



# Online human activity recognition employing hierarchical hidden Markov models

Parviz Asghari<sup>1</sup> · Elnaz Soleimani<sup>1</sup> · Ehsan Nazerfard<sup>2</sup>

Received: 9 January 2019 / Accepted: 27 June 2019  
© Springer-Verlag GmbH Germany, part of Springer Nature 2019

## Abstract

In the last few years there has been a growing interest in Human activity recognition (HAR) topic. Sensor-based HAR approaches, in particular, has been gaining more popularity owing to their privacy preserving nature. Furthermore, due to the widespread accessibility of the internet, a broad range of streaming-based applications such as online HAR, has emerged over the past decades. However, proposing sufficiently robust online activity recognition approach in smart environment setting is still considered as a remarkable challenge. This paper presents a novel online application of Hierarchical Hidden Markov Model in order to detect the current activity on the live streaming of sensor events. Our method consists of two phases. In the first phase, data stream is segmented based on the beginning and ending of the activity patterns. Also, on-going activity is reported with every receiving observation. This phase is implemented using Hierarchical Hidden Markov models. The second phase is devoted to the correction of the provided label for the segmented data stream based on statistical features. The proposed model can also discover the activities that happen during another activity - so-called interrupted activities. After detecting the activity pane, the predicted label will be corrected utilizing statistical features such as time of day at which the activity happened and the duration of the activity. We validated our proposed method by testing it against two different smart home datasets and demonstrated its effectiveness, which is competing with the state-of-the-art methods.

**Keywords** Online activity recognition · Streaming sensor data · Activity segmentation · Hierarchical hidden Markov models · Smart homes · Internet of things

## 1 Introduction

Nowadays, the prominent enhancement of sensor technology, alongside with the remarkable progress on Machine Learning methods has opened up a wide variety of real-world applications through combining different types of

sensor and Machine Learning techniques for data acquisition and processing, respectively. Various sensors have been utilized in research projects. Numerous studies have been done using visual sensors such as cameras which receive the sequence of images or videos as the sensor data. Despite the successful performance of vision based solutions, non-visual sensors are still required to solve the existing limitations. Non-visual sensors can be installed on the human's body or in the environment as wearable and ambient sensors. Exploiting the network of heterogeneous sensors has been of interest in many recent studies with the aim of enhancing the performance as well.

Human activity recognition (HAR) as an active research area in Machine Learning is of great importance. In many of the real world applications, system's decision is made based on the current human activity which is inferred by its HAR module from sensors data.

Smart Environment is a remarkable real world application which aims to monitor the behavior of individuals in the environment. This environment is enriched with

---

✉ Parviz Asghari  
parviz.asghari@aut.ac.ir

✉ Elnaz Soleimani  
elnaz.soleimani@aut.ac.ir

Ehsan Nazerfard  
nazerfard@aut.ac.ir

<sup>1</sup> Ambient Intelligence Research Lab, Department of Computer Engineering and Information Technology, Amirkabir University of Technology, 424 Hafez Ave., Tehran, Iran

<sup>2</sup> Department of Computer Engineering and Information Technology, Amirkabir University of Technology, 424 Hafez Ave., Tehran, Iran

sensors, actuators and processing units which can also be embedded in the objects. Monitoring systems as a use case of Smart Environments, are deployed to predict, detect and neutralize the abnormal behaviors of individuals in environments such as metro stations.

Ambient Assisted Living (AAL) approaches use Activity Recognition methods in order to monitor resident's behavior and support them so that improve the quality of their lives. HAR has also been exploited in game consoles, fitness and health monitoring on smartphones' applications. Many research laboratories have paid considerable attention to the Smart Environment related topics and consequently, several research projects such as CASAS (Cook et al. 2009), MAVhome (Cook et al. 2003), PlaceLab (Intille et al. 2005), CARE (Kröse et al. 2008), and AwareHome (Kidd et al. 1999) have been deployed. Activity recognition in this context denotes recognizing the daily activities of a person who is living in the environment. Activities of Daily Living (ADLs) is a terminology defined in the healthcare domain and it refers to people's daily self-care activities (Katz et al. 1970).

## 1.1 Activity modeling

In general, there exist two main approaches to model the activity; data-driven and knowledge-driven approach. Data-driven models, analyze large sets of data for each activity by means of Machine Learning and data mining techniques. Another approach is to acquire sufficient prior knowledge of personal preferences and tendencies by applying knowledge engineering and management technology methods. Activity recognition task contains different challenges such as labeling the activities, interrupted activities, sensors with non-valuable data and activity extraction. However, online recognition is noted as one of the most dominant challenges since the main requirement in many modern real-world applications is to recognize user's behavior based on the online stream of data.

In the present study, the stream of data is generated from a set of binary environmental sensors' output. The majority of online activity recognition approaches are fed with the data which preprocessed using the sliding window technique. The main drawback of this approach is the optimization problem of window size parameter itself.

To address this issue, we propose a method to detect the activity pane on data stream by discovering the sequential pattern of the sensors data in the beginning and termination of each activity. Detecting the activity boundaries, we can recognize the activity using a proper segment of streaming sensor data. Since the beginning of an activity is detected, the model keeps reporting the identified activity till the end of it gets recognized by its sequential pattern. While the model is reporting the current activity, it is capable of

recognizing other activities interrupting the ongoing one. Having activity's segment identified, statistical features of the segment including activity duration will be taken as the input to correct predicted label of the activity. Consequently, we get ride of the non-valuable sensor data by ignoring all the intervals but activity panes and tuning the window size parameter.

Our model consists of a set of Hidden Markov Models implemented in a hierarchical manner in order to recognize the beginning, estimate on-going activity, detect end and class of the activities. Our evaluation results demonstrated the superiority of the proposed model in comparison with some existing ones and competitive performance compared to the state-of-the-art models.

