

# Context Transitions: User Identification and Comparison of Mobile Device Motion Data

**Tom Lovett**

Department of Computer Science  
University of Bath/Vodafone Group R&D  
Bath, UK, BA2 7AY  
tom.lovett@vodafone.com

**Eamonn O’Neill**

Department of Computer Science  
University of Bath  
Bath, UK, BA2 7AY  
eamonn@cs.bath.ac.uk

## Abstract

In this paper, we study a time-critical facet of context-awareness: context transitions, which we model as changes in specific context types over time, e.g., activity or location. We present results from a user-centred field study involving participant interviews and motion data capture from two mobile device sensors: the accelerometer and magnetic field sensor. The results show how the participants subjectively interpret their daily context transitions with variable granularity, and a comparison of these context transitions with mobile device motion data shows how the motion data poorly reflect the identified transitions. The results imply that care should be taken when representing and modelling users’ subjective interpretations of context, as well as the objective nature of context sensors. Furthermore, processing and usability trade-offs should be made if real-time on-device transition detection is to be implemented.

## Introduction

Mobile devices such as smart-phones can be a rich source of user data in context-aware systems, due in part to increasingly sophisticated sensory hardware and devices’ everyday proximity to their users. Faster CPUs, greater storage capabilities and better connectivity can also allow these devices to perform context-inference in real-time with a reduced effect on usability and resource demand.

However, a mobile device’s ability to become an integral component in a context-aware system is limited by the availability of its key resources, e.g., power, CPU and connectivity. These in turn can limit the performance of context inference techniques, especially if the data inputs to the system do not reflect users’ context well, and the inference techniques are more sophisticated as a result.

The necessary management of these trade-offs have led us to study the notion of context transitions. At the simplest level, a context transition is a binary indicator of a

significant change in a user’s context, e.g., their activity or location. Aside from the reduced impact on device resources that detecting a binary context transition may have (rather than classifying the context itself), transition detection could be used for multiple time-dependent applications, e.g., triggering notification delivery, improving resource management efficiency and bootstrapping more sophisticated yet costly learning-based inference methods.

In this paper, we detail our findings from a user-centred field study involving participant interviews and motion data capture from two mobile device sensors: the accelerometer and magnetic field sensor. We show how users interpret their own significant context transitions, as well as the granularity of each over time. Furthermore, by comparing the outputs from the motion sensors against the recorded transitions, we show that changes in motion sensed by the device poorly reflect users’ identified context transitions. From this comparison, we argue that care should be taken when representing and modelling users’ subjective interpretations of context, as well as the objective nature of context sensors. Furthermore, processing and usability trade-offs should be made if real-time on-device transition detection is to be implemented.

## Background and Related Work

The concept of activity transitions has previously been studied by Ho and Intille (2005), who show that users respond favourably to interruptions at the moment of activity transition rather than at random intervals. The authors also show that using multiple on-body accelerometers with supervised learning techniques can achieve good classification performance. We extend the generality of activity transitions to context transitions, which – in this paper – include location transitions as well as activity transitions. Moreover, we analyse how mobile device motion sensors – which are typically located at a

single point on-body – reflect context transitions without extensive pre-processing or learning.

In other relevant work on mobile context-awareness, the SenSay system by Krause et al. (2006) performs online context classification and learning with multiple mobile sensors including, as a corollary, the identification of context transitions over time. Although good classification performance is achieved, the system requires exterior wearable hardware to perform well which, in turn, affects its scalability and usability.

In contrast, mobile devices are ubiquitous and portable, but have limited resources, e.g., power and CPU, that constrain the use of online feature extraction and learning. Könönen et al. (2010) show that devices can perform automatic feature selection using statistical classification, but the impact on device resources is significant. Similarly, Kwapisz et al. (2011) show good performance in basic activity recognition using mobile device accelerometers, but the system requires offline processing and learning. Patterson et al. (2003) demonstrate good inference of user travel behaviour from GPS sensors using unsupervised learning.

