SMARTPHONE-BASED USER ACTIVITY RECOGNITION METHOD FOR HEALTH REMOTE MONITORING APPLICATIONS

Igor Bisio, Fabio Lavagetto, Mario Marchese and Andrea Sciarrone
University of Genoa, DYNATECH, Genoa, Italy

Keywords: Remote Monitoring, Activity Recognition, Accelerometer, Decision Trees, Windowed Decision, Android Smartphones.

Abstract: In the framework of health remote monitoring applications for individuals with disabilities or particular pathologies, quantity and type of physical activity performed by an individual/patient constitute important information. On the other hand, the technological evolution of Smartphones, combined with their increasing diffusion, gives mobile network providers the opportunity to offer real-time services based on captured real world knowledge and events. This paper presents a Smartphone-based Activity Recognition (AR) method based on decision tree classification of accelerometer signals to classify the user's activity as Sitting, Standing, Walking or Running. The main contribution of the work is a method employing a novel windowing technique which reduces the rate of accelerometer readings while maintaining high recognition accuracy by combining two single-classification weighting policies. The proposed method has been implemented on Android OS smartphones and experimental tests have produced satisfying results. It represents a useful solution in the aforementioned health remote applications such as the Heart Failure (HF) patients monitoring mentioned below.

1 INTRODUCTION

In the framework of health remote monitoring technologies for individuals with disabilities or particular pathologies, quantity and type of physical activity performed by an/a individual/patient constitute important information for the medical staff that monitors its state of health. An important case is represented by people suffering from HF: continuous monitoring of biometric parameters, such as body weight and the physical activity really performed, allow defining specific therapies that can significantly improve the quality of life. On the other hand, the technological evolution of smartphones, combined with their increasing diffusion, gives mobile network providers the opportunity to offer more advanced and innovative services. Among these are the so-called context-aware services. Examples of context-aware services are user profile changes as a result of context changes, user proximity-based advertising or media content tagging, etc. In order to provide context-aware services, a description of the smartphone environment must be obtained by acquiring and combining context data from different sources and sensors, both external (e.g. cell IDs, GPS coordinates) and internal (e.g. battery power, accelerometer measurements). An example of context-aware applications is (Keally, 2011), for remote health care services and (Boyle, 2006) for monitoring patients affected by chronic diseases.

In general the monitoring of the physical activity represents a very useful tool to develop effective therapies. For this reason, the proposed AR method is designed to distinguish four different user activities by periodically classifying accelerometer signals frames using a decision tree approach. The method employs a novel efficient windowing technique which reduces the signals frame acquisition rate and groups sets of consecutive frames in windows representing user state. In order to maintain high recognition accuracy, the reduced frequency of accelerometer readings is compensated by weighting each single-frame classification with a combination of two different sets of weights, which takes into account each frame's instant of occurrence and classification confidence. Experimental tests show accurate results while preserving battery life.

Bisio I., Lavagetto F., Marchese M. and Sciarrone A..
SMARTPHONE-BASED USER ACTIVITY RECOGNITION METHOD FOR HEALTH REMOTE MONITORING APPLICATIONS.
DOI: 10.5220/0003905502000205
In *Proceedings of the 2nd International Conference on Pervasive Embedded Computing and Communication Systems* (PECCS-2012), pages 200-205
ISBN: 978-989-8565-00-6

The contributions of this position paper will be further developed. In fact, the application of the proposed approach requires an extensive experimentation in cooperation with a medical staff and a sample group of patients suffering from HF. Moreover, before the real deployment of such a smartphone-based user activity recognition method, the recognized movements' set should be enriched with other typical activities such as climbing/down the stairs and indoor/outdoor cycling, treadmill.

2 RELATED WORK

In recent years a significant amount of work has been proposed concerning context-aware services relying on accelerometer data. Application fields are diverse and include remote health-care (Ryder, 2009), social networking and Activity Recognition (Miluzzo, 2008). Differently from the approach proposed in this paper, some of the proposed methods are designed to work with ad hoc sensors worn by the users (Keally, 2011). Other methods are developed for commoner devices such as smartphones. For example, Nokia's N95 is a popular choice (e.g., (Miluzzo, 2008), (Wang, 2009) and (Ryder, 2009)) have been implemented on it), but newer devices have received some attention as well: an iPhone version of (Miluzzo, 2008) has been developed and (Ryder, 2009) has also been implemented on Android smartphones. Other methods stand in between, requiring both custom hardware and off-the-shelf devices. For example, the method described in (Keally, 2011) employs an Android smartphone and ad-hoc wearable sensors.