The remainder of the paper is organized as follows. Section 2 outlines the previous researches that have been done in this domain. Section 3 describes the proposed method and implementation details. Evaluation and experimental results are presented in Sects. 4 and 5 summarizes the results of this work and draws conclusions.

## 2 Related work

Smart environments obtain knowledge from physical platform and its residents in order to modify setting based on residents interactions' pattern with environment. These systems optimize numerous goals such as monitoring environment, assisting residents and control resources. The first idea that uses sensor technology for creating smart environment backs to 1990. Moser's neural network house modifies environment conditional statues based on user behavior pattern (Moser 1998). The next decades witnessed a considerable progress in this field which obtained by exploiting HAR techniques.

Several publications have appeared documenting research works on Human activity recognition. Generally, sensor-based and vision-based approaches, have been documented in the literature as two main categories of HAR. This paper is concentrating on the sensor-based HAR. The common basis for the majority of attempts of this type is to process the sequence of sensors' events and detect the corresponding activity, based on the discovered pattern. Furthermore, activity recognition has been vastly used in various applications and environments. Hence, it is conspicuous to have a robust Activity Recognition Model that is invariant to the environmental settings such as sensors' structure and placement in the smart environment.

To address this requirement, authors in Cook (2012) propose a robust model of the sensors which extracts different features such as the duration of the activity. In this generalized model, the problem is independent of the sensor environment and can be implemented for different users.

Besides, to evaluate the models, having enough correctly labeled data is inevitable. Nevertheless, manual labeling of the sensors' data, is prohibitive due to being highly time-consuming and often inaccurate.

Moreover, the way that annotation has been done is often ignored while it makes bias on the data. In the most common approach of annotating, residents of the smart environment are asked to perform an activity and then annotation will be done based on activated sensors. Nonetheless, this approach may not be practical in all situations. A labeling mechanism presented in Szewczyk et al. (2009) is an example of existing solutions for annotating sensor data automatically.

Much researches on activity recognition have been done using pre-segmented data; which means the beginning and end of the activities is pre-determined in the dataset (Sánchez et al. 2008; Fleury et al. 2009). Such approaches are far unrealistic compared to real-world setting and are not applicable in the online applications as the beginning and end of the activities is not determined when it comes to the stream of data. Researchers in Hong and Nugent (2013) and Okeyo et al. (2014) develop methods for sensor stream segmentation which brings the activity recognition based systems closer to those of actual world.

Despite all the researches on HAR, still opened challenges such as overlapping or concurrent activities, have yet to be solved. Overlap is noted as the phenomenon that different activity classes activate the same set of sensor events which makes overlapping activities hard to discriminate only based on the types of sensor events that they have triggered (Wen et al. 2015; Helaoui et al. 2011).

A trajectory propagation algorithm is proposed in Zhang et al. (2014) that uses concurrent activation events to recognize most probable occupants trajectories and locations. More abstract details of an occupants behavior can be extracted by using the concurrent activation events of a binary motion sensor network.

Riboni and et al. propose an unsupervised ontological method for interleaved activities (Riboni et al. 2016). Another unsupervised technique is introduced by Zhang and et al., which is able to learn the occupants moving patterns of resident activities' routine that may subject to change through time (Zhang et al. 2017). This decomposition based learning approach overcomes the limitation that a lot of other supervised learning approaches have faced.

AALO is a novel Activity recognition system for single resident smart homes that utilises Active Learning in the presence of Overlapped activities (Hoque and Stankovic 2012). It can recognize the overlapped Activity recognition-activities by preprocessing step and item-set mining phase. A key limitation of this research is that it is not capable of recognizing the overlapping activities which happen in the same location. In addition, performing on the online data stream is still lacking in this study.

Moreover, researchers in Malazi and Davari (2018) suggest a two-phase method based on emerging pattern which can recognize complex activities. In the first phase, this method extracts emerging pattern for distinct activities. In the second phase, it segments streaming sensor data, then uses time dependency between segments in order to concatenate the relevant segments. Segments concatenation lets the method recognize complex activities. In Quero et al. (2018), Quero et al. propose a Real-Time method for recognizing interleaved activities based on Fuzzy Logic and Recurrent Neural Networks.

Authors in Cook et al. (2013 and Gjoreski and Roggen (2017) study the problem of handling the large proportion of available data that are not categorized in predefined classes and addressed it by discovering patterns in them and segmenting it into learnable classes. These kinds of data usually belong to the sensors that are not exclusively involved in the predefined class of activities. In Gayathri et al. (2015), Gayathri et al. propose a hierarchical hybrid approach based on Markov logic network (MLN) in order to detect abnormal activities in dementia care systems. Several Machine Learning approaches have been examined in the domain of Activity Recognition. Ensemble methods (Jurek et al. 2014), Non-parametric models (Sun et al. 2014), Temporal frequent pattern mining (Nazerfard 2018), Support vector machines-based models (Sánchez et al. 2008), Recurrent neural networks (Singh et al. 2017), and probabilistic models like hidden Markov model and the Markov random field (Kabir et al. 2016; Kasteren et al. 2010; Yan et al. 2016; Jalal and Kamal 2018) have been exploited in the literature. Nonetheless, less attention has been paid to the domain of Online Activity Recognition which deals with processing stream of sensor data contrary to the conventional approaches that utilize pre-segmented data.

Most of the presented solutions for streaming data processing are based on sliding window technique (Okeyo et al. 2014; Krishnan and Cook 2014; Yala et al. 2015; Al Machot et al. 2016, 2018). The sliding window approach, briefly named as *windowing*, mainly considers the temporal relation or number of sensors for framing data. One of the key bottlenecks of this approach is fine-tuning the window size. One basic solution is to employ constant pre-determined window size (Kasteren et al. 2010; Krishnan and Cook 2014). Though, as the number of activated sensors are varied in different activities, applying dynamic window size has been noticed by many researchers (Wan et al. 2015; Espinilla et al. 2018; Al Machot et al. 2016; Okeyo et al. 2014).

