Mobile Based Prompted Labeling of Large Scale Activity Data

Ian Cleland[1], Manhyung Han[2], Christopher Nugent[1], Hosung Lee[2], Shuai Zhang[1], Sally McClean[3], and Sungyoung Lee[2]

[1] Computer Science Research Institute and School of Computing and Mathematics, University of Ulster, Newtownabbey, Co. Antrim, Northern Ireland, BT37 0QB
{i.cleland,cd.nugent,s.zhang}@ulster.ac.uk
[2] Dept. of Computer Engineering, Kyung Hee University, Korea
{smiley,hslee,sylee}@oslab.khu.ac.kr
[3] Computer Science Research Institute and School of Computing and Information Engineering, University of Ulster, Coleraine, Northern Ireland, BT52 1SA
si.mcclean@ulster.ac.uk

Abstract. This paper describes the use of a prompted labeling solution to obtain class labels for user activity and context information on a mobile device. Based on the output from an activity recognition module, the prompt labeling module polls for class transitions from any of the activities (e.g. walking, running) to the standing still activity. Once a transition has been detected the system prompts the user, through the provision of a message on the mobile phone, to provide a label for the last activity that was carried out. This label, along with the raw sensor data is then stored locally prior to being uploaded to cloud storage. The paper provides technical details of how and when the system prompts the user for an activity label and discusses the information that can be gleaned from sensor data. This system allows for activity and context information to be collected on a large scale. Data can then be used within new opportunities in data mining and modeling of user context for a variety of applications.

1 Introduction

The ubiquitous nature of smart phones within our everyday lives provides new opportunities to collect real time context information, such as activity, location and social interactions, from a large number of users [1]. This large amount of data has the potential to be used in a number of application areas such as activity promotion, self management of long term chronic health conditions, context aware services and life logging [2]. The automatic recognition of activities is performed through the application of machine learning techniques to data gleaned from low level sensors [3]. The training of these algorithms, from a data driven perspective, relies largely on the gathering, pre-processing, segmentation and annotation of the sensor data into distinct classes [4]. The data must therefore be correctly labeled prior to being used as a training set in a machine learning paradigm.

C. Nugent, A. Coronato, and J. Bravo (Eds.): IWAAL 2013, LNCS 8277, pp. 9–17, 2013.

Collecting this data from a larger population under free living conditions may have the potential to improve the generalization abilities of any activity recognition (AR) models developed through provision of a larger quantity of representative data for training purposes. Such data sets should include data from a variety of sensors, recorded during a wide range of activities and contexts from a large number of users, over an extended period of time (months or years). Most importantly the data should also include accurate ground truth labels that represent user activities [5].

The use of smart phones can be viewed as one possible manner in which this large amount of data may be captured unobtrusively. Many handsets now have a range of in-built sensors, large memory storage, fast processing and low power communications, which meet the requirements of the range of data to be collected [6]. Furthermore, unlike many devices used as part of a research study, many potential subjects already own mobile phones, are accustomed to carrying them and always keep them charged [1]. Unfortunately, using mobile devices to gather data on a large scale can also prove difficult. In particular the integrity of the user annotation can be questionable. For example, users may forget to label a section of valuable data or may complete the labeling inaccurately. Nevertheless, a large scale fully annotated data set is recognised as being the key step to improve and increase the widespread adoption of AR applications [1], [6].

This paper presents an overview of a mobile based prompted labeling application aimed at overcoming the challenges associated with collecting annotated activity data on a large scale. In order to set the context of this work, a brief review of related material is presented. Following this the system architecture of the proposed prompt labeling application is described and the paper concludes with a discussion of the data which can be collected and analyzed.

2 Background

Although a large amount of research has focused on the ability to accurately detect a range of physical activities, very few studies have provided a detailed description of how the ground truth of data sets, for the purposes of a data driven approach to AR, have been acquired. To date the majority of AR studies have used data collected under structured or semi structured conditions, from a small number of participants (1-20 subjects). Participants often perform a set of preplanned tasks which are completed within a controlled environment [7], [8], [9], [10]. In this case, the ground truth is often recorded by a human observer and sensor data are then annotated offline according to the observer. This is viewed as being necessary as it allows researchers to capture the ground truth, when labeling data, in an effort to create highly accurate data sets. Data collected in this manner may not, however, be truly representative of completing the task in a free living environment. Furthermore, labeling and processing data is this manner can be a laborious and time consuming task for researchers. Boa and Intillie asked participants to complete a list of planned activities and to note the time at which they started and completed each activity [8]. Again this process of continuously noting the time at which an activity is commenced and completed is fine for short term laboratory based studies, however, would not be feasible in the long term under free living conditions.

