Crowdsourcing Router Geolocation

Emile Aben

emile.aben@ripe.net

SEE3 | Sofia - Bulgaria | 2014-04-15

Router Geolocation?

- What?
  - “normal” IP geolocation looks only at the edge
  - router geolocation = figuring out the rest

- Why?
  - Detect sub-optimal paths in traceroutes
  - Does a forward path traverse a specific country/region
    - In case of events?
    - Structurally?
  - Bulk analysis

Background

- Tons of interesting RIPE Atlas traceroutes
- Hard to put them on a map
- Naive router geolocation: Use Maxmind (or any other geoloc DB): Doesn’t work!

IP	Geoloc
2001:2000:3018:50::1	EU
89.221.34.63	IT
4.69.148.30	US
83.217.227.13	ES
141.136.110.174	FR
173.194.39.215	Mountain View,CA,US
184.105.223.246	Fremont,CA,US

Background

- Tons of interesting RIPE Atlas traceroutes
- Hard to put them on a map
- Naive router geolocation: Use Maxmind (or any other geoloc DB): Doesn’t work!

IP	Geoloc	Hostname
2001:2000:3018:50::1	EU	**sfia**-b2-v6.telia.net
89.221.34.63	IT	xe-1-0-2.**sofia**1.**sof**.seabone.net
4.69.148.30	US	ae-11-11.car2.**Sofia**1.Level3.net
83.217.227.13	ES	xe-0-2-0-2.r00.**sofibu**01.bg.bb.gin.ntt.net
141.136.110.174	FR	xs-3-3-0.**sof**10.ip4.tinet.net
173.194.39.215	Mountain View,CA,US	**sof**01s01-in-f23.1e100.net
184.105.223.246	Fremont,CA,US	10ge1-1.core1.**sof**1.he.net

Idea

- Find ways to geolocate Internet infrastructure better
- Ask the experts (you!) to participate
- Make collected data publicly available
  - so also for geoloc providers

- **Not** a competing service to existing geoloc
  - their data can be enhanced with router geoloc

Prior Art (RFC1925, rule 11)

- Existing router geoloc bits-and-pieces
  - rocketfuel (undns), IXmaps, ...
  - Problem: Unmaintained and/or complex and/or limited scope
- ‘Visual traceroute’
  - Typically use edge geolocation service
- IETF draft google-self-published-geofeeds
  - Complementary
- CAIDA geoloc project
  - Cooperating

Self Published Geofeeds

- Format:
  - Prefix,Country,Region,City,Postal:

```
193.0.24.0/21,GR,GR-I,Athens,117 45
2001:67c:64::/48,GR,GR-I,Athens,117 45
```

- Self-published by site
  - Currently you’ll have to know where these feeds are
- Potential template?

Proposed Geoloc Method

- Combine data-sources:
  - Existing edge geolocation
  - Hostnames from reverse DNS
    - 1.13 billion reverse DNS records in IPv4
    - Users could tag naming schemes
  - RTTs allow for some triangulation / speed-of-light constraints
  - IXP IPs/prefixes (when not remote-peering)
  - DNS LOC records
- Probabilistic answer: ie. 95% Athens,GR

Detail: RTT Constraints

- Signal propagation bound by speed of light
- In fiber ~ 100 km per 1 ms (round trip)

- One day of RIPE Atlas traceroutes:
  - 84122 IPs (v4/v6) seen
  - 40975 IPs within 10ms from the source = within 1000km

- Problem: High latency last mile
  - Would need to account for that

Detail: DNS LOC record (RFC 1876)

- DNS record to map geographic location to a hostname

- `nbg-s1-rou-1001.DE.eurorings.net.    IN LOC 49 27 12.690 N 11 3 56.416 E 10.00m 1.00m 10000.00m 10.00m`

- Found 16 domains using it:
  - Western Europe incumbent telcos
  - Research & Education networks

Detail: Probabilistic answer?

- Crowdsourced info can be conflicting
  - UserA: **ams-ix**.br2.**sof**2.example.com is in **Amsterdam,NL**
  - UserB: **ams-ix**.br2.**sof**2.example.com is in **Sofia,BG**

- Overlapping city names
  - **Bakel, NL** vs. **Bakel, SN**
  - 5 cities named **San Jose (US,PH,CR)**

- A probabilistic answer could capture ambiguity

Proposed Geoloc Method

- Combine data-sources:
    - Existing edge geolocation
    - Hostnames from reverse DNS
        - 1.13 billion reverse DNS records in IPv4
        - Users could tag naming schemes
    - RTTs allow for some triangulation / speed-of-light constraints
    - IXP IPs/prefixes (when not remote-peering)
    - DNS LOC records
- Probabilistic answer: ie. 95% Athens,GR

