"Stop Questioning Me!": Towards Optimizing User Involvement during Data Collection on Mobile Devices

Nicholas Micallef
Interactive and Trustworthy
Technologies Research Group
Glasgow Caledonian University
Glasgow, United Kingdom
nicholas.micallef@gcu.ac.uk

Lynne Baillie
Interactive and Trustworthy
Technologies Research Group
Glasgow Caledonian University
Glasgow, United Kingdom
lynne.baillie@gcu.ac.uk

Mike Just
Interactive and Trustworthy
Technologies Research Group
Glasgow Caledonian University
Glasgow, United Kingdom
mike.just@gcu.ac.uk

Gunes Kayacik
Interactive and Trustworthy
Technologies Research Group
Glasgow Caledonian University
Glasgow, United Kingdom
gunes.kayacik@gcu.ac.uk

MobileHCI '13, Aug 27 – 30, 2013, Munich, Germany.
ACM 978-1-4503-2273-7/13/08.
http://dx.doi.org/10.1145/2493190.2494426

Abstract
Current methods of behavioral data collection from mobile devices either require significant involvement from participants to verify the 'ground truth' of the data, or approximations that involve post-experiment comparisons to seed data. In this paper we argue that user involvement can be gracefully reduced by performing more intelligent seed comparisons. We aim to reduce the participant involvement to the 'most interesting' temporal slots, both during the experiment and in post-experiment verification. We carried out a 2 week study with 4 users, consisting of an initial opportunistic gathering of mobile sensor data. Our findings suggest that by using such a method we can significantly reduce user involvement.

Author Keywords
Mobile; Sensing; HCI;

ACM Classification Keywords
H.5.2. Information interfaces and presentation: User Interfaces – Input devices and strategies, evaluation;

General Terms
Human Factors; Measurement; Experimentation;

Introduction

An open problem in the study of mobile device sensing relates to how active the device owner should be in the collection and analysis of sensor data [1]. We define mobile sensing as the collection of raw sensor data related to a user's activity. This data is then used to derive contextual information about the user's behavior. For example, accelerometer sensor readings may allow one to determine if a user is walking or running. This contextual information is also important because it will help in determining whether it is a good/optimal time to ask for user feedback.

Sensors are often used to infer some behavior of the owner, and the owner is typically best positioned to validate sensor readings, especially for uncommon situations (e.g. establishing that on Tuesday from 6:30-7pm the user was at the grocery store). However, such methods can be intrusive, sometimes requiring heavy user involvement [1], and they can also be very dependent on the user's willingness and ability to actively participate. This heavy user involvement is exposed in projects such as the CenceMe project [2], in which users were asked to annotate their actions every 15 to 30 minutes for an entire week. In SoundSense [3] users were asked to provide a textual description or select ignore if a new sound is detected. Similarly, in Phoneprioception [4] users were prompted every 30 minutes with specific questions about where their phone was at a specific time. An interesting result from this research is that users labeled on average 52% (males) and 48% (females) of the data. Such a high ratio of user involvement to collected data provides the motivation for discovering the most interesting timeslots to ask questions to users.

The other extreme uses opportunistic sensing where data collection is simply carried out in the background, but without explicit user feedback [5]. Here, a low amount of explicit user involvement is required [1]. However, such opportunistic data analysis can lead to less accurate inferences due to false interpretations in lieu of participant feedback. The EU project OPPORTUNITY [6] collected a data set of naturalistic human activities in a sensor rich environment: a room simulating a studio flat where subjects performed daily activities. 12 participants undertook common daily living activities in this closed environment, producing an average of 2 hours of data each, which resulted into 25 hours of sensor data. They labeled ground truth by annotating activities during the recording and by verifying them afterwards. The above study demonstrates that undertaking complete verification of ground truth can be difficult and time consuming.

Our research aims to optimize the usability of ground truth data collection during mobile sensing by reducing the number of questions that are asked to users by means of identifying the most interesting temporal slots about which to ask questions. In this paper we hypothesize that techniques extracted from analyzing opportunistic data would help in this regard. For example, we argue that by automatically performing seed comparisons during data collection, we can more effectively determine the points at which such seed comparisons are insufficient, and for which real-time user input is desirable. Similar improvements can reduce the user involvement in confirming activities post-experiment. Our study consisted of a 2 week pilot experiment with 4 users, starting with an initial opportunistic gathering of raw sensor data, followed by

Attribut-es	Type	Data Sources
At home	Location	20 minutes of Wi-Fi access points
At the office	Location	20 minutes of Wi-Fi access points
On break	Location	20 minutes of Wi-Fi access points
Commuti-ng	Location	20 minutes of Wi-Fi access points
Walking	Physical Activity	20 minutes of Accelerometer
Stationar-y	Physical Activity	20 minutes of Accelerometer
Stairs Up	Physical Activity	20 minutes of Accelerometer
Stairs down	Physical Activity	20 minutes of Accelerometer

Table 1. Seed ground truth attributes collected for each participant

an analysis of the data that includes some seed comparison methods (see, Table 1).

