

PEER-REVIEWED JOURNAL ON THE INTERNET



Efficient out-of-home activity recognition b complementing GPS data with semantic inf by Igor da Penha Natal, Luís Correia, Ana Cristina Garcia, and Leandro Fernande



Abstract

Smartphones have become an indispensable human device due to their increasing functionalities and decreasing prices. Their embedded sensors, including global positioning system (GPS), have opened opportunities to support human activity recognition, both indoor (in assisted living, for instance) and outdoor. This paper proposes a minimalist activity recognition model for out-of-home environments based on a smartphone. The only sensor used is GPS, whose data is enriched with semantic knowledge extracted online from the Internet, and with brief user's profile data collected off-line. We conducted an experiment for 20 days with 22 subjects in their day to day life, with identification of 13 selected activities, of which three were performed in movement. Experimental results show that the approach has a high activity recognition performance. This demonstrates that an adequate combination of information with different levels of semantic content can produce an efficient non-invasive solution to monitoring human activity in out-of-home environments.

Contents

- 1. Introduction
- 2. Related work
- 3. The MOut-HAR model
- 4. The HAR-Brazilian dataset
- 5. Experimental evaluation
- 6. Discussion and conclusions

1. Introduction

Automatic human activity recognition (HAR) is an important element in effective environment surveillance and security in general. There is a great deal of effort in automatic recognition of specific activities related to users' movement, such as walking, jogging, and riding on a bus (Difrancesco, *et al.*, 2016; Huang, *et al.*, 2013; Hung, *et al.*, 2014; Mousavi, *et al.*, 2017). The two most used mobile operating systems, Android and iOS, offer application programming interfaces (APIs) (Android; Apple) with movement detection based on global positioning system (GPS) and accelerometers, the latter being especially useful for recognition of indoor activities (Foerster, *et al.*, 1999).

Detecting human activity is desirable but on the other hand, there are costs involved in automatically identifying it. Mostly, the costs are related to carrying, using, spending the device's energy, and processing information from sensors. Using wearable sensors or other gadgets leads to extra expenses for users (Chon and Cha, 2011). Environment sensors, such as cameras, vibration, and temperature sensors, are already handy and available. However, they require more complex information processing, such as machine learning for face and movement recognition.

<u>By Title</u> <u>Other Journals</u> FONT SIZE

<u>By Issue</u>

By Author

Browse







Email this article (Login required)

Email the

author (Login required)

ABOUT THE AUTHORS

Igor da Penha Natal Universidade Federal Fluminense (UFF) Brazil

Luís Correia Universidade de Lisboa

Ana Cristina Garcia Federal University of Rio de Janeiro State (UNIRIO)

Mobile devices, such as smartphones and tablets, usually have a set of built-in sensors that have also been used for HAR.

Mobile devices are equipped with sensors allowing the acquisition of surrounding information, including sound, light, position, and movement. These widely used devices open many possibilities to deal with HAR and with assisted living issues for both indoor and outdoor environments. The sensors' activity log of mobile devices make possible to extract information that can be used in a virtually unlimited amount of applications, such as personal time management, surveillance, danger identification, city planning, as well as supporting health care assistive technology (*e.g.*, Lee, *et al.* [2017]; ben Khalifa, *et al.* [2017]; Mousavi, *et al.* [2017]; Difrancesco, *et al.* [2016]). Since smartphones are already part of our daily lives, using the information they provide does not impose any additional cost for the users. The challenge is to handle battery consumption for collecting and sending sensors' information (Carroll and Heiser, 2010).

For HAR, the most used sensors are accelerometer, gyroscope, and GPS. The accelerometer and gyroscope are effective when it comes to physical activity such as, but not limited to, walking, running, going upstairs, and downstairs (Ignatov, 2018; Cao, et al., 2018; Hassan, et al., 2018; Ogbuabor and La, 2018). Unfortunately, these sensors consume a large amount of energy when used continuously, leading users to abandon applications that rely on this type of usage. Consequently, there is an opportunity for an energetically efficient data collection protocol in mobile devices. The problem is to define the minimum amount of sensors' data that provide enough surrounding information to infer what is going on with the user, identifying short (*e.g.*, in bank), moderate (*e.g.*, having lunch), and long *e.g.*, working) duration activities. Several studies have focused only on GPS sensor data (Difrancesco, et al., 2016; Wan and Lin, 2013; Liao, et al., 2006; Huang, et al., 2010; Boukhechba, et al., 2015), indicating that GPS might be enough for HAR. Nevertheless, previous research still requires data collection in very small time intervals (Zhou, et al., 2016; Shin, et al., 2015; Wan and Lin, 2013; Boukhechba, et al., 2015) or a great deal of contextual information (Difrancesco, et al., 2016; Liao, et al., 2006). Others mention the possibility of considering user profile information (Boukhechba, et al., 2015; Chen and Nugent, 2009; Gomes, et al., 2012), but do not use it.

In such a scenario, the HAR function can benefit from additional sources of information coming from the environment and the Web. The proximity of places already classified in public maps as points of interest (POIs) can be used to define a set of possible activities related to the affordances of the places. For instance, it is expected that the primary affordance of a restaurant is to have lunch/dinner (Huang, *et al.*, 2010; Boukhechba, *et al.*, 2015). Of course, there are other affordances, such as work, that is also common. But it is not expected that people go to pray in a restaurant, although it is always possible, but not probable. Consequently, this type of semantic information may be used to distinguish among possible stationary activities, *i.e.*, those a user might perform in a given place while not moving. Along these lines, user-supplied information on his/her profile can also be used to enrich further the data over which activity recognition is made.

The motivation for our work is to investigate if semantic information may complement scarcity of sensor data in an out-of-home HAR model based on a smartphone in such a way that the HAR system can identify a variety of activities regardless of having a short, moderate, or long duration. Therefore, the approach we explored is that the use of semantic information consisting of POI identification, obtained online, and user's profile information, obtained off-line, can enrich sparse GPS data well enough to infer users' activities with a high accuracy that was only possible with multiple sensors. The energy savings are significant due to using a single sensor and moderate consumption.

Besides the energy issue, since the approach proposed in this paper does not require the use of additional sensors (*e.g.*, accelerometer, gyroscope, etc.), it opens up a potential application for HAR over legacy GPS data if we can complement it with the needed semantic information about the user. Therefore, in terms of further contributions, the high accuracy obtained by our technique with minimal collected sensor data can also be extended to the analysis of GPS data sparingly collected in the past.

To test the feasibility of our proposal, we took as base the publicly available Google Play Services Activity Recognition API (Android), with an adaptation of the work of Trebilcox-Ruiz (2016). This API detects four types of movement, namely, *Walking, Running, Automotive*, and *Cycling*, and also a *Stationary* position. Our goal is to enlarge the set of detected activities by including the detection of several others that are performed while the subject is stationary. To do that, we use the semantic enrichment of POI information obtained from the categories they are associated with Google Places (Google), and user profile collected before the training phase. Training data feeds an automatic inference engine, which produces the identification of the activity from the data stream collected. We compared six standard machine learning techniques, such as Decision Tree, Support Vector Machine, Random Forest, to evaluate the robustness of our approach. *Leandro Fernandes* Universidade Federal Fluminense As in similar work, the model assumes that people have routines. Therefore user profile is useful, and live in an urban environment where labeled POIs are available. In reviewed work, own datasets are the rule. However, they are not publicly available. We collected experimental data from 22 individuals during the period of 20 days in their day to day life, with total freedom for users, from which 13 activities were identified. It is important to emphasize that during the data collection period, there were no dropouts by the volunteers and no complaints about battery consumption. Our dataset is available on GitHub [1]. Results obtained are comparable to state-of-the-art models (Difrancesco, *et al.*, 2016; Huang, *et al.*, 2010; Boukhechba, *et al.*, 2015), for a significant number of activities, without the need to use elaborate algorithm configurations.

The main contributions of this work are:

- >MOut-HAR: A Minimalist Out-of-home HAR model using only the GPS sensor, online POI data, and off-line user profile data, with competitive accuracy results compared to state of the art.
- HAR-Brazilian dataset: A dataset with actual user's labeled activities of daily tasks to allow independent work in HAR.

In the <u>next section</u>, we give an overview related to work in human activity recognition. In <u>Section 3</u> we present the MOut-HAR model. In <u>Section 4</u>, we detail the dataset collected for use in the experiment described in <u>Section 5</u> together with results evaluation and comparison to other work. The article <u>concludes</u> with a discussion of the approach and identification of interesting further lines of work.

2. Related work

HAR has become an active area in the research community, drawing wide attention around the world. Many researchers and institutes have carried out HAR projects. This section presents relevant work related to the key concepts of HAR based on mobile devices. Other works in this section show the methodologies for semantic enrichment of location data.