A Paper Contribution. The proposed method is implemented on an Android smartphone and it takes into account the limitations of mobile devices. Even though processing power and battery capacity are improving, they still remain limited and valuable and must not be employed excessively for background added-value services giving priority to voice calls. Other projects have also tackled such problem: (Wang, 2009) proposes a system that manages sensor duty cycles and reduces energy consumption by shutting down unnecessary sensors. The proposed method follows a similar approach, but it also adds a window-based mechanism, which represents the main contribution of the paper.

3 APPLICATIVE SCENARIO: HEART FAILURE PATIENT MONITORING

The Heart Failure (HF) is a chronic disease that alternates intense and weak phases and requires repeated and frequent hospital treatments. The use of automatic instruments for a remote and ubiquitous monitoring of biological parameters, relevant with respect to the HF pathophysiology, offers new perspectives to improve the patients' life quality and the efficacy of the applied clinical treatment. In more details, HF is a disease represented by the limited capacity of the heart to provide a sufficient blood flow needed to meet all the body's necessities. HF usually causes a significant quantity of symptoms such as shortness of breath, weight gain due to excessive fluids, leg swelling, and exercise intolerance. This illness condition can be diagnosed with echocardiography and blood tests and the consequent treatment commonly consists of continuous lifestyle measures, drug therapy or, in very critical cases, surgery.

Currently, in medical practice, well-known patients' management models are focused on "manually handled" remote monitoring approaches: nurses daily interrogate, through a phone call, patients about their weight and the physical activity they have done. The achieved results of this practice show that this continuous remote monitoring approach improves the quality of life of these patients, prevents the progression to HF advanced stages and reduces the use of hospitalization.

The presented smartphone-based AR method associates modern context-aware capability of the recent smartphone platforms, obtained by implementing specific algorithmic solutions, with the any-time and any-where communication capability commonly offered by them. This joint usage will allow reducing the patients' involvement in the monitoring process without impacting its effectiveness. In particular, the method proposed is going to be applied in an real experimental campaign, in cooperation with a medical staff.

4 THE PROPOSED ACTIVITY RECOGNITION METHOD

The proposed activity recognition method is designed to distinguish four different user activities: *Sitting*, *Standing*, *Walking* and *Running*.

The algorithm periodically collects raw signals from the smartphone accelerometer. Sensed signal consists of a sequence of triples representing acceleration measurements along three orthogonal axes (produced at a variable rate). F (frame duration) seconds' worth of signal is acquired every T [s] (frame acquisition period), i.e., the accelerometer is switched off for $T-F$ seconds in order to reduce the overall energy consumption. With respect to constant accelerometer signals acquisition, a decrease in signals acquisition rate may lead to a less precise knowledge of the activity. Therefore, a windowing technique based on a single-frame classification weighting mechanism is employed, as described in Section 4 C.

For every frame a set of distinctive features is computed. Such feature vector is used by a decision tree classifier to classify the frame as *Sitting*, *Standing*, *Walking* and *Running*.

A groups of W consecutive frames are organized in windows, with consecutive windows overlapped by O frames. Every completed window is assigned to one of the four considered classes, based on a decision policy that takes into account each frame's instant of occurrence and classification confidence.

A Raw Accelerometer Signal. The smartphone employed in this work is an HTC Dream, which mounts a 3-axial accelerometer built by Asahi Kasei Corporation. Each signal sample (also called data in the following) produced by such integrated chip represents the acceleration (in m/s^2) measured on three orthogonal axes. In more detail, facing the phone display, the origin is in the lower-left corner of the screen, with the x axis horizontal and pointing right, the y axis vertical and pointing up and the z axis pointing outside the front face of the screen.

B Frame Classification. In order to be classified, a feature vector is associated to each individual frame composed by M samples. As in (Miluzzo, 2008), the features employed for single-frame classification are the mean (μ), standard deviation (σ) and number of peaks of the measurements (*P*, computed as reported in eq. 1) along the three axes x, y and z of the accelerometer.

$$P_j = \sum_{m=1}^{M} \rho_m, \text{ where}$$
$$\rho_m = \begin{cases} 1 \; if \;\; (j_{m+1} - j_m)(j_m - j_{m-1}) < 0, |j_m| \geq \varepsilon \\ 0 \; otherwise \end{cases} \quad (1)$$

j is a generic variable representing the accelerometer signal along the three axes and j_m is the m-th sample of the frame. ε in equation (1) is a threshold employed to define a signal peak. Thus the feature vector is $\{\mu_x, \mu_y, \mu_z, \sigma_x, \sigma_y, \sigma_z, P_x, P_y, P_z\}$.

