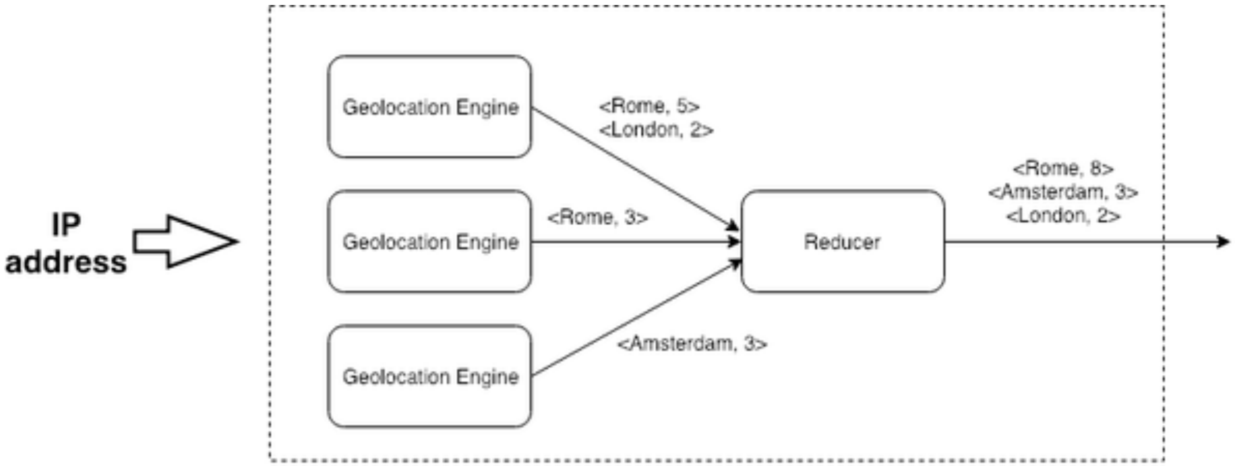
Mechanism and Performance Evaluation of RIPE IPmap Active Geolocation

Massimo Candela

RIPE IPmap

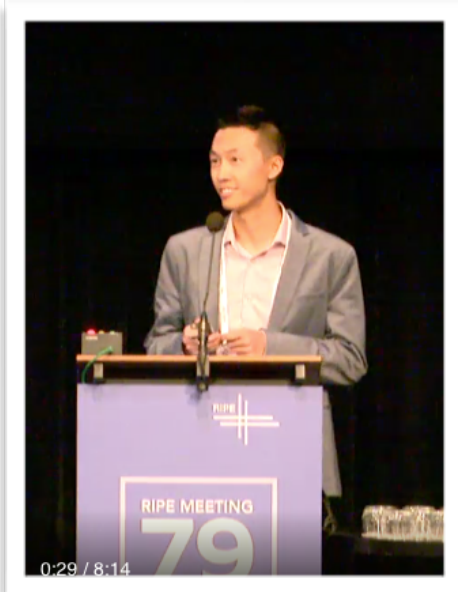
- RIPE IPmap is a multi-engine geolocation platform operated by RIPE NCC
- One of its engines, *single-radius*, uses active geolocation to infer the geographic coordinates of target IP addresses



