Country-in-the-Middle:

Measuring Paths between People and their Governments

Alisha Ukani

GMI-AIMS-5 Workshop

February 12, 2025

Governments are worried about foreign surveillance

Google-Facebook ditch plans to dock giant data cable in Hong Kong

How Russia Took Over Ukraine's Internet in Occupied Territories

When clients in one country access their government's websites:

- What other countries are involved (Countries-in-the-Middle),
- With what frequency,
- And why?

Methodology

1. Collect government websites for 11 different countries across different regions of the world and manually validate them
2. Find 10 RIPE Atlas probes in each of those 11 countries
3. Run traceroutes from each vantage point to 100 government websites for the corresponding country
4. Geolocate IP addresses to countries using ipinfo.io and validate results using speed-of-light constraints and hoiho

We collect >9,000 traceroutes across 11 countries

Challenges

- IP geolocation
- Collecting government websites
- Finding vantage points
- Interpreting traceroute responses
- Measuring the whole path
- Dealing with anycast

Paper	Validates Geolocation	Validates VP Geolocation	Considers Unreliable Responses	Considers Unreachable Traceroutes	Considers Anycast Sites
Gupta 2014 [15]	No	Yes	No	No	No
Fanou 2015 [13]	Yes	Yes	Yes	No	N/A
Shah 2016 [43]	No	No	Yes	No	No
Edmundson 2018 [11]	No	No	No	No	No
Gueye 2018 [14]	No	No	No	No	No
Candela 2021 [6]	No	No	No	Yes	No
Current Work	Yes	Yes	Yes	Yes	Yes

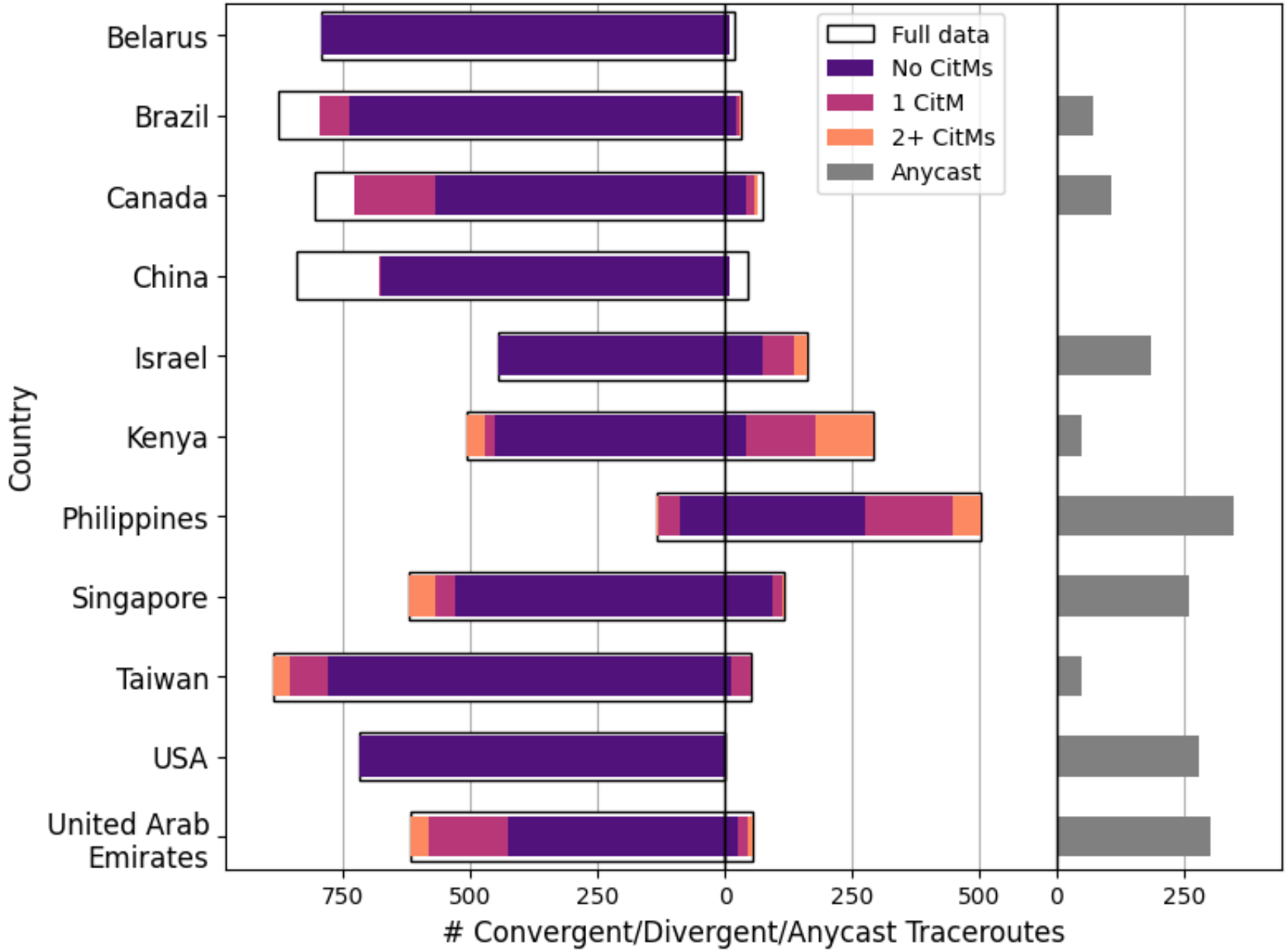
Validation Efforts

Probe issues:

- **Self-reported geolocation:** discarded 2 probes from China, resulting in 198 traceroutes discarded
- **Empty results:** one Brazil probe launched 93 traceroutes but returned empty paths

Path issues:

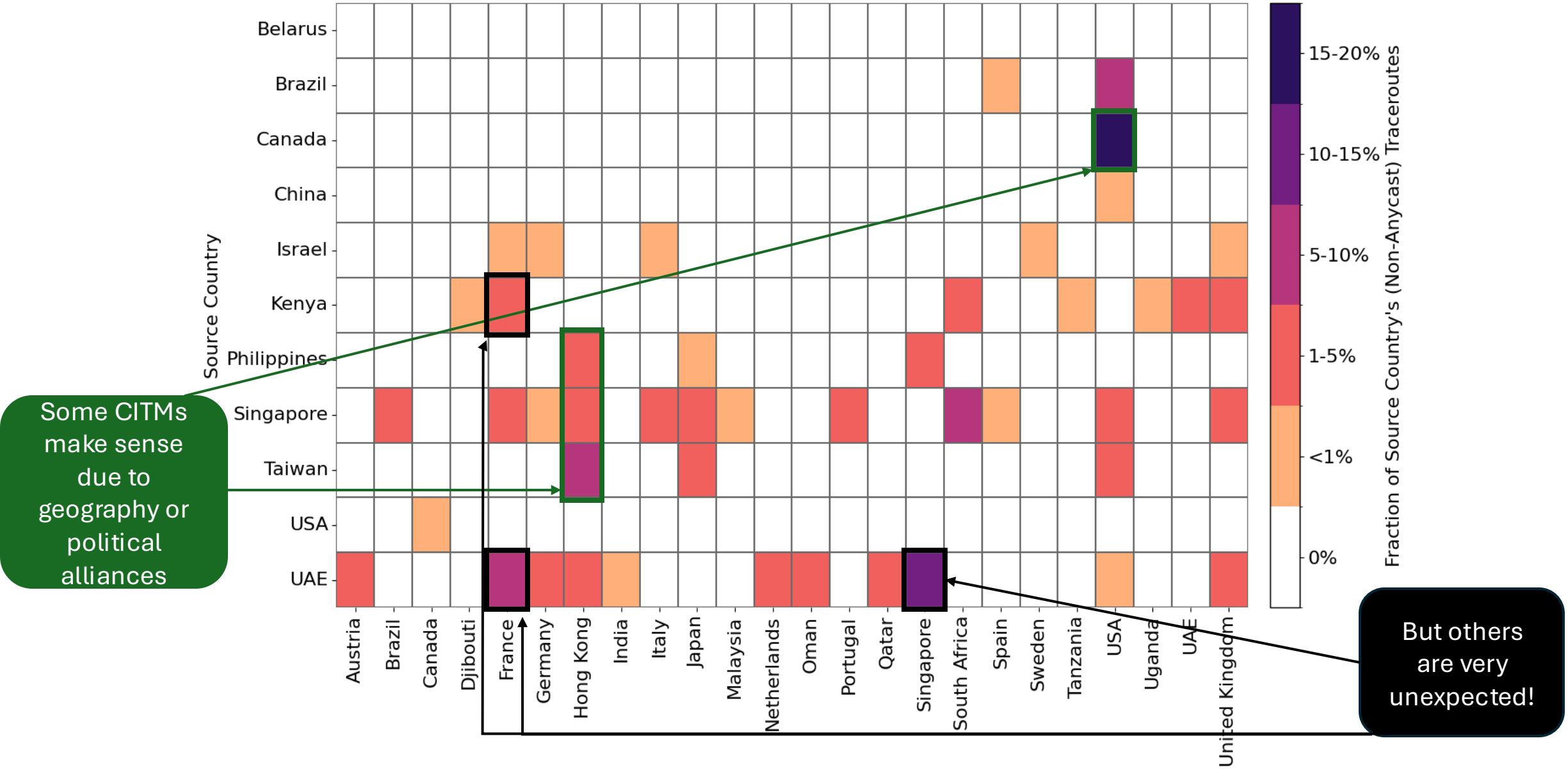
- **Geolocation:** found 55 traceroutes across 4 countries where a hop's geolocation violated speed-of-light constraints
- **IP squatting:** found 41 cases of IP squatting on DoD address space