Once a feature vector has been computed for a given frame, it is used by a classifier in order to associate the corresponding frame to one of the classes described at the beginning of Section 4. The employed classifier is a decision tree (Ross, 1993), a commonly used classifier in similar AR works such as (Miluzzo, 2008), (Ryder, 2009). Using the Weka workbench (Hall, 2009), several decision trees were designed and compared based on their recognition accuracy. A decision tree was trained for every combination of two and three of the users employed in the dataset creation (see Section 5 A). In order to evaluate the classifiers' performance, a separate test set was used for each combination.

C Windowed Decision. The rate of accelerometer readings must be compatible with the energy resources of the smartphone. Windowed decisions, defined below, guarantee satisfactory performance while saving the energy resource. In details, groups of W consecutive frames are organized in windows. The window size (i.e., the number of frames in each window) affects the state associated with the user and must be set carefully. Small windows ensure a quicker reaction to actual activity changes, but are more vulnerable to occasionally misclassified frames. On the other hand, large windows react more slowly to activity changes but provide better protection against misclassified frames. Consecutive windows are overlapped by O frames. Employing heavily-overlapped windows provides a better knowledge of the activity but may also imply consecutive windows bearing redundant information, while using slightly-overlapped windows could lead to signal sections representing meaningful data falling across consecutive windows. The parameters employed are: Δ, minimum time that must elapse between consecutive windowed decisions; W, number of periods in each window; O, number of periods shared by consecutive windows; N, pause between two consecutive signal acquisitions, expressed in number of frames. In practice, it is equivalent to considering a frame-acquisition period $T = (N+1) \cdot F$, where F is the frame duration expressed in seconds. Such parameters are tied by the following expression: $(W-O) \cdot T \geq \Delta$ and represented in Fig 1.

Each windowed decision assigns the current window to one of the four considered classes, based

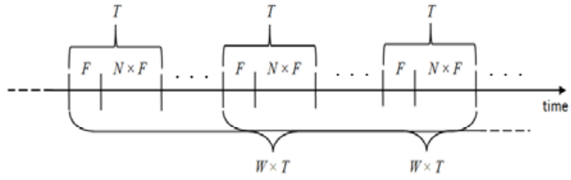

Figure 1: Diagram representing the raw accelerometer signal acquisition.

on a given decision policies. Four different decision policies were proposed, evaluated and compared, as detailed in the following.

1) Majority Decision. The simplest windowed decision policy is a majority-rule decision: the window is associated to the class with the most frames in the window. Such decision mechanism is employed in some earlier work on AR, e.g., (Toth, 2008). While it is clearly simple to implement and computationally inexpensive, the majority-rule windowed decision treats all frames within a window in the same way, without considering when the frames occurred or the single frame classifications' reliability.

2) Time-Weighted Decision. A first alternative to the majority-rule decision is the time-weighted decision. In a nutshell, it implies giving different weights to a window's frames based solely on their position in the window and assigning a window to the class with the highest total weight.

This way a frame will have a greater weight the closer it is to the end of the window, under the assumption that more recent classifications should be more useful to determine the current user activity. In order to determine what weight to give to frames, a weighting function $\Omega(t)$ was designed according to the criteria that $\Omega(0)=1$ and $\Omega(t_1) \geq \Omega(t_2)$ for any $t_1 \leq t_2$. It is worth noticing that t is non-negative and $t=0$ represents the time at which the most recent frame occurred.

If T_f is the instant associated with a frame and T_d is the instant at which the windowed decision is made, then the frame will be assigned a weight equal to $\Omega(T_d - T_f)$. Two different weighting functions were compared: *a gaussian* $\Omega_g(t) = e^{\frac{-t^2}{2k_g^2}}$ and a negative exponential $\Omega_e(t) = e^{-k_e t}$.

For each function type five different functions were compared by choosing k_g and k_e based on reference instant T_r and forcing $\Omega_i(T_r) = \omega,\ i = g, e$, where ω is one of five linearly-spaced values between 0 and 1.

