SENSOR-BASED SINGLE-USER ACTIVITY RECOGNITION

Blaž Kažič, Jan Rupnik
Artificial Intelligence Laboratory
Jožef Stefan Institute
Jamova cesta 39, 1000 Ljubljana, Slovenia
e-mail: blaz.kazic@gmail.com, jan.rupnik@ijs.si

ABSTRACT

The focus of this work is to explore the possibilities of recognizing three common user activities (sitting, walking and running) with accelerometer data from smartphones. Among five common machine learning algorithms, Naïve Bayes classifier proved to be the best choice. Classification accuracy of more than 90% was achieved when phone is carried in a pocket. It is shown that this method is appropriate and that the phone's orientation information is not needed. Finally, the classification of one day-long data set is presented.

1 INTRODUCTION

Over the previous decade, we have witnessed the significant rise of smartphone technology; the built-in sensors of such phones have also improved significantly. Because contemporary smartphones are highly programmable, their cheap and powerful sensors represent a variety of new research opportunities in the field of mobile context awareness. Data from the built-in sensors (e.g. gyroscope, accelerometer, digital compass, GPS, microphone and camera) have already led to the development of some interesting applications in the fields of healthcare [1], social networks [2], business and environmental monitoring [3] and transportation [4].
This article discusses the possibility of activity recognition of the smartphone user, with data obtained by the Apple iPhone's built-in accelerometer sensor. Knowing the users' activity could be useful information in profiling their preferences and behavior. This information may be used for offering personalized recommendations for points of interest, services or products (e.g., targeted advertising).
In this stage, recognizing three main user postural behaviors (sitting, walking and running) has been researched. Sitting also includes all standstill behaviors (like standing), and running also includes all behaviors that include sudden movements (jumping). Different activities are distinguished based on the data from the accelerometers' sensors, since accelerometer is one of the most common sensors and has already been used in many studies (even before they were commonly built into smartphones) [5]. Using only accelerometer data also has another advantage, as it is not particularly demanding on the battery, in comparison to other sensors (especially GPS). Power consumption can be a difficult obstacle in conducting research with smartphones, so energy efficient approaches must be considered [6].

2 METHODOLOGY

In this work, supervised machine learning is used for recognition of users' activity; therefore, labeled training data from accelerometer sensors had to be collected.

2.1 Application for Collecting Data

A brief survey of the two most popular application markets, Apple's *App Store* and Google's *Android Marketplace*, shows that many applications have already been developed to extract sensor data from the phone. Consequently, instead of developing new piece of software for the task an iPhone application called SensorLogger was used. This application records a phone's sensor data for a later review, or streams it to other devices via wireless networks as UDP broadcast packets.

2.2 Collecting Data

With the help of this application, a training set spanning five hours of user activity was collected by seven users. Data from the accelerometers was recorded at a sampling rate of 30 Hz. The data was then divided into nonintersecting 10 second intervals (windows). For each window, several features were computed. A set of features of one window represents one sample. In total, 1750 samples were collected, as summarized in Table 1.

Person	Sitting	Walking	Running
a	80	77	80
b	79	116	94
c	79	78	59
d	92	86	68
e	68	92	86
f	166	79	78
g	68	63	62
Sum	632	591	527

Table 1: *Training data set*

2.3 Orientation Problem

Most current smartphones have tri-axial (3d) accelerometers, i.e. sensors detecting acceleration in the x, y and z directions; sensor orientation depends on phone orientation. While information for x, y and z accelerations can be extremely useful in the case of body worn sensors, where the orientation of a sensor is fixed [9], such regime of operation cannot be expected in our case. Since a phone is a portable device, it is obvious that its position varies from person to person. Figure 1 presents accelerometer data from two different persons who were performing the same activity (sitting). It is clear that the two sets of data differ significantly, which means that the two orientations are different.

One easy solution to the orientation problem is the use of magnitude of each (x, y, z) accelerometer signal. Figure 1 also shows that although the readings from x, y and z axes were significantly different, the magnitude signal was very similar, which indicates that such information may be used as a feature for classification. The magnitude in both cases is stationary in time at approximately $1 m/s^2$, which corresponds to the sensor measuring the force of gravity while being static.