Lee and Kwan (2018) use a smartphone accelerometer and GPS to classify physical activities the GPS trajectories helped to visualise the errors and improved the accuracy of the activities. They only considered the activities: *Jogging, Walking, Sitting*, and *Standing*. The Random Forest and Gradient Boosting algorithms presented an accuracy of 99.03 percent and 99.22 percent, respectively. GPS and accelerometer data were collected in an one-second window. No energetic efficiency analysis was performed, and only semantic attributes of accelerometer were used.

The work of Shafique and Hato (2016) uses data from smartphone accelerometer and orientation sensors to classify six different travel modes: *Walk, Bicycle, Car, Bus, Train,* and *Subway.* Data was obtained from a population of 50 participants from Kobe, Japan, using Android smartphones during November 2013. A Random Forest classifier was used, and results of 95 percent of accuracy and above were obtained, depending on the data collection frequency. A variable amount of data was collected from different subjects.

Shin, *et al.* (2015) use smartphone accelerometers for movement detection and network location to detect longitude and latitude they were able to recognize traveling modes among *Walk, Car, Train*, and *Bus* using a handcrafted classifier with 82 percent accuracy. The experimental data was composed of 495 samples from 30 users in the city of Zurich. The limitation of this work is the samples used and the results of accuracy.

Wan and Lin (2013) consider as activity what is performed by individuals while not using motorized transportation the latter is considered as a connector between activities. Therefore, it is assumed that a speed change of approximately 10 km/h defines a change point in data segmentation. Predefined time and location criteria can also determine change points. Wan and Lin provided an analysis of the time interval to collect GPS data to identify the most efficient balance between quality and energy consumption. According to their results, intervals from 15 to 75 seconds provide reasonably accurate activity location and less than five percent of activity missing. However, activities were manually identified off-line.

Zheng, *et al.* (2008) aim to infer transportation modes including *Driving, Walking, Taking a bus*, and *Riding a bicycle* from raw GPS logs based on supervised learning they identify a set of features, such as the heading change rate, stop rate, and velocity change rate, which are more robust to traffic conditions than the features used by previous approaches. Those features improved inference accuracy. They also propose a graph-based post-processing method where multiple users' change points are converted into nodes using a density-based clustering algorithm. Edges between different clusters are defined based on user-generated GPS trajectories. Subsequently, the probability distribution of different transportation modes on each edge, as well as the transition probability between consecutive edges, can be summarized from users' GPS logs. They collected data by 65 users over ten months. Their results are 76 percent in terms of Recall and Precision. Van Dijk (2018) aims to classify GPS points as activity points or travel points GPS points were created artificially containing various levels of noise. The attribute information is supplemented with information derived from moving multiple spatial windows over preceding and succeeding points. The Random Forest algorithm presented accuracy results of 99.4 percent with a time window of 30 seconds and 99 percent with a time window of 60 seconds. The work only distinguishes between a static activity or a movement point and uses artificially generated GPS points.

Attempts at integrating semantic information about POIs in activity recognition have been reported. In one of the earliest works in this area, Liao, *et al.* (2006) got POI information from geographical databases and used a relational Markov network to identify significant places for a subject based on his/her pattern of movements. Activities are associated with locations. A set of seven activities was considered: *Work, Sleep, Leisure, Visit, Pick-up, On/Off car,* and *Other.* Data from four subjects was collected during one week, and location identification results achieved over 90 percent of accuracy.

Huang, *et al.* (2010) define the attractiveness of POI (depending on its size and popularity) together with its category to predict user activities from his/her trajectories obtained by GPS six activities are identified: *Dining, Shopping 1, Shopping 2, Entertainment, Public facilities*, and *Others*. Their data collected had 8,089 samples from 10 volunteers. Accuracy ranged from 70 percent to 100 percent, from multi-activities in clusters of POIs to single activity in an isolated POI. The number of samples and accuracy of multi-activities is the work limitations.

In Boukhechba, *et al.* (2015), GPS data obtained from an individual mobile device together with POI information was used to identify *Moves, Stops*, and *Moving* activities, and then to identify the activities performed while stopped and while moving. Only data of one user was collected for accuracy assessment. Results are considered by the authors to be successful in inferring user's activities of long duration, but of limited success in some of the short duration. *Home* was not defined, and the model only recognized *Staying* @ an address, which prompted authors to suggest that the model could learn the home address. They suggested including the user's profile to improve results.

Difrancesco, *et al.* (2016) also use GPS and POI information to identify outdoor activities of schizophrenia patients. Six areas of activity were defined, associated with different POIs: *Employment, Shopping, Sports, Social, Recreational*, and *Others*. Data collected from five patients showed Recall in the order of 77 percent and precision in the order of 95 percent. In their work, the inference was performed by user-defined rules, namely using the correspondence of POI to user activity. The limitations of the work is the dependence on users annotate their activities in a diary, recall results bellow 80 percent, and a specific group (schizophrenic patients) limiting the number of volunteers.

Some works (Furletti, *et al.*, 2013; Chon and Cha, 2011; Hjorth and Gu, 2012; Chon, *et al.*, 2012) proposing methodologies for the semantic enrichment of location data. Furletti, *et al.* (2013) presented a data enrichment methodology derived from GPS, transforming points of latitude and longitude into information of POIs. Chon and Cha (2011) used the context associated with using the smartphone for semantic information. Hjorth and Gu (2012) used the context between place, ambient images, and geographic locations. Chon, *et al.* (2012) presented a framework that combines signals based on location and user trajectories, along with various visual and audio place. Zhai, *et al.* (2019) proposed an approach to identify urban functional regions by capturing the full geographical information.

The integration of user profile information to better identify activities is scarce in the literature. To the best of our knowledge, it was only considered in two studies (Chen and Nugent, 2009; Gomes, *et al.*, 2012). In both cases, the approach is to associate specific activity profiles to each user. In Chen and Nugent (2009), the profile is an instance of the ontology describing activity obtained from user preferences. On the other hand, in Gomes, *et al.* (2012), the profile is a specific joint distribution probability associating sensor data with activity. There is no previous work integrating user profile characteristics that need to be obtained externally. In such a case, the information must be supplied by each user to the activity recognition system.

Safizadeh and Latifi (2014) present a method for bearing fault diagnosis using the fusion of two primary sensors: an accelerometer and a load cell they propose a novel condition-based monitoring (CBM) system consisting of six modules: sensing, signal processing, feature extraction, classification, high-level fusion, and decision-making. Waterfall-based high-level sensor fusion is used to derive information that would not be available on a single sensor. Results demonstrate that the load cell is powerful in detecting the healthy ball bearings from the defective ones, and the accelerometer is useful in detecting the location of the fault.

The reviewed works have limitations in terms of the number and type of recognized activities, presenting moving activities (Lee and Kwan, 2018; Zhou, *et al.*, 2016; Shagfique and Hato, 2016; Shin, *et al.*, 2015), or stationary activities (Huang, *et al.*, 2010; Boukhechba, *et al.*, 2015; Difrancesco, *et al.*, 2016), in a total of four to six activities. The papers

Boukhechba, *et al.* (2015), Chen and Nugent (2009), and Gomes, *et al.* (2012) suggest the user's profile data can be attributes that improve activity recognition. However none of them actually uses that information. Works solely based on GPS sensor (Wan and Lin, 2013; Liao, *et al.*, 2006; Huang, *et al.*, 2010; Boukhechba, *et al.*, 2015; Difrancesco, *et al.*, 2016) used one to 10 volunteers to obtain results. The works of Difrancesco, *et al.* (2016), Shin, *et al.* (2015), Huang, *et al.* (2010), and Zheng, *et al.* (2008) presented limited results, between 70 percent and 82 percent (values of accuracy, recall or precision, depending on the works). The works of Boukhechba, *et al.* (2015), Zhou, *et al.* (2016), Shin, *et al.* (2015), van Dijk (2018) and Wan and Lin (2013) presented an interval time to collect the sensor data below 60 seconds, which entails a high energy consumption. The papers of Shin, *et al.* (2015), Liao, *et al.* (2006), Huang, *et al.* (2010), and Boukhechba, *et al.* (2015) used small datasets, containing less than 10,000 instances. The works of Lee and Kwan (2018), van Dijk (2018), Zheng, *et al.* (2008), and Liao, *et al.* (2006) used more than three semantic attributes.

3. The MOut-HAR model

The MOut-HAR (Minimalist Out-of-home Human Activity Recognition) Model is presented in <u>Figure 1</u>, where blue and yellow boxes stand for processes of activity recognition. In our model two phases are used, on top of Figure 1 is the generation phase, and on the bottom is the execution phase. Processes with the same name are identical in the two phases. The processes of functioning are as follows:





Note: Larger version of Figure 1 available here.