There is a lot of related work on topics such as semantic location detection [7, 8, 9] and physical activity recognition [10] in the areas of ubicomp/percomp communities. The focus of this research is not to directly improve upon such techniques. Rather, we want to show that by using input sensors such as location and phyiscal activity (but we are not limiting ourselves to just these two sensors) we could identify interesting temporal slots which would optimize the level of feedback required by the user.

Methodology

We focus on the following two activities to identify the most interesting temporal slots in which to ask questions: user location (e.g. where was the user from 6-10pm), and physical activity of the user (e.g. were there changes in physical activity from 1-2pm). As we continue this work, we plan to extend our methods with sensors for light and noise that can improve our determination of location and physical activity.

We undertook a first study with 4 participants in which we collected opportunistic sensor data for 2 weeks. The participants were instructed to carry their phone with them throughout the duration of the study. These participants were aged 26 - 35 and used the following phones: Samsung Galaxy S2, Samsung Galaxy Ace and Galaxy Nexus (2). Specific seed ground truth features related to each user's location and physical activity (Table 1) were collected with the aim of analyzing the opportunistic data together with these seed ground truth features. These seed ground truth features were collected in 3 ways: from the participants themselves (e.g., users highlighted time periods when they were walking); from the experimenter (e.g., experimenter went to specific locations to collect Wi-Fi access points), and from external sources, we used tables of existing data collected and categorized by others (e.g. RICE lab data). The analysis in this instance was carried out offline, but we plan to automate this process for our second study in order to augment the real-time data collection process to find the most interesting temporal slots for soliciting user input from the collected data. The 'future work' section will explain how the second study (starting soon) will build on the findings of this first study by evaluating the discovery of these interesting temporal slots to optimize user involvement in the collection of ground truth during mobile sensing.

Location: Lists of Wi-Fi access points were collected every 5 minutes. The patterns and frequencies related to user location can be quickly inferred from processing opportunistic data. For example, to have a better understanding of a user's location we analyze the names of the Wi-Fi access points and if the names are semantically descriptive (e.g., “ABC-Shopping-Mall”) then we can infer with a certain level of confidence that at that time the user was “at the shopping mall”. This is not always the case since there might be instances where the user seems to be at a shopping mall but he is actually at a nearby office. In our analysis we also found it helpful to consider the time of the day that the user was in a particular location. For example, if we see a repetitive pattern of Wi-Fi access points from 6pm till 8am on each day then we infer that the user was “at home” during that time.

We also compare the seed ground truth attributes (as listed in Table 1) to the processed opportunistic data.

U1	U2	U3	U4	Ave
1) Frequently recurring lists of Wi-Fi access points				
2	2	2	2	2
2) Average no of hours/day covered by recurring patterns				
18.6	11.3	18	18.8	16.9
3) Average no of hours/day covered by seed attributes				
21	12.3	18.6	21.8	18.4
4) % of time covered by recurring patterns				
77.4%	47.2%	75.0%	78.5%	69.5%
5) % of time covered by seed attributes				
87.5%	51.4%	77.5%	90.6%	76.8%
6) No of times in which uncommon Wi-Fi access points are met during a week				
13	40	28	19	25

Figure 1. Summary results for user location.

For location, a similarity value is calculated to distinguish amongst two sets. This value defines the probabilistic similarity of the opportunistic sensor set to the seed ground truth attributes to which it has been compared. To obtain the similarity value for location, the list of Wi-Fi networks that were discovered during a 30 minute timeframe are compared to the lists of Wi-Fi networks attached to the collected seed ground attributes. The number of matched Wi-Fi networks is then used to decide whether that time frame can be labeled as matching any of the seed ground truth attributes. For example, if from 6-6:30pm four of the five Wi-Fi networks that are attached to the "at home" seed ground truth attribute are discovered, then it can be inferred that during that time the user was "at home".