Authors in Al Machot et al. (2016) and Al Machot et al. (2018) present a novel probabilistic method to determine the window size. A different window size is initialized regarding each class of activity based on prior estimation and it is getting updated by the upcoming sensor events.

Krishnan and Cook (2014) consider temporal dependency among sensors as a criterion for sliding windowing so that those sensors, which get activated at a certain time interval, are examined as one activity. Their other proposed solution is a sensor-based method which recognizes group of sensors as a window that was continuously activated together by their Mutual Information measure. Authors in Yala et al. (2015) extend the presented model of Krishnan and Cook (2014) and improved its performance by altering computation of Mutual Information.

Researchers in Kabir et al. (2016) present a multi-stage classification method. The first stage is to cluster the activities using a Hidden Markov Model based on location data and then in the next stage, another HMM classifies the exact activity using a sequence of sensor data. The major weakness of this method is that it makes no attempt to specify the boundaries of activities which negatively affects the performance.

Another method to tackle online recognition is introduced by Li et al. (2018). They propose cumulative fixed sliding windows for real time activity recognition. Their segmentation method consists of several fixed time length windows which have overlapped with each other. These overlapping windows considered as the whole a window, and its information is used to detect the on-going activities.

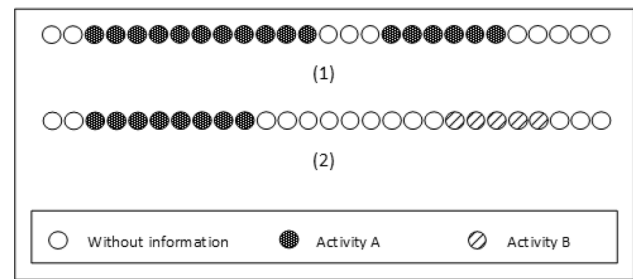
A novel incremental activity inference engine is introduced in Hao et al. (2018) to predict and recognize activities in real-time. It provides cognitive assistance to elderly people suffering from cognitive impairment with abnormal behavioral patterns. An ontological clustering method is also proposed to improve accuracy.

### 3 Proposed approach

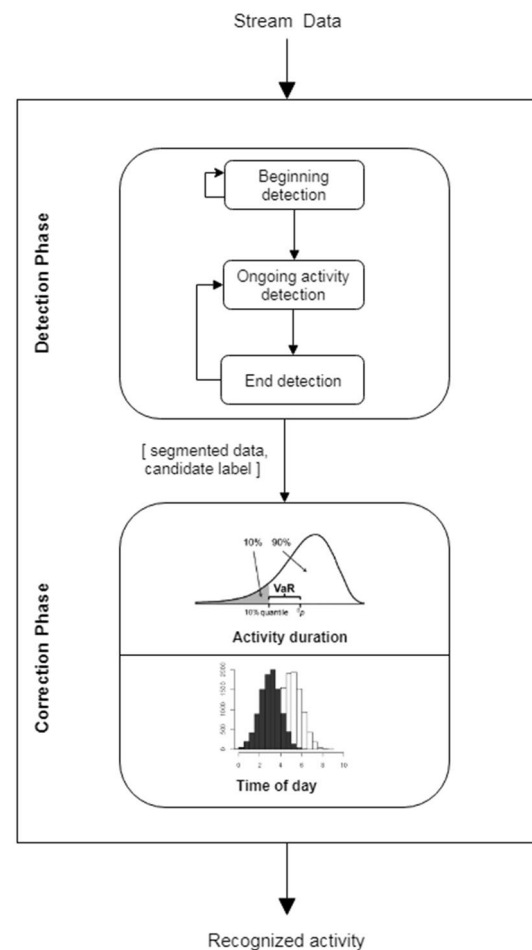
The Windowing technique is among the most widely used solutions to process the stream of sensor data. However, one practical question that arises when dealing with this approach is how to determine the proper window size. There is still some controversy surrounding window size tuning while it has a great impact on the performance as the decision-making process is halting until the model receives a complete window of data.

Moreover, the selected window passed to the model often contains inappropriate data which belongs to none of the predefined classes. Dynamic sliding window size may seem an appropriate solution. However, as it is shown in Fig. 1, not only different classes of activity need different window sizes, but also distinct samples of one class require distinct window sizes.

This paper seeks to address these problems by presenting a new model which works based on the dynamic segmentation of the data stream. Our proposed model breaks



