Human-Activity Recognition (HAR) Everywhere: Embedded, Mobile, Desktop, Home & Cloud

T. Scott Saponas, Brian R. Meyers and A.J. Bernheim Brush
Microsoft Research
One Microsoft Way, Redmond, WA 98052, USA
{ssaponas, brianme, ajbrush}@microsoft.com

ABSTRACT
Human Activity Recognition (HAR) has been a topic of investigation for multiple decades. Researchers have explored a variety of approaches including a service running in the background on a mobile phone, cameras in the environment, RFID readers on people, and many more. In this note, we advocate that for HAR to make the leap from interesting research area to a mainstream technology that enhances our day-to-day experiences with computing it must utilize many parts of a person's computing ecosystem: from low-power embedded hardware that runs continuously and can be mounted directly in a phone to a service in the Cloud that tracks a person's activities over the long-term and computes current activity from every available input. We provide a brief discussion of how such a system might be constructed and used.

1. Introduction
Current research has succeeded at building a collection of vertical proof-of-concept Human Activity Recognition (HAR) systems, but we have yet to see HAR bridge the gap to become an integrated part of people's experiences on the go or in the home. In order to make this transition, we believe that future HAR systems must adopt a multi-tiered approach integrating the full gambit of computing platforms from embedded devices to the Cloud. In this note, we briefly review the state-of-the-art, describe two key challenges for HAR on mobile phones, address why we believe it is critical for HAR systems to be structured in a tiered manner, and give brief examples of the architecture of such a structure.

2. State of the art
Research in HAR has set the stage for a revolution; our computing environments are going to transform into smart systems that understand people's context and deliver experiences that anticipate people's needs. For example, in the not-to-distant-future, when I glance at my mobile phone resting in my car's cup holder, the screen will illuminate and having determined that I am currently driving toward work show the traffic between my current location and work, my estimated travel time and the time of my first meeting. Researchers have explored a variety of approaches to gather important context for this scenario (i.e. am I currently driving, in which direction am I driving, what are my likely destinations [5]) and many others. We have seen systems that can infer activity based on GPS traces [5], sensors placed around a home [11], accelerometers around the body [1], custom belt-worn sensor platforms [3], RFID reader built into a glove [8], and using the accelerometer on a mobile phone [10]. We have also seen that these sensing solutions can provide enough accuracy to drive high-value applications such as tracking one's fitness-related activities throughout their day [4]. However, despite all this work, we still do not have many examples of activity recognition being adopted in people's everyday computing systems.

3. HAR on Mobile Phones
A requirement for applications of HAR on mobile phones is that they integrate seamlessly into a person's use of their phone. We derive two key challenges from this requirement: (1) HAR algorithms must be able to work in a **placement independent** fashion (i.e. they are accurate regardless if a person carries their mobile phone in their pocket, purse, or backpack); (2) HAR software is able to **operate continuously** without reducing the usable battery life of the phone. In this section, we briefly review existing work addressing these challenges.

To our knowledge, no existing work directly addresses the general problem of (1). A subset of this challenge is making phone-in-pocket based HAR algorithms robust to the orientation of a person's phone in their pocket. In this space, Blanke and Schiele have demonstrated the ability to distinguish moving events in an office environment using inertial sensing that factors out phone orientation [2]. In a different vein of work, Kunze and Lukowicz have explored recognizing activities using an accelerometer placed in different places around the upper leg [6]. One of their key findings is that adding a gyroscope to the accelerometer can make it a more tractable problem.

A potential approach to addressing (2) that has been proposed by Priyantha, *et al*. is to apply the concept of offloading to a subset of HAR algorithms on mobile phones [9]. In their approach, several key dedicated sensors such as an accelerometer are connected to a low-power processor. These sensors can handle such tasks as step counting, low-level activity inference (still, walking, jogging, in a car), and even detection and collection of audio for speaker detection. The low-power subsystem can run continuously and only wake up the main application processor at key times (when the history buffer is full, more processing power is needed, or to turn on more expensive sensors such as GPS or GSM radio tower logging). Using a low-power subsystem in this manner greatly reduces the amount of power needed to sense continuously.

4. HAR in the Cloud
While offloading some HAR to a low-power microcontroller and being able to recognize activities regardless of phone placement addresses many of the challenges of continuous HAR on phones, it is only one part of the larger HAR problem space. To start with,

we need to be able to fuse data collected on a person's phone with the larger set of data from all the other sources (perhaps a depth sensing camera part of a home-entertainment system or a motion sensor in their office space) and make that available to all their devices and services. We need adequate secure storage to handle this continuous stream of data and we need sufficient computation power to learn patterns that might well span multiple months and many Gigabytes of data. In addition, the results of all this analysis must be available to a wide variety of devices in order to deliver the experience we are discussing.

