User Activity Understanding from Mobile Phone Sensors

Yuki Arase
Microsoft Research Asia
yukiar@microsoft.com

Fei Ren
Beihang University
faylon.renfei@gmail.com

Xing Xie
Microsoft Research Asia
xingx@microsoft.com

ABSTRACT
Context acquisition is an important technology for ubiquitous computing. An ideal approach would be easy to deploy and non-intrusive to people's life. Mobile phones equipped with advanced sensors are preferable platform owing to their user-friendliness and freedom from extra costs to deploy. In this study, we propose to use a mobile phone to detect user contexts. We formally define the concept of context and then describe applications that leverage people's long-term activity, which can be inferred from their contexts.

Author Keywords activity understanding; context acquisition; context-awareness; mobile phone; sensor

ACM Classification Keywords H.1.2 [Models and Principles]: User/Machine Systems –Human information processing

General Terms Algorithms; Measurement

INTRODUCTION
Context acquisition is a core technology for ubiquitous computing. The goal of context acquisition is to understand users, i.e., to detect and interpret what they are doing in short-term and long-term. To detect contexts, we need to embed sensors to people's life. For user convenience, context acquisition approaches should satisfy the following three requirements [6]: to be *Easy to deploy*, *Easy to use*, and *Non-intrusive*. Systems that require expensive cost to deploy or taking instructions to use may face difficulty to obtain a large number of users. Non-intrusiveness is also important, because it would be annoying if a system disturbs people's daily life. Especially, as a kind of non-intrusiveness, the systems should not impress users as if they are always monitored by an outsider.

Most conventional approaches situate extra devices around a user, such as attaching RFID tags to home appliances to detect the user's contexts related to their usage and putting multiple sensors to the user's body to detect his/her movements. Since there has not been any familiar device to ordinary people, researchers had to embed something unfamiliar to their life. These devices inevitably need costs to deploy and users tend to feel them as intrusions, since such devices are not domestic to them. Therefore, these approaches cannot satisfy the requirements.

However, a change has brought by the advancement of mobile phones. Recent mobile phones are equipped with sensors that are effective for context acquisition, i.e., a GPS receiver, digital compass, acceleration sensor, light sensor, and proximity sensor. They also have WiFi, camera, and microphone, which are also useful for context acquisition. Therefore, they do not need any extra costs to detect contexts. Considering users are willing to carry their mobile phones (a survey conducted by Chipchase et al. showed that people consider a mobile phone as one of essentials to take when outside the home, as well as a key and wallet [1]), we can expect that the mobile phones naturally fit people's life without intruding them, i.e., they can satisfy all the requirements for context acquisition techniques.

A challenge to use a mobile phone for context acquisition is that we cannot fix the position and orientation of the phone. This makes context detection more difficult, since outputs of sensors vary depending on how a user holds his/her phone even for the same context. For example, accelerations caused by the user's movements are different on body parts and the sound captured by a microphone becomes muffled when the phone is in pockets and bags.

In this poster, we present our study on context acquisition using mobile phone sensors. First we formally define the concept of context. Then we present applications that leverage users' long-term activities, which can be derived from their contexts detected by mobile phones.

DEFINITION OF CONTEXT
Context modeling and representation is still an open problem [4]. Since we detect contexts using sensors, we define them based on dimensions of sensor outputs, which are independent with each other. In our definition, a user's context is composed of four elements, each of which has a hierarchical structure according to interpretation levels:

- Time: represents when a context arises
  - Timestamp obtainable by an inside clock
- Location: represents where the context arises
  - Latitude/longitude obtainable by GPS
  - Area inferable by WiFi and a GSM cell tower
- Environment: represents the environmental condition of the location
  - Noise volume obtainable by a microphone
  - Illumination obtainable by a light sensor
- Physical movement: represents the user's body movement, including the movement caused by outer factors, e.g., riding a car.
  - Accelerometer values obtainable by an acceleration sensor
  - Moving speed obtainable by GPS

UbiComp'10, September 26–29, 2010, Copenhagen, Denmark.
ACM 978-1-60558-843-8/10/09.

We note that this definition is general so as to be applicable to other context acquisition techniques and can be extended when more sensors are available.

Table 1 shows hierarchical interpretation examples of each element. By combining these basic elements, we can infer the user's context. For the example of Table 1, we can see this user is drinking at a bar considering that it is Friday after work, the location corresponds to the bar of the user's favorite, and the user is staying in a noisy environment. To detect contexts, we first need to detect the elements, interpret them accordingly, and then infer the context by combining the interpretations of all elements.

