View metadata, citation and similar papers at core.ac.uk
brought to you by CORE
provided by Universidad Carlos III de Madrid e-Archivo

This is a postprint version of the following published document:

Baldominos,A., Barrio, C. del, y Saez, Y. (2016). Exploring the Application of Hybrid Evolutionary Computation Techniques to Physical Activity Recognition. In *Proceedings of the 2016 on Genetic and Evolutionary Computation Conference Companion*, pp. 1377-1384.

DOI: https://doi.org/10.1145/2908961.2931732

© 2016 Copyright held by the owner/author(s). Publication rights licensed to ACM.

Exploring the Application of Hybrid Evolutionary Computation Techniques to Physical Activity Recognition

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

Carmen del Barrio
University Carlos III of Madrid
Avda. de la Universidad, 30
28911 Leganes, Spain
100283137@alumnos.uc3m.es

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

ABSTRACT

This paper focuses on the problem of physical activity recognition, i.e., the development of a system which is able to learn patterns from data in order to be able to detect which physical activity (e.g. running, walking, ascending stairs, etc.) a certain user is performing.

While this field is broadly explored in the literature, there are few works that face the problem with evolutionary computation techniques. In this case, we propose a hybrid system which combines particle swarm optimization for clustering features and genetic programming combined with evolutionary strategies for evolving a population of classifiers, shaped in the form of decision trees. This system would run the segmentation, feature extraction and classification stages of the activity recognition chain.

For this paper, we have used the PAMAP2 dataset with a basic preprocessing. This dataset is publicly available at UCI ML repository. Then, we have evaluated the proposed system using three different modes: a user-independent, a user-specific and a combined one. The results in terms of classification accuracy were poor for the first and the last mode, but it performed significantly well for the user-specific case. This paper aims to describe work in progress, to share early results an discuss them. There are many things that could be improved in this proposed system, but overall results were interesting especially because no manual data transformation took place.

Keywords

physical activity; activity recognition; hybrid evolutionary computation; particle swarm optimization; genetic programming; metaheuristics; classification

1. INTRODUCTION

Nowadays, there are plenty of fields in which human activity automatic recognition is very useful or necessary. While this discipline is used extensively in the fields of health and wellness, applications can also be found in different areas such as ubiquitous computing, entertainment, personal daily activities logging, or sports and professional monitoring.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first p age. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

Human population's lifestyle is getting more sedentary every day, and there is growing research interest in studying the correlation between physical inactivity and diseases, from non transmissible (like heart disease, hypertension or diabetes), to psychological issues like depression.

The World Health Organization (WHO) has highlighted some important key facts [45] on this problem, reflecting the seriousness of this generalized sedentary attitude and its relation to some health problems. Some of these show that the lack of physical activity is one of the 10 top risk factors for death and is also significantly related to other diseases. In addition, it also states that one in four adults is not physically active enough, and that ratio increases up to 80% in the adolescent population. However, WHO admits that there is a significant effort for reverting this situation, and in particular 56% of WHO Member States have defined policies for addressing insufficient physical activity and all of them have agreed to reduce these ratios by 10% by 2025.

More interestingly, the WHO provides recommendations about the minimum levels of physical activity required for a healthy lifestyle. They provide hints about duration and frequency of physical activity, but also about intensity and type, meaning that not all exercises have the same health implications. Due to this emerging trend, activity recognition systems are presented as a good solution for satisfying the need for automatic methods able to quantify routines and human activity patterns. There are many research related to physical activity motivation [42, 43], the improvement of diagnosis and treatment of neurological, degenerative and respiratory disorders, such as Parkinson disease [12], multiple sclerosis [20] or chronicle lung diseases [23].

Moreover, recent development in sensor miniaturization make possible the recompilation of many features of human motion unobtrusively, bringing new horizons for the application of automatic activity and context recognition, including the use of wearable sensors and ubiquitous computing. However, all this potentially useful information have no practical sense unless using the right tools for extracting the knowledge from these big data sources.

Finally, while this is a relatively new field, physical activity recognition has numerous approaches of very diverse nature. While most works so far have applied well-known machine learning techniques, approaches involving evolutionary computation are not abundant; however, as these techniques have been successfully applied to many other fields, we consider it interesting to follow this research line.

This paper explores the application of hybrid evolutionary computation techniques to the field of physical activity recognition, structured as follows: section 2 provides the context of the research topic and introduces some of the related work in the field of physical activity recognition, then section 3 thoroughly details our proposal based on evolutionary computation, with results of a preliminary evaluation following in section 4. Finally, section 5 provides conclusive remarks and future lines of work.

2. STATE OF THE ART

Nowadays, human activity monitoring is a widely extended practice, mostly due to the pervasiveness of sensor technologies that provide a natural interaction with affordable devices in the form of wearable devices.

