Feature Set Optimization for Physical Activity Recognition Using Genetic Algorithms

Alejandro Baldominos*
U. Carlos III de Madrid
Avda. de la Universidad, 30
28911 Leganes, Spain
abaldomi@inf.uc3m.es

Yago Saez
U. Carlos III de Madrid
Avda. de la Universidad, 30
28911 Leganes, Spain
ysaez@inf.uc3m.es

Pedro Isasi
U. Carlos III de Madrid
Avda. de la Universidad, 30
28911 Leganes, Spain
isasi@ia.uc3m.es

ABSTRACT

Physical activity is recognized as one of the key factors for a healthy life due to its beneficial effects. The range of physical activities is very broad, and not all of them require the same effort to be performed nor have the same effects on health. For this reason, automatically recognizing the physical activity performed by a user (or patient) turns out to be an interesting research field, mainly because of two reasons: (1) it increases personal awareness about the activity being performed and its consequences on health, allowing to receive proper credit (e.g. social recognition) for the effort; and (2) it allows doctors to perform continuous remote patient monitoring.

This paper proposes a new approach for improving activity recognition by describing an activity recognition chain (ARC) that is optimized by means of genetic algorithms. This optimization process determines the most suitable and informative set of features that turns out into higher recognition accuracy while reducing the total number of sensors required to track the user activity. These improvements can be translated into lower costs in hardware and less intrusive devices for the patients. In this work, for the assessment of the proposed approach versus other techniques and for replication purposes, a publicly available dataset on physical activity (PAMAP2) has been used.

Experiments are designed and conducted to evaluate the proposed ARC by using leave-one-subject-out cross validation and results are encouraging, reaching an average classification accuracy of about 94%.

Categories and Subject Descriptors

I.2.6 [**Artificial Intelligence**]: Learning—*concept learning*; I.2.8 [**Artificial Intelligence**]: Problem Solving, Control Methods, and Search—*heuristic methods*; J.3 [**Life and Medical Sciences**]: Health

*Alejandro Baldominos is the corresponding author.

GECCO '15 Companion, July 11 - 15, 2015, Madrid, Spain

ISBN 978-1-4503-3488-4/15/07…$15.00

DOI: http://dx.doi.org/10.1145/2739482.2768506

Keywords

physical activity; activity recognition; activity recognition chain; genetic algorithms; feature selection; classification

1. INTRODUCTION

The positive effects of physical activity in health have been extensively explored throughout the medical literature. Recent studies have explored the benefits of physical activity in healthy ageing and the reduction of risk factors of chronic diseases [15], its benefits in school-aged children and youth [28], the association between the lack of physical activity in teenagers and mental illnesses [8], the correlation between the lack of exercise and the risk of developing alcohol use disorders [18], the benefits of sport practice in adolescent wellbeing [31] and even the relation between physical activity with the prevention of certain types of cancer [44, 54, 27], while in the last cases not conclusive results have been found. Other disorders associated to the lack of physical activity have also been thoroughly reviewed [30].

Interestingly, in the case of study of physical activity and prevention of prostate cancer cited above [27], authors reported "inconsistencies due to misclassification of physical activity", thus revealing research interest in the development of accurate systems for correctly classifying physical activity performed by the population. It should be noted that not all kinds of physical activity have the same effects on physiology and health. A distinction between different activities could be established, for instance, by the group of muscles activated by performing each of these activities. Another classification of physical activity could be determined by their metabolic equivalent (MET) or energy required to perform an activity, while the adequacy of this metric has been put into question [11].

Previous studies had shown that awareness about the benefits of physical activity and its relevance to the risk of heart disease was disappointing [24]. Since then, other studies have examined the interventions carried out around the world to promote and increase physical activity [26], as well as the correlates of physical activity [25, 7, 22, 20], showing increasing interest in understanding why some people is more active than other as well as effort to create awareness about the benefits of physical activity.

In any case, there is a clear interest of recognizing the activity that a subject is performing, not only because of medical reasons, but also in order to provide self-awareness of the activity level and to be able to potentially give the subject proper credit for his/her effort, being both key requirements for developing technologies promoting physical activity [17].

Activity recognition is a relatively new field with increasing interest given the availability of commodity hardware (i.e. smartphones, wearable devices, etc.) which are each day owned by more people and contains a set of sensors (accelerometers, gyroscopes, GPS, etc.) able to provide the basic information required to perform activity recognition.

A proof of the increasing interest in activity monitoring can be found in the appearance of an important number of inventions and patents in the recent years which aim at monitoring and quantifying physical activity of final users [52, 56] and more recently also at recognizing the activity being performed by the user [29] providing in some cases features of customized personal training [32, 53]. In addition, related health care cell phone apps also gain popularity in the last years and the main manufacturers included specific hardware into the cell phones for sensoring (e.g. Bosch BMA280 accelerometer or the InvenSense's six-axis MPU-6700). The beneficial effects of activity monitor-based counseling on physical activity and health have been recently explored by Vaes et al. [51].

This paper proposes an approach for performing supervised learning of a classification model able to predict the activity of a new user in real-time. To improve the results feature selection will be carried out by means of genetic algorithms, a widely used technique [42]. A tradeoff between classification accuracy and the number of required sensors is achieved trying to increase the classification accuracy while reducing as much as possible the number of sensors.