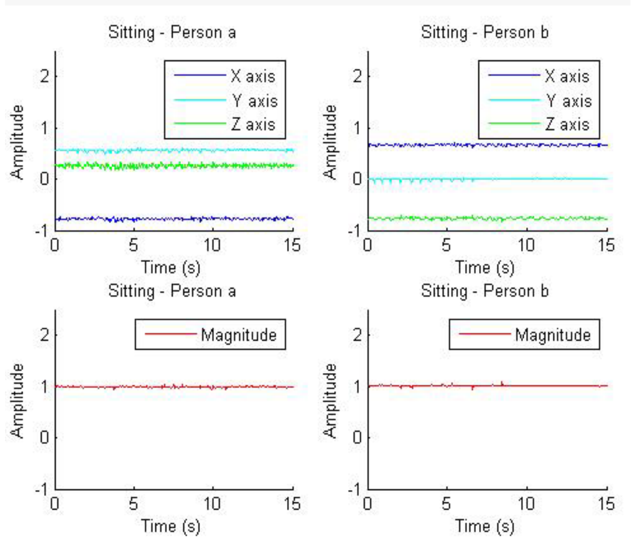

Figure 1: *Accelerometer data from two different persons performing same activity (sitting)*

2.4 Feature Extraction

During the research, many different features were considered. Some of them are common when processing accelerometer signals (e.g. mean and standard deviation). Others, such as dominant frequency [5], are calculated based on a signal preprocessed with a discrete Fourier transform (DFT).

However, using all available features as an input in a classifier is not always appropriate. If a feature does not provide any new information that would improve classification, it can be irrelevant, redundant or even distracting. To achieve the best classifications results, the number of features should be as low as possible, retaining only the most relevant ones [8].

To find the best set of features, the following three subsets were analyzed;

- **All Features:** mean, standard deviation, variance, median, root mean square, skewness, kurtosis, 25 percentile, 75 percentile, inter-quartile, mean crossing rate, dominant frequency, DFTs energy, spectral entropy, xy correlation, xz correlation, yz correlation.
- **Simplified Features:** mean, standard deviation, 75 percentile, dominant frequency, xy correlation
- **Mean & StdDev:** mean, standard deviation

The first set, *All Features*, contains features that have already been considered in some of the related literature on activity recognition research [9]. We omit the definitions of used features due to the lack of space and refer the reader to [9][5]. The second set, *Simplified Features*, includes some of the most popular features for activity recognition [5]. The third set contains only two features: mean and standard deviation. Different studies [7] show that only with these two features it is possible to classify user behavior to some degree of accuracy. The benefits of using only two features are energy efficiency and ease of computation, which make them highly appropriate for the use in systems with low computation power.

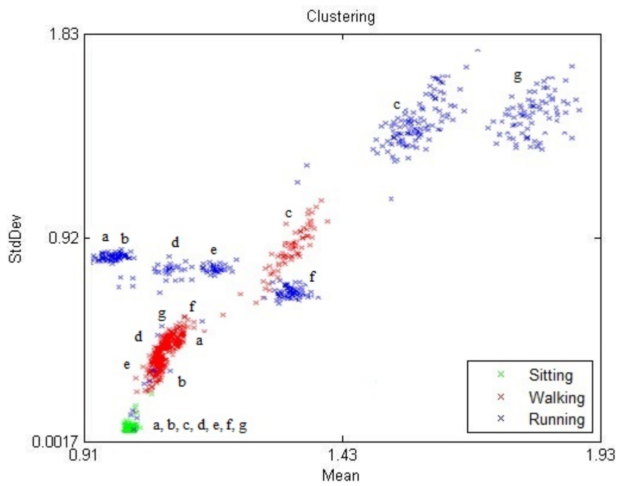

Figure 2: *Mean vs. Standard deviation clustering*

Figure 2 shows clustering of activities from different users by using only Mean and Standard deviation as features. The boundary between sitting and movement (walking or running) is quite obvious, but less so between walking and running. This is mainly because the smartphones were worn differently and also because people walk and run differently. The clusters with larger variance and larger average force magnitude (top right region) correspond to the users (c and g) wearing the phone in more loose pockets, causing more phone movement. In contrast, users who wore phone more tightly to their body (a, b, d, e and f) caused less vibrations and less dispersed signals. While the walking and running are linearly separable for each user individually, they are no longer linearly separable for all users simultaneously.

A particularly useful feature to distinguish between walking and running is the dominant frequency. From Figure 3, it is clearly seen that the average dominant frequency of persons in the test group for walking is 2 Hz, and the average dominant frequency for running is 2.9 Hz. This is also very interesting information from which the approximate speed of the user could be calculated.