A significant part of this kind of devices are aimed towards fitness and wellness, in what is known as the "quantified self in healthcare" [15, 44], where they monitor physiological factors such as physical activity, calory consumption, heart rate, sleep quality or corporal posture. Some specific brands and devices have been developed with this purpose in mind, e.g. Fitbit, Jawbone, Nike+ FuelBand, Samsung Galaxy Gear Fit or Misfit, while major manufacturers have also released apps for smartphones and smartwatches.

Besides this, inertial measurement units (IMUs) provide some parameters which, based on the hypothesis by which biomechanics of the human body can be decomposed into segments, are valuable resources for the analysis of corporal kinetics. Moreover, the cost of these devices and their sensors (accelerometers, gyroscopes, magnetometers, GPS...) have significantly decreased over the last years, providing affordable means for both daily usage and extensive research.

In the last decades, the field of human activity recognition (HAR), whose main objective is to identify the actions of one or more subjects by means of computer systems based on observations made on those agents, has experimented an increasing research interest; thus has also unveiled some design, implementation and evaluation challenges.

Automatic activity recognition usually requires dealing with big dimensionality data characterized by high variability and mostly free of constraints, so it turns out to be an attractive research field for the application of artificial intelligence and more specifically machine learning techniques.

Recently, Bulling et al. [7] have proposed a taxonomy to classify HAR depending on different criteria:

- Depending on its execution mode, the system can work offline (recognition is performed after data recording) or online (sensors data are processed in streaming).
- Depending on its generalization ability, the system can either be user-independent or user-specific.
- Depending on the type of the activities to be recognized, they can be periodic, sporadic or static. Most physical activities are inherently periodic.
- Depending on the system model, it can be either stateless (only considers sensors signals) or stateful (it observes a model of the environment).

The activity recognition system proposed in this work can be either user-dependent or user-independent, and its performance is measured in both cases. It aims to detect periodic activities and is stateless. Also, while it is trained offline, it could be used in online mode to perform activity recognition over streaming data.

Moreover, literature often makes a distinction between data obtained from ad-hoc body sensors and data obtained from generic wearable devices or smartphones. The latter case is specially interesting as it has more applications in real-life scenarios while, at the same, time poses new challenges: the device is not always placed the same way, sensors are often not as accurate as when body-worn, different brands and models exist, etc. In the latest years, some works have explored this field, such as those by Abdullah et al. [6]; Su, Tong and Ji [38] or Reiss, Hendeby and Stricker [26]. A survey on online activity recognition using smartphones devices is provided by Shoaib et al. [36].

On the other hand, there is the case of body-worn sensors, which is the focus of this paper, an approach extensively reviewed in the machine learning literature, when an extensive variety of techniques have been applied and evaluated including multilayer perceptron [8], ensemble methods combining SVM, ANN and 1-NN [47], fuzzy finite automata [1], Hough transformation along with random projection trees [46], online multitask learning [39], Naive Bayes and k-NN [14] or C4.5 decision trees along with classifiers [41] also featuring a review of the topic's literature in 2011 and 2012.

More recently, performance of deep learning has been explored [17, 21], as well as of weakly supervised learning, which has proven to perform well [37]. The pros of using these approaches is the lack of need for annotation during the data gathering stages. Finally, besides of classification itself, we have explored feature selection for physical activity recognition using genetic algorithms [5], and Monte Carlo Schemata Search [4], obtaining highly competitive results with decision trees ensembles, i.e. random forests.

As said before, there is a broad field of application of physical activity recognition. To mention a few, Seiter et al. [35] apply these techniques to stroke rehabilitation patients; Altini et al. [3] and Chen et al. [9] apply activity recognition and clustering to estimate energy expenditure; and Alshurafa et al. [2] propose gamification in order to reward physical activity.

Given the difficulty of acquiring physical activity data from body-worn sensors, which may require establishing a protocol and conducting pilots, some authors have relied on and published their own datasets, like Reiss and Stricker [30] where the PAMAP2 dataset is described. This dataset, used in this work, is included in the UCI ML repository and was presented along with a baseline classifier benchmark [30].

Finally, as indicated above, there are few works related to the application of evolutionary computation techniques, most of which use genetic algorithms for optimizing classifier weights within ensembles [10, 13]. However, these works do not focus specifically on physical activity, but rather on activity recognition in homes. Genetic algorithms are also commonly used for feature selection [5, 11, 34].

Regarding hybrid evolutionary computation techniques applied to classification in a medical environment, it is worth mentioning the work done by Tan et al. [40]: they proposed a two-phase hybrid evolutionary technique to extract classification rules used in clinical practice for a better understanding and prevention of undesired medical events. In this work genetic programming is applied to evolving nominal attributes for free structured rules and a genetic algorithm is used to optimize the numeric attributes for getting a short and comprehensive set of classification rules. These candidate rules were then used in a second phase for optimizing the order and number of rules in the evolution for finally building accurate and concise rule sets.