This paper is structured as follows: section 2 discusses the state of the art in the field of activity recognition and presents some of the related work, section 3 describes the methodology followed for the development of this work, thoroughly describing each step in the activity recognition chain. Experimental setup and results are later discussed in section 4 and finally conclusive remarks are future lines of work are provided in section 5.

2. ACTIVITY RECOGNITION SYSTEMS

As discussed before, activity recognition is a field broadly explored in recent research works. The ubiquity of sensors has eased the task of researchers to acquire data from physical activity, in most cases from accelerometers and gyroscopes located within cell phones.

Recently, Bulling et al. [10] have examined the different characteristics of human activity recognition systems (which is a more general concept than that of physical activity recognition systems) according to different criteria:

- According to its execution mode, the system can work either in an offline (the sensors data is available beforehand, and the activity recognition is performed once all the signals have been logged) or online (the sensors data is processed in real time) manner.
- According to its generalization ability, the system can either be user independent or user specific.
- According to the type of recognition, the system can either process continuous data (by identifying each activity or gesture within the stream of signals data) or isolated data (where the beginning and the end of the activity are defined beforehand).
- According to the type of the activities to be recognized, they can be periodic, sporadic or static. In most cases, physical activity is periodic.
- According to the system model, it can be either stateless (if the system only considers the sensors signals) or stateful (if the system considers also a model of the environment).

The physical activity recognition system aimed at this paper is trained offline with previously gathered data (while it can be used online in order to recognize physical activity performed by users in real time), is intended to be user-independent, it processes segmented data, where this segmentation proccess is part of the activity recognition chain, is aimed at detecting periodic activities (as it is often the case of physical activities, in contrast to sporadic gestures) and is stateless, as only data provided by sensors placed in the human body is considered.

Additionally, recent works often differentiate between data obtained from ad-hoc body sensors and from smartphones. The latter case is specially interesting because of the ubiquity of smartphone devices and because it enables performing activity recognition using non-intrusive technologies for the user, while on the other hand this approach tends to be less accurate because sensors (the smartphone itself) is not always wore in the same location and is not pointing always to the same direction [58]. Works focusing on activity recognition in smartphones have proliferated during the recent years. Some works focus on discussing the classification algorithms which can be used for this task [9, 48], or the most suitable feature sets for the classification task [34] while others look for specific techniques in order to reduce computational cost and energy expenditure [5, 47], a key issue when using smartphones as battery life is often affected when sensors are used and computing is performed. In some cases, public datasets have been released, as in the work from Anguita et al. [6], where a baseline benchmark using support vector machines (SVM) is also provided. A recent survey of the state-of-the-art in online physical activity recognition using smartphones has been addressed by Shoaib et al. [46].

On the other hand, this paper focuses in physical activity recognition using body-worn sensors, an approach which has also been extensively reviewed in the literature. Many works have focused on proposing techniques for the task of activity recognition, including fuzzy finite automata [1], ensemble methods combining support vector machines, artificial neural networks, and 1-nearest neighbors [57] or combining C4.5, multilayer perceptron and logistic regression [12], Hough transformation along with random projection trees [55], online multitask learning [49] or Naive Bayes and k-nearest neighbors [23]. A review of the topic's literature from the years 2011 and 2012 along with a proposal using C4.5 and AdaBoost is provided by Ugulino et al. [50], while this work was published before the PAMAP2 dataset used in this work was released.

In other cases, context-aware recognition systems are used in order to include information about the environment and complement the data extracted from sensors. This is the case of recent works, such as the ones from Gu et al. [21] and Alvarez et al. [4] where activity recognition is assisted using WiFi signals.

Activity recognition is often used for specific applications,

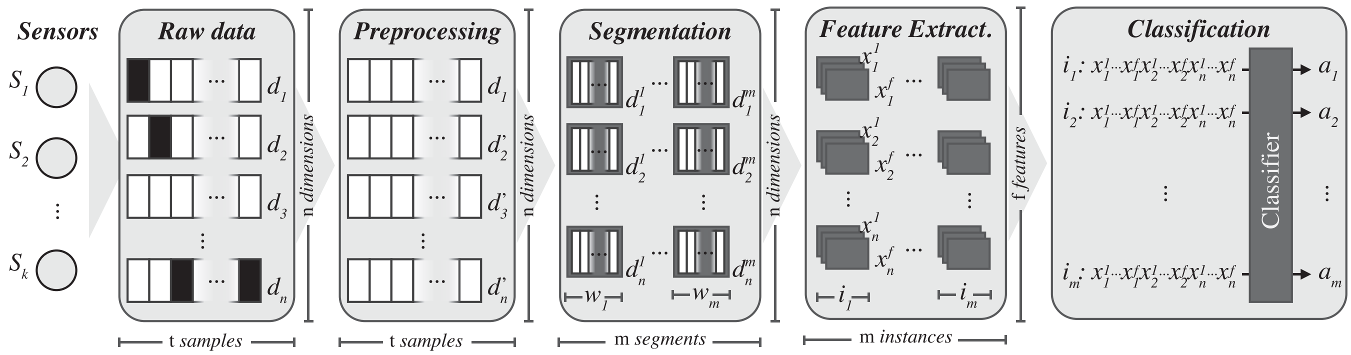

Figure 1: Steps involved in the activity recognition chain (ARC), from data acquisition to classification.

as in the case of Seiter et al. [45] where these techniques are used with stroke rehabilitation patients, the works from Altini et al. [3] or Chen et al. [13] where activity recognition and clustering is used to estimate energy expenditure, or the work from Alshurafa et al. [2] where games are used to reward physical activity.