In order to allow the collection of data in a more free living environment, researchers have utilized video cameras [11]. The subsequent video recording is reviewed offline to identify what activity was being performed at a particular time. Similar techniques have been used within smart environments to label the onset/ completion of object interactions [12]. Using groups of labelers sourced from the crowd is viewed as one way of dealing with the labour intensity of the task. Lasecki *et al.* [13] used activity labels, generated by groups of crowd sourced labelers to annotate activities from video data. All of the aforementioned labeling methods are labor intensive and time consuming and some approaches, in particular video annotation, can have implications with data privacy. Furthermore, the need to install or issue video cameras for recoding the activities reduces the scalability of the approach.

Alternatively on a larger scale, users are often asked to annotate their own data using a mobile interface. This usually requires the user to start and stop the data capture process manually [14]. When using the application the user is then asked to label the activity they have just or are about to complete. Although this method is relatively accurate at segmenting the activity it requires the user to explicitly start and stop recording. Tapia *et al.* [15] used a technique based on the experience sampling method to trigger self reported diary entries every 15 minutes. Multiple choice questions were answered by the user to select which of the 35 activities users were completing. Due to the intermittent nature of the labels it was found to be difficult to detect short time activities. The process of continually labeling can become laborious for users, particularly when performed over an extended period of time. Furthermore, this can result in the user providing incorrect labels for the data or simply not engaging with the system at all. In order for the user to input a label, some interaction with the mobile device is required. This may interrupt the user during the activity, which in turn may impact upon the activity that the person is undertaking, thus impacting overall on the data recorded. In an attempt to address the issue of interaction voice recognition has been used for the purposes of annotation [16]. The mobile device listens for key words such as "start activity" and "stop activity" to start and stop the recording. Voice recognition is then used to label the activity, with the user again saying keywords, such as "standing" or "walking". Nevertheless, having the smart phone continuously listening for keywords can consume battery power and may hamper the usability of the application. Additionally, inaccuracies of voice recognition can lead to mislabeling of data.

Our approach uses prompted labeling, driven by an underlying mobile based AR module, in an effort to improve the process of collecting and annotating data. Users can annotate their everyday activities through use of a personalized mobile application. When the user is detected as standing still, a prompt is provided to enable the user to label the activity they were previously completing. In this manner the sensor data for the respective activity is segmented and saved automatically and an activity label is supplied by the user after the activity has been finished thus maintaining the integrity of the data.

3 Prompted Data Labeling

The proposed mobile application is based upon the principle of prompts to label a user's context and activity data. At periodic times throughout the day, the application will prompt the user to confirm which activities they have just completed. In addition to user reported data, additional information gleaned from the mobile device, such as automated activity classifications, GPS latitude and longitude, accelerometry data and bluetooth interactions will also be recorded. This additional data will aid in further contextualizing the annotated data sets with the intention of improving the validity of labelling. An overview of the system architecture of the proposed application is shown in Fig. 1.

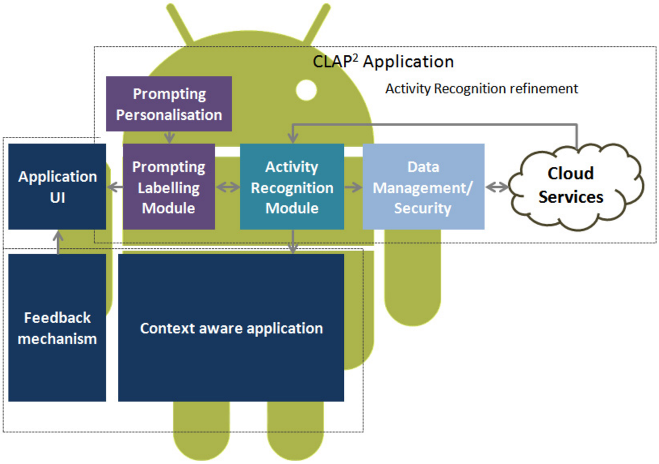

Fig. 1. An overview of personalized mobile application for prompt labelling. The prompted labeling module sits on top of an existing activity recognition module and periodically prompts users to label their activity. The architecture includes mobile services to support the secure transmission and processing of data in addition to the collection of other sensory data available from the mobile platform.

