On the importance of context detection for crowdsourced mobile network measurements

Valentin Radu

Joint work with Dr. Mahesh K. Marina and Konstantinos Balampekos

SCONE, 5 November 2014
What is the problem?

- Exponential growth of mobile data traffic
  - Rapid adoption of smartphones and tables
  - Greater dependence on cloud services and video applications
- Pressure on mobile operators to monitor services.
- Discrepancy between the estimated coverage of operators and perceived user experience.
- Policy makers need to track the effect of public spending to improve mobile coverage.
Solutions

- Drive testing
  - Expensive
- Modelling the coverage
  - Limited accuracy
- Network-side passive analysis
  - No location precision,
  - No access to other sensors
- Crowdsourcing
Crowdsourcing

• Advantages
  – Cheap – smartphones are prevalent
  – True user experience

• Disadvantages (yet):
  – No control over context
    (indoor vs outdoor, bag vs in hand)

• Has good perspectives to be part of future monitoring systems.
Indoor-Outdoor

• Existing crowdsourcing solutions (e.g., Open Signal, MobiPerf) lack the indoor-outdoor detection capability.
  - We think this information is essential.

• Gember et al. (IMC'12) revile that moving the phone from hand to pocket can cause up to 79% difference in throughput, so indoor vs outdoor is expected to be even more important.
Signal strength (RSSI)

• We use RSSI as measurement metric because:
  – Intuitive metric for users, targeted by operators and regulators
  – Correlates well with throughput and with mobile battery drain
  – Unlike throughput, signal strength can be measured passively
  – Widely adopted by most crowdsourced mobile measurement systems.
Dataset

- OpenSignal dataset
  - Central London (58 square km) – 16k post codes
  - 8 million records
  - 3 years
  - Four major UK operators (O2, Vodafone, EE, 3)
  - 2G & 3G
OpenSignal dataset

Source: www.opensignal.com/
Indoor-Outdoor context detection

- GPS fix as outdoor and WiFi based location as indoor
  - The only reasonable method for post facto indoor-outdoor detection.
  - In controlled experiments accuracy 75-80%.
Findings

- OpenSignal, in Cell ID 10508514

Ofcom:
- > -91.7 dBm excellent
- < -105.5 dBm poor

Validation in controlled experiment
Context mix

- A general rule in the dataset – average RSSI:

Top 25 cell IDs based on # of measurements.

Biased towards indoor measured values since people spend ~80% of their time indoors.
Context differences

- Across the entire dataset – difference of average signal strengths (outdoor - indoor)

At least 150 measurements for each context.
At post code level

Poor < $-105.5\text{dBm}$

< Fair $\leq -100.9\text{dBm}$

< Average $\leq -96.3\text{dBm}$

< Good $\leq -91.7\text{dBm}$

< Excellent

Over 1500 measurements per context.
Validation

- Controlled experiment with indoor-outdoor ground truth.
- More than 5000 measurements across 18 distinct cell sectors.
5 dBm difference between outdoor and indoor in more than 60% of the cases.
Context-aware measurements

- Indoor-Outdoor context matters for more accurate coverage maps.
- Using the energy-hungry GPS is not ideal for crowdsourcing.
  - Also not always reliable, with 75-80% accuracy.
- We need better systems for crowdsourcing with:
  - Context detection on the phone – relying on other sensors
  - Energy efficient and more reliable.
Conclusions

- Distinguishing between measurements taken in different contexts (indoor-outdoor) is important
  - In the vast majority of cases the average RSSI in outdoors is higher than the one in indoors.
  - This can be as high as 18 dBm, with the median of 6 dBm.
- Aggregating across contexts is biased towards indoor measured values in large crowdsourced datasets.
Thanks you!

Remember, remember the 5th of November!

IS FOR VALENTIN ;)

valentin.radu@ed.ac.uk