As acquiring data from body-worn sensors is an expensive task, some works have proposed their own datasets. It is the case of the work from Reiss and Stricker [38] where the PAMAP2 dataset used for this work is introduced and a first benchmark using standard machine learning techniques is provided.

Finally, the application of genetic and evolutionary algorithms to the field of activity recognition has not been extensively explored, while there are some works which are relevant to the topic. These works involve in most cases the use of genetic algorithms for fusion weight selection of classifiers within ensembles, such as in the works of Fatima et al. [19] and Chernbumroong et al. [14], while these works do not focus specifically in physical activity, but rather activity recognition within homes. Also, genetic algorithms are also used for feature selection, as described in the works from Cilla et al. [16] and Saputri et al. [43].

3. METHODOLOGY

Works on human activity recognition usually follow a common sequence of steps, namely an activity recognition chain (ARC) [10], a general-purpose framework for obtaining data, building and evaluating activity recognition systems. As shown in figure 1, these steps involve data acquisition from sensors, signal preprocessing, signal segmentation, feature extraction and finally training of a classifier. This section describes each of these steps of the ARC.

3.1 Data Acquisition

The first step of the ARC involves acquiring physical activity data. This process usually involves the setting of a protocol for a set of subjects to perform a sequence of established activities while wearing certain sensors (or carrying their smartphones or other devices). Nevertheless, for this paper the PAMAP2 Physical Activity Monitoring dataset is used [41, 36, 38, 37, 40, 39, 35], which is publicly available at UCI Machine Learning Repository.

This dataset contains labeled information about physical activity for nine subjects wearing a heart rate unit and three Colibri wireless Inertial Measurement Units (IMUs) located over the wrist of the dominant arm, on the chest and on the dominant side's ankle respectively while performing the next set of activities:

- **lying**: lying quietly while doing nothing, small movements - e.g. changing the lying posture - are allowed.
- **sitting**: sitting in a chair in whatever posture the subject feels comfortable, changing sitting postures is allowed.
- **standing**: consists of standing still or standing still and talking, possibly gesticulating.
- **ironing**: ironing 1-2 shirts or t-shirts.
- **vacuuming**: vacuum cleaning one or two office rooms (which includes moving objects, e.g. chairs, placed on the floor).
- **ascending stairs**: was performed in a building between the ground and the top floors, a distance of five floors had to be covered going upstairs.
- **descending stairs**: was performed in a building between the ground and the top floors, a distance of five floors had to be covered going downstairs.
- **normal walk**: walking outside with moderate to brisk pace with a speed of 4-6 km/h, according to what was suitable for the subject.
- **nordic walk**: walking performed outside on asphaltic terrain, using asphalt pads on the walking poles (it has to be noted that none of the subjects was very familiar with this Nordic sport activity).
- **cycling**: was performed outside with a real bike with slow to moderate pace, as if the subject would bike to work or bike for pleasure (but not as a sport activity).
- **running**: jogging outside with a suitable speed for the individual subjects.
- **rope jumping**: the subjects used the technique most suitable for them, which mainly consisted of the basic jump (where both feet jump at the same time over the rope) or the alternate foot jump (where alternate feet are used to jump off the ground).

Table 1: PAMAP2 attributes extracted from IMUs

1	temperature (°C)
2-4	3D-acceleration data (ms^{-2}), scale: ±16g, resolution: 13-bit
5-7	3D-acceleration data (ms^{-2}), scale: ±6g, resolution: 13-bit
8-10	3D-gyroscope data (rad/s)
11-13	3D-magnetometer data (μT)
14-17	orientation (invalid data)

These activities are performed by all subjects under a fixed protocol [41], which consists in the execution of the previous activities in an established order, with all subjects spending the same time in the same exercise.

In total, 9 subjects (8 males and 1 female) take part in the data acquisition step, aged 27.22 ± 3.31 years and having a BMI of 25.11 ± 2.62 kgm^{-2}, one being left-handed and the remaining being right-handed [36]. Some subjects may show a slight deviation with this protocol due to problems with the hardware setup such as connections losses or system crashes, which mostly causes differences in the timing of the activities or leads to unavailability or loss of information for some activities. However, subject 9 is an extreme case as his data completely differs from the specified protocol, and thus he is ignored for the experiments carried out in this paper.

The IMUs worn by the subjects generate a total of 51 attributes, as individual IMUs provide 17 attributes each, as shown in table 1. Besides, each instance also contains a timestamp (in seconds) and the heart rate (in bpm), so the total number of attributes is 53 plus the activity, which is the class.

The IMUs have a sampling frequency of 100Hz, so one activity second provides 100 instances for the dataset. Meanwhile, the heart rate monitor has a sampling frequency of 9Hz, and as a result heart rate information is not available in about 91% of the dataset instances.