3. PROPOSAL

Human activity recognition involves some fixed stages to be performed, which have received the name of activity recognition chain (ARC) [7], a framework for building and evaluating activity recognition systems, which also involves the early stages of acquiring and processing data. The different steps comprising the ARC are shown in figure 1.

In this section, we will first explain the steps followed during the stages of data acquisition and signal preprocessing. Later, we propose an evolutionary computation based system which would perform the other three stages (segmentation, feature extraction and classification) by itself.

3.1 Data Acquisition

The earliest stage in the ARC involves data acquisition. For this process, we could set up a protocol for recording the activity of a sample of subjects wearing on-body sensors. However, as we specified before, this process is difficult and some researchers have published their physical activity datasets. Using these datasets saves time and enables a comparative evaluation to be performed. In this paper, we have used the PAMAP2 Physical Activity Monitoring dataset, which is publicly available at UCI ML Repository and whose holders, Reiss and Stricker, have published many different papers describing the data and the protocol followed for obtaining it [27, 29, 30, 31, 32].

This dataset provides information about physical activity performed by nine subjects wearing three Colibri wireless inertial measurement units (IMUs) located in the dominant arm's wrist, the chest and the dominant side's ankle, as well as a heart rate monitor. Every record is labeled with one of the next activities: lying quietly, sitting, standing still, ironing, vacuum cleaning, ascending stairs, descending stairs, walking, nordic walking, cycling, running and rope jumping. All activities are carried out according to a predefined protocol [33], which defines both the ordering of the activities and the time to be spent in each exercise.

As indicated above, nine subjects (8 males and 1 female) took part in the data acquisition stage. They were aged 27.22 $\pm$ 3.31 years and had a BMI of 25.11 $\pm$ 2.62 kgm^{-2}, one being left-handed and the remaining right-handed [28].

When it comes to dealing with real-world data, it often happens that noise is captured or some information is lost. In this case, some subjects show a deviation from the protocol due to hardware problems causing mismatches in the timing or leading to loss of information for some activities. Subject 9 is an extreme case of this effect, as his data completely differs from the specified protocol; for this reason we have ignored the data regarding this subject in this paper.

The PAMAP2 dataset comes with 53 dimensions, comprising 17 attributes from each IMU (described in table 1), the heart rate in beats per minute (bpm) and a timestamp in seconds. Besides, each instance contains the label with the corresponding activity. Each IMU has a sampling frequency of 100Hz, while the heart rate (HR) monitor features a frequency of 9Hz. For this reason, the value for this dimension is unavailable in around 91% of the instances.

Table 1: PAMAP2 attributes extracted from IMUs

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

3.2 Signal Preprocessing

As already said, raw signal data are noisy and some information is not correct or unavailable due to hardware problems. For this reason, the signal preprocessing stage becomes so important, as it outputs clean data which is ready to be input in our system. This section describes the different procedures carried out for signal preprocessing.

Firstly, the timestamp was removed, as it is a useless feature for performing activity recognition. Leaving this feature could indeed tweak the classifier's performance, as it could learn about the protocol timings used for signal recording.

Secondly, all the information about orientation and the second accelerometer was removed, as the dataset holders describe it as being invalid [25] and, in the case of the accelerometer, they recommend using only the first one. This results on a deletion of 21 features, 7 per IMU.

Later, missing values for the HR monitor were filled by estimating their real values, which were computed as the previous available value. This is considered a valid approach, as data values do not likely change too much within a tenth of a second. For other sensors, missing values are extremely exceptional and those instances have been removed.

Finally, all instances labeled as *transition*, which correspond to periods after one activity ends and before the next one starts, have been removed.

After preprocessing is completed, the dimensionality is reduced to 31 features, plus the label.

3.3 Classification

In this section we detail the proposal of an evolutionary computation system which would carry out the stages of data segmentation, feature extraction and classification.

The system architecture comprises three modules, as depicted in figure 2. First, clustering is performed by means of Particle Swarm Optimization (PSO), which searches for the optimal centroids in order to increase affinity within the instances of each cluster. Later, a Hybrid Evolutionary Algorithm (HEA) is in charge of the training process by optimizing classifiers where each individual is represented by means of a decision tree.

Finally, the evaluation process takes both the decision trees and the centroids as inputs and provides predictions when new data are introduced to it. This program acts like an ensemble, where the most suitable decision tree is based on the cluster which is assigned to each test instance.

3.3.1 Clustering with PSO

This module performs the first stage of model training, where input data is clustered by means of particle swarms minimizing the Euclidean distance of every instance to the centroids of the cluster where it belongs.