- Sensing: The raw data provided by GPS and the clock of the smartphone is appropriately transformed to provide the required information about the environment.
- User Profiling: The user fills in her profile data off-line.
- Signal Processing: Sensor data is pre-processed, generating the basic attributes for the dataset and feature extraction.
- Feature Extraction: This process is performed to obtain symbolic information from the basic attributes, adding enhanced attributes to the dataset. It aims to minimize the data content while maximizing the information delivered.
- HAR Recognizer Generation: This process receives the attributes in the dataset and applies machine learning techniques to generate a trained model.
- HAR Recognizer: This process receives new sensing and user profile data and performs a classification using the trained model, to recognize an activity.

In our model, data from different sources of information is used, therefore it is important to have a data fusion model. Some of the most widely used

models for sensor data fusion are Joint Directors of Laboratories (JDL) (Llinas and Hall, 1998) and Waterfall (Markin, *et al.*, 1997). The models aim to standardize the processing of raw data in an application that needs to use the information acquired with a certain level of similarity. Waterfall is presented as a simpler model in comparison to JDL, although with similar capacity in the task of data fusion. The Waterfall data fusion model was proposed by Markin, *et al.* (1997). This architecture emphasizes the processing functions on the lower levels, taking the flow of data from the data level to the level of decision making. For this work, an abstraction of the Waterfall model was used to integrate different information sources presented in Figure 2.



Figure 2: Waterfall model work flow. Our system considers a simplified versic comprised of blue-colored boxes. Yellow-colored boxes mean that different app implementation were tested.

Note: Larger version of Figure 2 available here.

Activity is a broad concept that encompasses a variety of meanings and scope. Studying, shopping, having lunch, walking, going by car, or having coffee are activities characterized by different levels of abstraction. As defined in Ranasinghe, *et al.* (2016), activities happen within a time span. In some cases, activities can be performed in an interleaved mode or can be composed hierarchically. Through time, activities follow a sequential or concurrent mode. For the temporal description of activities, a model based on characteristics of Allen's interval algebra (Allen, 1990; Nebel and Bürckert, 1995) was used. This allows us to represent relations between temporal events as exemplified in <u>Figure 3</u>. In this paper, we only consider sequential activities, further described in <u>Section 4</u>.



Figure 3: Examples of composite, sequential, concurrent, and interleaved mode activities.

Note: Larger version of Figure 3 available here.

3.1. Sensing

The sensing process presented in <u>Figure 1</u> is identical in the generation and execution phases. It collects data according to the requirements of MOut-HAR.

Energy consumption is an issue for smartphone users. Unfortunately, the extensive use of multiple sensors decreases battery life. In the work of Wan and Lin (2013), several time intervals for GPS data collection were tested: 5, 15, 30, 45, 60, 75, 90, 105, 120, 135, 150, 165, and 180 seconds. The best compromise between data collection interval and number of activities identified with high accuracy was 75 seconds for 167 recognized activities. However, the activities were identified off-line, and they have a high degree of detail. For instance, Wan and Lin assume that Waiting transport, Getting on the bus, and Getting off the bus are different activities. We, on the other hand, use Waiting transport and Going by bus. Therefore, in our work, the number of activities recognized and the detailing is smaller than in Wan and Lin (2013). So new tests were performed to set the appropriate time interval for our scenario. In order to achieve suitable energy consumption in the sensing module of our model, we carried out an experimental analysis to identify the longest time between consecutive GPS readings with an acceptable data liveliness as location accuracy depends on it. The experiment was dedicated to verifying energy consumption using the application for data collection. No further tools were considered to reduce energetic consumption, only those the system itself makes available by default.

3.2. Signal processing

The signal processing process in $\underline{Figure 1}$ is identical in the generation and execution phases. Its role is to gather the operations to pre-process data, which are detailed next.

Raw GPS and clock-related attributes were discretized into a few categories each. This allowed the use of well-established supervised learning techniques to perform the activity classification task. In order to guide discretization, sampled values were pre-processed using the k-means algorithm with three different configurations for clustering: k = 3, 6, and 9. The results were similar. For the sake of simplicity, k = 3 clusters were used. The attributes obtained from GPS and system clock are described in Table 2. In this table, the collected samples are presented in the first column. The second column shows the discretization range, while the third column shows the assumed discrete value.

Table	1.	Someon'e	onormy	consumption	and	information	lose	at dif	foront	time
rante		Sensor s	energy	consumption	and	mormation	1 1088 i	at un	rerent	unne

Reading Time Interval	Energy Consumption	Missing R
Every 10 minutes	4.04%	0.0%
Every 5 minutes	5.57%	8.39
Every 3 minutes	7.19%	15.09
Every 1 minute	12.32%	30.09

Note: Larger version of Table 1 available here.

Attribute	Range	Disc
Day	Monday-Friday	V
	Weekend & Holiday	V
	05:30-12:14	Ν
Start/End time (hh:mm)	12:15-18:17	A
	18:18-22:32	
	$t \le 18:49$	
Duration $(t \text{ in } mm:ss)$	$18:49 < t \le 39:53$	\mathbf{N}
	t > 39:53	
	d = 0	
Distance from previous point	$0 < d \leq 234$	
(d in meters)	$234 < d \le 453$	\mathbf{N}
	d > 453	
	p = 0	
Pace	0	
(p in km/h)	5.45	\mathbf{N}
	p > 33.22	

Table 2: The discretization of GPS and clock attribut

Note: Larger version of Table 2 available here.

From the successive locations measured by the sensors, the distance the subject moves is continuously computed, as well as the resulting Pace, defined by Distance/Duration. While the Pace does not vary, we accumulate the distance in the current activity. Once a change in Pace is detected, the current activity ends, and a new one starts. Consequently, one activity is immediately succeeded by another one with nothing in between (similarly to Fox, et al. [1995]). Therefore, the End time of activity $A^{(t)}$ coincides with the *Start time* of activity $A^{(t+1)}$.

3.3. Feature extraction and user profiling

The feature extraction and user profiling processes presented in Figure 1 are also identical to those in the generation and execution phases. Their roles are detailed in this section.

GPS sensor data provides a sequence of locations. Pace was derived from the distance between GPS locations and time information, providing hints for segmenting the flow, and determining the activity window. From the location, we retrieve the POI. We associate the POI with locations visited for a significant time lapse, *i.e.*, more than two consecutive GPS points collected in the same place. Each POI is of a type and intrinsically offers a set of activities affordances. The type of POI is directly retrieved from Google Places [2]. Figure 4 illustrates mapped POI information obtained by our mobile app (the app is described in Section 4). The affordances of a POI considers the activities that the POI can be expected to carry (Kim and Martinson, 2016). It represents the semantic information about the place, which is a key attribute in determining the activity there performed (Liao, et al., 2006; Difrancesco, et al., 2016). In the current work, each place is identified by a name, latitude, longitude, and category to which it belongs.

For each category of places, characteristics were observed in order to establish the associated local affordances. These relationships were defined according to Furletti, et al. (2013) previous work and taking into account the experience of authors and volunteers. However, a direct relationship between the user's location and POI is not observed in all cases. To illustrate this lack of relation, we analyzed the activities registered in each POI in the experimental data collected, presented in Table 3. For example, the Shopping activity was registered with the following POI categories: Financial, Food, and Shopping. See all relationships in Table 4. In summary, affordances reduce the set of probable activities that can be performed in a certain location. Instead of retrieving affordances from a predefined list, we claim it should be learned from actual observation, as we did with our collected dataset.

The extraction of GPS features (Feature Extraction in Figure 1) retrieves values for the online semantic attribute POI category. The selection of POIs was based on the algorithms presented in Furletti, et al. (2013), in which the semantic enrichment of the trajectories is done first by checking whether there are POIs linked to the given GPS coordinates. We retrieve the POI type as the one having the highest-ranking among the set of types assigned by Google Places to the closest open POI in a 200-meter radius.



Figure 4: Examples of POIs located in a map. This image was obtained from Gc Point of Interest Auto Map [3].

Note: Larger version of Figure 4 available here.

Table 3: Relation between POI type according to Google Places and PO.	l categ
POI (information taken from Google Places)	
Library, School, University	
Aquarium, Movie_theater, Museum, Night_club, Park, Stadium	
ATM, Bank, Finance	
Bakery, Bar, Cafe, Food, Restaurant	
Lawyer_office	
Church, Place_of_worship	
Establishment	
Book_store, Clothing_store, Convenience_store, Department_store, Electronics_st	ore,
Furniture_store, Grocery_or_supermarket, Hardware_store, Home_goods_store,	
Jewelry_store, Pet_store, Pharmacy, Shoe_store, Shopping_mall	
Airport, Bus_station, Parking, Taxi_stand	
Beauty_salon, Car_dealer, Car_rental, Car_repair, Car_wash, Courthouse, Dentist, I	Doctor
Fire_station, Gas_station, General_contractor, Gym, Hair_care, Health, Hospital, La	aundry
Local government office, Lodging, Null, Pharmacy, Police, Post office, Veterinary c	are

Note: Larger version of Table 3 available here.