3.2 Signal Preprocessing

Signal's preprocessing step is very relevant because the signal samples recovered typically contain noise and other features that are worth being taken into consideration. This phase will receive the original samples as an input and will return a new sequence of samples. For this dataset, the next actions will take part of the preprocessing phase:

- Removing the timestamp, as it is an identifier which would add bias to the classifier (actually, knowing the time at which an activity was carried out is probably enough to accurately predict the activity itself, as all subjects follow a fixed protocol).

- Removing the information of orientation (a total of 12 attributes of the dataset, 4 attributes per IMU) as the authors state that these attributes are invalid, or not relevant, for this data collection [33].

- Completing the missing values (available in the original samples as NaN values) by estimating their real values. In all cases, this data is computed as the same as the previous available value. This is considered a good approximation as values are not expected to change significantly within one hundredth of a second, and for all attributes except the heart rate the number of missing values is very small (and only happens due to communication failures). In the case of the heart rate, there are more missing values because of the lowest sampling frequency, but again it has been considered that the heart rate will not change significantly within one tenth of a second.

- Removing instances labeled as *transition*, as they do not correspond to any activity, but the time after one activity ends and before the next starts.

After the preprocessing stage, the dataset dimensionality is reduced to 40 features plus the activity.

3.3 Signal Segmentation

The raw data is in the time domain, i.e., each instance represents the signals provided by the sensor in a given point in time. While this format is suitable for training a model, it is expected that higher accuracy can be achieved in the classification if instances themselves capture temporal information. To do so, in the segmentation phase the signals will be converted into the frequency domain, for which a Discrete Fourier Transform (DFT) will be applied over the dataset.

For this process, a sliding window of size 512 (corresponding to 5.12 seconds of data) is defined, and for each of these windows the Fast Fourier Transform (FFT) is computed[1]. This computation will return for each signal its transform in the frequency domain, composed of 512 different values.

The window of samples, for which the FFT is computed, is moved once at a time, and only instances belonging to the same class are considered for the same window (i.e., a window will never contain instances belonging to two different physical activities).

3.4 Feature Extraction

Once each signal is transformed into the frequency domain, the resulting signal must be processed in order to extract features to compose the new dataset. In this case, each window will generate a new instance in the dataset.

To extract the new features, a statistical summary of the 512 values generated after computing the DFT will be calculated, leading to the generation of 7 attributes for each signal: the mean, the median, the standard deviation, the maximum value, the minimum value, the 25% percentile and the 75% percentile. As a result, each instance in the original dataset is replaced by a new new one containing 280 features plus the activity.

As it can be seen, the dataset dimensionality has significantly increased in size, by a factor of 7. This high dimensionality will lead to higher training times and could potentially affect the model accuracy negatively due to overfitting. In order to reduce the number of features in the resulting dataset, a genetic optimization approach will be used.

In particular, a local optimization of the feature set for each user will be pursued which will later be used to obtain an approximation of the best feature set for all the users. For this local optimization, a binary chromosome of size 280 is defined, where each gene represents a feature and whether it is considered for training the model (1) or it is not (0). A genetic algorithm is setup with a population of 50 individuals,

[1]It should be noted that a window with a size power of two has been chosen in order to be able to use the FFT algorithm, thus increasing computational performance.

Table 2: Number of features for each value of τ

τ	# of features	τ	# of features
1	280	5	94
2	264	6	35
3	233	7	7
4	166	8	0

a crossover rate of 35% (with random crossover point) and a mutation rate of $1/12^{th}$ (preliminary experimentation had shown no significant difference in the results with alternative setups). The fitness function is defined as the accuracy achieved by a Random Forest classifier in the test set.

While feature optimization is performed in a per-user basis, leading to the obtainment of 8 different chromosomes, this paper looks for a set of features that is applicable to all users. To achieve this set, a threshold τ is defined so that an attribute a_i is chosen only if at least τ out of eight bits for the corresponding genes in all chromosomes are 1, i.e.:

$$a_i = \begin{cases} 1, & \text{if } \sum_{n=\{1..8\}} g_i^n \geq \tau \\ 0, & \text{otherwise} \end{cases}$$

where a_i is the i-th attribute and g_i^n is the i-th gene of the best individual for the n-th fold.

Table 2 shows the number of resulting features once that this process is applied for different values of the parameter τ, excluding the class. It must be remarked than as the value of τ grows, the resulting feature set is always a subset of the set for smaller values of the threshold.

It can be realized that the feature set for $\tau = 1$ contains all the features, and also the value for $\tau = 8$ is the empty set ($\varnothing$), which means that there is no attribute for which the value for its corresponding gene is either 0 or 1 for all chromosomes.

3.5 Classification

The problem of activity recognition belongs to the field of supervised learning, where a model is to be trained using a set of instances from which the class (in this case the activity) is known a priori.

For building the classification model, standard machine learning techniques addressed in the literature are used. In a first approach, before feature optimization takes place, both Naive Bayes, C4.5 and Random Forest are used in order to establish baseline results. As it will be seen in the next section, Random Forest significantly outperforms Naive Bayes and C4.5 for all users, so this technique is used for the rest of the work (both the fitness computation and the evaluation of the final models).