Convergent = website is hosted in the source country

Divergent = website is hosted in another country

CITMs to Convergent Websites (Validated Data)



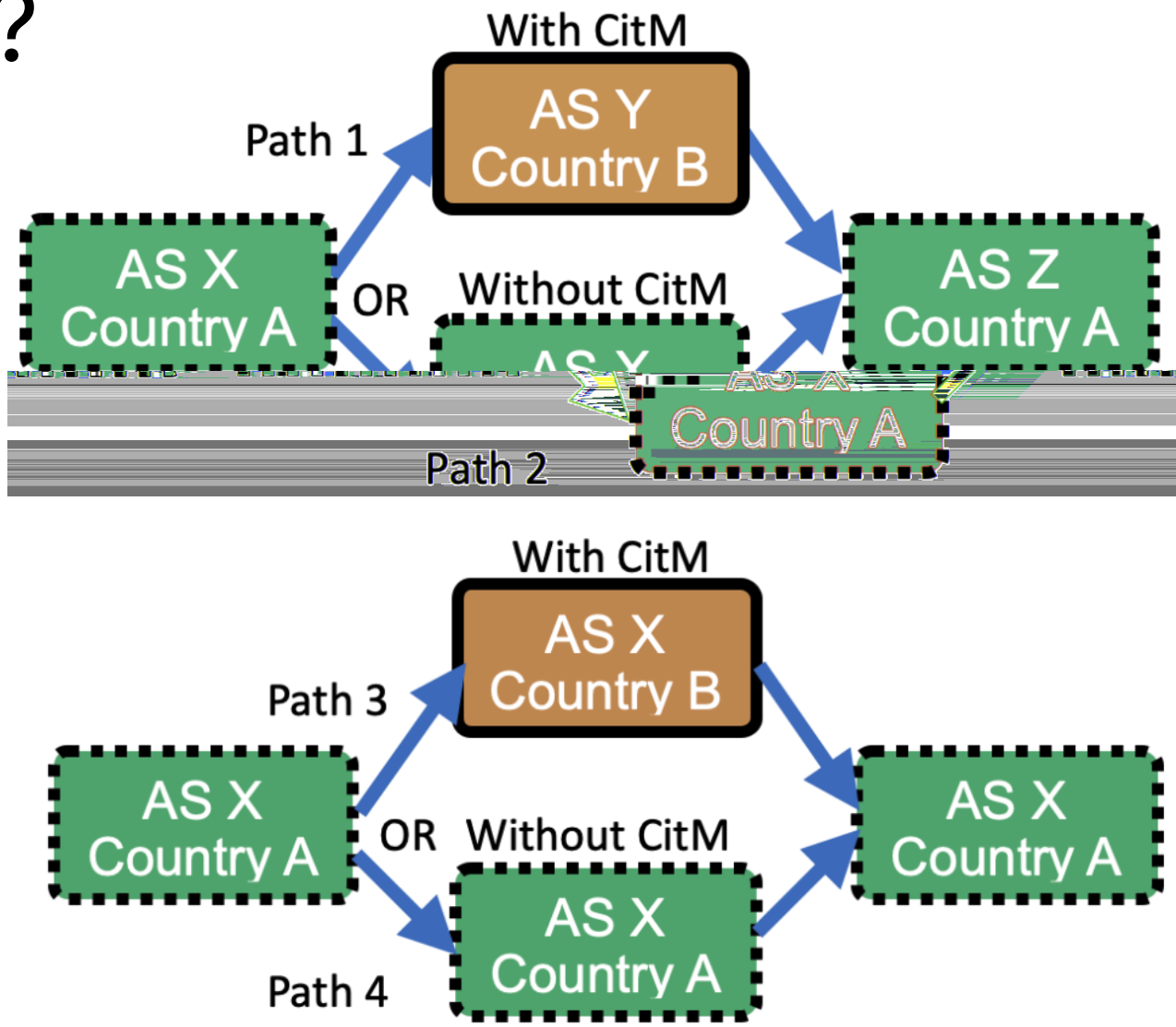
Some CITMs make sense due to geography or political alliances

But others are very unexpected!

Can CITMs Be Avoided?

- 2 of our Taiwan probes are hosted by PCCW Global, but one consistently has Hong Kong as a CITM but the other doesn't
- Liquid Telecom sometimes routes traffic starting and ending in Kenya through the U.K.

Takeaway: we observe cases where CITMs can be avoided, but that this is not a priority in routing decisions



Conclusion

- Identifying CITMs is an interesting but tricky problem, and deserves a higher standard of rigor and validation
  - We also need to continue to improve on geolocation (especially for anycast websites), IP squatting identification, and dealing with unlabeled traceroute hops
  - Country-level studies can also sidestep many of these challenges
- Tackling these challenges results in interesting examples of CitMs!