In (Amato and Straccia, 1999), user profile attributes are collected to enrich the dataset and to refine the results of suggested books in a digital library. In Yu, et al. (2019), sensor data obtained from mobile devices is used to infer user profile, namely Age, Gender, and Personality traits. Both these works use only those attributes that have any significant effect on improving outcomes. Taking this into account, for this work, we conducted a pilot study with the GPS attributes and the respective timestamp. From this study, it was possible to note that most of the wrong classifications were between two pairs of activities: Going by car versus Going by bus, and *Working* versus *Studying*. User profile data collected for the pilot study consisted of: *Age, Gender, Profession, Address,* and if the user *Has* a car. However, Age, Gender, and Address are information that has not made a significant difference in terms of accuracy in the pilot study, which combined the possibility user direct identification led to discarding it. Retaining the profession information helped to improve the refinement between working and other activities (e.g., In bank and Studying). And retaining information about the user having a car was important to improve the accuracy of distinguishing Going by car versus Going by bus. Therefore, for the proposed model, user profiling is an off-line manual process where the user introduces little personal data. This is done only once when the user first runs the app, and the data is reduced to Profession (textual input) and Has car (binary choice).

As will be demonstrated in <u>Section 5</u>, for many tasks, semantic attributes bring meaningful information that efficiently complements the basic attributes. The values of semantic attributes assumed in our experiments are indicated in <u>Table 5</u>.

	Table 4:	Relation	s betweer	1 POI	categor	ries an	d activ	vities o	observed	in our	da
	Dinning	Going by bus	Going by car	In bank	Having lunch	Praying	Recreation	Shopping	Studying	Taking coffee	
Education					33				943		1
Entertain.							954				
Financial	2	-	-	835	51		_	19	-		4
Food	1,002			68	1,692	59		25		800	1
Law											
Religion	360	-				901			-		
Services	_	-	-		_	_			-	-	
Shopping	-	-	-	-	_	-	-	1,585	-	_	
Transport	-	-	124		-			-	-		7
Others	575	6,774	13,986	-	3,343	26			2,161	352	

Note: Larger version of Table 4 available here.

Ta <u>ble 5: Values of</u>	the semantic att
Attribute	Value
	Education
	Entertainmen
	Financial
	Food
POI category	Law
	Religion
	Services
	Shopping
	Transport
	Others
	Academic
	Admin. technic
Profession	Bank clerk
	Businessman
	Lawyer
	Professor
Has car?	Yes
	No

Nome: Jo	hn Snow		Nome: John	Snow
Atividade Atividade Latitude: Longitude	Almoçando Comprando Deslocamento a pé Deslocamento de ónibus Em casa Esperando transporte Estudando Igreja Jantando No Banco Recreação Tomando café Trabalhando	•	Atividade De Atividade Re Latitude: Longitude:	slocamento a pé gistrada: Deslocar 65.9667 -18.5333 SALVAR
	CONTATOS			CONTATOS

Figure 5: The interface of the app developed to collect annotated data in o

Note: Larger version of Figure 5 available here.

3.4. HAR recognizer generation & HAR recognizer

HAR recognizer generation and HAR recognizer presented in Figure 1 are similar in the generation and execution phases. Techniques and metrics are described in this section.

HAR recognizer generation and HAR recognizer processes of our model are combined in our implementation. In fact, both are realized together by a machine learning process that learns to identify activities from the features previously obtained. Six algorithms were tested to extract data knowledge and then to classify the activities: J48, Multilayer Perceptrons (MLP), Support Vector Machines (SVM), Random Forests (RF), K-Nearest Neighbours (KNN), and Naive Bayes (NB). The algorithms were trained and evaluated in terms of accuracy, precision, recall, and F-measure (see <u>Section 5</u> for details). We consider that the learning process forms the HAR recognizer generation process and that the usage of the trained model provides the HAR recognizer process, performing the activity recognition.

4. The HAR-Brazilian dataset

We have collected real out-of-home human activity data in order to build a dataset for our research. After submitting our project to the university ethics committee, we sent invitations to university students and workers to participate in our experiment. Each volunteer signed an agreement and installed a mobile app developed for Android smartphones by us. The Android platform was chosen based on the profile analysis of the volunteers of a previous study. The mobile app collects raw GPS data (*Longitude* and *Latitude*) and system date and time every three minutes (see Section 3.1 for details). The user interface (written in Portuguese) allows the volunteer to annotate when he/she starts a new activity. As illustrated by Figure 5 (left), the user may associate an optional user name to his/her records (*Nome* field) and has to select the started activity using a combo box (*Atividade* field). Feedback is given to the user indicating the last recorded activity and its GPS coordinates (Figure 5, right). The *Salvar* button saves the information, while the *Contato* button allows the communication between the user and the app developers.

We collected data from 22 subjects for 20 days, during March and April 2018. Subjects were from both genders, aged from 18 to 56, and lived either in Lavras, Ituiutaba, or Uberlândia, both cities in the state of Minas Gerais, Brazil. The volunteers completed an initial questionnaire that allowed us to create the initial individual user profiles. This questionnaire includes only four questions regarding age, gender, profession, and car

ownership. $\underline{\text{Table 6}}$ summarizes the distribution of subjects' characteristics in our experiment.

Profession	Count	Male	Female	 Has car
Academic	11	6	5	4
Admin. technician	2	-	2	2
Bank clerk	1	1	-	-
Businessman	1	1	-	1
Lawyer	2	1	1	2
Professor	5	2	3	4
Total	22	11	11	13

Table 6: User profile distribution	Table 6:	User	profile	distri	bution
------------------------------------	----------	------	---------	--------	--------

Note: Larger version of Table 6 available here.

	Table 7: Activities recognized in the case study.
Type of activity	Activity
Moving	Going by bus - Going by car - Walking
Stationary	Dinning - Having lunch - In bank - Praying - Recreation - Sho
	Taking coffee - Waiting transport - Working

Note: Larger version of Table 7 available here.

Activity	Count	Shorter	Longer	Average
Dinning	1,939	12	78	28.4
Going by bus	6,774	12	81	29.7
Going by car	14,110	9	66	22.7
Having lunch	5,119	12	51	24.3
In bank	903	6	33	12.8
Praying	986	21	66	45.3
Recreation	954	18	90	60.8
Shopping	1,629	15	87	38.4
Studying	3,104	27	126	59.7
Taking coffee	1,152	9	33	15.4
Waiting transport	804	3	30	12.7
Walking	1,378	6	66	18.9
Working	8,456	21	180	66.2
Total	47,308	-	-	-

Table 8: Number of occurrences and duration (in minutes) of each activity

Note: Larger version of Table 8 available here.

Taking into account the population of volunteers, we configured the experiment to recognize 13 activities. Only out-of-home activities were used. They were divided into two categories: stationary and moving activities, as listed in Table 7. The set of activities and their distribution in terms of the duration (in minutes), average time, standard deviation (SD), and the number of occurrences were analyzed to grasp the global distribution of activities better. Statistics are presented in Table 8.

5. Experimental evaluation

We have evaluated the performance of our slim HAR approach by analyzing its results under six different classification algorithms. More specifically, we have used the Weka [4] implementation of the following techniques as *HAR Recognizer Generation* and *HAR Recognizer* steps of the our model (yellow boxes in Figure 1): J48, MLP, SVM, RF, KNN, and NB.

Attribute	Type	Sample 1	Sample 2	Sample
Day		Weekend	Weekday	Weekday
Start time	Basic	Night	Morning	Afternoo
End time	attributes	Night	Afternoon	Afternoo
Total duration		Long	Long	Low
Distance from previous point	Enhanced	Null	Null	High
Pace	GPS	Null	Null	Medium
POI category	attributes	Food	Lawyer	Others
Profession	Profile	Businessman	Lawyer	Bank cler
Has car?	attributes	Yes	Yes	No
Activity	Annotation	Dinning	Working	Going by h

Note: Larger version of Table 9 available here.

We have performed three tests considering different views of the dataset presented in Section 4:

- Test 1 The dataset view includes only basic attributes (*i.e.*, Day, Start time, End time, and Total duration) and user profile attributes (*i.e.*, Profession and Has car?).
- Test 2 The dataset view includes only basic attributes and enhanced GPS attributes (*i.e.*, *Distance from previous point*, *Pace*, and *POI category*).
- *Test 3* The dataset includes all attributes, *i.e.*, basic attributes, enhanced GPS attributes, and user profile attributes.

Tests 1, 2, and 3 aim to verify the influence of enhanced GPS data and user profile on the task of activity recognition. They have been performed using leave-one-out *k*-fold cross-validation, with k = 10. An additional test was performed to assert the resilience of trained classification models to changes on the characteristics of the population:

 Test 4 — The dataset presented in <u>Section 4</u> is used for training the classification models while the dataset presented in our previous work (da Penha Natal, et al., 2017) is used for testing.

Test 4 is feasible because the new dataset was designed with the same attributes as the previous one. However, the user profile attribute *Profession* extends the number of cases from three to six. Some decrease in performance could be expected due to this difference.