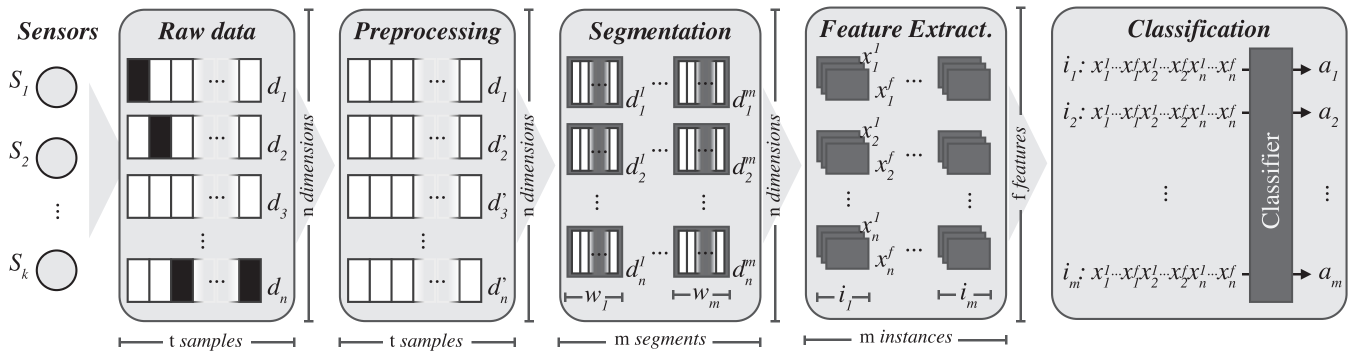

Figure 1: Steps involved in the activity recognition chain (ARC), from data acquisition to classification.

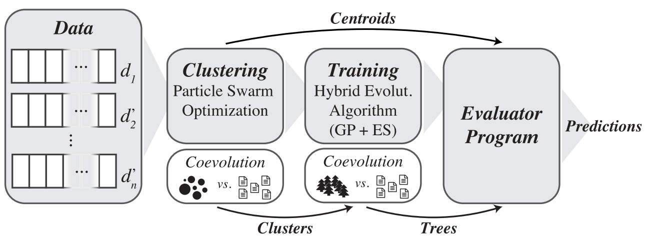

Figure 2: Proposed architecture for data classification using hybrid evolutionary computation.

In PSO, particles are characterized by their location and speed. In our proposal, inspired by works such as [19] and [18], every particle's location represents a cluster configuration (a set of centroids) with each coordinate representing one centroid's dimension. Particles' locations are randomly initialized in the range [0,1] because, as we will see later, input data is normalized before entering the system.

Moreover, we have introduced a co-evolutionary approach, where the swarm compete against some training subsets. The swarm tries to minimize the sum of the distances between each instance and its corresponding centroid, i.e., the K-Means heuristic. Also, this fitness function is used for the training sets, but in this case it must be maximized.

For updating velocity, in this paper we are using the standard formulation for the PSO update rule, but adding one more factor to modify the particles' velocity: the average distance between each particle's centroid and the instances associated to it ($D(t)$ in equation 1, where w_d is a weight). In this equation, the first addend corresponds to inertia, the second to the individual memory, the third to the collective memory and the last one is the newly introduced factor. By doing this, we increase the velocity in those dimensions where clusters are far from their instances.

$$\begin{aligned} v(t+1) = w_i \cdot v(t) + \varphi_1 \cdot r(o,a_1) \cdot (x^{max}(t) - x(t)) + \\ + \varphi_2 \cdot r(o,a_2) \cdot (G(t)(t) - x(t)) + w_d \cdot D(t) \end{aligned} \tag{1}$$

To reduce the risk of premature convergence, the weight w_i is reduced over time in order to reduce the effect of inertia, whereas the weight w_d is increased over time.

3.3.2 Decision Tree Training with HEA

The second module is where the learning phase of decision trees takes part. For this aim, a hybrid evolutionary algorithm is proposed. This hybrid approach comprises genetic programming, well suited for decision trees representation, and evolutionary strategies, very useful to search for the real values at the decision tree nodes. In this part of the system, the hybrid evolutionary algorithm is executed using the instances in each cluster as the training set, building a different decision tree adapted to each cluster.

The trained decision trees perform, in each of their inner nodes, a comparison between an attribute and a real value, resulting in a binary tree where the right branch is chosen if the condition is satisfied, and the left branch is applied otherwise. By definition, leafs are labeled with the classes, in this case physical activities.

Individual initialization is performed randomly by picking an attribute and a real value both from an uniform distribution for inner nodes, and a random activity for leafs. Each inner node, in addition to its own value, also stores metadata indicating a variance value, which are used by the evolutionary strategy for mutation, and which are initialized in all cases to a value of $\frac{1}{3}$.