5. Mechanics of a Tiered Approach

Once you have all the activity data flowing to the Cloud, the next step is the design of the service necessary to support such a system. Consider the model necessary to determine if a particular person is heading to work or not. Although this information could be learned on a single device you certainly do not want to have to retrain the device if you upgrade to a new phone, nor do you want to train your phone, your car, your bicycle computer or any other device that might want to know if you are heading to work or not. These models, once learned, are the heart of a HAR system and should be generalized and then made available to any of your devices. Abstracting and generalizing this kind of data so that regardless of the specific make and model of sensor it can use data that has already been learned by a previous system is a critical and unsolved research area.

In addition to merely learning and generalizing activity models, the Cloud must resolve conflicting sensor readings, acting as a mediator for separate sensor systems. For example, if the Cloud components receive information that the accelerometers are moderately certain that they are in an automobile, but the GPS reports being in the middle of a body of water within a three meter error, it falls to the Cloud to disambiguate and perhaps tell the accelerometers that they are wrong and to begin learning a model for piloting a boat.

6. Conclusion

The purpose of this note is to highlight key challenges facing HAR's adoption in future commercial systems. We believe that HAR on mobile phones is inherently an important part of this equation and as such highlight two important HAR challenges in this area: placement independence and continuous operation. We also believe that no one device will be able to determine all of the context necessary to deliver the experiences we want HAR to enable. To this point, we suggest that future HAR solutions have a tiered approach leveraging all parts of the computing ecosystem from embedded hardware to the Cloud. We do not mean to suggest that there are not many other important unanswered research questions in core HAR; alternatively, we highlight these challenges as being particularly important for enabling HAR's move from research deployments to the masses.

7. REFERENCES

[1] Bao, L. & Intille, S.S. 2004. Activity recognition from user-annotated acceleration data. Proc. Pervasive '04. pp. 1-17.

[2] Blanke, U., & Schiele, B. 2008. Sensing Location in the Pocket. Adjunct Proc. UbiComp '08.

[3] Choudhury, T., Borriello, G., Consolvo, S., Haehnel, D., Harrison, B., Hemingway, B., Hightower, J., Klasnja, P., Koscher, K., LaMarca, A., Lester, J., Landay, J., Legrand, L., Rahimi, A., Rea, A., & Wyatt, D. 2008. The Mobile Sensing Platform: An Embedded System for Capturing and Recognizing Human Activities, *IEEE Pervasive Computing*, vol. 7, no. 2. pp. 32 – 41.

[4] Consolvo, S., McDonald, D.W., Toscos, T., Chen, M.Y., Froehlich, J., Harrison, B., Klasnja, P., LaMarca, A., LeGrand, L., Libby, R., Smith, I., & Landay, J.A. 2008. Activity sensing in the wild: a field trial of ubifit garden. Proc. CHI '08. pp. 1797-1806.

[5] Krumm, J. A Markov Model for driver turn prediction. Society of Automotive Engineers (SAE) 2008 World Congress, April 2008, Paper 2008-01-0195.

[6] Kunze, K. & Lukowicz, P. 2008. Dealing with sensor displacement in motion-based onbody activity recognition systems. In Proceedings of the 10th international conference on Ubiquitous computing (UbiComp '08).

[7] Liao, L., Fox, D., & Kautz, H. 2007. Extracting Places and Activities from GPS Traces Using Hierarchical Conditional Random Fields. Int. J. Rob. Res. 26, 1 (January 2007), 119-134.

[8] Patterson, D.J., Fox, D., Kautz, D., & Philipose, M. 2005. Fine-Grained Activity Recognition by Aggregating Abstract Object Usage. Proc. ISWC '05. pp. 44-51.

[9] Priyantha, B., Lymberopoulos, D., and Liu, J., LittleRock: Enabling Energy Efficient Continuous Sensing on Mobile Phones. IEEE Pervasive Computing Magazine, April- June 2011.

[10] Saponas, T., Lester, J., Froehlich, J, Fogarty, J., Landay, J. 2008. iLearn on the iPhone: Real-Time Human Activity Classification on Commodity Mobile Phones. University of Washington CSE Tech Report UW-CSE-08-04-02.

[11] Tapia, E.M., Intille, S.S., and Larson, K. 2004. Activity recognition in the home setting using simple and ubiquitous sensors. Proc. Pervasive '04. pp. 158-175.