RIPE IPmap architecture schema. Source: labs.ripe.net

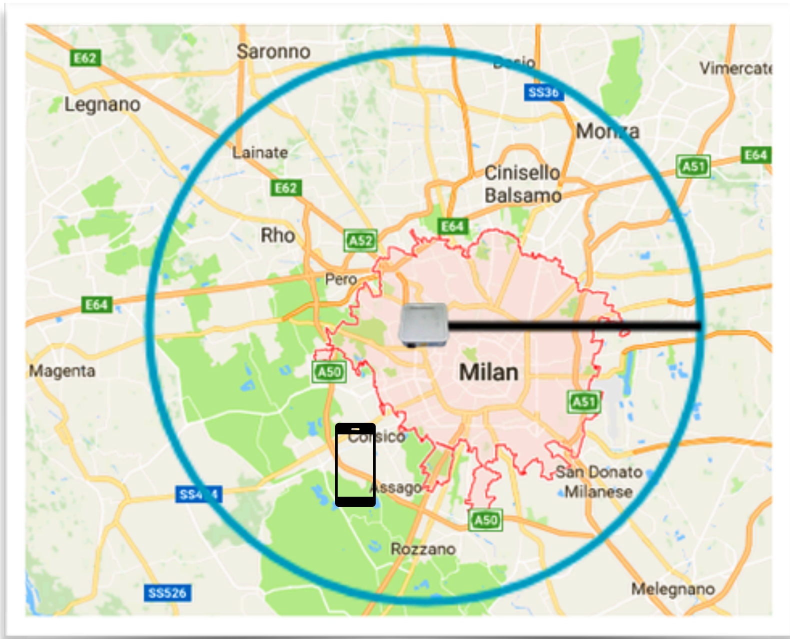
More info…

- I developed single-radius while I was at RIPE NCC
- We got some good feedback about it's accuracy but:
  - In "[Tracing Cross Border Web Tracking]" (IMC2018) reports accuracy at country (99.58%) and continent level (100%)
  - We needed a formal study dedicated to single-radius (especially city-level geolocation)
  - The geolocation approach was not formally described
- CAIDA decides to study single-radius
  - Ben Du presents preliminary results at RACI (RIPE 79)
  - What better way to understand how it works than having the developer in the team?



Active IP geolocation

- Deriving the geographical location of a connected device by means of latency measurements
  - Devices (landmarks) of known position ping a target device of unknown position
  - Latencies are converted to distances with a conversion factor expressed in km/ms



Single-radius methodology

- Initial selection of probes
- Issue Ping measurements
- Collect the Ping measurement with lowest RTT <= 10ms
- Use the probe originating such Ping measurement as center of a circle of radius RTT/2 * 2/3c
- Return to the user the closest 100 cities from the probe
    - Only if they are inside the circle
    - Ranked based on distance, number of facilities, and population

Initial selection of probes

- Single-radius tries select probes that are topologically, if not geographically, near the target t by selecting probes either in AS(t), the AS of target.

- Single-radius creates a list of cities C and a list of ASes A as follows.

  1. Add AS(t) to A
  2. Add to C the cities where AS (t) has a probe
  3. Add to A the ASes neighbors (BGP distance 1) of AS(t)
  4. Add to C the cities with IXPs where AS(t) is present (idea of Vasileios G.)
  5. Add to A the ASes present at the IXPs identified in step 4
  6. Add to C all the cities corresponding to the facilities where AS(t) is present
  7. Add to A the ASes peering at facilities identified in step 6

Single-radius evaluation - Goals

- **Evaluate accuracy**
  - Compare accuracy with MaxMind and NetAcuity
- Evaluate effectiveness of Probe selection mechanism
- **Evaluate coverage**
- Evaluate Router-Level Consistency
- **Provide suggestions to RIPE NCC and to the community**