We are interested in trade-offs between methods necessary for good context transition detection and the impact they may have on resources if implemented on-device. Our work in this paper aims to understand how users perceive their own context transitions, and how their mobile device motion sensors reflect these transitions prior to processing or learning.

## Approach

Our research question asks: to what extent do mobile device motion data reflect users' context transitions?

To answer this, we must formally model and identify typical user context transitions, before quantitatively comparing mobile device motion data with the transition data when undertaking the context transitions. This section details our approach to answering our research question.

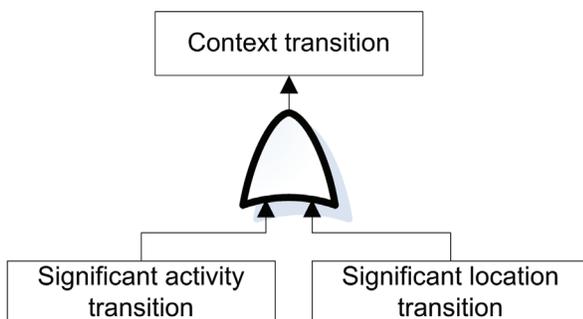


Figure 1: *Simple context transition model. A context transition occurs when a user-defined significant activity transition OR location transition occurs.*

## Modelling Users' Context Transitions

We model users' context transitions as significant changes in user context over time as defined by the users themselves. This can be further broken down into significant changes in user activity or location, each of which constitutes a significant context transition (see Figure 1).

There are several challenges to consider with this model; for example, users will have subjective, time-varying interpretations of their own context transitions, and an objective transition, e.g., any change in context, may be insignificant to the user. To gain insight into users' context transitions, we performed a series of interviews with a sample of users from different environments.

## Identifying Typical Context Transitions

We recruited 14 participants from two different environments: an office (8 participants) and a university (6 participants). We asked them to describe their typical day through significant changes in their activity or location and – immediately following the interview – we asked them to choose a subset of context transitions from their identified sets that could be performed in an ethnographic field study. Once these sequences were chosen, the participants were asked to perform them whilst carrying a mobile device.

## Mobile Device Motion Data Capture

Depending on the capabilities of each participant's personal device, we either installed an application on the device or provided them with a device that had the application preinstalled. Each device was Android-based, version 2.1 or greater, and the application was designed to continuously log data from the following sensors over time:

1. Tri-axis accelerometer, which senses the second order difference in position of the device along each orthogonal axis.
2. Tri-axis magnetic field sensor, which senses the ambient magnetic field strength of the device along each orthogonal axis.

To capture ground truth, i.e., what each participant's context was and the timestamp of each transition, we observed the participants throughout their sequences whilst recording each context and their respective transition times on a separate device.

## Motion Data Pre-processing

In order to compare the motion data with the participants' context transitions, we performed a small amount of pre-processing on the raw sensor data. The accelerometer directly senses motion by measuring the second-order finite difference in position along each orthogonal axis

over time. We interpret the data as a time series and – to remove the force of gravity and produce a relative output over time – we recursively pass the raw accelerometer data through a low pass filter at each time step of the series,  $k$ , with a smoothing factor,  $\alpha$ , of 0.8:

$$\mathbf{y}_k = \alpha \mathbf{x}_k + (1 - \alpha) \mathbf{y}_{k-1} \quad (1)$$

Here  $\mathbf{x}_k$  is the column vector in  $\mathbf{R}^3$  containing the force data on each axis at time step  $k$ , and  $\mathbf{y}_k$  is the filter output at  $k$ .  $\mathbf{y}_{k-1}$  is the filter output from the previous time step.

The magnetic field sensor measures the ambient magnetic field along each orthogonal axis over time. We again interpret the data as a time-series, and the relative output at each time step of the series is calculated using the same low pass filter in Equation (1) with  $\alpha = 0.8$ .  $\mathbf{x}_k$  is then the column vector in  $\mathbf{R}^3$  of magnetic field strength on each axis. The magnetic field sensor does not directly measure motion so, at each time step  $k$ , we calculate the second-order difference quotient:

$$\Delta^2(\mathbf{y}_k) = \frac{\mathbf{y}_{k+2} - 2\mathbf{y}_{k+1} + \mathbf{y}_k}{\Delta t^2} \quad (2)$$

Where  $\Delta t$  is the difference in time between our (regular) time steps  $k$ .

