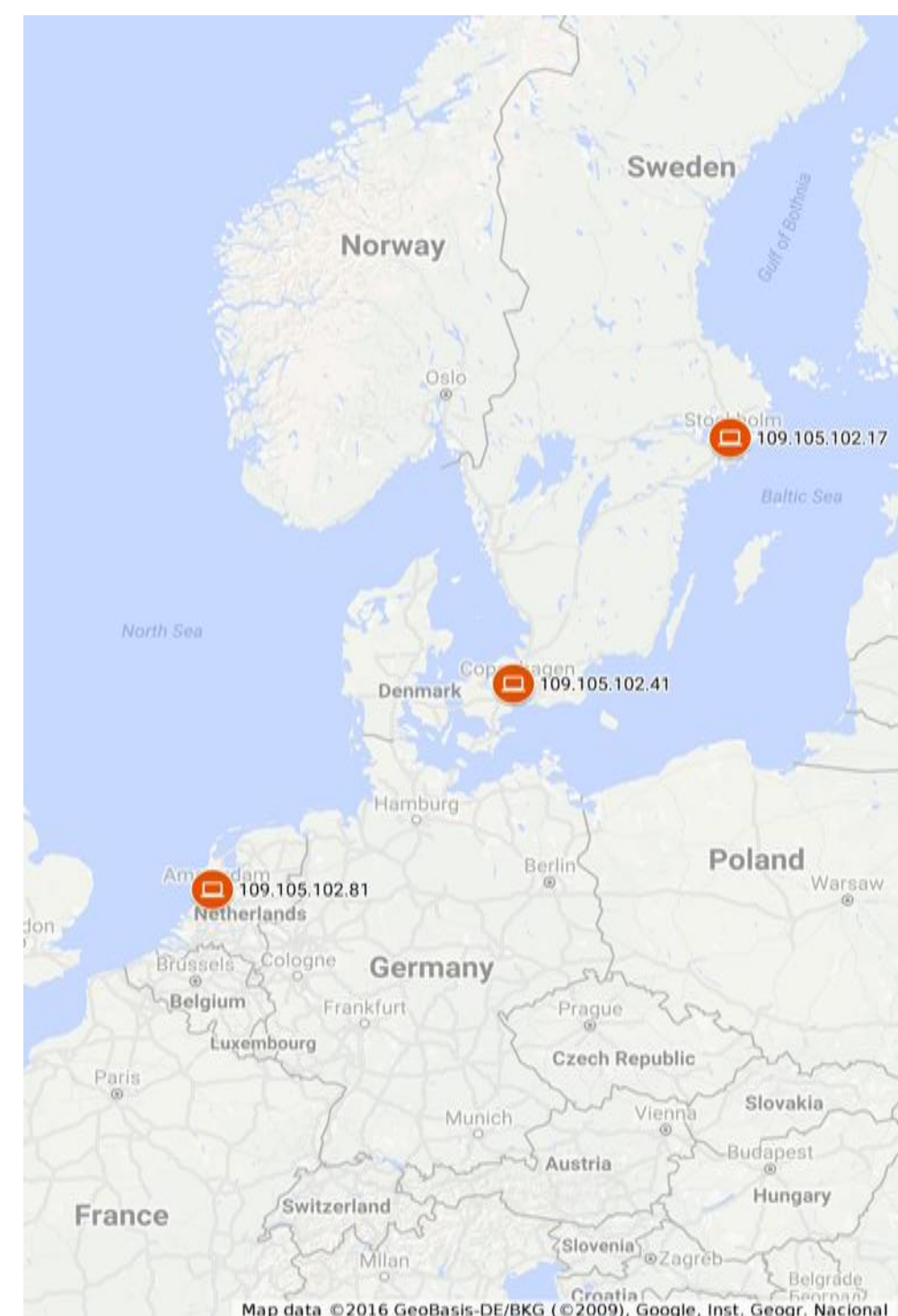
### Applications

- Mapping the Internet
- Identifying inflated paths
- Understanding impact of paths on exposure to foreign governments
- Evading paths traversing targeted nation-states