In order to enhance user engagement and compliance of the application it is important that the prompted labeling module is to be incorporated within an application which provides some incentive through appropriate feedback mechanisms. This type of application could include any context aware application such as an activity monitor, calorie counter or context aware services. A suite of mobile services will be developed to ensure the secure processing and transmission of all data collected from the users. These services will be responsible for managing security, efficient transmission of data and interfacing with cloud services. A brief description of these components is provided in Table 1.

Table 1. Provides a description of the main components of the system architecture

Component Name	Component Description
Prompt labeling module	This component contains a splash screen which allows the user to select a label for their activity data from a predefined list.
Activity recognition module	The activity recognition module attempts to detect changes in activity class to the standing still activity. From this a prompt is then initiated.
Data management	This module ensures the data is appropriately structured and formatted to ensure efficient transfer and storage. In this respect the security of the sensitive data is crucial, therefore efficient cryptography protocols shall be employed.
Cloud services	Cloud services provide the appropriate infrastructure to support data storage analysis and mining of the large data set.
Context aware application	The prompter sits upon a context aware application which enhances user engagement by providing tailored feedback (e.g. activity levels, calorie counting and context aware services)

3.1 Activity Recognition Module

The AR model used within this work, originally developed by Han *et al.* [17], utilizes multimodal sensor data from accelerometery, audio, GPS and Wi-Fi to classify a range of everyday activities such as walking, jogging and using transport. The activity recognition is performed independently of the position or orientation of the smart phone. This approach increases the practicality and usability of the system as the phone can be carried at any location and the AR is not affected by the user interacting with the device. Data from the accelerometer is used to detect transitions between ambulatory activities to activities which involve the use of transport. Accelerometer data, sampled at 50Hz, is computed into time and frequency domain features which are used as inputs to a Gaussian Mixture Classifier. Audio data is used in the classification if there is a need to classify between transportation activities (riding a bus or subway). Only using the audio when necessary allows the power consumption on the Smartphone to be minimized. GPS and Wi-Fi signals are then used to validate the classification between activities. Speed information, derived from GPS is used to determine whether a user is walking, running or standing still. The Wi-Fi signal is used to differentiate between bus and subway activities, as very few public or private wireless networks are available within the subway system.

3.2 Prompted Labeling Module

The prompted labeling module (PLM) prompts the user to label the activity they have just completed. Based on the output from the AR module the PLM polls for class transitions from any of the activities (e.g. walking and running) to the standing still activity. Once a transition has been detected the PLM prompts the user, through the provision of a message on the mobile phone, to provide a label for the last activity that was carried out. The raw data from the accelerometry sensors are then stored to the mobile device before being transmitted to the cloud for storage and further processing. By prompting the user to label the activity it is possible to verify that the activity has been correctly identified by the AR module. In this way the trustworthiness of the AR module can be validated in addition to providing a fully annotated data set. Fig. 2 presents an example of interaction with the prompt labeling screen on the mobile device.

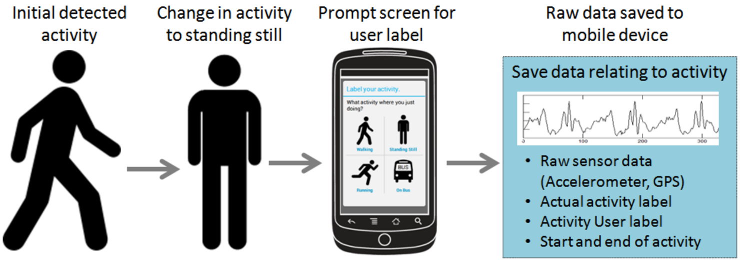

Fig. 2. An example of user interaction with the prompt labeling screen. The AR module detects a change in class from the original activity to standing still. The prompt is then issued for the user to label the previous activity. Raw sensor data, for this activity, is then saved to the mobile device before being uploaded to the cloud for further processing and storage.

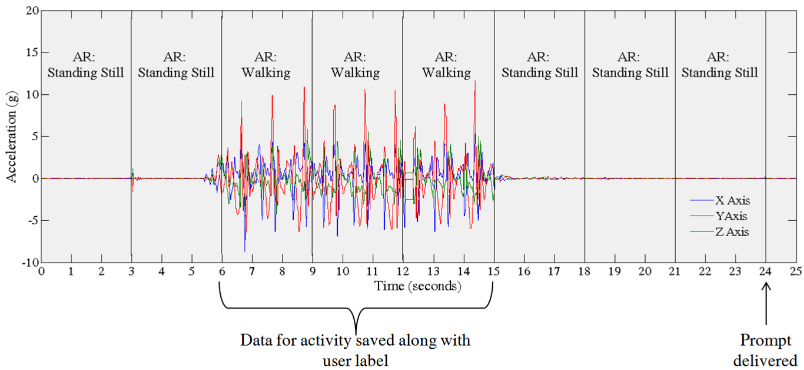

Fig. 3. Illustrates how activities are detected from the raw accelerometer signal by the AR module. Activates are detected every 3 seconds, three consecutive detections are used to label the activity. The prompt is initiated when the AR module detects a change in class from one activity to standing still.