The genetic operators used are the following ones: selection by tournament, reproduction through trees crossover by exchanging two subtrees of the parents and mutation both of the values of the nodes (attribute in inner nodes or the activity in leafs) and of whole subtrees. Additionally, we have introduced a special stage for mutating the inner nodes real values which is based on evolutionary strategies, where individuals whose fitness differ less than 5% from the fittest individual's fitness are chosen and their real values in inner nodes are mutated following a (1+1)-ES.

In addition, when the decision trees are built, a pruning method is applied to them, in order to get rid of useless rules and branches that may cause noise. The pruning method first removes those branches for which not a single instance from the training set has gone through. Later, if a subtree is found such that all its leafs belong to the same class, it is replaced by a leaf labeled with this class.

A naive fitness function for GP trees could be the number of correct classified instances. However, since all trees are randomly initialized, a competitive co-evolutionary approach with training subsets organized by difficulty is proposed. The co-evolutionary approach is an elegant idea inspired by biological co-evolution of different species. Under this scheme, evaluation function of species which belong to one ecosystem are affected by others' performance, and vice versa. For more information about co-evolution see [16], [22], [24]. In this work, the fitness for each individual is defined as the average training accuracy of the classification over all training subsets, defined as the sum of true positives and true negatives divided by the number of issued predictions, a value which must be maximized. Meanwhile, the fitness for each training set is again the average accuracy of all decision trees over it, and in this case it must be minimized. This co-evolutionary approach is intended to lead to generating training sets which are more difficult to classify and, as a result, also better decision trees.

3.3.3 Evaluator Program

The resulting model must be seen as an ensemble of decision trees, with one tree associated to each cluster. Let be t_c the tree associated to cluster c. For generating a prediction, we have implemented two different modes of operation with this ensemble. The first one allocates the instance to its corresponding cluster and applies the corresponding decision tree, i.e. if instance belongs to cluster c, it will return the output of t_c.

Meanwhile, the second approach incorporates fuzziness instead of computing a membership function for each cluster, this meaning that an instance will belong to all clusters to a greater or lesser extent, instead of only one as in the previous case. Thus, an instance have a higher membership value for those clusters whose centroid is closer. Later, all decision trees will be executed over the instance and their outputs will be weighted to conform the final output of the ensemble. The resulting class k_i is then computed following equation 2, where d_{ic} is the distance between centroid c and instance i, C is the set of all cluster centroids and $\eta(t_c, k)$ is a binary function which outputs 1 if tree t_c returns the class k and 0 otherwise.

$$k_i = \arg\max_k \sum_{c \in C} \left(\frac{1}{d_{ic}} \times \eta(t_c, k) \right) \quad (2)$$

4. EVALUATION

Our proposal has been evaluated by conducting experiments in order to check its performance. We have tried three different execution modes, which simulate diverse scenarios of an activity recognition system working in production:

- Subject-independent: a user has just arrived and we have no information about physical activity performed by him/her. To simulate this behavior, we use leave-one-subject-out (LOSO) cross validation, so that we have 8 different test sets (one per subject) and each one have a training set associated comprising the data from the remaining 7.
- Subject-specific: a model is trained for each user, something which could happen when many data about the user is available beforehand. In this case, we use standard 10-fold cross validation for each subject.
- Combined: there is some previous data from the user but not enough as to build a subject-specific system. Here, we combine the data from the 8 subjects and then perform standard 10-fold cross validation over this combined data.

All values has been first normalized in the interval [0,1], prior to entering into the system. When in the subject-specific case the maximum and minimum values of each subject are considered for normalization, whereas in the subject-independent and the combined approaches the global maximum and minimum for all subjects is used.

Table 2: Parameters used for particle swarm optimization

Intermediate iterations	50
Full iterations	10
Particles	25
Clusters	5
Subsets population size	5
Subset sampling size	0.01
w_i	0.08-0.16
w_d	0.12-0.22
φ_1	0.3
φ_2	0.18

Table 3: Parameters used for genetic programming

Initial tree depth	min: 2, max: 4
Trees population size	200
Tournament size	30
Maximum GP generations	100
Subtree mutation rate	0.01
Node mutation rate	0.02
Maximum ES generations	100
ES sample size fitness range	0.05
Subsets population size	5
Subset sampling size	0.01

4.1 Experimental Setup

This section describes the experimental setup for both the PSO and the genetic programming.

4.1.1 Particle Swarm Optimization Parameters

The parameters for the clustering module using PSO are shown in table 2.

Intermediate iterations are those for which the subsets population remains the same. Once these iterations are completed, the swarm has adapted to the datasets but without too much specialization. Each full iteration comprises 50 intermediate iterations, and the subsets population changes from one to another.