Proposed Method - Crowdsource

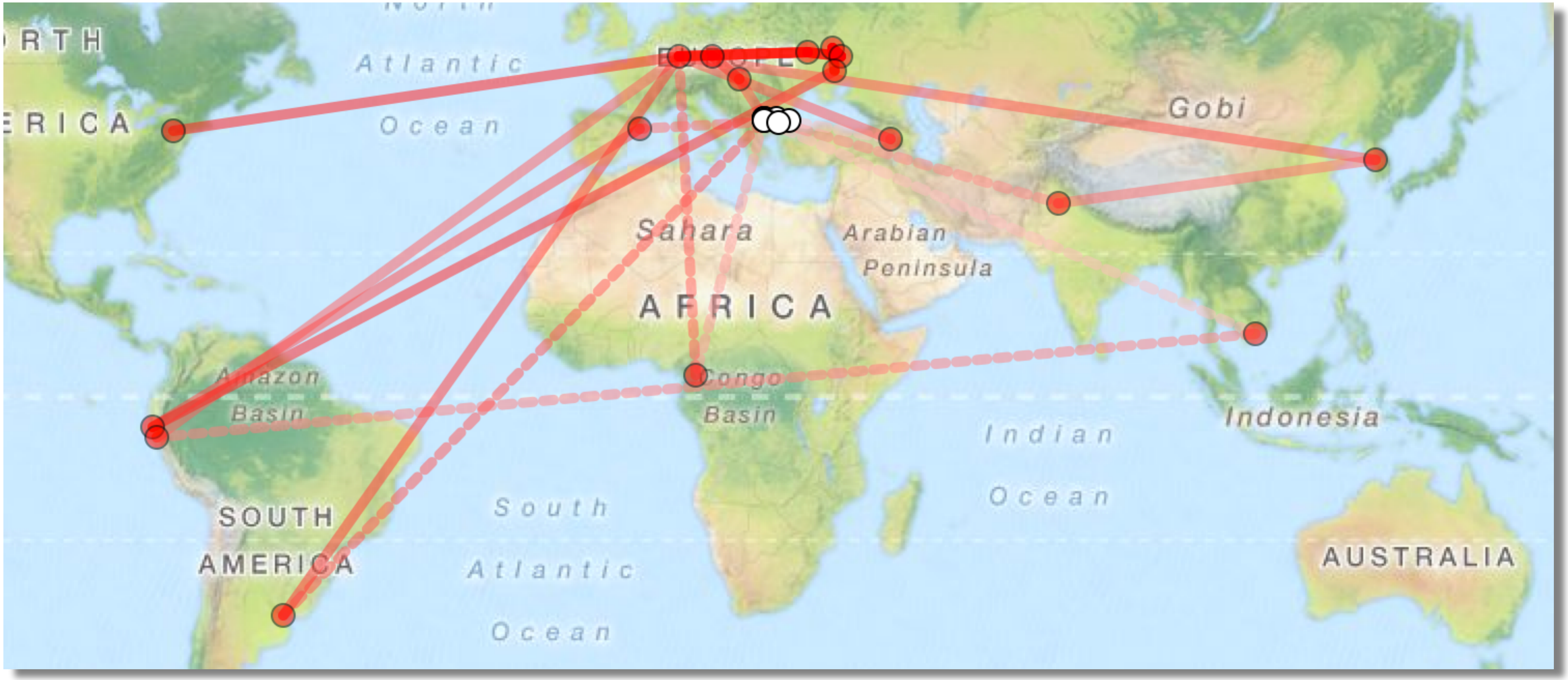
- Combine data-sources:
    - Existing edge geolocation
    - Hostnames from reverse DNS
        - 1.13 billion reverse DNS records in IPv4
        - Users could tag naming schemes
    - RTTs allow for some triangulation / speed-of-light constraints
    - IXP IPs/prefixes (when not remote-peering)
    - DNS LOC records
- Probabilistic answer: ie. 95% Athens,GR

Prototype Traceroute Visualisation

- Having a computer guess just based on hostname, doesn’t work very well yet (BG -> UA):

Prototype Traceroute Visualisation

- Having a computer guess just based on hostname, doesn't work very well yet (BG -> UA):



Ambiguity in Hostnames

- Lots of people use IATA-airport codes, but
  - **atm** - Altamira,BR (IATA) or ATM link?
- Mixed naming schemes
  - **fra**07s29-in-x10.1e100.net (IATA) vs. **ea**-in-f99.1e100.net
- Almost IATA-schemes
- Different languages
  - Wien vs. Vienna
- Different abbreviations
  - **nyc** vs. **nyk** for New York