4. RESULTS

In order to evaluate the activity recognition system, a leave-one-subject-out (LOSO) cross validation procedure is used for the experiments, where 8 different experiments take place, each one taking a different subject as the test set and the remaining 7 subjects as the training set.

In the first place, several models are learnt and tested with each user using the complete set of features (composed of 280 attributes as described in the previous section). In order to establish a baseline comparison, three different machine learning techniques are used: Naive Bayes, C4.5 and Random Forest. Results are shown in table 3, where it can be seen that Random Forest outperforms its competitors.

Once these baseline results are achieved, the next task involves the optimization of the feature set using genetic algorithms as described in section 3.4. As discussed before, the feature set is individually optimized in a per-user basis, i.e., a model is optimized for each fold in the LOSO cross validation, leading to the results shown in table 4.

As it can be seen, the results are quite close to 100%. Nevertheless, the main handicap of this approach is that the set of selected features is different for each user. Therefore, it is optimized for each different test subject and fails to provide a fixed set of features for a user-independent classification model, not accomplishing the objective of reducing the total number of sensors.

To attain a subject-independent set of features, the procedure described in section 3.4 is carried out. The resulting number of features for each value of τ was already discussed in table 2. The results for each different value of τ (excluding $\tau = 0$ and $\tau = 8$) are shown in table 5. Each result for each subject and value of τ is the average of 30 different executions, each one training the model with a random sample of 10% of the training set.

It can be seen that results slightly increase for small values of τ. The optimum value is achieved when $\tau = 4$, but specially interesting is the case when $\tau = 5$, because the difference in the average accuracy is small and the number of features is almost reduced by a factor of three from the original set of features generated after computing the signal segmentation and the first feature extraction process. Not only the difference in accuracies is small, but running a two-sample t-test for each subject with the significance level set to $\alpha = 0.05$ reveals that there is not statistical difference[2] between the values of $\tau = 4$ and $\tau = 5$ for subjects 5, 6, 7 and 8 (i.e., the null hypothesis stating that samples for both values of τ are random samples coming from normal distributions with equal means and equal but unknown variances fails to be rejected); and with the significance level set to $\alpha = 0.01$ the same happens for subject 2.

These results can be compared to those obtained in other works performing activity classification over the PAMAP2 dataset. A benchmark is first provided by Reiss and Stricker [38] when the dataset is introduced. In that work, it can be seen that different classification tasks are evaluated and different experimentation setups are used. The ones that are comparable to the results obtained in this paper are those referring to the "all activity" recognition tasks using LOSO cross validation. In their work, the highest accuracy is obtained using k-nearest neighbors, getting a value of 89.24%, more than five points below the best result obtained in this work (94.64%). While other authors, such as Soria et al. [47] have also used the PAMAP2 dataset for evaluating their proposals, the accuracy is aggregated with other datasets, making further comparison impossible.

Concluding, results show that the proposed activity recognition system achieves a very high accuracy for all subjects, using an unbiased LOSO cross-validation experimentation. Also, optimization using a genetic algorithm allows reducing

[2]Using $\alpha = 0.05$, homoscedasticity is not met using Levene's test for subject 4, and normality cannot be established using Lilliefors test for subjects 3 ($\tau = 4$), 5 ($\tau = 5$) and 8 ($\tau = 4$). In these cases, a Wilcoxon rank sum test is used, resulting in failing to reject the null hypothesis for subjects 5 and 8.

Table 3: Classification accuracy for each fold using the whole feature set

	Subj. 1	Subj. 2	Subj. 3	Subj. 4	Subj. 5	Subj. 6	Subj. 7	Subj. 8	Avg.
Naive Bayes	73.29%	65.79%	94.19%	96.66%	90.91%	88.33%	91.73%	89.50%	**86.30%**
C4.5	78.31%	80.48%	74.70%	84.39%	71.27%	89.59%	72.19%	82.85%	**79,23%**
Random Forest	85.82%	89.73%	92.83%	96.02%	92.83%	96.50%	97.35%	97.03%	**93.51%**

Table 4: Classification accuracy after optimizing the feature set for each fold

Subj. 1	Subj. 2	Subj. 3	Subj. 4	Subj. 5	Subj. 6	Subj. 7	Subj. 8	Avg.
97.91%	97.10%	99.16%	99.88%	98.42%	99.31%	99.66%	99.68%	**98.89%**

the total set of features from 280 to 94, and in consequence, decreasing the total number of sensors wore by the subjects and their costs.

5. CONCLUSIONS AND FUTURE WORK

This paper has explored the field of human activity recognition. Experts consider physical activity as a key aspect for human health, preventing numerous diseases and boosting energy. Not all physical activities have the same effect on health, as these effects may vary depending on the muscles used for the activity, the energy cost required to complete the activity, etc. For this reason, a system able to accurately recognize the physical activity performed by a subject sounds quite promising, both due to the fact that it may increase awareness about the subject's health and because it would potentially be able to track this activity with medical purposes (e.g. during specific treatment).

This paper has discussed the development of an activity recognition system by thoroughly describing its activity recognition chain (ARC), which explains how the data is obtained, preprocessed, segmented, how features are extracted and finally how a classification model is trained. For this paper, the PAMAP2 dataset has been used, which is publicly available for download, thus easing replication. Basic preprocessing is performed and the DFT is applied over a sliding window of the data in order to convert the signals from the time domain to the frequency domain. Then, features are extracted by performing a statistical summary of the signal computed by the DFT.