The AR module detects an activity based on three seconds (150 samples) of data. Three consecutive detections (9 seconds) are then used to label the activity. This is carried out in order to limit the number of detection errors. Once the AR module detects a change from the current activity to the standing still activity for 9 seconds the previous activity data from the sensors is saved to memory. This process, from the perspective of raw accelerometry data is shown in Fig. 3. Currently, the prompt is initiated every time the AR detects a transition from an activity to standing still. It is envisioned that when the application is rolled out on a larger scale the user will be able to set how many prompts they receive per day in order to improve adoption of the system.

Currently data recorded by the system is stored directly to the local memory of the Smartphone, in the form of a text file. Fig. 4 shows the structure of this file. Data includes date and time stamp, raw accelerometer values (X, Y and Z axis), GPS latitude and longitude in addition to the Class label from the AR module and the prompted user label (named AR Label and User Label respectively). It is envisioned that this data could then be encrypted before being transmitted and stored in the cloud.

Date/Time	Sample No.	Accel (X axis)	Accel (Y axis)	Accel (Z axis)	GPS (Lat)	GPS (Long)	AR Label	User label
201306141415	1	-5.152806193	3.587482	2.759922	54.68812	-5.88404	Walking	Walking
201306141415	2	-1.186659648	1.129002	-4.368226	54.68812	-5.88404	Walking	Walking
201306141415	3	-1.186659648	1.129002	-4.368226	54.68812	-5.88404	Walking	Walking
201306141415	4	0.886066667	3.421133	1.504522	54.68812	-5.88404	Walking	Walking
201306141415	5	0.886066667	3.421133	1.504522	54.68812	-5.88404	Walking	Walking
201306141415	6	2.60190622	1.719733	2.988136	54.68812	-5.88404	Walking	Walking
201306141415	7	1.487437767	0.281056	-0.683005	54.68812	-5.88404	Walking	Walking
201306141415	8	1.487437767	0.281056	-0.683005	54.68812	-5.88404	Walking	Walking
201306141415	9	-0.170149054	0.3915	-2.339131	54.68812	-5.88404	Walking	Walking
201306141415	10	-0.170149054	0.3915	-2.339131	54.68812	-5.88404	Walking	Walking
201306141415	11	-0.267032794	-0.10193	-1.685765	54.68812	-5.88404	Walking	Walking
201306141415	12	-0.267032794	-0.10193	-1.685765	54.68812	-5.88404	Walking	Walking
201306141415	13	0.163866438	-0.77931	-0.574205	54.68812	-5.88404	Walking	Walking
201306141415	14	0.163866438	-0.77931	-0.574205	54.68812	-5.88404	Walking	Walking
201306141415	15	0.288782327	-0.25574	0.478368	54.68812	-5.88404	Walking	Walking
201306141415	16	0.288782327	-0.25574	0.478368	54.68812	-5.88404	Walking	Walking
201306141415	17	0.667406954	0.356161	0.643228	54.68812	-5.88404	Walking	Walking
201306141415	18	0.667406954	0.356161	0.643228	54.68812	-5.88404	Walking	Walking
201306141415	19	0.061443988	0.43449	0.824545	54.68812	-5.88404	Walking	Walking
201306141415	20	0.061443988	0.43449	0.824545	54.68812			
201306141415	21	0.295442565	0.464337	1.710036				
201306141415	22	1.423894523	-0.41485					
201306141415	23	1.423894523						

Fig. 4. Shows an example of data recorded by the prompted labeling module. Recorded data includes, Date and time stamp, sample number, Accelerometer data from each axis, GPS latitude and longitude, in addition to the Class label from the AR module and prompted user label.