Following this pre-processing at  $k$ , the post-filter vectors are normalised to provide a scalar second-order output,  $\Delta^2(y_k)$ , for each sensor.

## Binary Conversion

To convert the motion data into a binary output, we use a latch. A latch is a comparator that outputs a binary value when its input crosses a threshold, and we use it to compare motion data between time steps,  $\Delta^2(y_k)$ . The latch outputs true when this second order difference crosses the threshold from low to high or vice versa.

We use a time-varying threshold that reacts to the motion data over time using a general three-parameter logistic function:

$$\phi_k = \frac{A}{1 + e^{-B(t-C)}} \quad (3)$$

Where  $\phi_k$  is the threshold at time step  $k$ . The other parameters are set as follows:  $t = \Delta^2(y_k)$ ;  $A = 2(E[\Delta^2(y_n)] + \sigma_n)$ ;  $B = -\sigma_s$ ; and  $C = E[\Delta^2(y_s)]$ .

$E[\Delta^2(y_n)]$  is the expected second-order output for noise, i.e., small, insignificant fluctuations in sensor data due to inherent noise and small movements by the user.  $\sigma_n$  is its standard deviation.  $E[\Delta^2(y_s)]$  is the expected second-order output for signal, i.e., significant changes in sensor data

caused by user motion, and  $\sigma_s$  is its standard deviation. To estimate these parameters, we recorded – from 3 participants – 30 minutes of noise motion data from a mobile device in each participant’s pocket whilst they were working at a desk, and 30 minutes of signal motion data from a mobile device in each participant’s pocket whilst they were walking.

Using a logistic function with these parameters causes the latch threshold to rise toward its upper asymptote when  $\Delta^2(y_k)$  moves below the signal expected value, and drop toward 0 when  $\Delta^2(y_k)$  moves above the signal expected value.

## Results

Figure 2 shows the context transition sequences for each of the 14 participants. Each cell represents a change in either activity or location from its predecessor and the vertical cell dividers mark the transitions. Double vertical cell dividers denote the two transitions either side of walking between adjacent contexts, and single dividers denote a direct transition between adjacent contexts. The mean number of context transitions per participant was 16, with mean sequence duration of 1 hour. The activity and location descriptions in Figure 2 show how each participant described their context transitions, as well as the granularity at which they perceive them.

Our regular time step,  $k$ , was 1 second. For the latch threshold parameters, our measured values were:  $E[\Delta^2(y_n)] = \text{sample mean}$ ,  $\bar{x}_n = 0.11 \text{ ms}^{-2}$  (accelerometer) and  $1.67 \mu\text{Ts}^{-2}$  (magnetic field sensor);  $\sigma_n = \text{unbiased standard deviation estimate}$ ,  $\hat{\sigma}_n = 0.05 \text{ ms}^{-2}$  (accelerometer) and  $0.81 \mu\text{Ts}^{-2}$  (magnetic field sensor);  $E[\Delta^2(y_s)] = \bar{x}_s = 4.42 \text{ ms}^{-2}$  (accelerometer) and  $13.60 \mu\text{Ts}^{-2}$  (magnetic field sensor); and  $\sigma_s = \hat{\sigma}_s = 2.79 \text{ ms}^{-2}$  (accelerometer) and  $10.69 \mu\text{Ts}^{-2}$  (magnetic field sensor). From these data, and the fact that they are normalised (thus always positive), we can estimate signal-to-noise ratios (SNRs) for the motion data from each sensor, i.e., the ratio of the expected value of signal to the standard deviation of noise: 88.4 (accelerometer) and 16.8 (magnetic field sensor).

For quantitatively analysing how mobile device motion sensor data reflect the participants’ context transitions, we use time-dependent binary classification, i.e., classification of a context transition must be made within a specified time window. As such, sensor performance can be measured by its true positives, false positives and false negatives.