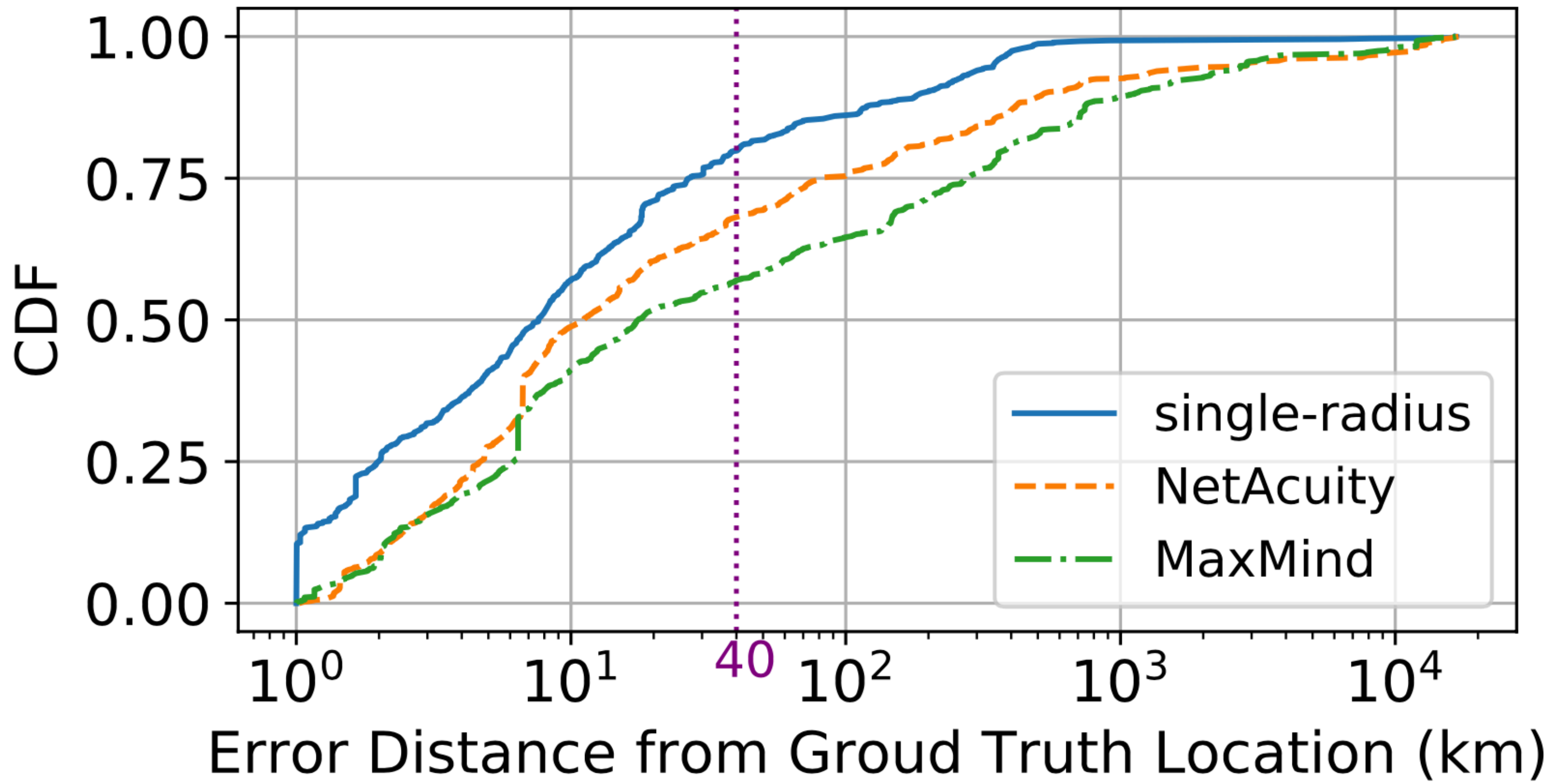
Datasets

- Ground Truth dataset
  - Composed of IPs of known locations
  - We cannot use RIPE Atlas
  - We remove IPs without metadata or offline (final selection: 968 IPs)

Cont.	Ark		NLNOG		M-Lab		Ark-prox		Total	
	IP	ASN	IP	ASN	IP	ASN	IP	ASN	IP	ASN
AF	26	25	8	7	7	6	0	0	41	34
AS	26	21	28	20	14	10	90	17	158	57
EU	60	50	359	323	55	16	35	17	509	375
NA	80	59	89	71	69	22	4	2	242	140
OC	9	7	11	11	5	3	56	32	81	46
SA	16	15	11	11	5	5	21	5	53	29

- `Coverage dataset
  - We use CAIDA’s MANIC Interconnection Dataset (16,245 IPs)