Once features are extracted, baseline results are obtained by applying standard machine learning algorithms in order to learn classification models, using leave-one-subject-out (LOSO) cross validation for preventing biased results. An average classification accuracy of 93.5% is achieved.

Finally, a genetic algorithm is used to optimize the feature set. A local per-user approach is used for this optimization obtaining accuracies up to 99% and even higher, and reaching an uniform level of forecasting more suitable for every individual. Later, all feature sets are aggregated by defining a threshold τ for all individuals, allowing to generate a subject-independent classification model with still very high levels of accuracy. This method reveals the relevant sensors for the classification, reducing features and economizing the sensors to wear without lacking on the activity detection. This genetic-based optimization increases the accuracy up to 94.6% but more interestingly, it is able to achieve accuracies of 94% with only 94 features, a third part of the original set of 280 features.

Some improvements and further experiments could be performed over this work, and are proposed here as future work. For instance, the genetic algorithm could be used to select sensors (the first step in the ARC) rather than features. This would reduce the search space increasing the speed during the GA execution, on the other hand leading to lower average accuracies. While the genetic algorithm presented in this paper performs local optimization of the feature set for each subject, it would be interesting to perform global optimization of the feature set for all subjects, while doing so would increase the fitness computation significantly. Both tasks could be performed simultaneously in order to achieve a tradeoff between the attained accuracy and the optimization time.

6. ACKNOWLEDGEMENTS

This research was partially funded by European Union's CIP Programme (ICT-PSP-2012) under grant agreement no. 325146 (SEACW project).

7. REFERENCES

[1] H. Ali and D. Amalarethinam. Activity recognition with fuzzy finite automata. In *Proc. WCCCT'14*, pages 222–227, 2014.

[2] N. Alshurafa, W. Xu, J. Liu, M.-C. Huang, B. Mortazavi, C. Roberts, and M. Sarrafzadeh. Designing a robust activity recognition framework for health and exergaming using wearable sensors. *IEEE Journal of Biomedical and Health Informatics*, 18(5):1636–1646, 2013.

[3] M. Altini, J. Penders, R. Vullers, and O. Amft. Estimating energy expenditure using body-worn accelerometers: a comparison of methods, sensors number and positioning. *IEEE Journal of Biomedical and Health Informatics*, 19(1):219–226, 2014.

[4] A. Alvarez, J. Alonso, and G. Trivino. Human activity recognition in indoor environments by means of fusing information extracted from intensity of WiFi signal and accelerations. *Information Sciences*, 233:162–182, 2013.

[5] D. Anguita, A. Ghio, L. Oneto, X. Parra, and J. L. Reyes-Ortiz. Human activity recognition on smartphones using a multiclass hardware-friendly support vector machine. In *Ambient Assisted Living and Home Care*, volume 7657 of *Lecture Notes in Computer Science*, pages 216–223. 2012.

[6] D. Anguita, A. Ghio, L. Oneto, X. Parra, and J. L. Reyes-Ortiz. A public domain dataset for human activity recognition using smartphones. In *Proc. ESANN'13*, pages 24–26, 2013.

[7] A. E. Bauman, R. S. Reis, J. F. Sallis, J. C. Wells, R. J. Loos, and B. W. Martin. Correlates of physical

Table 5: Classification accuracy for each fold using the reduced feature set

	Subj. 1	Subj. 2	Subj. 3	Subj. 4	Subj. 5	Subj. 6	Subj. 7	Subj. 8	Avg.
$\tau = 1$	88.93%	88.39%	93.16%	96.17%	93.78%	95.89%	96.76%	96.44%	**93.69%**
$\tau = 2$	88.93%	89.23%	94.44%	97.35%	94.06%	96.01%	96.39%	96.43%	**94.06%**
$\tau = 3$	88.77%	88.43%	93.38%	97.47%	94.71%	96.51%	97.45%	96.07%	**94.10%**
$\tau = 4$	88.43%	91.78%	92.73%	98.98%	96.02%	96.92%	98.21%	94.02%	**94.64%**
$\tau = 5$	86.60%	88.20%	91.32%	99.49%	96.42%	96.89%	98.32%	94.41%	**93.96%**
$\tau = 6$	77.68%	80.00%	71.30%	89.88%	89.72%	93.71%	94.34%	88.10%	**85.59%**
$\tau = 7$	71.04%	67.42%	58.42%	62.87%	74.46%	68.22%	69.47%	65.85%	**67.22%**

activity: why are some people physically active and others not? *The Lancet*, 380(9838):258–271, 2012.

[8] S. J. H. Biddle and M. Asare. Physical activity and mental health in children and adolescents: a review of reviews. *British Journal of Sports Medicine*, 45(11):886–895, 2011.

[9] M. F. A. bin Abdullah, A. F. P. Negara, S. Sayeed, D.-J. Choi, and K. S. Muthu. Classification algorithms in human activity recognition using smartphones. *International Journal of Computer and Information Engineer*, 6:77–84, 2012.