For both sensors, we use the F1 score as a performance measure, which incorporates precision and recall. As the transition classifications are time-dependent – and to account for minor time deviations in ground truth recording, as well as the  $2\Delta t$  latency required for Equation

1	A	Work	Coffee	Work	Shop	Shop	Lunch	Work								
	L	<i>Office</i>	<i>Kitchen</i>	<i>Office</i>	<i>Shop1</i>	<i>Shop2</i>	<i>Café</i>	<i>Office</i>								
2	A	Coffee	Seminar	Coffee	Meeting	Work	Meeting	Meeting	Lunch							
	L	<i>Café</i>	<i>Hall</i>	<i>Café</i>	<i>Lab</i>	<i>Office</i>	<i>Kitchen</i>	<i>Library</i>	<i>Café</i>							
3	A	Lecture	Read	Eat	Lecture	Train	Lunch									
	L	<i>Theatre1</i>	<i>Library</i>	<i>Café1</i>	<i>Theatre2</i>	<i>Hall</i>	<i>Café2</i>									
4	A	Work	Coffee	Eat	Work	Meeting										
	L	<i>Desk</i>	<i>Kitchen</i>	<i>Kitchen</i>	<i>Desk</i>	<i>Tables</i>										
5	A	Lunch	Work	Coffee	Teach	Work	Cycle	Cycle	Cycle	Work						
	L	<i>Café1</i>	<i>Office</i>	<i>Café2</i>	<i>Theatre</i>	<i>Office</i>	<i>Bike shed</i>	<i>Town</i>	<i>Bike shed</i>	<i>Office</i>						
6	A	Work	Meeting	Work	Cycle	Chat	Work	Call	Work							
	L	<i>Desk1</i>	<i>Room1</i>	<i>Desk1</i>	<i>Sheds</i>	<i>Desk2</i>	<i>Desk1</i>	<i>Room2</i>	<i>Desk1</i>							
7	A	Work	Meeting	Work	Toilet	Exercise	Work	Meeting	Work	Queue	Lunch	Work	Drive	Work		
	L	<i>Desk</i>	<i>Desk</i>	<i>Desk</i>	<i>Toilet</i>	<i>Gym</i>	<i>Desk</i>	<i>Room</i>	<i>Desk</i>	<i>Café</i>	<i>Café</i>	<i>Desk</i>	<i>C. Park</i>	<i>Desk</i>		
8	A	Work	Coffee	Meeting	Work	Meeting	Work	Exercise	Lunch	Work	Drive	Chat	Drive	Exercise	Toilet	Work
	L	<i>Desk1</i>	<i>Kitchen</i>	<i>Tables</i>	<i>Desk1</i>	<i>Room</i>	<i>Desk1</i>	<i>Gym</i>	<i>Café</i>	<i>Desk1</i>	<i>C. Park</i>	<i>Desk2</i>	<i>C. Park</i>	<i>Gym</i>	<i>Toilet</i>	<i>Desk1</i>
9	A	Meeting	Work	Coffee	Work	Lunch	Work	Travel	Work							
	L	<i>Tables</i>	<i>Desk</i>	<i>Kitchen</i>	<i>Desk</i>	<i>Café</i>	<i>Desk</i>	<i>B. Stop</i>	<i>Desk</i>							
10	A	Meeting	Work	Call	Work	Printing	Work	Meeting	Work	Travel	Work					
	L	<i>Tables</i>	<i>Desk</i>	<i>Desk</i>	<i>Desk</i>	<i>Printer</i>	<i>Desk</i>	<i>Room</i>	<i>Desk</i>	<i>Bike Shed</i>	<i>Desk</i>					
11	A	Work	Drink	Work	Exercise	Lunch	Work	Talk	Talk	Meeting	Work	Present	Shop	Work		
	L	<i>Desk1</i>	<i>Kitchen</i>	<i>Desk1</i>	<i>Gym</i>	<i>Café</i>	<i>Desk1</i>	<i>Desk1</i>	<i>Desk2</i>	<i>Tables</i>	<i>Desk1</i>	<i>Building</i>	<i>Shop</i>	<i>Desk1</i>		
12	A	Meeting	Work	Meeting	Meeting	Work	Snack	Work	Lecture	Work						
	L	<i>Café</i>	<i>Office</i>	<i>Room1</i>	<i>Room2</i>	<i>Office</i>	<i>Café</i>	<i>Office</i>	<i>Theatre</i>	<i>Office</i>						
13	A	Relax	Drive	Drive	Refuel	Shop	Drive	Drive	Shop	Drive	Drive	Work				
	L	<i>Home</i>	<i>Car park</i>	<i>Petrol St.</i>	<i>Petrol St.</i>	<i>Shop1</i>	<i>Shop1</i>	<i>Shop2</i>	<i>Shop2</i>	<i>Shop2</i>	<i>Home</i>	<i>Home</i>				
14	A	Work	Tutor	Talk	Work	Bank	Shop	Work	Coffee	Toilet	Work					
	L	<i>Office1</i>	<i>Building</i>	<i>Office2</i>	<i>Office1</i>	<i>Bank</i>	<i>Shop</i>	<i>Office1</i>	<i>Café</i>	<i>Toilet</i>	<i>Office1</i>					