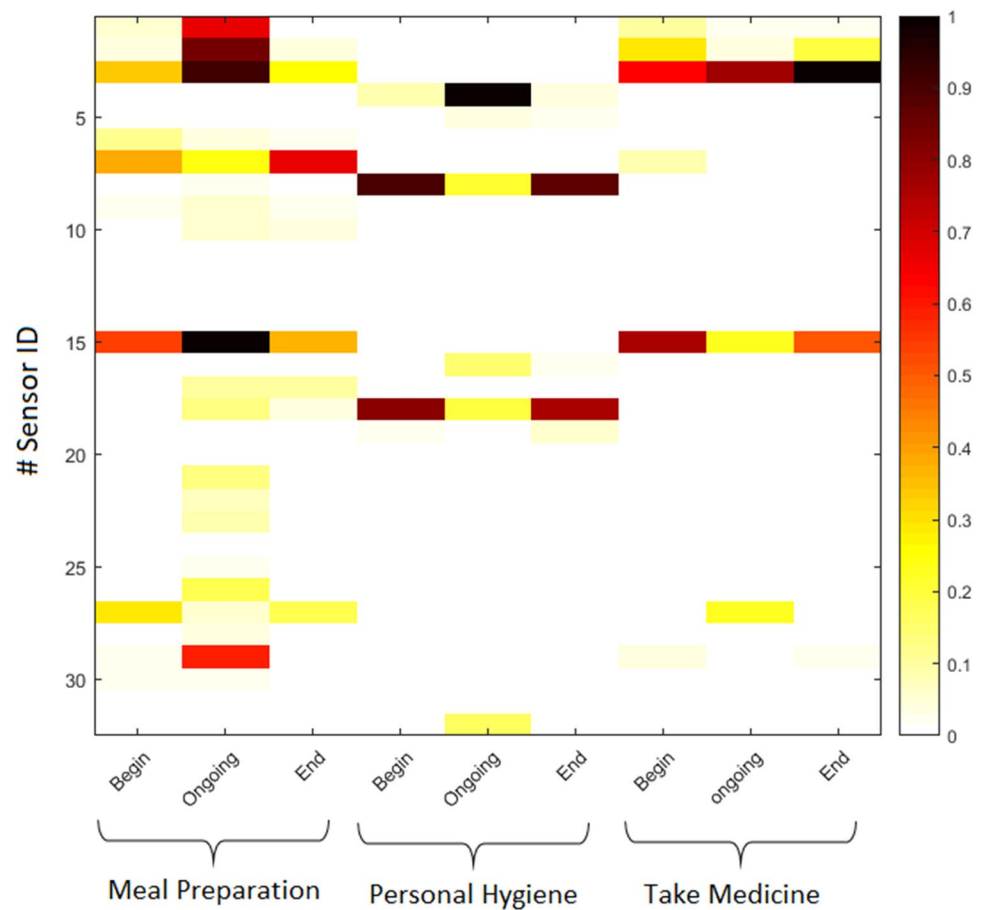
**Fig. 1** (1) distinct samples with different sizes (2) different classes have different sizes



**Fig. 2** Abstract structure of the proposed method

down the problem of Online Activity Recognition into two sub-problems as it is depicted in Fig. 2. In the initial stage, the model detects the activity pane by recognizing sensor events' patterns of when the activities start and finish and makes the prediction while receiving data. Having the beginning and end of the activities determined, in the next stage the model proceeds to refine the prediction by classifying the activities into the predefined classes and the *other* class.

**Fig. 3** Normalized frequency of sensor activation in the beginning, end and duration of *Meal Preparation*, *Personal Hygiene* and *Take Medicine* activities



Predefined classes proposed by medical specialists (Katz et al. 1970), included Daily Living Activities such as eating, bathing, sleeping, etc. *Other* class label is devoted to the activities that do not fit into predefined classes. Following sections explain each stage in detail.

The main contributions and innovations of this paper are as follows:

- Description of a novel method based on the occurrence pattern of activities, that addresses window size problem of previous sliding window methods.
- Providing an estimation of the on-going activity due to each sensor observation, which delivers Real-time Activity Recognition.
- Recognizing interrupted activities even those which occur in the same location.

### 3.1 Detection phase

The first step to recognize the activity on the stream of sensor data is to detect its occurrence. The idea is to investigate the sensor events in the beginning and end of the activities to extract corresponding patterns of start and end. As it can be inferred from Fig. 3, the set of activated sensors, for each

activity is almost unchanged in most of the occurrences. This assumption can be justified by considering locally limited functional zone for each activity. *Meal preparation* as an example happens in the kitchen, thus it should be kitchen's sensors that get activated during this activity.

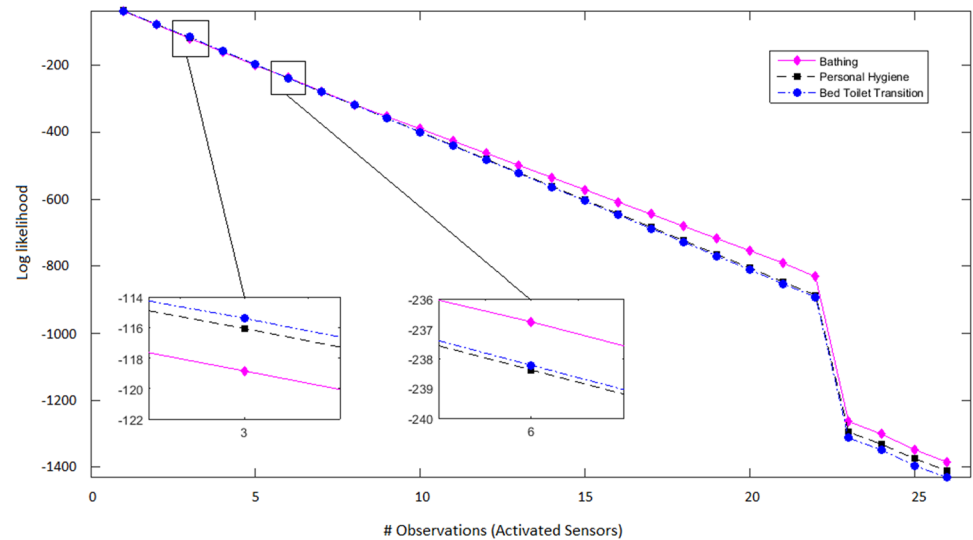
The Number of observations which are taken into account for beginning detection purpose, affects the model performance. Table 1 compares different values for this parameter and corresponding results. Based on this comparison, we opted for considering three sensor events in our model. Indeed, the number of considered sensors should be small enough to promise the recognition feasibility of the next activity. To exemplify, when the resident opens or closes the door, only a few sensors get activated; hence waiting for more observations is out of options.