3) Score-Weighted Decision. A second kind of windowed decision policy requires assigning to each frame *a* score representing how reliable its classification is. *In our work we implemented the method proposed in (Toth, 2008), not reported for the sake of brevity, and we applied it to the windowed based AR method.* The basic idea is that the closer a frame's feature vector is to the decision boundary, the more unreliable the frame's classification will be, under the hypothesis that the majority of badly classified samples lie near the decision boundary.

The distance of a feature vector from the decision boundary is given by the shortest distance to the leaves with class label different from the label associated to the feature vector. The distance between a feature vector and a leaf is obtained by solving a constrained quadratic program. Using separate training data for each leaf, an estimate of the correct classification probability conditional to the distance from the decision boundary is produced. Such estimate is computed by using the leaf's

- probability of correctly and incorrectly classifying training set samples (obtained in terms of relative frequency), and
- probability density of the distance from the decision boundary conditional to correct and false classification (obtained through kernel density estimation).

The classification score is finally given by the lower bound of the 95% confidence interval for the estimate of the correct classification probability conditional to the distance from the decision boundary. The confidence interval lower bound is used instead of the correct classification probability conditional to the distance from the decision boundary estimate because the latter may remain close to 1 even for large distances. However, a large distances may not imply a reliable classification probably due to an unknown sample located in a region of the feature space insufficiently represented in the training set. On the contrary, passed a certain distance (which varies with every leaf), the confidence interval lower bound decreases rapidly.

4) Joint Score-/Time-Weighted Decision. Another windowed decision policy is given by combining the temporal weights and the classification scores into a single, joint time-and-score weight. Fusion is obtained simply by multiplying the corresponding time weight and classification score, since both are between 0 and 1.

Table 1: Employed dataset.

	Sitting	Standing	Walking	Running
Frames	3702	3981	3822	3711
Duration [min]	246.8	265.4	254.8	247.4

5 PERFORMANCE INVESTIGATION

A Dataset. The dataset employed in the experiments was acquired by four volunteers. Each volunteer acquired about 1 hour of data for each of the classes described in Section 4, producing a total of almost 17 hours of data, as shown in Table 1. The phone was kept in the user's front or rear pants pocket (the last one is not used for the Sitting class since users will not keep the smartphone in a back pocket while sitting), as suggested in (Bao, 2004) and training data was acquired accordingly. Furthermore, the acquisition of training data was performed keeping the smartphone with the display facing towards the user or away from him and keeping the smartphone itself pointing up or down. For every combination of two and three users, the dataset was then divided into a training set for classifier training and a distinct test set for performance evaluation purposes.

B Parameters Setting. In order to determine the best values for parameters W, O, N and ω, an additional ad hoc sequence, not included in the dataset used for classifier training and testing, was acquired by a fifth volunteer. Such sequence is made of just over an hour of raw accelerometer signal and it is referred to all four considered user activities. Activities are performed in random order and their labels are used as ground truth. At first, single-frame classification is performed on the sequence, producing recognized-class labels and classification scores. After that, windowed decision accuracy is evaluated (as described in Section 5 C) for all admissible combinations of W, O and N, i.e., parameter values respecting the equations in Section 4 C. Δ was set to 60 [s] and W, O and N were evaluated in the following intervals: $W \in [3,9]$, $O \in [0, W-1]$, $N \in [0,14]$ fixed in an empirical way. Therefore, 411 different $\{W,O,N\}$ triples were evaluated.

C Results. Considering the single-frame results of all the evaluated classifiers, the one with the best accuracy produced a 98% correct test set classification average. In the following, the related confusion matrix (with percentages) has been reported.

Table 2: Confusion matrix in case of single-frame classification (%).

	Sitting	*Standing*	*Walking*	*Running*
Sitting	99	0	1	0
Standing	0.27	98.68	0.82	0.23
Walking	0	0.05	98.85	1.1
Running	0.28	0.1	4.4	95.22

As described in Section 5 B, windowed decision was applied to an *ad hoc* sequence using 411 different parameter configurations. Furthermore, all six decision policies described in Section 4 C (and also listed in Table 3) were compared for each parameter configuration, using five different values for ω (as described in Section 4 C) and two different values for T_r (i.e., 60 s and 120 s) for each policy. The results can be summed up in Tables 3. It reports the Recognition Accuracy (%) defined as the average correct detection over all considered classes and the Reading Time (%), which is the percentage of time dedicated to accelerometer signal reading with respect to the continuos reading (strictly related to the energy consumpion). From Table 2, the Recognition Accuracy is really outstanding in case of frame-based approach. Concerning the windowed approaches, which allow to save energy, the time-based frame classification weighting doesn't seem to improve performance compared to the majority decision significantly, while employing classification-score weighting, by itself or combined with time-weighting, led to significant improvements in windowed decision accuracy.