Figure 2: Context transition sequences as identified and undertaken by each of the 14 participants (one per table). Each table contains the participant-identified activity descriptions (A, upper row) and location descriptions (L, lower row, italicised). Double vertical separators represent the activity 'walking' between adjacent contexts, i.e., **two** context transitions. Single vertical separators represent a single transition between adjacent contexts.

(2) – we count a true positive if the sensed motion crosses the latch threshold within a 10 second lag window of an unclassified recorded ground truth transition.

A false positive occurs if a transition is classified with no recorded ground truth transition in the lag window, and a false negative occurs if no transition is classified during the lag window following a recorded ground truth transition.

Figure 3 shows the F1, precision and recall performance for both sensors across each of the 14 participants. The rounded sample means and sample standard deviations  $s$  for the accelerometer are: F1 = 0.14 ( $s = 0.13$ ); precision = 0.21 ( $s = 0.16$ ); and recall = 0.60 ( $s = 0.17$ ). For the

magnetic field sensor, they are: F1 = 0.06 ( $s = 0.03$ ); precision = 0.11 ( $s = 0.05$ ); and recall = 0.88 ( $s = 0.13$ ).

## Discussion

Here we discuss the qualitative and quantitative results of our study, prior to examination of their potential implications.

### User-defined Context Transitions

The context transition sequences listed in Figure 2 give an insight into how users perceive their own significant transitions.