Table 9 illustrates four data samples collected from subjects in our experiments. Here, columns one and two indicate, respectively, the name of the attribute and its type regarding the source of information. Notice that the attribute in the last line of Table 9 (*Activity*) is assumed to be known only during the training of classification models. Columns three to six correspond to samples collected from the subjects.

The metrics used for evaluation of the results are: accuracy, precision, recall, and F-measure. These values can be computed from the confusion matrix M, where element M_{ij} indicates the number of instances of class (*Activity*) A_i that were classified as class (*Activity*) A_j . More specifically, the above-mentioned metrics are defined as:

$$\begin{aligned} \text{Accuracy} &= \frac{TP + TN}{TP + TN + FP + FN},\\ \text{Precision} &= \frac{TP}{TP + FP},\\ \text{Recall} &= \frac{TP}{TP + FN}, \quad \text{and}\\ \text{F-measure} &= 2 \cdot \frac{\text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}}, \end{aligned}$$

where *TP* (True Positives) is the number of instances of an activity A_i correctly classified as A_i , *TN* (True Negatives) is the number of instances of activities other than A_i that are not classified as A_i , *FP* (False Positives) is the number of instances classified as A_i that are not activity A_i , and *FN* (False Negatives) is the number of instances of activity A_i that are not classified as A_i .

5.1. Results of Tests 1, 2, and 3

The results of Test 1 are presented in <u>Table 10</u>. They correspond to a scenario where semantic GPS attributes are not considered. However, it is important to note that this scenario makes indirect use of conventional GPS coordinates since *Start/End time* and *Duration* are estimated as function of the displacement of the subject in space (see <u>Section 3.1</u>). Therefore, one can associate the performance of the HAR system using this limited kind of data to the performance of conventional attempts that would be efficient in terms of energy consumption (*i.e.*, they are still performing sparse readings from one sensor), but weak in terms of quality of activity recognition. In this scenario, J48, MLP, RF, and KNN present the best results in all metrics evaluated, but all results are under 70 percent in terms of accuracy. This is far from the results, with around 90 percent of accuracy presented by more sophisticated approaches using similar inputs (see <u>Section 2</u>).

The scenario of Test 2 explicitly incorporates information about the displacement of the subject (*Distance from the previous point* and *Pace*) and the POI category but does not include data from the user's profile. As can be observed in Table 11, the contextualization provided by the enhanced GPS information leads to improvements ranging from 19.23 percent to 34.62 percent in accuracy, 15.62 percent to 31.18 percent in precision, 19.23 percent to 34.62 percent in recall, and 16.24 percent to 34.12 percent in F-measure. Again, J48, MLP, RF, and KNN present the best results in all evaluated metrics. Even so, the maximum of 80.00 percent of accuracy is still well below what is presented in the recent literature.

From a careful inspection of results achieved in Tests 1 and 2, it is possible to conclude that user profile attributes and enhanced GPS data alone may be sufficient to assist in determining some activities. For instance, the confusion matrices produced for these tests show that *Lunching, Working, Going by car*, and *Going by bus* were well classified in Test 1, while *Taking coffee, Recreation, Walking*, and *Shopping* presented good classification results in Test 2. However, as can be seen in Tables <u>10</u> and <u>11</u>, user profile attributes and enhanced GPS data by themselves are not very effective in general.

In Test 3 we used all the collected information together, leading to results that exceed 91 percent of accuracy, precision, recall, and F-measure with J48, MLP, SVM, RF, and KNN. This shows that the fusion of the presented attributes in our slim energetic-saving HAR approach is as accurate as of the more sophisticated solutions presented in the literature for the identification of out-of-home activities. Table 12 summarizes the results achieved in Test 3. A direct comparison of results in Table 12 to those of Table 10 and Table 11 shows improvements in all metrics, accuracy, precision, recall and F-measure in the order of 40 percent. This improvement illustrates how important data enrichment is, even if only with a few semantic attributes. It enables simple yet effective off-the-shelf classification algorithms to reach state-of-the-art performance in human activity recognition.

Algorithm	Accuracy	Precision	Recall	F-n
J48	0.671	0.685	0.671	(
MLP	0.660	0.664	0.660	(
SVM	0.640	0.630	0.640	(
\mathbf{RF}	0.670	0.684	0.670	(
KNN	0.670	0.684	0.670	(
NB	0.569	0.574	0.569	(

Table 10: Results using only basic attributes and profile attri

Note: Larger version of Table 10 available here.

Algorithm	Accuracy	Precision	Recall	F-mea
J48	0.800	0.792	0.800	0.78
MLP	0.794	0.784	0.794	0.77
SVM	0.781	0.760	0.781	0.72
\mathbf{RF}	0.800	0.791	0.800	0.78
KNN	0.800	0.792	0.800	0.78
NB	0.766	0.753	0.766	0.73

Table 11: Results using only basic attributes and enhanced GPS attr

Note: Larger version of Table 11 available here.

Algorithm	Accuracy	Precision	\mathbf{Recall}	\mathbf{F} -
J48	0.924	0.924	0.924	
MLP	0.921	0.921	0.921	
SVM	0.918	0.919	0.918	
\mathbf{RF}	0.924	0.924	0.924	
KNN	0.924	0.924	0.924	
NB	0.844	0.845	0.844	

Table 12: Results using all attributes (Test 3)

Note: Larger version of Table 12 available here.

5.2. Results of Test 4

The dataset presented in da Penha Natal, *et al.* (2017) includes 1,998 samples collected from 10 subjects during 10 days of November 2016. Subjects are from both genders, aged from 25 to 40, and lived either in Niterói or Campos dos Goytacazes, both cities in the state of Rio de Janeiro, Brazil. They are students, teachers, or administrative technicians, a subset of the professions in the new dataset. It is important to emphasize that subjects are from different universities and cities having different geographic and demographic characteristics than Lavras, Ituiutaba, and Uberlândia. Consequently, some differences between the two populations can be expected.

The results of Test 4 are presented in <u>Table 13</u>. As in previous scenarios, the algorithms J48, MLP, SVM, RF, and KNN have similar behavior for this test, with accuracies above 97 percent. The efficiency of the training dataset is demonstrated through recall and precision, ranging from 91.2 percent to 97.5 percent. We conclude that the contextualization provided by enhanced GPS data and user profile makes the trained classification models resilient to changes in the population.

Figure 6 illustrates the accuracy results of Tests 1, 2, 3, and 4 side-byside. We observe that data fusion, in this case, brought better results than using only a single data source. Also, this shows that for future applications developers will not depend on a single classification method since the tested ones present similar behavior. The results in Table 13 (above 97 percent) indicate that the enriched data proposed in our approach can generate robust models. We were able to obtain state-ofthe-art results even when we apply a classification model obtained by training with different people from different cities.

Table 13: Results using the new dataset for training and the dataset from da Penha Na testing (Test 4).

Algorithm	Accuracy	Precision	Recall	F-measure
J48	0.975	0.975	0.975	0.974
MLP	0.975	0.975	0.975	0.974
SVM	0.974	0.974	0.974	0.973
RF	0.975	0.975	0.975	0.974
KNN	0.975	0.975	0.975	0.974
NB	0.912	0.916	0.912	0.911

Note: Larger version of Table 13 available here.



Figure 6: Accuracy (in percent) of the proposed HAR approach using differer algorithms in each one of the four tests discussed in <u>Section 5</u>

Note: Larger version of Figure 6 available here.

6. Discussion and conclusions

We have presented the MOut-HAR approach whose key insight is that the performance of activity recognition based only on GPS data may be improved by enriching positioning information with semantic information of the closest POI, and with information collected from the user's profile. In our approach, data fusion is base on a simplified Waterfall model, and off-the-shelf classification algorithms perform activity recognition.

Our experiment considers a dataset comprised of GPS samples collected from 22 different subjects, from three different cities, during 20 days, performing their usual daily tasks. The tasks include 10 stationary and three moving activities. Semantic enrichment of GPS samples was performed by identifying the closest POI in Google Places and by automatic annotation of its associated affordance. The datasets produced for this work is more complete and has more variability than others described in the literature.

