Measurement and Monitoring

Scott Kirkpatrick
Hebrew University

New Requirement: Measure from edge to center

- With widespread LTE deployment and “5G” imminent, where will the “capacity crunch” appear?
- How can “crowd-sourced” monitors be best used in an era of network neutrality and transparency
- Two goals
  - ensure level playing field
  - exploit new technologies — SDN/X/I and wireless

Activities in US & Europe

- What is network neutrality and how will it be enforced?
  - Clear principles in the US, little experience
    - No blocking, throttling or “paid prioritization”
    - Europe has net neutrality regulation in three countries, but allows “special services” that are end-to-end distinct from public internet access. Blocking is not uncommon.
    - Transparency is the tool to understand the limits of “normal network management”
- SMART 2014/0016 EC tender to provide geographic data base for public access to QOS and QOE measurements derived from “crowd-sourcing” of broadband, both home and mobile.
  - Tender asks for mechanisms to ensure contributions, protect privacy, establish funding needed to support data collection — a fully developed data collection campaign strategy
  - Focus is on precise characterization of edge performance including QOE.

Wireless “crowd-sourced” data capabilities

Scott Kirkpatrick, Danny Bickson, Orpaz Goldstein
Hebrew University
Alex Zaidelson
WeFi, Inc.

What do we see from WeFi's1-10M pocket "machine rooms?"

Boston (Nov 2014)
69M measurements
15629 IDs (~4%)
~12 IDs/Hectare
we see 23.2 TB
downloaded data
(15 TB wifi,
8 TB cellular/mixed)
est. total ~1 PB?

11 TB video, +
Facebook, Instagram,
Snapchat

Brooklyn (Nov 2014)
95M measurements
19350 IDs (~2%)
64 IDs/Ha.
we see 39 TB
(24 TB wifi,
15 TB cellular/mixed)
est. 2 PB/month?

24 TB video,
+ Facebook, Instagram

Los Angeles (Nov 2014)
503M measurements
56844 IDs (~2%)
~16 IDs/Ha.
we see 296 TB

est. 15 PB/month?

video/other breakdown and
motorist high frequency
measurements not fully
sorted yet.

Can we drill down without privacy concerns?

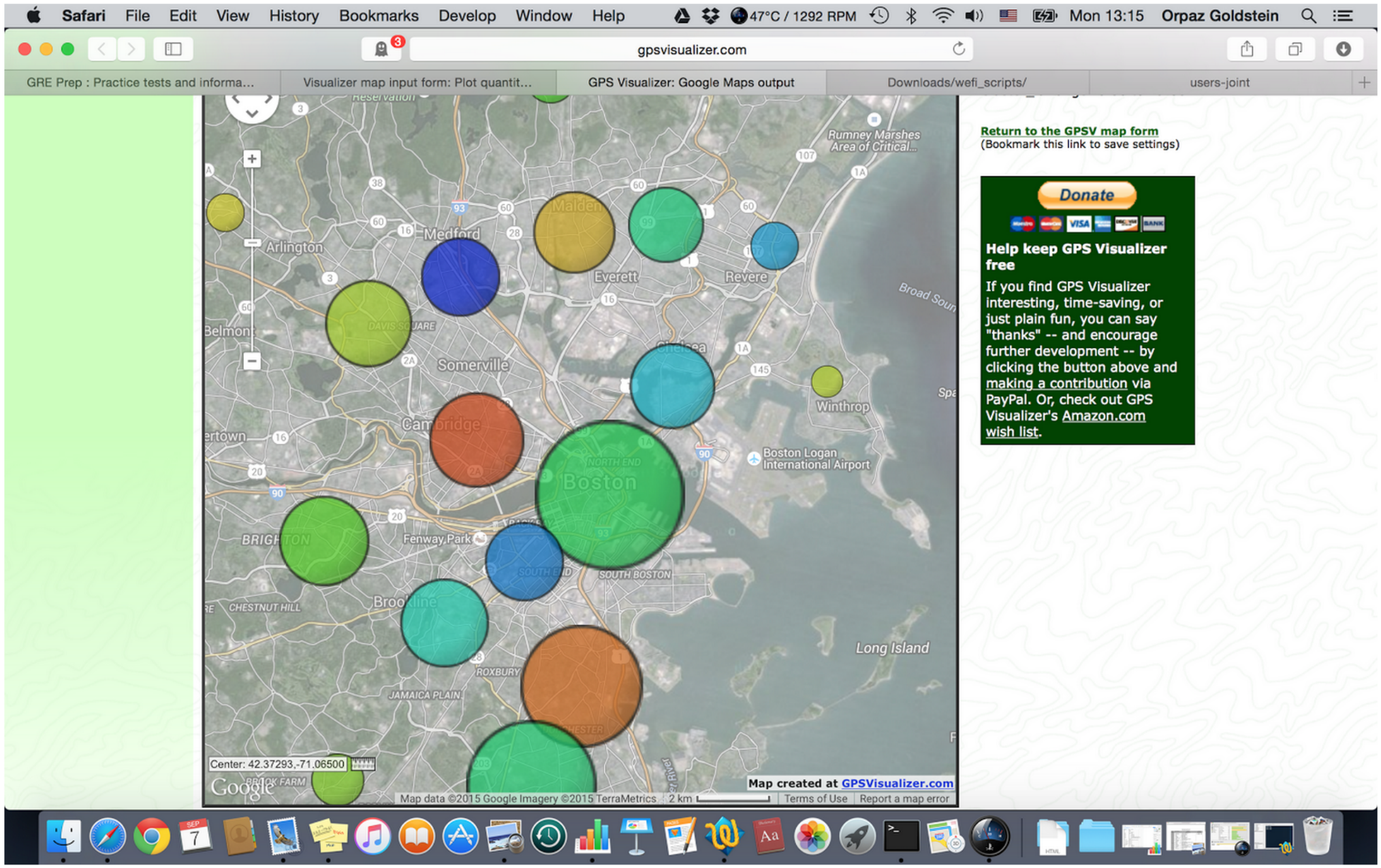
- Yes, groups that aggregate to safe levels are easily identified.