4.1.2 Genetic Programming Parameters

The parameters specified for learning decision trees through genetic programming are shown in table 3.

While some of these parameters are fairly intuitive, others may require further explanation. We have decided to initialize rather small trees, so that they grow by means of crossover and subtree mutation in order to better fit the training sets. The maximum ES generations are run in every GP generation, in order to evolve the real values of the decision trees. ES sample size refers to the percentage of the trees population over which ES is run. More generations for GP and ES will be tested in the future.

4.2 Results

The results in this section are shown separately for each of the execution modes, where the accuracy for each fold is shown as well as the average accuracy. Moreover, for each

Table 4: Accuracy in the subject-independent mode

Subject (fold)	Single	Weighted
1	29.18%	30.56%
2	45.28%	08.55%
3	37.07%	32.62%
4	33.47%	26.15%
5	29.52%	30.13%
6	52.37%	09.40%
7	26.80%	25.89%
8	24.68%	22.73%
Average	**34.80%**	**23.25%**

Table 5: Accuracy in the subject-specific mode

Subject	Single	Weighted
1	97.31%	97.83%
2	95.21%	94.26%
3	98.07%	93.89%
4	94.82%	90.81%
5	91.53%	89.04%
6	95.68%	89.14%
7	92.92%	87.46%
8	92.88%	90.71%
Average	**94.80%**	**91.64%**

Table 6: Accuracy in the combined mode

Fold	Single	Weighted
1	42.40%	41.48%
2	47.08%	45.79%
3	39.37%	37.30%
4	56.04%	55.05%
5	47.38%	46.03%
6	61.32%	48.97%
7	46.76%	45.18%
8	54.27%	51.25%
9	55.72%	51.00%
10	67.30%	63.49%
Average	**51.77%**	**48.55%**

execution modes the results are shown with the two modes of operation of the evaluator program, where *single* refers to the case where only one decision tree from the ensemble is considered, and *weighted* refers to the case where the results from all decision trees are aggregated and weighted according to the instance membership to each cluster.

Table 4 shows the results for the subject-independent case, where LOSO (leave-one-subject-out) cross validation has been used, and so each fold corresponds to the subject of the test set. The performance is low, with accuracies much below the baseline established by Reiss and Stricker (averaging 89,24%) [30] or our previous results (averaging up to 94,64%) [5]. This fact may be due to the lack of an actual segmentation stage. In this stage, both Reiss and Stricker and Baldominos et al. works had the data transformed from the time domain to the frequency domain by applying the DFT after the preprocessing stage. In this case, we have removed this segmentation and transformation, and have substituted it with a clustering phase using PSO. For the cases of the subject-independent execution, this system is not leading to significant results. These experiments are reflecting that the clustering stage is not suitable when data is coming from a new subject, not previously used for training, and it is introduced into the system.

On the other hand, table 5 displays the accuracies for the subject-specific working mode, where the same subject's data is used both for training and testing using standard cross validation. In this table, as each subject is tested with 10-fold CV, we have only included the average result for each subject. In this case, results are significantly better than before, and actually lead to competitive results. At this point, we must say that to the best of our knowledge there are no other works training a subject-specific model using the PAMAP2 dataset. Since this is a working paper, we expect that these results are going to improve in the future, as we will explain later when describing the combined working mode results. Still, it is remarkable that these results are fairly good when considering that no data segmentation or transformation was manually done, and all the process was done instead automatically by a co-evolutionary combination of the PSO and hybrid modules. It needs to be remarked that prediction at this point is made with training data coming from that same specific user, results vary when working with training data incoming from different users (see subject-independent and combined mode).

Finally, table 6 shows the results for the combined working mode, where information from all subjects are mixed and then standard cross validation is performed. It can be observed that results are better than the user independent case, but significantly lower than in the subject-specific case. These results can be compared by those obtained by Reiss and Stricker [30], which can be as high as 99.69%, an accuracy considerably better than the one provided by our system. That is why we said before that subject-specific classification could easily be improved. Again, this time the system is negatively affected by the subject-independent case, which is not able to properly handle.

Also, in all cases we can see that the weighted mode for the ensemble is not providing any advantage over the single working mode. This may happen because, if PSO leads to very different clusters, averaging their outputs may be introducing noise.

In summary, results show that while the system is learning (a naive classifier outputting always the most frequent activity would lead to an accuracy of about 11.68%), performance is rather low when different subjects are considered. In the subject-specific mode, the system attain a significantly better performance, which is competitive; and while it could probably be improved, results are pretty high considering that the whole segmentation–feature extraction–classification stages are automatically performed by the system, with no manual intervention.

5. CONCLUSIONS AND FUTURE WORK

In this working paper we are exploring the problem of physical activity recognition proposing a co-evolutionary system based on PSO plus an hybrid evolutionary algorithm.