4 Summary

The ability to collect contextual information, such as activity, location or social interactions, on a large scale is becoming increasingly important. Such data sets allow for a deeper understanding of a population's activity habits and allow information to be delivered in a context sensitive manner. Current methods of collecting contextual information, particularly activity data, are normally limited to small scale studies. This is partly due to issues surrounding the ability to obtain ground truth information to annotate such data. The current approach aims to address such issues, through the use of a mobile based context aware PLM which prompts the user to supply label information for their current activities. In turn this improves the validity of data labels, which can then be used to improve the accuracy of data driven activity models. Transmitting and storing this data within the cloud opens new possibilities to exploit cloud services in order to mine these big data sets further in order to provide a deeper understanding of activity trends within healthcare. Plans for future work involve the evaluation of the current solution in order to assess its ability to accurately label data in a free living environment.

Acknowledgement. The authors wish to acknowledge support from the EPSRC through the MATCH programme (EP/F063822/1 and EP/G012393/1). The views expressed are those of the authors alone.

References

1. Intille, S.S., Lester, J., Sallis, J.F., et al.: New Horizons in Sensor Development. Medicine & Science in Sports & Exercise 44, S24–S31 (2012)
2. Hamm, J., Stone, B., Belkin, M., Dennis, S.: Automatic Annotation of Daily Activity from Smartphone-Based Multisensory Streams. In: Uhler, D., Mehta, K., Wong, J.L. (eds.) MobiCASE 2012. LNICST, vol. 110, pp. 328–342. Springer, Heidelberg (2013)
3. Preece, S.J., Goulermas, J.Y., Kenney, L.P.J., et al.: Activity Identification using Body-Mounted sensors—a Review of Classification Techniques. Physiol. Meas. 30, R1–R33 (2009)
4. Avci, A., Bosch, S., Marin-Perianu, M., et al.: Activity Recognition using Inertial Sensing for Healthcare, Wellbeing and Sports Applications: A Survey, pp. 1–10 (2010)
5. Hossmann, T., Efstratiou, C., Mascolo, C.: Collecting Big Datasets of Human Activity One Checkin at a Time, pp. 15–20 (2012)
6. Lane, N.D., Miluzzo, E., Lu, H., et al.: A Survey of Mobile Phone Sensing. IEEE Communications Magazine 48, 140–150 (2010)
7. Krishnan, N.C., Colbry, D., Juillard, C., et al.: Real Time Human Activity Recognition using Tri-Axial Accelerometers, pp. 1–5 (2008)
8. Bao, L., Intille, S.S.: Activity Recognition from User-Annotated Acceleration Data. In: Ferscha, A., Mattern, F. (eds.) PERVASIVE 2004. LNCS, vol. 3001, pp. 1–17. Springer, Heidelberg (2004)
9. Parkka, J., Ermes, M., Korpipaa, P., et al.: Activity Classification using Realistic Data from Wearable Sensors. IEEE Transactions on Information Technology in Biomedicine 10, 119–128 (2006)
10. Mannini, A., Intille, S.S., Rosenberger, M., et al.: Activity Recognition using a Single Accelerometer Placed at the Wrist Or Ankle. Med. Sci. Sports Exerc. (2013); E-Published ahead of Print
11. Plotz, T., Chen, C., Hammerla, N.Y., et al.: Automatic Synchronization of Wearable Sensors and Video-Cameras for Ground Truth Annotation–A Practical Approach, pp. 100–103 (2012)
12. Cruciani, F., Donnelly, M.P., Nugent, C.D., Parente, G., Paggetti, C., Burns, W.: DANTE: A video based annotation tool for smart environments. In: Par, G., Morrow, P. (eds.) S-CUBE 2010. LNICST, vol. 57, pp. 179–188. Springer, Heidelberg (2011)
13. Lasecki, W.S., Song, Y.C., Kautz, H., et al.: Real-Time Crowd Labeling for Deployable Activity Recognition, pp. 1203–1212 (2013)

14. Kawaguchi, N., Watanabe, H., Yang, T., et al.: HASC2012corpus: Large Scale Human Activity Corpus and its Application (2012)
15. Tapia, E.M., Intille, S.S., Larson, K.: Activity recognition in the home using simple and ubiquitous sensors. In: Ferscha, A., Mattern, F. (eds.) PERVASIVE 2004. LNCS, vol. 3001, pp. 158–175. Springer, Heidelberg (2004)
16. Harada, S., Lester, J., Patel, K., et al.: VoiceLabel: Using Speech to Label Mobile Sensor Data, pp. 69–76 (2008)
17. Han, M., Lee, Y., Lee, S.: Comprehensive Context Recognizer Based on Multimodal Sensors in a Smartphone. Sensors 12, 12588–12605 (2012)