- Example — office workers (seen in same place 7 hours a day, at least 10 days in a month), drivers, Uber drivers and their customers…

- Localize by place — either by clustering or by coarse-graining (results consistent, clusters find more).

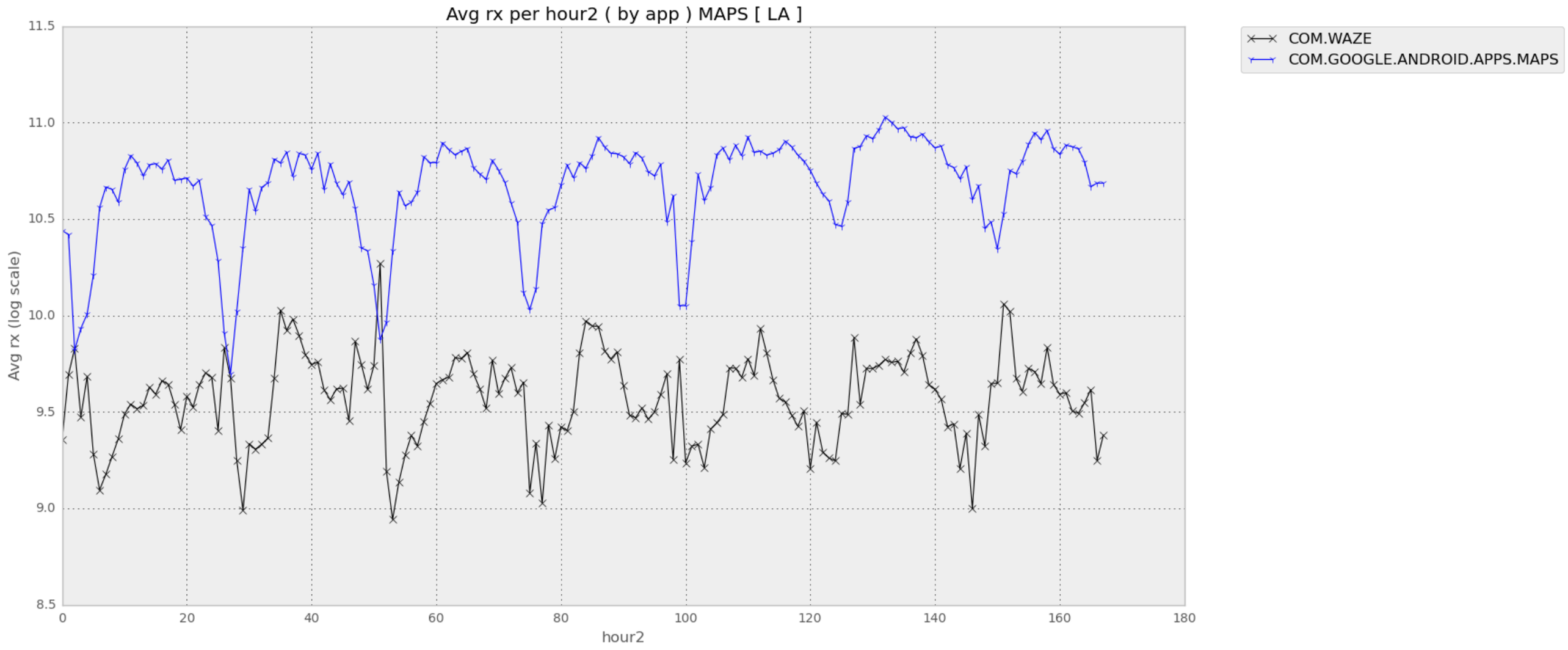
Office workers in Boston, for example

After filtering people who were at least 10 days within a month in one place at least 7 hours we get: 1544 unique id's left.
Their cluster:



Use application performance to inquire about network issues

- Example: two sources of map data, Google and Waze in LA, diurnal variation of download performance



Use network character to study application mobility