**Table 1** Proposed model accuracy % obtained using  $\beta$  observations

Dataset	$\beta$			
	2	3	4	5
HomeA	78.7	97.0	92.9	93.2
HomeB	84.3	96.4	94.2	91.0



**Fig. 4** Log likelihood trend of candidate classes during the observation procedure



On the other hand, there exist some classes of activities such as *Personal Hygiene* and *Bathing*, which share a common set of activated sensors in the beginning and the end. Therefore, recognizing these activities requires more information like sensor activation sequences. Once the beginning of an activity gets recognized, the model utilizes the upcoming sequence of sensor events to recognize the activity class itself and as it receives more events, its prediction gradually becomes more confident. This trend can be seen in Fig. 4.

As Fig. 4 shows, at first the most probable activity is *Bed to Toilet transition*, while after receiving more observations, estimated activity changes to *Bathing*. These activities have overlap due to their similar sensors. Therefore, there is a need for more information to produce a reliable prediction.

Numerous Machine Learning methods such as Support Vector Machine, Neural Networks, Decision Tree, and Probabilistic Graphical Models have been applied in activity recognition field. Hidden Markov Model is one of the most popular models in the literature for sequential data processing as it does not suffer from labor-intensive calculations like Neural Networks. Besides, prior knowledge can be simply leveraged in probabilistic models which are also more robust to the noise.

Taking into account the aforementioned reasons, we chose HMM to recognize the activities. Suppose  $X_t$  represents the hidden state vector and  $Y_t$  represents vector of observations. Assuming  $K$  possible hidden states we have  $X_t \in \{1, \dots, K\}$ . In our case, observations are the sequence of sensor events. Equation (1) highlights the calculation to obtain the most probable class based on the observations:

$$P(X_t|y_{1:t}) = P(y_t|X_t) \left[ \sum_{X_{t-1}} P(X_t|x_{t-1})P(x_{t-1}|y_{1:t-1}) \right] \quad (1)$$

The Hidden Markov Model approach is not well suited to process long sequences and this led us to employ an extension of HMM called Hierarchical Hidden Markov Model in our study. Hierarchical-HMM is suitable for the problems that contain multilevel dependencies in terms of time and follow a hierarchical structure in their context.

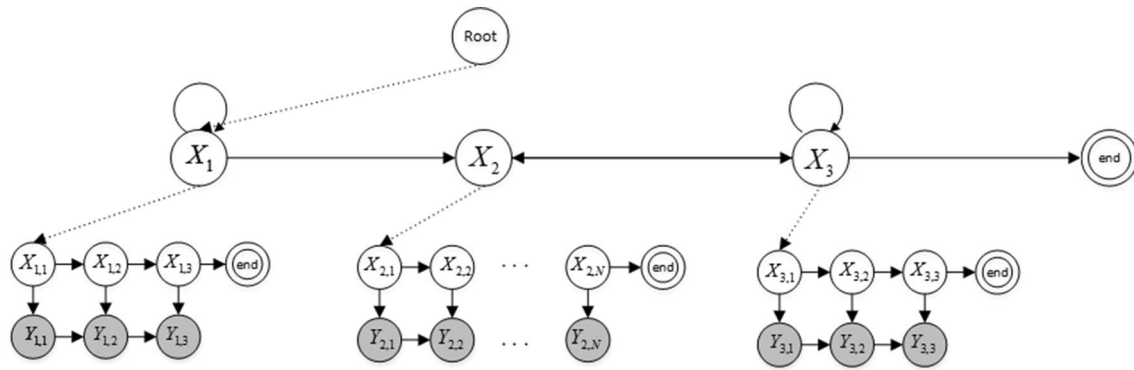
Hierarchical-HMM is represented by a triple  $(\zeta, \gamma, \theta)$  while  $\zeta$ ,  $\gamma$ , and  $\theta$  stand for model structure, set of observations and model's parameters respectively. Model structure defines number of levels  $d$  and children-parent relations on each level. States in the lowest level, are the only generative ones which generate the observations. Parameter set  $\theta$  is defined as following (Nguyen et al. 2005):

$$\theta = \{B(y|p), \pi^{d,p^*}, A_{ij}^{d,p^*}, A_{i,end}^{d,p^*} | \forall (y, p, d, p^*, i, j)\}, \quad (2)$$

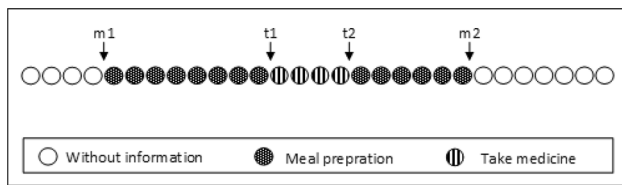
where  $p^*$  is an abstract state in level  $d$  with children set  $ch(p^*)$ ,  $B(y|p)$  is the probability of generating observation  $y$  by generative state  $p$ ,  $\pi^{d,p^*}$  is the prior probability of  $ch(p^*)$ ,  $A_{ij}^{d,p^*}$  is the probability of inhering child  $i$  to  $j$ , and  $A_{i,end}^{d,p^*}$  is the probability of termination of  $p^*$  in state  $i$ . The proper value set of  $\theta$  is obtained by Maximum Likelihood technique.

Figure 5 demonstrates the transition diagram of our implemented Hierarchical-HMM. In this diagram we have 2 type of transition, vertical and horizontal. Vertical transitions are shown by dotted arc, and occurs between different level of HHMM. Horizontal transitions happen among same level node, when task of one node finishes.