[10] A. Bulling, U. Blanke, and B. Schiele. A tutorial on human activity recognition using body-worn inertial sensors. *ACM Computing Surveys (CSUR)*, 46(3):33:1–33:33, 2014.

[11] N. M. Byrne, A. P. Hills, G. R. Hunter, R. L. Weinsier, and Y. Schutz. Metabolic equivalent: one size does not fit all. *Journal of Applied Physiology*, 99(3):1112–1119, 2005.

[12] C. Catal, S. Tufekci, E. Pirmit, and G. Kocabag. On the use of ensemble of classifiers for accelerometer-based activity recognition. *Applied Soft Computing*, In Press, 2015.

[13] S. Chen, J. Lach, O., M. Altini, and J. Penders. Unsupervised activity clustering to estimate energy expenditure with a single body sensor. In *Proc. BSN'13)*, 2013.

[14] S. Chernbumroong, S. Cang, and H. Yu. GA-based classifiers fusion for multi-sensor activity recognition of elderly people. *IEEE Journal of Biomedical and Health Informatics*, 19(1):282–289, 2015.

[15] W. Chodzko-Zajko. Exercise and physical activity for older adults. *Universal Access in the Information Society*, 3(1):101–106, 2014.

[16] R. Cilla, M. A. Patricio, J. García, A. Berlanga, and J. M. Molina. Recognizing human activities from sensors using hidden markov models constructed by feature selection techniques. *Algorithms*, 2(1):282–300, 2009.

[17] S. Consolvo, K. Everitt, I. Smith, and J. A. Landay. Design requirements for technologies that encourage physical activity. In *Proc. SIGCHI Conf. on Human Factors in Comput. Syst. (CHI'06)*, pages 457–466, 2006.

[18] L. K. Ejsing, U. Becker, J. S. Tolstrup, and T. Flensborg-Madsen. Physical activity and risk of alcohol use disorders: results from a prospective cohort study. *Alcohol and Alcoholism*, 50(2):206–212, 2014.

[19] I. Fatima, M. Fahim, Y.-K. Lee, and S. Lee. A genetic algorithm-based classifier ensemble optimization for activity recognition in smart homes. *Applied Soft Computing*, In Press, 2015.

[20] J. G. Godino, C. Watkinson, K. Corder, S. Sutton, S. J. Griffin, and E. M. van Sluijs. Awareness of physical activity in healthy middle-aged adults: a cross-sectional study of associations with sociodemographic, biological, behavioural, and psychological factors. *BMC Public Health*, 14(421), 2014.

[21] Y. Gu, L. Quan, and F. Ren. Wifi-assisted human activity recognition. In *Proc. APWiMob'14*, pages 60–65, 2014.

[22] K. B. Guntera, K. R. Ricea, D. S. Wardb, and S. G. Trosta. Factors associated with physical activity in children attending family child care homes. *Preventive Medicine*, 54(2):131–133, 2012.

[23] P. Gupta and T. Dallas. Feature selection and activity recognition system using a single triaxial accelerometer. *IEEE Transactions on Biomedical Engineering*, 61(6):1780–1786, 2014.

[24] A. Haase, A. Steptoe, J. F. Sallis, and J. Wardle. Leisure-time physical activity in university students from 23 countries: associations with health beliefs, risk awareness, and national economic development. *Preventive Medicine*, 39(1):182–190, 2004.

[25] P. C. Hallal, L. B. Andersen, F. C. Bull, R. Guthold, W. Haskell, and U. Ekelund. Global physical activity levels: surveillance progress, pitfalls, and prospects. *The Lancet*, 380(9838):247–257, 2012.

[26] G. W. Heath, D. C. Parra, O. L. Sarmiento, L. B. Andersen, N. Owen, S. Goenka, F. Montes, and R. C. Brownson. Evidence-based intervention in physical activity: lessons from around the world. *The Lancet*, 380(9838):272–281, 2012.

[27] I. L. IM, H. Sesso, J. Chen, and R. Paffenbarger. Does physical activity play a role in the prevention of prostate cancer? *Epidemiologic Reviews*, 23(1):132–137, 2001.

[28] Janssen and LeBlanc. Systematic review of the health benefits of physical activity and fitness in school-aged children and youth. *International Journal of Behavioral Nutrition and Physical Activity*, 7(40), 2010.

[29] P. R. Kahn, A. Kinsolving, M. A. Christensen, B. Y. Lee, and D. Vogel. Human activity monitoring device with activity identification, 2015.

[30] J. A. Knight. Physical inactivity: associated diseases and disorders. *Annals of Clinical & Laboratory Science*, 42(3):320–337, 2012.

[31] A. Merglen, A. Flatz, R. E. Bélanger, P.-A. Michaud,

and J.-C. Suris. Weekly sport practice and adolescent well-being. *Archives of Disease in Childhood*, 99:208–210, 2014.

[32] D. Morris, I. Kelner, F. Shariff, D. Tom, T. S. Saponas, and A. Guillory. Personal training with physical activity monitoring device, 2015.

[33] A. Reiss. *Personalized Mobile Physical Activity Monitoring for Everyday Life*. PhD thesis, Technical University of Kaiserslautern, 2013.

[34] A. Reiss, G. Hendeby, and D. Stricker. A competitive approach for human activity recognition on smartphones. In *Proc. ESANN'13*, 2013.