First, in this proposal, we have described the activity recognition chain (ARC), a sequence of stages typically followed for performing human activity recognition. The first stage involves data acquisition, and in this case we have used the PAMAP2 dataset, which gathers physical information about 12 different activities collected from 9 subjects (while we have omitted one of them) wearing 3 IMUs and one heart rate monitor; and is available in UCI Machine Learning Repository. Later, we have performed a basic preprocessing which essentially is in charge of cleaning data and handling missing values.

The remaining stages of the ARC; segmentation, feature extraction and classification, are directly managed by the system we have proposed. This system comprises three different modules. The first module clusters the training data using particle swarm optimization (PSO). Meanwhile, the second one evolves a population of decision trees for each cluster using genetic programming combined (GP) with evolutionary strategies (ES) for mutating real values in their inner nodes. In both modules a co-evolutionary approach is applied, so that the models themselves compete with the datasets, and thus potentially leading to classifiers optimized to solve more difficult problems.

The last module is the one in charge of computing predictions over new data. This module uses the information from the clustering (the centroids indeed) in order to decide the working mode of the ensemble: either a single decision tree corresponding to the cluster is chosen, or the outputs from all decision trees are weighted based on the degree of membership of the instance with the cluster.

The work in progress system has been evaluated using three different experimental modes: subject-independent, which uses leave-one-subject-out cross validation; subject-specific, where training and testing are done over the same user with standard cross validation; and combined, where the data from all users are merged and training and testing is performed with standard cross validation.

Results do not compete with the state-of-the-art. In fact, results are rather low for the subject-independent and the combined cases, from which we conclude that the system is negatively affected when data from different users are considered, and this may be because of the lack of a proper segmentation stage, which the PSO algorithm is not really replacing. Anyway, the system shows a learning behavior as it outperforms a simple classifier predicting always the majority class by a factor of 3 in the case of subject-independent and almost 5 in the combined case. Experiments with more than 9 subjects would be ideally to introduce more diversity into the available data and to compare results.

In subject-specific experiments, however, results are fairly good, averaging over 90% and almost 95% when only one classifier from the ensemble was chosen. This is specially significant as no data transformation, segmentation or feature extraction took place, and still the system was able to achieve a high accuracy. This case would be interesting in a production system which has already obtained enough data from the user as to learn a specific model for him/her.

This paper is describing an ongoing research to deal with the problem of physical activity recognition using biologically-inspired techniques. We are presenting these preliminary results in order to discuss them properly and see if it is worth to continue working on this research direction.

Since this research line is quite new, there are many possibilities for further work. An example involves carrying out data transformation (e.g. a DFT) before the hybrid evolutionary computation system performs feature extraction and classification. Another option is that, given the expressive power of decision trees, additional operators are included into the genetic programming scheme allowing more powerful programs by enabling arithmetical operations and aggregation between different attributes. Finally, given the encouraging results of subject-specific evaluation, another option that we are considering as future work is to perform an hyper-optimization of the parameters in order to check how other experimental setups could affect the results.

6. ACKNOWLEDGEMENTS

This project was partially funded by European Union's CIP Programme (ICT-PSP-2012) under grant agreement no. 325146 (SEACW project), and is supported the Spanish Ministry of Education, Culture and Sport through FPU fellowship with identifier FPU13/03917.

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 J. Biomed. Health Inform.*, 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 J. Biomed. Health Inform.*, 19(1):219–226, 2014.

[4] A. Baldominos, P. Isasi, Y. Saez, and B. Manderick. Monte carlo schemata searching for physical activity recognition. In *Proc. INCoS'15*, pages 176–183, 2015.

[5] A. Baldominos, Y. Saez, and P. Isasi. Feature set optimization for physical activity recognition using genetic algorithms. In *Proc. Comp. GECCO'15*, pages 1311–1318, 2015.

[6] 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. *Intl. J Comp. Inf. Engineer*, 6:77–84, 2012.

[7] 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.

[8] C. Catal, S. Tufekci, E. Pirmit, and G. Kocabag. On the use of ensemble of classifiers for accelerometer-based activity recognition. *Appl. Soft Comp.*, 37:1018–1022, 2015.

[9] 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.

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

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

[12] L. Dugdill, D. Crone, and R. Graham. *Physical activity and health promotion: evidence-based approaches to practice*. Oxford: Wiley-Blackwell, 2009.

[13] I. Fatima, M. Fahim, Y.-K. Lee, and S. Lee. A genetic algorithm-based classifier ensemble optimization for activity recognition in smart homes. *KSII Trans. Internet Inf. Syst.*, 7(11):2853–2873, 2013.