Additionally, there are two types of node in HHMM, abstract nodes which produce a sequence of observations, and production nodes which produce single observation. In Fig. 5, nodes  $X_1$ ,  $X_2$ , and  $X_3$  are abstract nodes that each of them considered as a distinct HMM. In this model,  $X_1$  is the responsible sub-HMM for detecting the beginning of an



**Fig. 5** The transition diagram of the implemented Hierarchical-HMM



**Fig. 6** An illustration on interrupted activities: The subject has been preparing meal since time  $m1$ . At time  $t1$  he decided to interrupt so as to take his medicine. At time  $t2$  he resumed and kept preparing meal till  $m2$

activity.  $X_1$  is an auto regressive HMM which consist of 3 observation. After receiving 3 sensor observations, if beginning of an activity is detected, the control will pass through the  $X_2$  in order to recognize the class of on-going activity.

In  $X_2$  sub-HMM with occurring each observation, the corresponding activity will be recognized. This process continues until the termination of activity which detects by  $X_3$ . At the end of the cycle, control will be given back to the root node. The output of this process  $S$ , is the segmentation of data which holds the related information.

Sometimes, subject interrupts the current activity and commences a new activity. There exist some techniques which can recognize the latter activity independently of the first one (Nef et al. 2015). Thought, their discrimination is limited to the activities which happen at different locations. Since our hierarchical detection technique is location invariant the proposed model is capable of recognizing interrupted activities that take place in a common location. Figure 6 demonstrates an example of the interrupted activities.

### 3.2 Correction phase

With the completion of the Detection phase, our model can move to the next stage. In this stage, for each class, Joint Probabilistic Distribution Function (Joint PDF) of its time length and time of day gets extracted. Figure 7 exemplifies

**Table 2** Accuracy of the proposed model employing different threshold parameter  $\alpha$

Dataset	$\alpha$				
	0.02	0.04	0.06	0.08	0.10
HomeA	59.7	62.1	63.0	65.2	64.8
HomeB	54.0	53.0	60.0	58.0	57.4

how time of the day as an statistical feature can ease the discrimination. The extracted PDFs now can be used to measure the level of belonging of a data segment to each of the classes. As Eq. (3) reveals, if this probability passes the threshold  $\alpha$  for each class  $C_k$ , the data can be labeled as  $C_k$ . Otherwise, it is labeled as Other class. A fine tuning process has been done to achieve the proper value for  $\alpha$ . Table 2 details this trial.

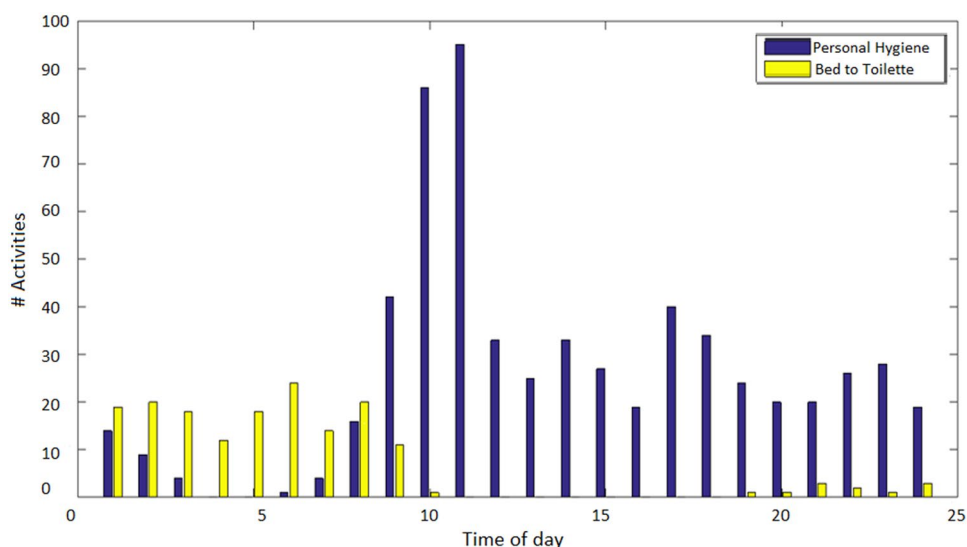
$$P(S \in C_k) = \begin{cases} 1 & f(T_{C_k} = t_s) \geq \alpha \\ 0 & \text{Otherwise} \end{cases} \quad (3)$$

In summary, this model first detects the activity occurrence utilizing the beginning pattern of activities. Next, it recognizes the class of ongoing activity based on activated sensors until the end of the activity gets detected. To improve the discrimination performance between similar activities, our model also exploits the time of day in which the activity is occurring as an statistical feature.

## 4 methods

The performance of our proposed approach is assessed by conducting 2 experiments on 2 different datasets to recognize the current activity on the stream of data. We compared our approach with a set of most common Machine Learning models and reported the results.

**Fig. 7** Choosing time of day as a feature facilitates classification task. *Bed to Toilet transition* and *Personal Hygiene* activities can be discriminated by this feature since *Bed to Toilet transition* occurrence time has low overlap with the *Personal Hygiene* activity



#### 4.1 Dataset description

We evaluated our proposed model on HomeA<sup>1</sup> and HomeB<sup>2</sup> datasets where each home is one bedroom apartment hosting a single older adult performing unscripted daily activities. Figure 8 shows a sample slice of this dataset. These 2 datasets have been collected using 32 and 30 sensors which have been deployed in 2 different houses depicted in Fig. 9 for the period of 5 months. Labeling of this dataset has been done later by human experts. Table 3 summarizes characteristics of data acquisition for HomeA and HomeB datasets. These datasets contain 11 classes of interest. Details on sample distribution of classes are highlighted in Table 6. Of datasets, 70% (3.5 months), 10% (2 weeks), and 20% (4 weeks) were dedicated for training, validation and testing respectively.