[35] A. Reiss, G. Hendeby, and D. Stricker. A novel confidence-based multiclass boosting algorithm for mobile physical activity monitoring. *Personal and Ubiquitous Computing*, 19(1):105–121, 2014.

[36] A. Reiss and D. Stricker. Towards global aerobic activity monitoring. In *Proc. PETRA'11*, 2011.

[37] A. Reiss and D. Stricker. Creating and benchmarking a new dataset for physical activity monitoring. In *Proc. 5th Intl. Conf. Pervasive Tech. related to Assistive Environments (PETRA'12)*, pages 108–109, 2012.

[38] A. Reiss and D. Stricker. Introducing a new benchmarked dataset for activity monitoring. In *Proc. Intl. Symp. Wearable Computers (ISWC'12)*, pages 108–109, 2012.

[39] A. Reiss and D. Stricker. Aerobic activity monitoring: towards a long-term approach. *Universal Access in the Information Society*, 13:101–114, 2013.

[40] A. Reiss, D. Stricker, and I. Lamprinos. An integrated mobile system for long-term aerobic activity monitoring and support in daily life. In *Proc. TrustCom'12*, pages 2021–2028, 2012.

[41] A. Reiss, M. Weber, and D. Stricker. Exploring and extending the boundaries of physical activity recognition. In *Proc. ICSMC'11*, pages 46–50, 2011.

[42] Y. Saeys, I. Inza, and P. Larrañaga. A review of feature selection techniques in bioinformatics. *Bioinformatics*, 23(19):2507–2517, 2007.

[43] T. R. D. Saputri, A. M. Khan, and S.-W. Lee. User-independent activity recognition via three-stage GA-based feature selection. *International Journal of Distributed Sensor Networks*, 2014, 2014.

[44] P. Schnohr, M. Grønbæk, L. Petersen, H. O. Hein, and T. I. Sørensen. Physical activity in leisure-time and risk of cancer: 14-year follow-up of 28,000 Danish men and women. *Scandinavian Journal of Public Health*, 33(4):244–249, 2005.

[45] J. Seiter, A. Derungs, C. Schuster-Amft, O. Amft, , and G. Trster. Activity routine discovery in stroke rehabilitation patients without data annotation. In *Proc. Rehab'14*, 2014.

[46] M. Shoaib, S. Bosch, O. D. Incel, H. Scholten, and P. J. Havinga. A survey of online activity recognition using mobile phones. *Sensors*, 15(1):2059–2085, 2015.

[47] L. M. Soria, L. Gonzalez-Abril, J. A. Ortega, and M. A. Alvarez. Low energy physical activity recognition system on smartphones. *Sensors*, 15(3):5163–5196, 2015.

[48] X. Su, H. Tong, and P. Ji. Activity recognition with smartphone sensors. *Tsinghua Science and Technology*, 19(3):235–249, 2014.

[49] X. Sun, H. Kashima, and N. Ueda. Large-scale personalized human activity recognition using online multitask learning. *IEEE Transactions on Knowledge and Data Engineering*, 25(11):2551–2563, 2013.

[50] W. Ugulino, D. Cardador, K. Vega, E. Velloso, R. Milidiú, , and H. Fuks. Wearable computing: accelerometers' data classification of body postures and movements. In *Advances in Artificial Intelligence - SBIA 2012*, pages 52–61. 2012.

[51] A. W. Vaes, A. Cheung, M. Atakhorrami, M. T. J. Groenen, O. Amft, F. M. E. Franssen, E. F. M. Wouters, and M. A. Spruit. Effect of 'activity monitor-based' counseling on physical activity and health-related outcomes in patients with chronic diseases: A systematic review and meta-analysis. *Annals of Medicine*, 45(5-6):397–412, 2013.

[52] C. A. Vock, P. Flentov, and D. M. Darcy. Activity monitoring systems and methods, 2013.

[53] K. L. White, M. L. Orenstein, J. Campbell, C. S. Self, E. Walker, M. Micheletti, G. McKeag, J. Zipperer, and M. Lapinsky. Activity recognition with activity reminders, 2015.

[54] Q. Xiao, H. P. Yang, N. Wentzensen, A. Hollenbeck, and C. E. Matthews. Physical activity in different periods of life, sedentary behavior, and the risk of ovarian cancer in the NIH-AARP diet and health study. *Cancer Epidemiology, Biomarkers & Prevention*, 22(11):2000–2008, 2013.

[55] G. Yu, J. Yuan, and Z. Liu. Propagative Hough voting for human activity detection and recognition. *IEEE Transactions on Circuits and Systems for Video Technology*, 25(1):87–98, 2015.

[56] S. G. J. Yuen, J. Park, and E. N. Friedman. Activity monitoring systems and methods of operating same, 2013.

[57] Y. Zheng, W.-K. Wong, X. Guan, and S. Trost. Physical activity recognition from accelerometer data using a multi-scale ensemble method. In *Proc. IAAI'10*, pages 1575–1581, 2010.

[58] Y. Zhu, C. Wang, J. Zhang, and J. Xu. Human activity recognition based on similarity. In *Proc. CSE'14*, pages 1382–1387, 2014.