[14] P. Gupta and T. Dallas. Feature selection and activity recognition system using a single triaxial accelerometer. *IEEE Trans. Biomed. Eng.*, 61(6):1780–1786, 2014.

[15] T. Hay. The rise of the 'quantified self' in health care. WSJ, 2013. [Online; published Aug 13].

[16] W. Hillis. Co-evolving parasites improve simulated evolution as an optimization procedure. *Physica D: Nonlinear Phenomena*, 42(1):228–234, 1990.

[17] W. Jiang and Z. Yin. Human activity recognition using wearable sensors by deep convolutional neural networks. In *Proc. MM'15*, pages 1307–1310, 2015.

[18] Y. Kao and S. Y. Lee. Combining k-means and particle swarm optimization for dynamic data clustering problems. In *Proc. IEEE ICIS'09*, volume 1, pages 757–761, 2009.

[19] Y. Lu, S. Wang, S. Li, and C. Zhou. Particle swarm optimizer for variable weighting in clustering high-dimensional data. *Mach. Learn.*, 82:43–70, 2009.

[20] A. V. Ng and J. A. Kent-Braun. Quantitation of lower physical activity in persons with multiple sclerosis. *Med. Sci. Sports Exerc.*, 29(4):517–523, 1997.

[21] F. J. Ordoñez and D. Roggen. Deep convolutional and LSTM recurrent neural networks for multimodal wearable activity recognition. *Sensors*, 16(1):115–139, 2016.

[22] F. M. P. Husbands. Simulated co-evolution as the mechanism for emergent planning and scheduling. In R. K. Belew and L. B. Booker, editors, *ICGA*, pages 264–270. Morgan Kaufmann, 1991.

[23] F. Pitta, T. Troosters, V. S. Probst, M. A. Spruit, M. Decramer, and R. Gosselink. Quantifying physical activity in daily life with questionnaires and motion sensors. *COPD Eur. Respir. J.*, 27(5):1040–1055, 2006.

[24] M. A. Potter. A coevolutionary approach to learning sequential decision rules. In *Proceedings of the Sixth International Conference on Genetic Algorithms*, pages 366–372. Morgan Kaufmann, 1995.

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

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

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

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

[29] A. Reiss and D. Stricker. Creating and benchmarking a new dataset for physical activity monitoring. In *Proc. PETRA'12*, pages 108–109, 2012.

[30] A. Reiss and D. Stricker. Introducing a new benchmarked dataset for activity monitoring. In *Proc. ISWC'12*, pages 108–109, 2012.

[31] A. Reiss and D. Stricker. Aerobic activity monitoring: towards a long-term approach. *Universal Access Inf. Soc.*, 13:101–114, 2013.

[32] 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.

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

[34] T. R. D. Saputri, A. M. Khan, and S.-W. Lee. User-independent activity recognition via three-stage GA-based feature selection. *Intl. J Distributed Sensor Netw.*, 2014, 2014.

[35] 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.

[36] 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.

[37] M. Stikic, D. Larlus, S. Ebert, and B. Schiele. Weakly supervised recognition of daily life activities with wearable sensors. *IEEE Trans. Pattern Anal. Mach. Intell.*, 33(12):2521–2537, 2011.

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

[39] X. Sun, H. Kashima, and N. Ueda. Large-scale personalized human activity recognition using online multitask learning. *IEEE Trans. Knowl. Data Eng.*, 25(11):2551–2563, 2013.

[40] K. Tan, Q. Yu, C. Heng, and T. Lee. Evolutionary computing for knowledge discovery in medical diagnosis. *Artif. Intell. Med.*, 27(2):129–154, 2003.

[41] W. Ugulino, D. Cardador, K. Vega, E. Velloso, R. Milidiú, , and H. Fuks. Wearable computing: accelerometers' data classification of body postures and movements. In *Adv. Artif. Intell. - SBIA'12*, pages 52–61. Springer, 2012.

[42] E. M. van Sluijs, A. M. McMinn, and S. J. Griffin. Effectiveness of interventions to promote physical activity in children and adolescents: systematic review of controlled trials. *The BMJ*, 335:703–707, 2007.

[43] L. Vinh, S. Lee, H. Le, H. Ngo, H. Kim, M. Han, and Y. K. Lee. Semi-markov conditional random fields for accelerometer-based activity recognition. *Appl. Intell.*, 35:226–241, 2011.

[44] K. White. The quantified self and corporate wellness. Edelman, 2013. [Online; published Feb 22].

[45] WHO. Physical activity – fact sheet n. 385. www.who.int/mediacentre/factsheets/fs385/en/, 2015. [Online; updated Jan].

[46] G. Yu, J. Yuan, and Z. Liu. Propagative Hough voting for human activity detection and recognition. *IEEE Trans. Circuits Syst. Video Technol.*, 25:87–98, 2015.

[47] 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.