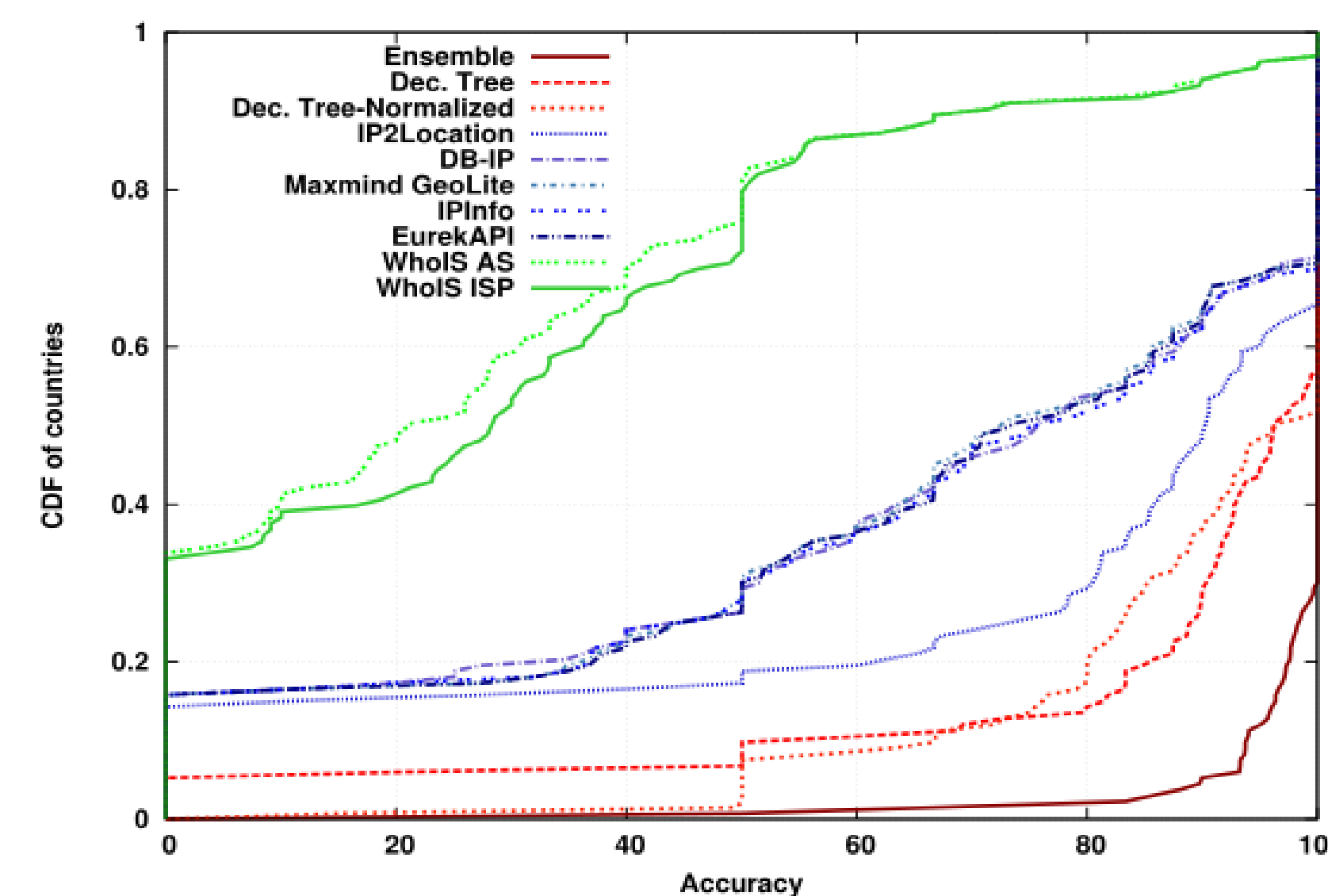
### Why is it hard to geolocate routers?

- Routing based on **IP addresses** and **ISP policies**.
- IP addresses do not encode **geographic** information
- Public sources of router geolocation data have low accuracy in general



### Accuracy:

- ML approach **accurately maps routers** in vast majority of countries
- **Reduces mismatches by 50%** compared to any single data source



### Ongoing/Future Work

- How classifier accuracy changes with time?
- Can we increase our location precision by adding more measurements?
- How do we compare with commercial services for geolocation?
- Are there any suspicious detours into other countries? How long do the detours last?
- Can we use routing announcements to avoid a specific country?