Due to the limitation of accessing real-time data in this domain, we have simulated data streaming using offline datasets. Each row in the datasets is a sample of a sensor event which has a time label, sensor ID, and sensor status. In every time slot, one sample is given as the input to the system, and the system makes predictions based on sensor events that have acquired until the current time slot. To delineate more, let us denote the *Eating* activity as *E*. It contains some events of triggered sensors while each of them is one distinct sample in the dataset:

$$E = \{S_{t_1}, S_{t_2}, S_{t_3}, \dots, S_{t_{end}}\}. \quad (4)$$

In the first time slot  $t_1$ , sample data  $S_{t_1}$  is passed to the system as input to make a prediction. In the next time slot  $t_2$ , the system received  $S_{t_2}$  as input and this process goes on like that. Consequently, data is treated in an online manner.

2009-11-04	22:25:55.703001	D021	CLOSE
2009-11-04	22:26:04.203001	M015	ON
2009-11-04	22:26:05.875001	D027	OPEN
2009-11-04	22:26:07.906001	M003	ON
2009-11-04	22:26:08.203001	M015	OFF
2009-11-04	22:26:11.828001	M003	OFF
2009-11-04	22:26:13.125001	M003	ON
2009-11-04	22:26:17.421001	M003	OFF
2009-11-04	22:26:21.109001	M003	ON
2009-11-04	22:26:25.031001	M003	OFF
2009-11-04	22:26:31.296001	M003	ON
2009-11-04	22:26:33.875001	D027	CLOSE
2009-11-04	22:26:35.312001	M003	OFF
2009-11-04	22:28:04.000001	M015	ON
2009-11-04	22:28:04.984001	M003	ON
2009-11-04	22:28:08.265001	M015	OFF
2009-11-04	22:28:08.984001	M003	OFF
2009-11-04	22:28:13.593001	M007	ON

**Fig. 8** Sample of sensor data during an activity in dataset HomeA and HomeB

#### 4.2 Baseline models

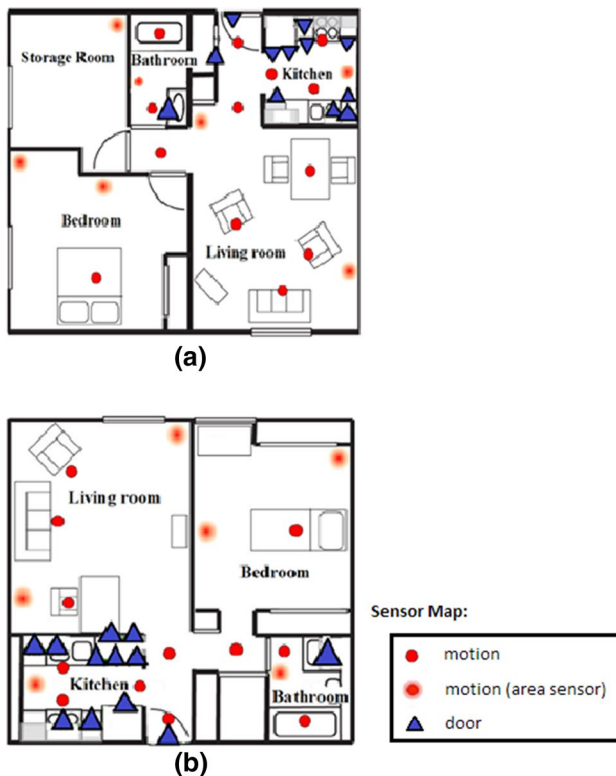
We have compared performance of our proposed model with several existing approaches:

- *SW* this model utilises constant window size. Sensors are equally contributing in window size calculation (Krishnan and Cook 2014).
- *TW* it employs a constant window size that is obtained based on time interval (Krishnan and Cook 2014).
- *SWMI* it considers window size as a constant value. Mutual Information of sensors and activity classes is

<sup>1</sup> <http://eecs.wsu.edu/~nazerfard/AIR/datasets/data1.zip>

<sup>2</sup> <http://eecs.wsu.edu/~nazerfard/AIR/datasets/data2.zip>





**Fig. 9** The architecture of **a** HomeA, **b** HomeB. Red circles represent motion sensors and blue triangles stand for Door/Cabinet sensors (color figure online)

**Table 3** Dataset Statistics

# Attribute	HomeA	HomeB
Motion sensors	20	18
Door/cabinet sensors	12	12
Residents	1 Person	1 Person
Sensor events collected	371925	274920
Timespan	5 months	5 months

taken into account for window size calculation (Yala et al. 2015).

- *SWTW* window size remains constant in this model and is calculated by time based weighting of the sensor events (Yala et al. 2017).
- *DW* in this model window size is variable for each activity and its value is calculated using probabilistic information of sensor events (Al Machot et al. 2016).
- *PWPA* it benefits from a 2-level window with a fixed size. First level contains the probabilistic information of the previous windows and activities. Second level includes the data of current step with possible activity as a augmented feature (Krishnan and Cook 2014).

### 4.3 Evaluation metrics

The performance of the proposed model has been evaluated by Classification accuracy and F1-measure. Classification accuracy is computed considering each test sample using its predicted label and corresponding ground truth label as highlighted in Eq. (5).

$$\text{Accuracy} = \frac{tp + tn}{tp + tn + fp + fn} \quad (5)$$

To calculate F1-measure, activities are considered despite the length of their windows. To exemplify, consider activity A1 and A2 containing 500 and 50 sensor event samples. Accuracy metric evaluates all 550 samples while F1-measure is calculated for those two activities as a whole. Equation (5) show the F-measure formula.

$$F1 - \text{measure} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (6)$$

Performance of the models has been compared in terms of F1-measure and Accuracy as it is reported in Tables 4, 5, and 7.