Physical activity: The device acceleration on the x, y and z axes (accelerometer data) is used to infer the physical activity of the user. Acceleration data was collected every 200ms. During interpretation, the raw accelerometer data is divided in 10 seconds intervals and the peaks for each of the x, y and z values in the intervals are counted. Research suggests that activities such as walking have a clear repetitive wave pattern in most of the axes [10]. Calculating the average number of peaks on the z axes in a set of 10 second time frames is already sufficient to determine when a user is stationary or when a user is walking, or going up or down stairs, because there are a much higher number of peaks in the latter rather than in the former. Then to differentiate amongst walking, climbing up or down the stairs, the average number of peaks for the y and x values are used. There are a much higher number of peaks in the y axis for walking when compared to stairs up. As regards to stairs down the frequency of the peaks is lower than that of walking. There is also a noticeable difference in the x axis amongst climbing up and down the stairs.

Results

In some cases, by opportunistically interpreting the semantic names of the Wi-Fi access points together with the time slots in which these access points are encountered we can infer that a user is at university every day of the week from 10am-5pm since the university name is consistently one of the top Wi-Fi access points during this time. Though, this method is not very helpful when the names of the Wi-Fi access points do not carry a semantic meaning. For this reason some form of controlled labeling is required to get more meaning from this kind of raw data. Item no. 3 in Figure 1 shows that when using location seed data (described in Table 1) we manage to label an average of 18.4 hours per day amongst all users. This means that during this time the users where either at home, on a break, at work or commuting.

From the analysis of Wi-Fi access points names and using known location labels we can learn that to minimize the number of questions that will be asked to the user during sensing the system should contain an algorithm which extracts regular patterns of similar sets of related Wi-Fi access points (e.g., those points that are often clustered together). Using this algorithm on the data collected in our experiment we demonstrated that by asking the question "Where are you now?" on four separate occasions during 2 weeks of sensing (which would be related to the seed ground truth location attributes listed in table 1) the system would be capable of labeling the users location for 76.5% of a week, as shown in item no. 5 in Figure 1.

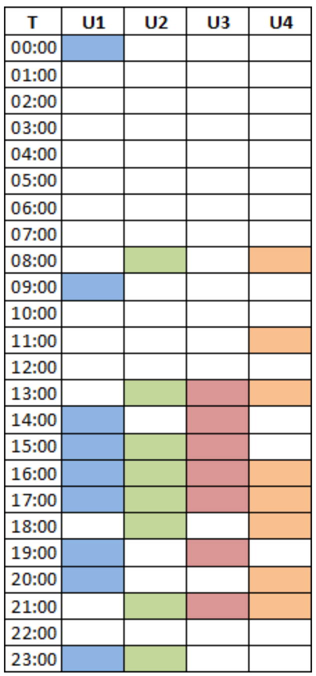

Figure 2. Time frames, which contain the most changes in physical activity.

By comparing the collected seed data for physical activity to the opportunistic accelerometer data we were able to measure the number of changes in physical activity for every user. For example, from the results that we collected from 1-2pm, user 1 is marked as stationary, then walking, then stationary again, and so on. Figure 2 shows the time frames which contain several changes in physical activity; notice that the afternoon is the time in which the users carry out most of their physical activities. For every user (except user 3) there is also an hour in the morning in which considerable changes in activity were detected. When comparing this to the list of Wi-Fi access points one could also infer that in most cases the user was commuting in this particular hour because the number of Wi-Fi access points encountered in this particular hour is much longer and diverse than in the other hours. By combining the "higher level" information from Figure 2 to the list of Wi-Fi access points described earlier one can determine when the users are most active during the time frames in which the same Wi-Fi access points are constantly appearing. In other words, the person might be at home but he may not always be sitting on the couch. Therefore, the "higher level" information about the changes in activity shown in Figure 2 combined with the list of Wi-Fi access points encountered throughout a day will help in identifying the most interesting temporal time frames in which the user should be queried for more information when he is in the most recurring locations.