Results confirm our expectations by showing that the performance of classification methods like J48, MLP, SVM, RF, and KNN is equivalent to the performance of more sophisticated state-of-the-art algorithms if both semantic information of POIs and user's profile are fused to sparse GPS samples. From the user's perspective, an advantageous consequence of the non-continuous use of a single smartphone sensor (GPS only, collected every three minutes) is the battery saving. Strong evidence in this respect is the fact that from the users initially enrolled for data collection there was not a single complaint about energy consumption and all of the completed the experiment.

attributes, NA: Number of	activiti	es, and I	Results: A	=accuracy	, R:
Related Work	\mathbf{NV}	\mathbf{ED}	\mathbf{TI}	NI	
Lee and Kwan (2018)	36	-	0.2	$\sim 10^6$	A
Zhou et al. (2016)	12	6 d	20	9,981	Ac
Shin et al. (2015)	30	-	1	495	Ac
Wan and Lin (2013)	1	$121 \mathrm{d}$	15	366,291	
Zheng et al. (2008)	65	10 mo		-	
Van Dijk (2018)	-	-	30-60	345,000	
Liao et al. (2006)	4	-	60	10,000	
Huang et al. (2010)	10	-	-	8,089	
Boukhechba et al. (2015)	1	-	5 - 10	1,368	
Difrancesco et al. (2016)	5	5 d	(10 m)	-	
MOut-HAR	22	$20 \mathrm{d}$	180	47,308	

Table 14: Summary of related work. NV: Number of voluntee
mo=months), TI: Time interval between data collection points (in
every 10 meters), NI: Number of instances, SI: Sensors used to colle
attributes, NA: Number of activities, and Results: A=accuracy, R:

Note: Larger version of Table 14 availa

In our previous work da Penha Natal, et al. (2017), the integration of POI and user profile was presented preliminarily. This paper builds on that introducing: (i) a slim model for out-of-home HAR (see Figure 1) with low energy consumption, whose hallmark is to compensate a low sensor usage with semantic information on the user and on the environment, to obtain a high performance in recognizing more than a dozen activities; (ii) an evaluation of the relative contribution of the different types of attributes (basic, enhanced GPS, and user profile) for the accuracy; and (iii) a more in-depth performance evaluation with accuracy, precision, recall, and Fmeasure, showing the competitiveness of the approach.

6.1. Comparison to related work

To better situate our contribution in the frame of related work, we assess the different approaches according to seven properties: scope, the number of recognizable activities, the number of subjects, used sensors, data acquisition schedule, the level of control on the activities, and accuracy. The main related works are summarized in Table 14 to guide our discussion.

Scope reflects the applicability range of the proposed solution. There are [2014]) or outdoor (*e.g.*, Difrancesco, *et al.* [2016]; Huang, *et al.* [2013]) activities, while others encompass both (e.g., Han, et al. [2014]). Our research is oriented towards out-of-home activities, encompassing both indoor (Shopping or Praying, for example) and outdoor (Walking or Waiting for transport, for example) activities that can be associated to the POI affordances, and users' Profession and Have car? information.

The number of recognizable activities considers the power of the recall. There are solutions for recognizing few (e.g., Hung, et al. [2014]), many (e.g., Difrancesco, et al. [2016]; Huang, et al. [2013]), or even to cover any kind of activity (e.g., Liao, et al. [2006]; Hung, et al. [2014]), generally in a very abstract manner. In terms of the context, we classify the activities in scripts and open context. The former has users performing a script of predefined activities for a short time (e.g., Noo, et al. [2016]; Riboni and Bettini [2011]), while in the latter users do their daily activities during a long period (e.g., Liao, et al. [2006]). The datasets used in our work were produced in an open context with a set of 13 activities, although this number is not a restriction. See column NA in the Table 14.

The number of subjects considered in the experiment is important to reflect the robustness of the activity identification model to process heterogeneous user data. Each subject provides an independent data source that is collected and used for activity recognition. There are models that were generated with observation from a single user (*e.g.*, Mousavi, *et* al. [2017]), few users (up to five, e.g., Liao, et al. [2006]; Difrancesco, et al. [2016]), or from a large groups of users (more than ten, e.g., Riboni and Bettini [2011]). The greater the number of observed subjects, the more reliable is the model, and the better are the chances to reproduce and scale the solution up. In our research, we used data collected from 22 subjects during the daytime over 20 days. See column NV in the Table 14.

Existing work also differs regarding the used sensors. Indoor activities are supported not only by a mobile device but also by a variety of sensors embedded in the environment, such as video capture (e.g., Mitchell, et al.

[2014]; Zhu, *et al.* [2010]) and RFID (*e.g.*, Park and Kautz [2008]). Outdoor activities rely on a variety of mobile device sensors such as accelerometers, GPS, luminosity, sound, etc. In some cases, wearable sensors may also be used. Mostly, researchers use both GPS and accelerometers (*e.g.*, Karantonis, *et al.* [2006]; Preece, *et al.* [2009]). Our proposal considers only GPS information. Sourcing information on the user's profile obtained off-line, and environmental conditions and POI's category obtained online increases the capability of activity recognition resolving many information ambiguities. See columns SI and SA in the <u>Table 14</u>.

Data acquisition schedule has a direct impact on the use of energy. Often data acquisition from sensors is performed continuously (*e.g.*, Krishnan and Cook [2014]) without accounting for the energy consumption required. We take energy consumption seriously by acquiring only GPS data every three minutes. This acquisition rate was designed after testing the trade-off between data liveliness and location accuracy on the one hand and GPS energy consumption and data loss on the other hand. See column TI in the Table 14.

The level of control of the activities performed by the subjects may limit the applicability of a technique to situations for which it was not originally conceived. For instance, some approaches use data obtained from users performing a pre-defined script of activities (*e.g.*, Riboni and Bettini [2011]; Noor, *et al.* [2016]), while others collect data in an ethnographic approach (*e.g.*, Mousavi, *et al.* [2017]), where subjects perform their usual day to day activities. The former approach imposes some limitations that are not observed on the latter. Our work uses the ethnographic approach, and it was tested with a dataset obtained from subjects of three different cities.

Accuracy reflects the correctness of the approach in recognizing the set of planned activities and how it scales to other, unplanned activities. Our tests showed results with a high accuracy generally improving the top state-of-the-art approaches (see <u>Section 2</u>) in demanding conditions as indicated in the *user activity control* and with values obtained in the classification of an independent test set. See column Result in the <u>Table 14</u>.

6.2. Future work

A possible direction of future work is the addition of user profile data with lower latency and the detection of seasonal activities (for instance, winter and summer activities). The detection capability of the proposed model may benefit semi-supervised approaches such as Levatić, *et al.* (2017) by automatically associating activity-oriented categories to unknown POIs.

We are currently exploring the integration of our model with nondeterministic representations of sequences of activities, like Hidden Markov Models, for instance. Further research will explore the integration of the *decision making* module, in Level 3 of the Waterfall model (see <u>Figure 2</u>), by creating applications that trigger actions. They can be particularly useful for assisted living, warning the user or a user support team.

As previously mentioned, this approach opens up the possibility of using legacy data to do HAR, provided the needed semantic information can be associated with that data. On the other hand, this work also should make us reflect on its misuses, such as population surveillance and control. It is an area where the need for public policies to establish strict regulations on the design and auditing is highly stringent.

About the authors

Igor Natal is a Ph.D. student in computing at Universidade Federal Fluminense (UFF). He holds a Master's in computer science at Universidade Federal do Pará. His areas of interest include human activity recognition, machine learning, programming and artificial intelligence. E-mail: igorpnatal [at] gmail [dot] com

Lus Correia is an associate professor at Universidade de Lisboa. He obtained Ph.D. in informatics (behaviour based mobile robots) from Universidade Nova de Lisboa. Currently he coordinates the Agents and Systems Modelling (MAS) at the Biosystems and Integrative Sciences Institute (BioISI) His research interests are in the area of artificial life, autonomous robots and self-organisation in multi-agent systems. E-mail: luis [dot] correia [at] ciencias [dot] ulisboa [dot] pt

Ana Cristina Bicharra Garcia is professor of information systems in the Department of Applied Informatics of the Federal University of Rio de Janeiro State (UNIRIO). She completed postdoctoral fellowships in 2002 at Stanford University and in 2013 at the MIT Sloan School of Management. She holds a Ph.D. in computer-aided civil engineering and a Master's in computer-aided civil engineering from Stanford University as well as a Bachelor's in civil engineering from the Federal University of Rio de Janeiro. Her areas of work and interest include artificial intelligence, collective intelligence, environmental intelligence, and human-computer interaction.

E-mail: cristina [dot] bicharra [at] uniriotec [dot] br

Leandro A.F. Fernandes is an qdjunct professor at the Instituto de Computaç ão at the Universidade Federal Fluminense (IC-UFF). He holds Ph.D. in computer science at Universidade Federal do Rio Grande do Sul (UFRGS). He completed postdoctoral fellowships in 2011 at UFRGS. His research interests are in the area of computer vision and image processing, in particular, he is interested in the subjects: extraction of invariant visual characteristics, detection of structures in multidimensional data, image-based metrology and geometric algebra (Clifford algebra in physics) and its applications in visual computation. E-mail: laffernandes [at] ic [dot] uff [dot] br

Acknowledgements

Thanks to CNPq-Brazil, FAPERJ, FCT and CAPES/COFECUB. This research was partially sponsored by CNPq-Brazil (grants 303.503/2015-7 and 311.037/2017-8) and FAPERJ (grant E-26/202.718/2018). L. Correia acknowledges support of FCT, Portugal (grant UID/Multi/04046/2013). I. Natal's scholarship was supported by the International Cooperation Program CAPES/ COFECUB at the Universidade de Lisboa. We thank Mel Todd for revising the English.