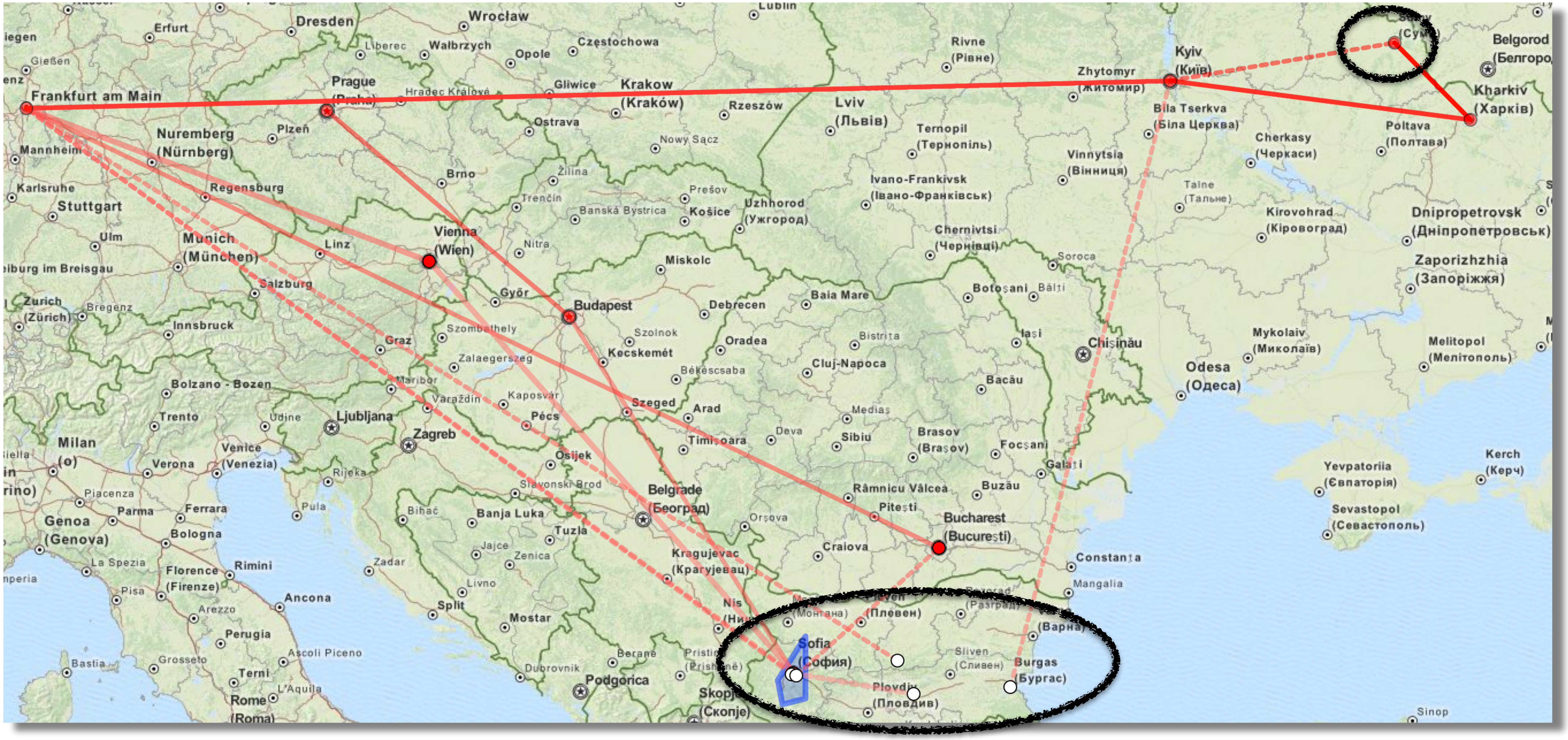
How to Improve: Crowdsource

- What you give:
  - Info on your network
  - Info on other networks

- What you get back:
  - Better router geolocation for everybody

Prototype Traceroute Visualisation

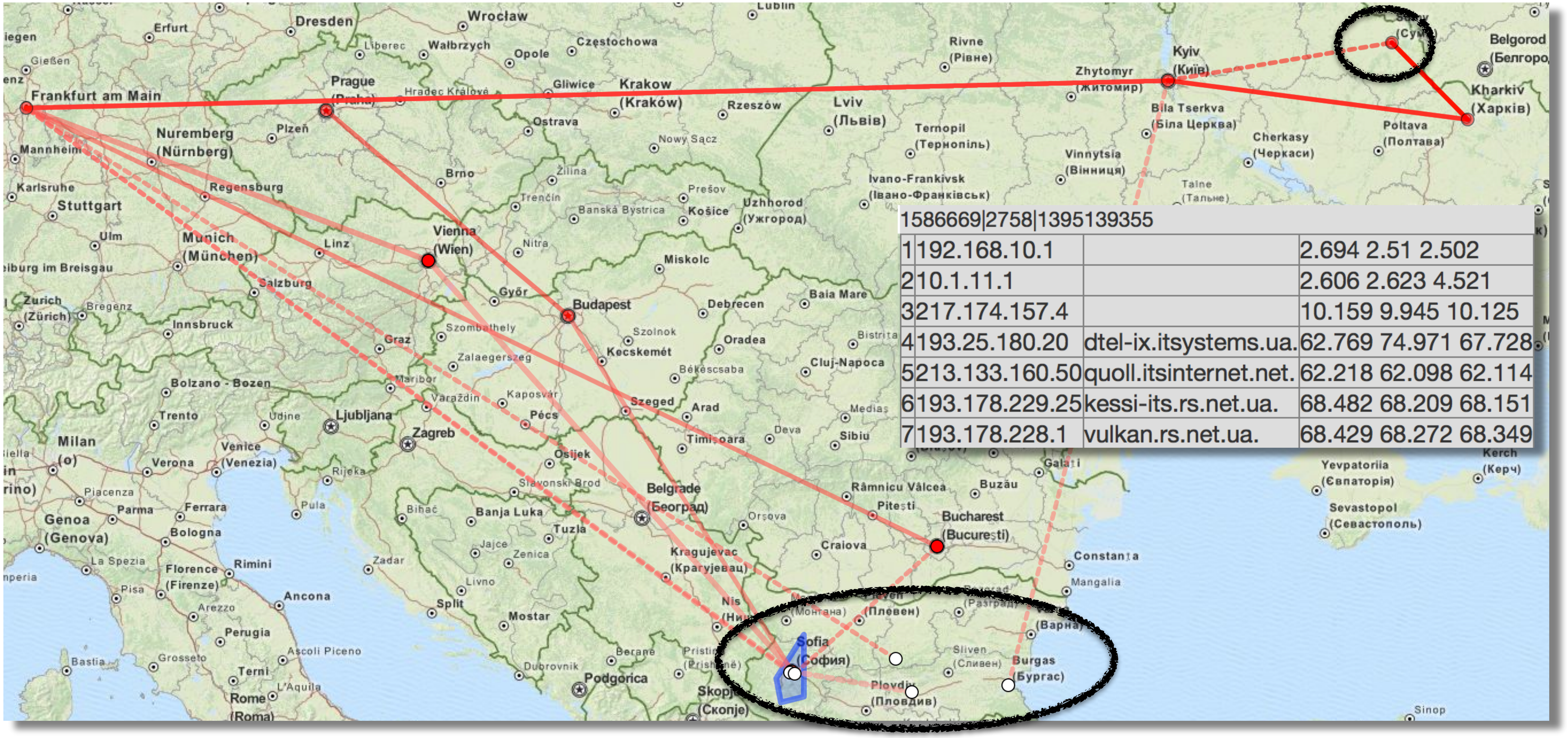
Destination



Sources

Prototype Traceroute Visualisation

Destination



1586669|2758|1395139355			
1	192.168.10.1		2.694 2.51 2.502
2	10.1.11.1		2.606 2.623 4.521
3	217.174.157.4		10.159 9.945 10.125
4	193.25.180.20	dtel-ix.itsystems.ua.	62.769 74.971 67.728
5	213.133.160.50	quoll.itsinternet.net.	62.218 62.098 62.114
6	193.178.229.25	kessi-its.rs.net.ua.	68.482 68.209 68.151
7	193.178.228.1	vulkan.rs.net.ua.	68.429 68.272 68.349

Sources

Open Questions

- How to crowd-source exactly?
  - Regular expressions: **^([a-z]{3})\d+.*\.1e100\.net**
    - Pro: Can capture everything
    - Con: Not exactly user-friendly
  - Tag to city: **sof = Sofia,BG**
    - Pro: More user-friendly/closer to how info is stored already
    - Con: Can be ambiguous

Conclusion

- Exploring this idea because:
  - Could give you better tools/viz in RIPE Atlas
  - Could give you data to build your own tools on
  - Could give geolocation providers data to make their data better

- Let us know what you think!

Questions?