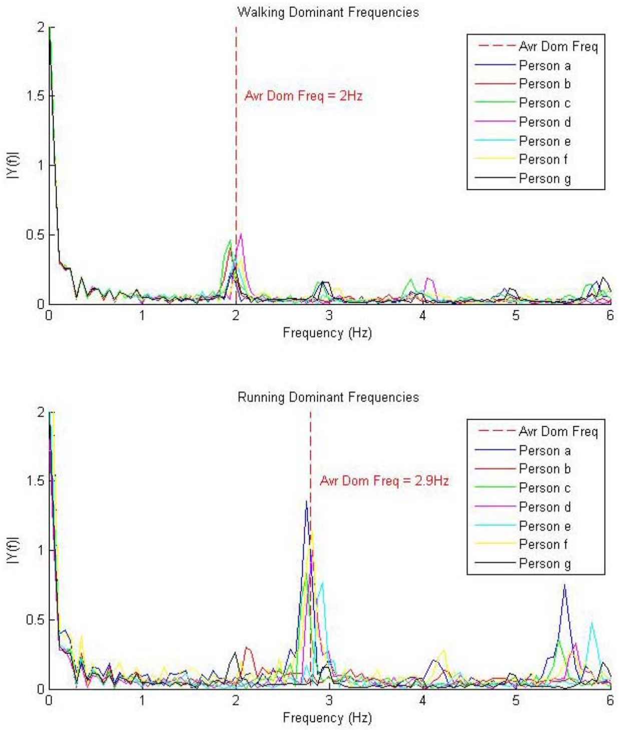

Figure 3: *Dominant frequency*

2.5 Feature and Classifier Selection

For testing and evaluating different sets of features and different classifiers, a Weka toolkit [10] was used. Five common machine-learning algorithms (*Decision Tree (J.48), Naive Bayes (NB), k-Nearest Neighbor (IBK), Support Vector Machine (SMO)* and *Neural Network (Multilayer perceptron))* were used to test the accuracy of the activity detection of all three previously mentioned feature sets.

All classifiers for every person in the data set have been tested, in a way that the person that was being tested was excluded from the training data. The results in Figure 4 are the average results for all classifiers and all features sets.

It is seen that *Naïve Bayes* in combination with the *Simplified feature* set has the most correctly classified samples (93.2%). The second best results (88.2%) are of the *Decision Tree* classifier, which is a tremendously popular choice in activity recognition research [11]. It is also seen that the *Simplified Feature* set generally gives better results than *All Features* set, as it was expected and discussed in previous section. *Mean & StdDev* set also gave some quite acceptable results, but considering the fact that smartphones are constantly becoming more powerful, there is no need for this kind of simplification. Based on these observations, we focused the evaluation on the *Naïve Bayes classifier* based on *Simplified features*.

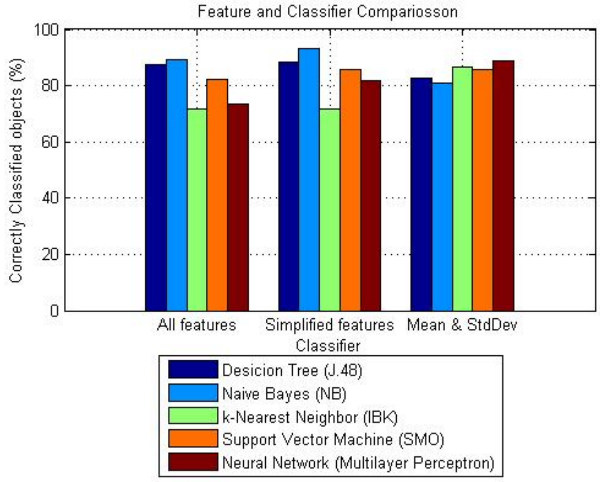

Figure 4: *Feature sets and Classifiers evaluation*

2.6 Naïve Bayes Classifier

For classification with Naïve Bayes, data from every feature set first had to be discretized in to several parts, so that every piece contains approximately the same amount of data. Bayes Theorem with Laplacian smoothing is then used to calculate the likelihood of one sample belonging to each class (activity). The sample is classified into classes with the highest calculated probability.

Finally, the classifier was tested on the set of data that was not in the test group. From the confusion matrix in Table 2, it is seen that the classifier is working properly. With 96.3% correctly classified samples, the result is slightly better than in the previous test; this is due to data discretization.

Labeled activity	Recognized activity		
	Running	Walking	Sitting
Running	92	2	0
Walking	0	116	5
Sitting	0	4	75

Table 2: *Confusion matrix*

3 RESULTS

The goal of this research was to recognize three common human physical activities (sitting, walking, running) with accelerometer data, regardless of the phones' orientation. In the previous section, it was shown that classification with Naïve Bayes classifier worked well with the labeled samples. In this section, data collected from one entire day is classified.

Twelve hours of data were collected on one working day by the author. The classifier was trained on a previously collected data set, which includes seven different persons (the data from the author was not part of the training set).

From Figure 5, the daily activity of the subject (the author) is clearly seen. It can be observed that there was some walking activity combined with sitting in the morning. At around 8am, the subject cycled to work, which usually takes half an hour. It can be seen that cycling is considered more similar to walking than running. Also it is evident that the subject's work involves sitting most of the time. Some walking activity is recognized during the lunch break around 1pm. At 5pm, the subject cycled back home, which is again recognized as walking. Later that day, the subject went for a short walk up the nearest hill. It is seen that during this trip walking is sometimes considered as running, which is understandable, since walking up or down a hill can cause more vibrations and shocks than a normal walk would, therefore classifier can recognize it as running.