Table 3: Performance of the proposed windowed decision approaches.

	Decision Approach	*RA(%)*	*Reading Time (%)*
Frame Based	*Single-frame classification*	98	100
Window Based	*Majority*	80	12.5
	Time Weighted (Gaussian / Exponential)	80 / 8 4.62	12.5 / 20
	Score Weighted	88.24	9.09
	Joint Score-/Time-Weighted (Gaussian / Exponential)	88.24 / 88.24	9.09 / 9.09

Overall, the best parameter configuration led to the mentioned 88.24% windowed decision accuracy: it was obtained $W = 5$, $O = 1$ and $N = 7$, joint score/time frame weighting using a Gaussian function and T_r = 120 [s]. Such decision policy led to an 8.24% increase in windowed decision accuracy compared to the classical majority-rule decision.

Furthermore, using N = 7 allows reducing the Reading Time by 87.5% with respect to the case of constant accelerometer signal acquisition (N = 0), thus reducing energy consumption while maintaining a satisfying recognition accuracy.

Table 4: Performance of alternative smartphone-based approaches in the literature.

Reference	***RA (%)***	***Recognized Classes***	***Sensor(s)***
(Miluzzo, 2008)	79	*Sitting, Standing, Walking, Running*	Accelerometer
(Wang, 2009)	90	*Still,Vehicle, Walking, Running*	Accelerometer, GPS
(Ryder, 2009)	96	*Outdoor Activities*	Accelerometer

6 CONCLUSIONS

In this paper a smartphone-based activity recognition method designed to distinguish four different user activities is described. It represents a useful solution for health remote monitoring applications in particular in case of patients affected by Hearth Failure. It is based on the classification of accelerometer signal frames using a decision tree mechanism. In order to limit the device energy consumption, the proposed method employs a windowing technique which reduces the frame acquisition rate and groups sets of consecutive frames in windows representing the user state. The proposed AR method has a good level of accuracy. Its recognized movements' set will be enriched with other typical activities such as climbing/down the stairs and indoor/outdoor cycling, treadmill, soon. After that, it will be applied in an experimental campaign, in cooperation with a medical staff, to measure the quantity and the type of physical activity of patients affected by Heart Failure.

REFERENCES

M. Keally, G. Zhou, G. Xing, J. Wu and A. Pyles, 2011. PBN: Towards Practical Activity Recognition Using Smartphone-Based Body Sensor Networks, *SenSys'11*, November 1–4, 2011, Seattle, WA, USA.

J. Boyle, M. Karunanithi, T. Wark, W. Chan, and C. Colavitti, 2006. Quantifying functional mobility progress for chronic disease management, Engineering in Medicine and Biology Society. *EMBS'06. 28th Annual International Conference of the IEEE 5916*—5919.

E. Miluzzo et al., 2008. Sensing meets mobile social networks: the design, implementation and evaluation of the CenceMe application, in *SenSys '08: In Proc. of the 6th ACM conference on Embedded network sensor systems*. ACM, November 5 - 7, pp. 337–350.

Y. Wang, J. Lin, M. Annavaram, Q. A. Jacobson, J. Hong, B. Krishnamachari and N. Sadeh, 2009. A framework of energy efficient mobile sensing for automatic user state recognition, *In Proceedings of the 7th international Conference on Mobile Systems, Applications, and Services (Kraków, Poland, June 22 - 25, 2009)*. MobiSys '09. ACM, New York, NY, 179-192.

J. Ryder, B. Longstaff, S. Reddy, and D. Estrin, 2009. Ambulation: a tool for monitoring mobility patterns over time using mobile phones, UC Los Angeles: *Center for Embedded Network Sensing*.

L. Bao and S. S. Intille, 2004. Activity recognition from user-annotated acceleration data, *In 2nd International Conference, PERVASIVE '04*, Vienna, Austria, April 21-23.

J. Ross Quinlan, 1993. C4.5: programs for machine learning, *Morgan Kaufmann Publishers Inc*.

M. Hall et al., 2009. The WEKA data mining software: an update, *SIGKDD Explorations*, Volume 11, Issue 1.

N. Toth, and B. Pataki, 2008. Classification confidence weighted majority voting using decision tree classifiers, *International Journal of Intelligent Computing and Cybernetics*, vol. 1, no. 2, April.