Accuracy

- We consider geolocations within 40 km of the true locations to be accurate
- 

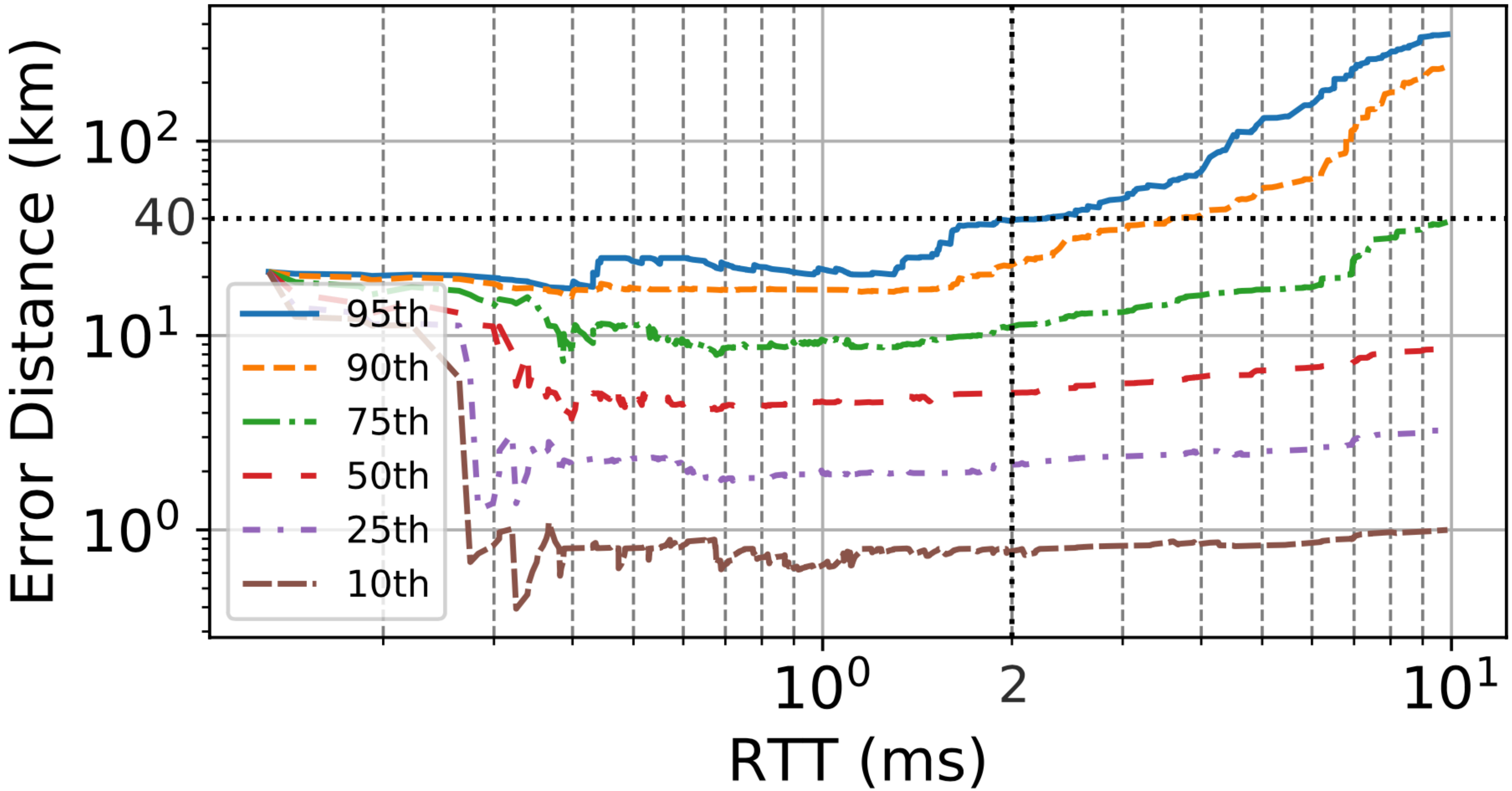
Accuracy

- Single-radius outperformed the other two services on our ground truth dataset.

- Single-radius achieved median, 75th, and 95th percentile error distances of 6, 26, and 344 km, respectively.
  In contrast, those of NetAcuity and MaxMind were 10, 80, and 2867 km, and 17, 278, and 2886 km, respectively.

- Over 80% of single-radius results were within the 40km error threshold

RTT threshold evaluation on accuracy



Coverage

- Single-radius provided results for 12,319 (78.5%) of the MANIC addresses.
- We studied the impact of the RTT thresholds on the coverage
  - Switching from the default 10 ms threshold to 5 ms drastically dropped coverage to 51.1%.
  - A 2-ms threshold, the MANIC coverage would further decrease to 3,916 (24.1%) according

Limitations

- Geographical bias of ground truth

- Ground truth size

Thanks to

- Ben Du
  c4du@caida.org
- Bradley Huffaker
  bradley@caida.org
- Alex C. Snoeren
  snoeren@cs.ucsd.edu
- kc claffy
  kc@caida.org

Questions?

Massimo Candela

massimo@ntt.net

Twitter: @webrobotics