Table 1. Example of context elements

Element	Interpretation hierarchy		
	Low	Middle	High
Time	10pm, Friday	Night	After work
Location	55.683,12.589	"ABC Bar"	Favorite bar
Environment	70db	Noise level: middle	Need laud voice for conversation
Physical movement	Accelerations of 3-axis	Mean, energy, FFT coefficients etc.	Staying

CONTEXT ACQUISITION BY MOBILE PHONES

According to a survey conducted by Cui et al., about 60% females carry their mobile phones in their hand bags, while 60% males carry their phones in their trousers pockets [3]. Due to this natural variation of placements, context acquisition by mobile phones is challenging. Although Reddy et al. [6] showed that users' transportation modes, such as walking and motors, can be detected with 94% accuracy by using a mobile phone, more research effort is needed to cover variety of contexts.

In addition to [6], we conducted a preliminary experiment to verify the capability of mobile phones for capturing a user's body movement. In our experiment, we used an acceleration sensor and GPS receiver of a mobile phone. As a result, we found that we can distinguish the user is riding a car or a bus with respectively 67% and 71% accuracies, even though these are quite similar body movements. Also we confirmed that we can count the user's steps fairly accurately. If the user holds the phone in hand or put in pockets of waist up, the accuracy was more than 96%. If the user puts the phone in his/her trousers pocket, the accuracy drops to about 80%. This is because the movement of legs affects the sensor. To detect more detailed contexts, e.g., turning right and left, a digital compass is available.

ACTIVITY UNDERSTANDING

On top of contexts acquired by mobile phone sensors, we can infer a user's long-term activities. We define the activity as a people's task for a certain period, such as having fun at a bar with friends. In other words, the context is a snapshot of a long-term activity, i.e., being at a bar at a specific time is a part of the partying activity.

We routinely conduct activities. Let us take a morning of a businessman as an example. He checks Twitter (http://twitter.com) to see his friends' status using his mobile phone while commuting. He drops by a Starbucks for morning coffee and takes an elevator to the floor of his office. Longer activity is going out with his friends every Friday night. We repeat these activities almost unconsciously, in spite that exceptional events happen, for example, the Starbucks may be closed and we may have an urgent work in Friday night.

We can mine these activities based on contexts detected by mobile phones, which are fundamentals of context-aware applications. For example, an application can help users efficiently conduct activities, by updating Twitter in back-end so that the user can check it without waiting for downloading the data and notifying when the Starbucks is closed and recommending other coffee shops. One benefit of this application is that it can prevent the user from being exposed to over-load information by updating information at a time the user needs it. It would be annoying to update Twitter during the user attending a business meeting. By considering the user's contexts to select and update information [5], we can also reduce the energy consumption of the mobile phone. Another application prompts the user to use stairs at the point where he always takes the elevator to have more exercise for his health [2].

SUMMARY

In this poster, we described our initial study on user activity understanding by mobile phones. We will work for detecting more diverse contexts and also study effectiveness of other sensors, e.g., a microphone and camera, and a combination of them. Moreover, we plan to study the long term activity detection.

REFERENCES

1. Chipchase, J., Persson, P., Piippo, P., Aarras, M., and Yamamoto, T. Mobile essentials: field study and concepting. *Proc. of the conference on Designing for User eXperience* (2005).
2. Consolvo, S., Libby, R., Smith, I., et al. Activity sensing in the wild: a field trial of ubifit garden. *Proc. of the conference on Human factors in computing systems* (*CHI '08*), (2008), 1797-1806.
3. Cui, Y., Chipchase, J., and Ichikawa, F. A cross culture study on phone carrying and physical personalization. *Proc. of the International Conference on Human-Computer Interaction* (*HCI International '07*), (2007), 483-492.
4. Dey, A. Context-aware computing. In J. Krumm Eds., Ubiquitous computing fundamentals, Chapman&Hall/CRC, (2009), 321-352.
5. Han, J.H., Xie, X., and Woo, W. Context-based real-time local hot topic browser for mobile user. *Proc. of the International Conference on Pervasive Computing* (*Pervasive '10*), (2010), poster.
6. Reddy, S., Mun, M., Burke, J., Estrin, D., Hansen, M., and Srivastava, M. Using mobile phones to determine transportation modes. *ACM Transactions on Sensor Networks 6*, 2 (2010), 1-27.