An important point to take from this discussion is that by combining the "higher level" information inferred from user location and physical activity one can have a better understanding of what are the most interesting time slots by defining the most appropriate time frames to ask questions, and how much value would be added if the users were queried about those actions. This will help us both during the experiment, because by using this data an algorithm can automatically decide when to ask questions, and during post-experiment evaluation by reducing the number of questions that will be asked to the user. This strategy can also be used to obtain more information about time slots which have uncommon lists of Wi-Fi access points, e.g., item no 6. in Figure 1 shows the total number of times that uncommon Wi-Fi access points were encountered during a week. This demonstrates that the average number is quite high and it wouldn't be reasonable to ask questions about all these locations. Therefore, if the user has a higher than usual amount of changes in physical activity in a specific hour in the morning and in this same hour there is a list of uncommon Wi-Fi access points then it would be valuable to explicitly ask the user what he was doing at that particular time (e.g. "Were you commuting?"). An important factor to determine whether a temporal time slot is important or not, and therefore query the user about it, is the frequency in which a particular scenario happens. Thus, if it is a scenario that happens repetitively then it is worth asking the question; "is this a one off event?". If the user confirms that this is a one off event then the system should not disturb the user with questions that add little value to the overall interpretation.

Conclusion

In this paper we show how user involvement can be gracefully reduced by performing more intelligent seed comparisons. By doing this comparison we can reduce the level of participant involvement to the 'most interesting' temporal slots, both during the experiment and in post-experiment activities, with the latter being an easier and less time consuming option to implement. Our results show that an algorithm which extracts regular patterns of similar sets of related Wi-Fi access points would minimize user involvement. But most importantly we found that by combining “higher level” information inferred from user location and physical activity we can determine the most appropriate time frames to ask questions, get more details about what's happening in the most recurring locations, and identify which are the most interesting temporal patterns amongst uncommon locations, thereby optimizing the level of involvement required of the user.

Future Work

For our future work we plan to extend our methods by adding environmental sensor inputs such as light, noise and magnetic field, and build on the findings of this first study by evaluating user involvement when collecting ground truth data during mobile sensing in two different ways. The first technique would be to conduct post experiment analysis (e.g. questionnaires or interviews) to retrieve the details about the user’s interesting temporal slots. The second technique would be to build a custom application in which users are asked to label interesting temporal slots 20/30 minutes after detection. Feedback from users will be collected to determine the level of intrusiveness when using these techniques.

Acknowledgements

The research leading to these results has received funding from the People Programme (Marie Curie Actions) of the European Union's Seventh Framework Programme (FP7/2007-2013) under REA grant agreement n° PIIF-GA-2011-301536. We would also like to thank the reviewers for their useful comments that helped to improve the paper.

References

[1] Lane, N., Miluzzo, E., Hong L., Peebles, D., Choudhury, T., Campbell, A. “A survey of mobile phone sensing”. In IEEE Communications Magazine, vol.48, no.9, pp.140-150, Sept. 2010.

[2] Miluzzo, E., Lane, N., Eisenman, S., Campbell, A. “CenceMe: injecting sensing presence into social networking applications”. In EuroSSC 2007.

[3] Lu, H., Pan, W., Lane, N., Choudhury, T., Campbell, T. “SoundSense: scalable sound sensing for people-centric applications on mobile phones”. MobiSys 2009.

[4] Wiese, J., Saponas, T. S., Brush, A. B. “Phoneprioception: Enabling Mobile Phones to Infer Where They Are Kept”. In ACM CHI 2013.

[5] Campbell, A., Eisenman, S., Lane, N., Miluzzo, E., Peterson, R. “People-centric urban sensing”. In IEEE WICON 2006.

[6] Roggen, D., Calatroni, A., Rossi, M., Holleczek, T., Forster, K., Troster, G., Lukowicz, P.,Bannach, D., Pirkl, G., Ferscha, A., Doppler, J., Holzmann, C., Kurz, M., Holl, G., Chavarriaga, R., Creatura, M., del Millan, R. “Collecting complex activity data sets in highly rich networked sensor environments”. In IEEE INSS 2010.

[7] Cheng, Y., Chawathe, Y., LaMarca, A., Krumm, J. “Accuracy characterization for metropolitan-scale Wi-Fi localization”. In MobiSys '05.

[8] Fox, D., Hightower, J., Lin L., Schulz, D., Borriello, G. “Bayesian filtering for location estimation”. In IEEE Pervasive Computing, vol.2, no.3, pp.24,33, 2003.

[9] Mao, G., Fidan, B., Anderson, B. “Wireless sensor network localization techniques”. In Computer Networks, Volume 51, Issue 10, 2007.

[10] Kwapisz, J., Weiss, G., Moore, S. “Activity recognition using cell phone accelerometers”. In SIGKDD Explor. March 2011, 74-82.