Notes

1. https://github.com/Prograf-UFF/HAR-BrazilianDataset.

<u>https://developers.google.com/places/?hl=pt-br</u>.

<u>3.</u> Image source: <u>http://www.webdesigndev.com/wp-content/uploads/2013/02/7-The-Wall1.jpg</u>.

4. http://www.cs.waikato.ac.nz/ml/weka/.

References

J.F. Allen, 1990. "Maintaining knowledge about temporal intervals," In: D.S. Weld and J. de Kleer (editors). *Readings in qualitative reasoning about physical systems*. San Mateo, Calif.: Morgan Kaufmann, pp. 361–372.

doi: <u>https://doi.org/10.1016/B978-1-4832-1447-4.50033-X</u>, accessed 31 October 2019.

G. Amato and U. Straccia, 1999. "User profile modeling and applications to digital libraries," In: S. Abiteboul and A.-M. Vercoustre (editors). *Research and advanced technology for digital libraries. Lecture Notes in Computer Science*, volume 1696. Berlin: Springer, pp. 184–197. doi: <u>https://doi.org/10.1007/3-540-48155-9_13</u>, accessed 31 October 2019.

Android, 2018. "ActivityRecognitionClient," <u>https://developers.google.com/android/reference/com/google/android/gms/locati</u> accessed 10 March 2019.

Apple, 2019. "CMMotionActivity," at <u>https://developer.apple.com/reference/coremotion/cmmotionactivity</u>, accessed 27 March 2019.

M. ben Khalifa, R.P.D.Redondo, A.F. Vilas and S.S. Rodríguez, 2017. "Identifying urban crowds using geo-located social media data: A Twitter experiment in New York City," *Journal of Intelligent Information Systems*, volume 48, number 2, pp. 287–308. doi: <u>https://doi.org/10.1007/s10844-016-0411-x</u>, accessed 31 October 2019.

M. Boukhechba, A. Bouzouane, B. Bouchard, C. Gouin-Vallerand and S. Giroux, 2015. "Online recognition of people's activities from raw GPS data: Semantic trajectory data analysis," *PETRA '15: Proceedings of the Eighth ACM International Conference on Pervasive Technologies Related to Assistive Environments*, article number 40. doi: <u>https://doi.org/10.1145/2769493.2769498</u>, accessed 31 October 2019.

L. Cao, Y. Wang, B. Zhang, Q. Jin and A.V. Vasilakos, 2018. "GCHAR: An efficient group-based context—aware human activity recognition on smartphone," *Journal of Parallel and Distributed Computing*, volume 118, part 1, pp. 67–80.

doi: https://doi.org/10.1016/j.jpdc.2017.05.007, accessed 31 October 2019.

A. Carroll and G. Heiser, 2010. "An analysis of power consumption in a smartphone," *USENIXATC'10: Proceedings of the 2010 USENIX Conference*, p. 21-21.

L. Chen and C. Nugent, 2009. "Ontology-based activity recognition in intelligent pervasive environments," *International Journal of Web Information Systems*, volume 5, number 4, pp. 410–430.

2019.

doi: <u>https://doi.org/10.1108/17440080911006199</u>, accessed 31 October 2019.

J. Chon and H. Cha, 2011. "LifeMap: A smartphone-based context provider for location-based services," *IEEE Pervasive Computing*, volume 10, number 2, pp. 58–67. doi: <u>https://doi.org/10.1109/MPRV.2011.13</u>, accessed 31 October 2019.

Y. Chon, N.D. Lane, F. Li, H. Cha and F. Zhao, 2012. "Automatically characterizing places with opportunistic crowdsensing using smartphones," *UbiComp '12: Proceedings of the 2012 ACM Conference on Ubiquitous Computing*, pp. 481–490. doi: <u>https://doi.org/10.1145/2370216.2370288</u>, accessed 31 October

I. da Penha Natal, R. de Avellar Campos Cordeiro and A.C.B. Garcia, 2017. "Activity recognition model based on GPS data, points of interest and user profile," In: M. Kryszkiewicz, A. Appice, D. Ślęzak, H. Rybinski, A. Skowron and Z. Raś (editors). *Foundations of intelligent systems. Lecture Notes in*

Computer Science, volume 10352. Cham, Switzerland: Springer, pp. 358–367. doi: https://doi.org/10.1007/978-3-319-60438-1_35, accessed 31

ooi: <u>https://doi.org/10.100//978-3-319-60438-1_35</u>, accessed 3. October 2019.

S.Difrancesco, P. Fraccaro, S.N. van der Veer, B. Alshoumr, J. Ainsworth, R. Bellazzi and N. Peek, 2016. "Out-of-home activity recognition from GPS data in schizophrenic patients," *Proceedings of the 2016 IEEE 29th International Symposium on Computer-Based Medical Systems (CBMS)*, pp. 324–328.

doi: https://doi.org/10.1109/CBMS.2016.54, accessed 31 October 2019.

F. Foerster, M. Smeja and J. Fahrenberg, 1999. "Detection of posture and motion by accelerometry: A validation study in ambulatory monitoring," *Computers in Human Behavior*, volume 15, number 5, pp. 571–583. doi: <u>https://doi.org/10.1016/S0747-5632(99)00037-0</u>, accessed 31 October 2019.

M.S. Fox, M. Barbuceanu and M. Gruninger, 1995. "An organisation ontology for enterprise modelling: Preliminary concepts for linking structure and behaviour," *WET-ICE '95: Proceedings of the Fourth Workshop on Enabling Technologies: Infrastructure for Collaborative Enterprises*, at <u>http://www.eil.utoronto.ca/wp-</u> <u>content/uploads/public/papers/org.pdf</u>, accessed 31 October 2019.

B. Furletti, P. Cintia, C. Renso and L. Spinsanti, 2013. "Inferring human activities from GPS tracks," *UrbComp '13: Proceedings of the Second ACM SIGKDD International Workshop on Urban Computing*, article number 5. doi: <u>https://doi.org/10.1145/2505821.2505830</u>, accessed 31 October 2019.

J.B. Gomes, S. Krishnaswamy, M.M. Gaber, P.A.C. Sousa and E Menasalvas, 2012. "MARS: A personalised mobile activity recognition system," *Proceedings of the 2012 IEEE 13th International Conference on Mobile Data Management*, pp. 316–319. doi: <u>https://doi.org/10.1109/MDM.2012.33</u>, accessed 31 October 2019.

Google, 2005. "Places," at <u>https://developers.google.com/places/</u>, accessed 29 October 2018.

M. Han, J.H. Bang, C. Nugent, S. McClean and S. Lee, 2014. "A lightweight hierarchical activity recognition framework using smartphone sensors," *Sensors*, volume 14, number 9, pp. 16,181–16,195. doi: <u>https://doi.org/10.3390/s140916181</u>, accessed 31 October 2019.

M.M. Hassan, M.Z. Uddin, A. Mohamed and A. Almogren, 2018. "A robust human activity recognition system using smartphone sensors and deep learning," *Future Generation Computer Systems*, volume 81, pp. 307–313. doi: <u>https://doi.org/10.1016/j.future.2017.11.029</u>, accessed 31 October 2019.

L. Hjorth and K. Gu, 2012. "The place of emplaced visualities: A case study of smartphone visuality and location-based social media in Shanghai, China," *Continuum*, volume 26, number 5, pp. 699–713. doi: <u>https://doi.org/10.1080/10304312.2012.706459</u>, accessed 31 October 2019.

L. Huang, Q. Li and Y. Yue, 2010. "Activity identification from GPS trajectories using spatial temporal POSs' attractiveness," *LBSN '10: Proceedings of the 2nd ACM SIGSPATIAL International Workshop on Location Based Social Networks*, pp. 27–30. doi: <u>https://doi.org/10.1145/1867699.1867704</u>, accessed 31 October 2019.

W. Huang, M. Li, W. Hu, G. Song, X. Xing and K. Xie, 2013. "Cost sensitive GPS-based activity recognition," *Proceedings of the 2013 10th International Conference on Fuzzy Systems and Knowledge Discovery (FSKD)*, pp. 962–966. doi: <u>https://doi.org/10.1109/FSKD.2013.6816334</u>, accessed 31 October 2019. 2019.

W.-C. Hung, F. Shen, Y.-L. Wu, M.-K. Hor and C.-Y. Tang, 2014. "Activity recognition with sensors on mobile devices," *Proceedings of the 2014 International Conference on Machine Learning and Cybernetics*, pp. 449–454.

doi: <u>https://doi.org/10.1109/ICMLC.2014.7009650</u>, accessed 31 October 2019.