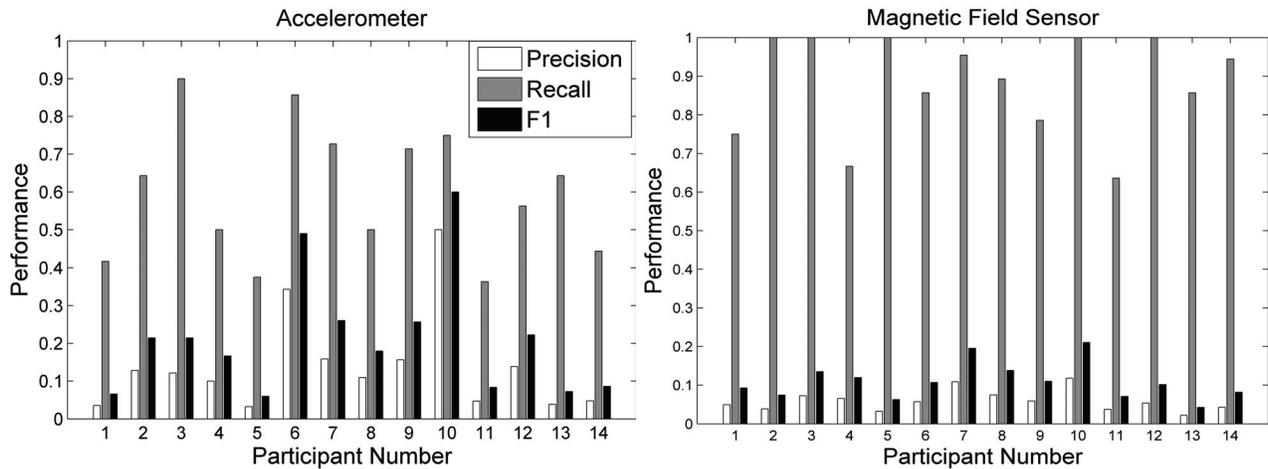


Figure 3: Comparing mobile device motion sensor data against participants' context transitions: accelerometer (left) and magnetic field sensor (right) across the 14 participants. Legend applies to both figures.

Of interest is the granularity at which some participants describe their significant transitions: for example, Participants 1, 2, 5 and 14 all describe the work activity at the coarse-grained office location, which contrasts with the fine-grained desk location given by other participants, e.g., 7. Thus some participants perceive the beginning of their work context to be when they enter a building, whereas others perceive it to start once they reach their desk.

There are some cases of multiple activity transitions in one location, e.g., Participant 10's work; call; work transitions at their desk, as well as cases of multiple

location transitions during a single activity, e.g., each walking activity; and various transit activities, e.g., Participant 5 cycling and Participant 13 driving. The walking contexts were somewhat unimportant to the participants: they tended to view them as intermediate periods between more significant contexts.

### Comparing Motion Data with Context Transitions

The mobile device motion sensor results in Figure 3 along with the mean F1 scores show that the raw motion sensor data poorly reflect the participants' identified context

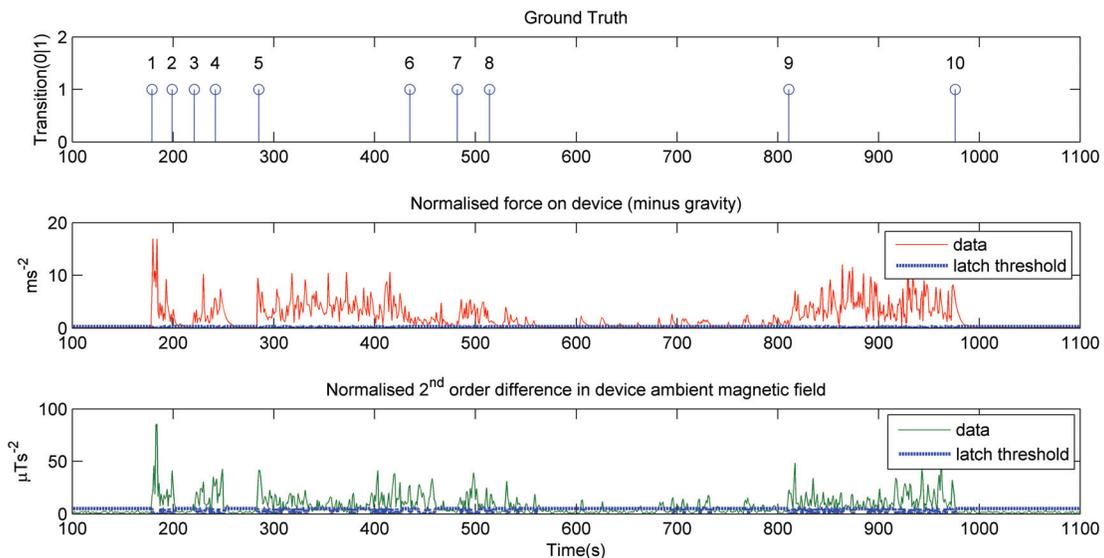


Figure 4: Plot showing a sample sequence of context transitions for Participant 1 (upper), compared with motion sensor data: normalised force on device from the accelerometer (middle), and normalised second order difference in ambient magnetic field from the magnetic field sensor (lower). The latch thresholds are also shown on each motion data subplot.

transitions. This is largely due to low precision, i.e., a high number of false positives.

