Mobile paper review

# Some guideline

1. We must address all the criteria listed by 1AE(primary reviewer). Failed to do so will get our paper rejected unless compelling reasons are provided as part of the resubmission.
2. We can send anonymous emails to 1AE if we need further clarification of any review.

# Major problems by meta review: accuracy & generalization

1. How the data is collected:
   1. The length of the data (3 hours) is short

It would be good to explain why the authors feel this is an appropriate amount of data to analyze.

* 1. When was it collected

*Note: we have timestamp*

* 1. Culture and type of user: limited area, only three cities in china. Especially for Fig 13 ~ 15 (speed vs. unique apps used, app switch frequency, number of concurrent apps), their results are not meaningful for understanding general patterns. The authors need to describe the limitation of their analysis and method.

*Note: Although some results are affected by this limitation, the speed estimation procedure can also be applied to other similar datasets. We only use this dataset as a use case.*

1. Generalization:
   1. The contexts users use their smartphones are blurred with moving speed.
      1. Location
      2. Time
      3. Only for users who frequently moved and use smartphones
      4. Difference between inter-cell moving and staying within a cell

*Note: This is kind of a limitation of the speed estimation procedure itself. We only have limited data that can be used as speed estimation. If we incorporate location and time, the trend might be very unclear.*

* 1. How to incorporate wifi data  
     How representative are the results presented in this work given the fact that they consider cellular app usage and traffic only?

The authors should discuss or demonstrate the implications of considering cellular traffic only on the results presented in this work.

*Note: Our speed lower bound estimation can be applied to wifi data as well. For indoor wifi access point, just think wifi access point as another type of tower with shorter range. However, we can no longer make voronoi map with wifi APs, as it will overlap with cellphone tower coverage. Also, vehicles with wifi enabled might be very hard to dealt with.*

*I’m not very sure if I can increase the speed estimation with wifi locations, as there are some data available with wifi traces, I can give it a try.*

*We are more like massive data analysis, in my opinion, wifi are only for small subsets as it is not very easy to collect a large scale wifi data.*

1. Prove speed estimation accuracy
   1. The authors should demonstrate that their approach to estimate intra-cell movement speed is accurate. It is possible that to some extent the insights presented in this work result from the inaccurate movement speed estimations.

*Note: This is a very good comment. I do have feelings that some of our results might be inaccurate due to the speed estimations.*

* + 1. Addressing this shortcoming might for instance include an evaluation of the approach on a data set that contains GPS and cellular data. There are several publicly available data sets that might be helpful at this stage (reality mining, nokia, lifemap, etc.)

*Note: This is a very valuable suggestion. I will run our code on these datasets and see which one is the best fit.*

* + - 1. *Reality mining: 94 subjects, no GPS, with Wifi*

*The Reality Mining study followed ninety four subjects using mobile phones pre-installed with several pieces of software that recorded and sent the researcher data about call logs, Bluetooth devices in proximity of approximately five meters, cell tower IDs, application usage, and phone status.*

* + - 1. *Nokia dataset: 1 subject, with GPS and Wifi*

*For each locale (see the traces included - downtown, ravenna, and kirkland), we drove around the areas for sixty minutes with a laptop, a GPS unit, and a Nokia 6600 cell phone. 802.11 scans were performed at 4Hz using an Orinoco 802.11 interface in the laptop. GPS readings were taken at approximately 1Hz using an external serial GPS unit. Finally, the GSM measurements were taken at 1Hz by the Nokia 6600 and relayed to the laptop via Bluetooth4. At all times we tried to navigate within areas in which GPS lock would not be lost as GPS forms the round truth location to be used to estimate beacon positions and Place Lab's accuracy.*

* + - 1. *Lifemap: 8 subject, with GPS and Wifi*

*We collected real traces from 8 graduate students over four weeks using HTC Hero, HTC Desire, and Samsung Galaxy S smartphones. SmartDC was running as a background service to automatically collect the user's mobility and to trace sensor usage time. Participants installed SmartDC on their primary phones. To collect the ground truth, the participants explicitly labeled the place names and kept a diary of places they had visited with the entrance and departure times.*

* + 1. Running a small custom study  
       Even a simple study with 20 users where ground truth and cell records were collected would be sufficient.
  1. Any assumptions in estimating moving speed are made without much explanation. These assumptions should be proved and discussed in detail.

*Note: These assumptions we made are only for speed estimations. However, the distance lower bound estimation can be applied to more advanced trajectory estimation method. I think we need to make it more clear that my method is helpful to filter out inaccuracy distance estimates. So it is possible to combine with other advanced methods to improve their performance.*

* + 1. users move with a constant speed (4.1)
    2. straight line trajectory (4.3)
  1. The classes of moving speed should be analyzed and discussed in detail.

*Note: Right now we classify estimated speed into 5 categories with 20km interval (just a random value). Do we need to use a classification algorithm to classify the moving speed? But we cannot really afford a larger number of classes (now we only have 5) due to lack of data, e.g., lack of high speed user data in some app category.*

# Other Major problems:

1. The authors should describe how the massage smartphone usage data were dealt. Any tools and environments?
2. How are our insights useful? Motivations.
3. Back up our insights with some evidence (a survey or a focus group or some observations studies)

# Minor problems:

1. Figs and listings one the same page with related text
2. Inconsistent writing of PBE in Sec 4.1 and Sec 4.2
3. Potential naming (variables) inconsistencies between Alg1 and Sec 4.2
4. Fig10a y-axis label: is it not supposed to be “bytes/sec” as explained in the corresponding section?
5. Sec 5.3: how do you define “a data access”. is it a single CDR in the data set?
6. What is the overall value / take-home message of Fig11 and Fig12? Is it not better to have relative values if the number of instances differ?
7. Fig14 and Fig15 should be a bit smaller to match the font size of the text

# Suggested References:

[1] Kyunghan Lee,Joohyun Lee,Yung Yi,Injong Rhee, and Song Chong. 2013.

Mobile Data Offloading: How Much can WiFi Deliver? IEEE/ACM Transactions

On Networking 21, 2 (2013), 536–551.

[2] Paul Baumann and Silvia Santini. 2014. How the availability of Wi-Fi

connections influences the use of mobile devices. In Proceedings of the

2014 ACM International Joint Conference on Pervasive and Ubiquitous

Computing Adjunct Publication - UbiComp ’14 Adjunct. (2014), 367–372.