A. Ignatov, 2018. "Real-time human activity recognition from accelerometer data using convolutional neural networks," *Applied Soft Computing*, volume 62, pp. 915–922. doi: <u>https://doi.org/10.1016/j.asoc.2017.09.027</u>, accessed 31 October

D.M. Karantonis, M.R. Narayanan, M. Mathie, N.H. Lovell and B.G. Celler, 2006. "Implementation of a real-time human movement classifier using a triaxial accelerometer for ambulatory monitoring," *IEEE Transactions on Information Technology in Biomedicine*, volume 10, number 1, pp. 156–

167. doi: <u>https://doi.org/10.1109/TITB.2005.856864</u>, accessed 31 October 2019.

D.I. Kim and E. Martinson, 2016. "Human centric spatial affordances for improving human activity recognition," *Proceedings of the 2016 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pp. 725–730.

doi: <u>https://doi.org/10.1109/IROS.2016.7759132</u>, accessed 31 October 2019.

N.C. Krishnan and D.J. Cook, 2014. "Activity recognition on streaming sensor data," *Pervasive and Mobile Computing*, volume 10, pp. 138–154. doi: <u>https://doi.org/10.1016/j.pmcj.2012.07.003</u>, accessed 31 October 2019.

B. Lee, C. Lim and K. Lee, 2017. "Classification of indoor-outdoor location using combined global positioning system (GPS) and temperature data for personal exposure assessment," *Environmental Health and Preventive Medicine*, volume 22, article number 29. doi: https://doi.org/10.1186/s12199-017-0637-4, accessed 31 October

doi: <u>https://doi.org/10.1186/s12199-01/-063/-4</u>, accessed 31 October 2019.

K. Lee and M.-P. Kwan, 2018. "Physical activity classification in free-living conditions using smartphone accelerometer data and exploration of predicted results," *Computers, Environment and Urban Systems*, volume 67, pp. 124–131. doi: <u>https://doi.org/10.1016/j.compenvurbsys.2017.09.012</u>, accessed 31

doi: <u>https://doi.org/10.1016/j.compenvurbsys.2017.09.012</u>, accessed 31 October 2019.

J. Levatić, M. Ceci, D. Kocev and S. Džeroski, 2017. "Semi-supervised classification trees," *Journal of Intelligent Information Systems*, volume 49, number 3, pp. 461–486. doi: <u>https://doi.org/10.1007/s10844-017-0457-4</u>, accessed 31 October 2019.

L. Liao, D. Fox and H. Kautz, 2006. "Location-based activity recognition," *NIPS'05: Proceedings of the Eighteenth International Conference on Neural Information Processing Systems*, pp. 787–794.

M.H.M. Noor, Z. Salcic, I. Kevin and K.I-K. Wang, 2016. "Enhancing ontological reasoning with uncertainty handling for activity recognition," *Knowledge-Based Systems*, volume 114, pp. 47–60. doi: <u>https://doi.org/10.1016/j.knosys.2016.09.028</u>, accessed 31 October 2019.

G. Ogbuabor and R. La, 2018. "Human activity recognition for healthcare using smartphones," *ICMLC 2018: Proceedings of the 2018 10th International Conference on Machine Learning and Computing*, pp. 41–46. doi: <u>https://doi.org/10.1145/3195106.3195157</u>, accessed 31 October 2019.

S. Park and H. Kautz, 2008. "Hierarchical recognition of activities of daily living using multi-scale, multi-perspective vision and RFID," 2008 IET Fourth International Conference on Intelligent Environments, pp. 1–4. doi: <u>https://doi.org/10.1049/cp:20081157</u>, accessed 31 October 2019.

S.J. Preece, J.Y. Goulermas, L.P. Kenney and D. Howard, 2009. "A comparison of feature extraction methods for the classification of dynamic activities from accelerometer data," *IEEE Transactions on Biomedical Engineering*, volume 56, number 3, 871–879. doi: <u>https://doi.org/10.1109/TBME.2008.2006190</u>, accessed 31 October 2019.

S. Ranasinghe, F. Al Machot and H.C. Mayr, 2016. "A review on applications of activity recognition systems with regard to performance and evaluation," *International Journal of Distributed Sensor Networks*, volume 12 (24 August). doi: <u>https://doi.org/10.1177/1550147716665520</u>, accessed 31 October 2019.

D. Riboni and C. Bettini, 2011. "COSAR: Hybrid reasoning for contextaware activity recognition," *Personal and Ubiquitous Computing*, volume 15, number 3, pp. 271–289. doi: <u>https://doi.org/10.1007/s00779-010-0331-7</u>, accessed 31 October 2019.

M. Safizadeh and S. Latifi, 2014. "Using multi-sensor data fusion for vibration fault diagnosis of rolling element bearings by accelerometer and load cell," *Information Fusion*, volume 18, 1–8. doi: <u>https://doi.org/10.1016/j.inffus.2013.10.002</u>, accessed 31 October 2019.

M.A. Shafique and E. Hato, 2016. "Travel mode detection with varying smartphone data collection frequencies," *Sensors*, volume 16, number 5, e716.

doi: <u>https://doi.org/10.3390/s16050716</u>, accessed 31 October 2019.

D. Shin, D. Aliaga, B. Tunçer, S.M. Arisona, S. Kim, D. Zünd and G. Schmitt, 2015. "Urban sensing: Using smartphones for transportation mode classification," *Computers, Environment and Urban Systems*, volume 53, pp. 76–86.

doi: <u>https://doi.org/10.1016/j.compenvurbsys.2014.07.011</u>, accessed 31 October 2019.

P. Trebilcox-Ruiz, 2016. "How to recognize user activity with activity recognition" (3 February), at <u>https://code.tutsplus.com/tutorials/how-to-recognize-user-activity-with-activity-recognition--cms-25851</u>, accessed 29 October 2018.

J. van Dijk, 2018. "Identifying activity-travel points from GPS-data with multiple moving windows," *Computers, Environment and Urban Systems*, volume 70, pp. 84–101. doi: <u>https://doi.org/10.1016/j.compenvurbsys.2018.02.004</u>, accessed 31

doi: <u>https://doi.org/10.1016/j.compenvurbsys.2018.02.004</u>, accessed 31 October 2019.

N. Wan and G. Lin, 2013. "Life-space characterization from cellular telephone collected GPS data," *Computers, Environment and Urban Systems*, volume 39, 63–70.

doi: https://doi.org/10.1016/j.compenvurbsys.2013.01.003, accessed 31 October 2019.

Z. Yu, E. Xu, H. Du, B. Guo and L. Yao, 2019. "Inferring user profile attributes from multi-dimensional mobile phone sensory data," *IEEE Internet of Things Journal*, volume 6, number 3, pp. 5,152–5,162. doi: <u>https://doi.org/10.1109/JIOT.2019.2897334</u>, accessed 31 October 2019.

W. Zhai, X. Bai, Y. Shi, Y. Han, Z.-R. Peng and C. Gu, 2019. "Beyond Word2vec: An approach for urban functional region extraction and identification by combining Place2vec and POIs," *Computers, Environment and Urban Systems*, volume 74, 12–12. doi: <u>https://doi.org/10.1016/j.compenvurbsys.2018.11.008</u>, accessed 31 October 2019.

Y. Zheng, Q. Li, Y. Chen, X. Xie and W.-Y. Ma, 2008. "Understanding mobility based on GPS data," *UbiComp '08: Proceedings of the 10th International Conference on Ubiquitous Computing*, pp. 312–321. doi: <u>https://doi.org/10.1145/1409635.1409677</u>, accessed 31 October 2019.

X. Zhou, W. Yu and W.C. Sullivan, 2016. "Making pervasive sensing possible: Effective travel mode sensing based on smartphones," *Computers, Environment and Urban Systems*, volume 58, 52–59. doi: <u>https://doi.org/10.1016/j.compenvurbsys.2016.03.001</u>, accessed 31 October 2019.

C. Zhu, Q. Cheng and W. Sheng, 2010. Human activity recognition via motion and vision data fusion, *Proceedings of the 2010 Conference Record of the Forty Fourth Asilomar Conference on Signals, Systems and Computers*, pp. 332–336.

doi: <u>https://doi.org/10.1109/ACSSC.2010.5757529</u>, accessed 31 October 2019.

Editorial history

Received 8 April 2019; revised 21 October 2019; accepted 31 October 2019.



This paper is licensed under a <u>Creative Commons Attribution-</u> NonCommercial-ShareAlike 4.0 International License.

Efficient out-of-home activity recognition by complementing GPS data with semantic information by Igor da Penha Natal, Lus Correia, Ana Cristina Garcia, and Leandro Fernandes. *First Monday*, Volume 24, Number 11 - 4 November 2019 https://journals.uic.edu/ojs/index.php/fm/article/view/9971/8187 doi: <u>http://dx.doi.org/10.5210/fm.v24i11.9971</u> 🚺 SHARE 🛛 📲 🏫 🔣 ...)

A Great Cities Initiative of the University of Illinois at Chicago University Library.

© First Monday, 1995-2019. ISSN 1396-0466.