The accelerometer scores are generally better than the magnetic field sensor's. The SNR of the accelerometer is greater than the SNR of the magnetic field sensor, though both are much greater than 1. These high SNRs and low F1 scores imply that the genuine motion sensed by the device is not indicative of the participants' context transitions. Figure 4, which shows a sample from Participant 1's context transition sequence, illustrates how the magnetic field sensor motion signal is less discernable than the accelerometer, and how it crosses its latch threshold more frequently between context transitions, i.e., false positives. This frequent latch triggering also results in fewer false negatives for the magnetic field sensor and, consequently, better recall than the accelerometer.

The highest F1 scores across the sensors are Participants 6, 7 and 10, whose context sequences are largely fine-grained and directly involve the mobile device itself (calling). The lowest F1 scores can be seen for Participants 1, 5, 11, 13 and 14. All participants in this set have coarse-grained contexts such as shop, teach, drive and cycle, which involve many fine-grained motion changes, e.g., stopping and starting due to queuing (shopping) or traffic (driving and cycling). An example of shopping can be seen in Figure 4 between transitions 8 and 9. Here, Participant 1 was moving between shop aisles, pausing to browse products and queuing for the checkout. These are arguably significant motion changes, and they manifest in accelerometer and magnetic field sensor changes, yet the participant does not consider them significant enough for a context transition. Other participants however, e.g., Participant 7, do consider finer-grained activities such as queuing as significant.

## Implications

The key implication of these results is that users' have subjective interpretations of their own context and variable perceptions of its granularity. This further implies that context dynamics and variability should be taken into account when modelling user context states and transitions.

Moreover, the dissimilarity between the motion data and context transitions necessitate the use of further processing and learning in order to achieve good context inference from the objective sensors.

However – as outlined in the introduction and background – devices have severe resource constraints that limit the extent to which online processing and learning can be applied. Indeed, as Könönen et al. (2010) show, it would be impractical to expect mobile devices to perform statistical classification without impacting heavily on device resources.

Therefore, to implement on-device context transition detection, tradeoffs have to be made between performance, cf. Krause et al. (2006) and Ho and Intille (2005), and device usability, cf. Könönen et al. (2010).

There are also implications for modelling and representing user context. The results illustrate the subjective and inconsistent nature of users' context interpretations, as well as the objective nature of the motion sensor data. This contrast – in addition to the dynamic nature of the context transitions – highlight the need for robustness when describing context semantics and designing context models.

## Conclusion and Future Work

In conclusion, we have presented a model for context transitions, i.e., changes in users' location or activity, and shown that users have subjective interpretations of their own context transitions' granularity. Furthermore, by comparing these transitions against motion data sensed by a mobile devices, we have shown that objective motion data do not reflect these subjective transitions well.

For future work, we will investigate techniques to infer context transitions from the motion data whilst optimising performance for usable implementation on mobile devices.

## Acknowledgements

Eamonn O'Neill's research is supported by a Royal Society Industry Fellowship at Vodafone Group R&D. We thank our field study participants for their time and efforts.

## References

- Ho, J., and Intille, S. S. 2005. Using Context-Aware Computing to Reduce the Perceived Burden of Interruptions From Mobile Devices. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* 909–918.
- Könönen, V.; Mäntyjärvi, J.; Similä, H.; Pärkkä, J.; and Ermes, M. 2010. Automatic Feature Selection for Context Recognition in Mobile Devices. *Pervasive and Mobile Computing* 6(2):181–197.
- Krause, A.; Smailagic, A.; and Siewiorek, D. 2006. Context-Aware Mobile Computing: Learning Context-Dependent Personal Preferences From a Wearable Sensor Array. *IEEE Transactions on Mobile Computing* 5(2):113–127.
- Kwapisz, J.R.; Weiss, G.M.; and Moore, S.A. 2011. Activity Recognition Using Cell Phone Accelerometers. *ACM SIGKDD Explorations Newsletter* 12(2):74–82.
- Patterson, D.; Liao, L.; Fox, D.; and Kautz, H. 2003. Inferring High-Level Behavior From Low-Level Sensors. In *UbiComp 2003: Ubiquitous Computing*, volume 2864 of LNCS. Springer Berlin / Heidelberg. 73–89.