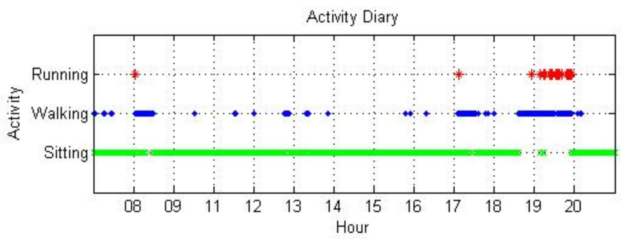

Figure 5: A*ctivity recognition during one day period*

Since the data was recorded constantly throughout entire day, with 10-second long samples, this was extremely energy consuming. Therefore, the data were filtered as if a 10-second sample had been recorded only every 3 minutes. The results can be seen on Figure 6. It is seen that all activities during the day are still reasonably recognized, as they were on Figure 6. Walking up a hill between 18:30 and 20:00 is now even better classified, since many of the previously running classified samples are no longer in the data set.

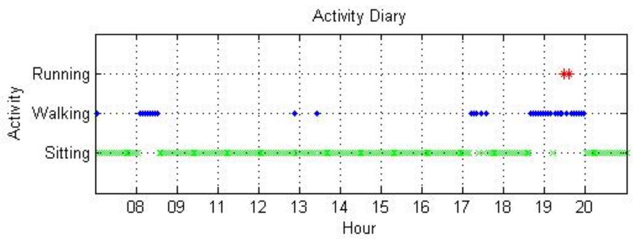

Figure 6: *Filtered activity recognition*

4 CONCLUSION

In this paper, human activity recognition of a single user is presented. The research indicates that with Naïve Bayed classifier and orientation independent features, it is possible to distinguish user behavior into three common activities: sitting, walking and running. The results show that such an approach has the potential and it can be extended into several directions.

First, it would be interesting to test the system on a larger data set, spanning a weak or a month of user activity. Increasing the number of recognizable human activities, such as standing, cycling and driving, could be the next step. To maximize the classification accuracy for a larger set of recognizable human activities, combination with GPS data could be considered. Finally, the effect of varying the sampling frequency and the window size on the classification performance could be investigated.

Acknowledgement

The authors acknowledge the European Commission for its support and partial funding and the partners of the research project: FP7-318452 MobiS.

References

[1] S. Consolvo. Activity Sensing in the Wild: A Field Trial of Ubifit Garden. *Proc. 26th Annual ACM SIGCHI Conf. Human Factors Comp. Sys*., 2008, pp. 1797-1806

[2] E. Miluzzo. Sensing meets Mobile Social Networks: The design, Implementation, and Evaluation of the CenceMe Application. *Proc. 6th ACM SenSys, 2008,* pp. 337-50

[3] M. Mun. Peir, the Personal Environmental Impact Report, as a Platform for Participatory Sensing Systems Research. *Proc. 7th ACM MobiSys, 2009,* pp. 55-68

[4] A. Thiagarajan. VTrack: Accurate, Energy-Aware Traffic Delay Estimation Using Mobile Phones. *Proc. 7th ACM SenSys*, Berkeley, CA, Nov. 2009.

[5] T. Huynh, B. Schiele. Analizing Features for Activity Recognition. *Proc sOc-EUSAI'05*, pp 159-163. 2005

[6] Z. Yan, V. Subbaraju, D. Chakraborty, A. Misra, K. Aberer. Energy-Efficient Continuous Activity Recognition on Mobile Phones: An Activity-Adaptive Approach. *Proc. 16th Annual ISWC '12 Symposium on Wearable Computers (ISWC),* pp. 17-24. 2012

[7] J. Yang. Toward physical activity diary: motion recognition using simple acceleration features with mobile phones. *Proc. 1st IMCE '09 International Workshop on Interactive multimedia for consumer electronics*, pp. 1-10. 2009

[8] C. M. Bishop, Pattern Recognition and Machine Learning (Information Science and Statistics). *Proc Spring-Verlag New York, Inc. Secaucus*, NJ, USA, 2006

[9] M. Zhang, A. A. Sawchuk. A feature selection-based framework for human activity recognition using wearable multimodal sensors. *Proc. BodyNets '11 Proceedings of the 6th International Conference on Body Area Networks*, pp 92-98. 2011

[10] M Hall, E. Frank, G. Holmes, B. Pfahringer, P. Reutmann, I. H. Witten. The WEKA Data Mining Software: An Update; *SIGKDD Eplorations*, Volume 11, Issue 1, 2009

[11] J. R. Kwapisz, G. M. Weiss, S. A. Moore. Activity recognition using cell phone accelerometers, *Proc. ACM SIGKDD Explorations Newsletter*, Volume 12, 2010