## 5 Experimental results and discussion

As can be inferred from 4 for HomeA dataset, *Personal Hygiene*, *Bathing*, *Sleeping*, *Take Medicine*, *Meal Preparation*, and *Eating Meal* classes can be highly discriminated. However, unsatisfactory results were reached for *Enter Home*, *Leave Home*, and *Sleeping* classes. This performance degradation is justifying given the fact that *Enter Home* and *Leave Home* both activate the same set of sensors and these two classes of activity often occur in sequence. Besides, sometimes the resident walks out of the house just for short moments which is not considered as Leaving home. Similar to other windowing-based methods such as SWTW, SWMI, and PWPA, our proposed method is facing problem in recognizing activities such as *Enter Home* and *Leave Home* which shares the same corresponding sensors and involves a few sensors. It makes their beginning and end pattern exactly similar. PWPA can acquire knowledge from previous activities which makes it able to detect *Enter Home* and *Leave Home* more precisely as *Enter Home* always happens after *Leave Home* activity and vice versa. Moreover, classifying these two classes of activities from the rest is quite challenging due to their short length which does not give enough time to the model to improve its estimation and makes their distribution more sensitive to noise compared with other activities. According to Table 6, the *Housekeeping* activity has only

**Table 4** Confusion matrix of the proposed model on HomeA dataset

	1	2	3	4	5	6	7	8	9	10	11	12
1 = Personal hygiene	0.88	0.00	0.00	0.02	0.00	0.00	0.00	0.02	0.00	0.00	0.00	0.08
2 = Leave home	0.00	0.14	0.00	0.00	0.29	0.00	0.00	0.00	0.00	0.00	0.00	0.57
3 = Enter home	0.00	0.30	0.14	0.00	0.00	0.00	0.03	0.00	0.00	0.00	0.00	0.54
4 = Bathing	0.08	0.00	0.00	0.67	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.25
5 = Meal preparation	0.00	0.00	0.00	0.00	0.52	0.00	0.08	0.00	0.04	0.00	0.00	0.36
6 = Napping	0.00	0.00	0.02	0.00	0.01	0.13	0.00	0.12	0.00	0.00	0.00	0.72
7 = Take medicine	0.00	0.00	0.00	0.00	0.13	0.00	0.55	0.00	0.05	0.00	0.00	0.27
8 = Eating food	0.00	0.00	0.01	0.00	0.01	0.00	0.00	0.28	0.00	0.00	0.00	0.69
9 = Housekeeping	0.00	0.00	0.00	0.00	0.13	0.00	0.26	0.00	0.00	0.00	0.00	0.61
10 = Sleeping in bed	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.40	0.02	0.58
11 = Bed to toilet transition	0.60	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.27	0.13
12 = Other	0.19	0.01	0.25	0.02	0.04	0.01	0.13	0.03	0.01	0.01	0.02	0.28

**Table 5** F1-measure of the proposed model on HomeB dataset

	1	2	3	4	5	6	7	8	9	10	11	12
1 = Personal hygiene	0.57	0.00	0.00	0.04	0.01	0.00	0.00	0.00	0.00	0.00	0.08	0.31
2 = Leave home	0.00	0.50	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.50
3 = Enter home	0.00	0.65	0.09	0.00	0.03	0.00	0.00	0.00	0.00	0.00	0.00	0.21
4 = Bathing	0.00	0.00	0.00	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
5 = Meal preparation	0.00	0.01	0.00	0.00	0.30	0.00	0.02	0.00	0.08	0.00	0.00	0.59
6 = Napping	0.00	0.00	0.00	0.00	0.00	0.13	0.00	0.05	0.00	0.00	0.00	0.82
7 = Take medicine	0.00	0.00	0.00	0.00	0.11	0.00	0.09	0.00	0.56	0.00	0.00	0.23
8 = Eating food	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.33	0.04	0.00	0.00	0.63
9 = Housekeeping	0.00	0.00	0.00	0.00	0.05	0.00	0.00	0.00	0.35	0.00	0.00	0.60
10 = Sleeping in bed	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.96	0.00	0.04
11 = Bed to toilet transition	0.13	0.00	0.00	0.03	0.00	0.00	0.00	0.00	0.00	0.00	0.56	0.28
12 = Other	0.14	0.01	0.05	0.01	0.00	0.00	0.00	0.00	0.01	0.01	0.05	0.71

**Table 6** Overall instances of each class in dataset HomeA and HomeB

Activity	HomeA	HomeB
Personal hygiene	704	782
Leave home	420	245
Enter home	417	245
Bathing	72	41
Meal preparation	554	245
Napping	564	138
Take medicine	473	81
Eating meal	334	208
Housekeeping	13	175
Sleeping in bed	210	207
Bed to toilet transition	96	236
Total	2603	3857

**Table 7** Performance comparison of models in terms of accuracy and F1-measure

Model	Accuracy		F1-measure	
	HomeA (%)	HomeB (%)	HomeA (%)	HomeB (%)
Proposed model	65.20	60	51	49
SW	58	48	51	49
TW	59	52	53	54
SWMI	64	54	60	57
SWTW	62	45	55	55
DW	59	55	58	58
PWPA	62	50	55	51

13 instances in the HomeA dataset which is dramatically low compared to the rest of the classes. This imbalanced distribution may have been the cause of the inadequate performance of the model.

In HomeB dataset, *Bathing* and *Sleeping* are well discriminated and detection of *Meal Preparation*, *Personal Hygiene*, *Eating*, and *Bed to Toilet transition* are satisfactory. The performance dropped for *Taking Medicine*, *Napping*, *Enter Home* and *Leave Home*. In HomeB there are times that the resident is eating meal on the couch which is a place for napping, and this act causes problem in recognition of both napping and eating activities. Another difference compared to HomeA is that, the *Housekeeping* activity recognition's accuracy increases dramatically due to the adequate samples in HomeB dataset.

The proposed model is capable of detecting overlapped activities such as *Meal Preparation* and *Taking Medicine* even though they happen in a common location. Whereas previous methods have difficulty to recognize overlapped activity in an online setting. Datasets' description implies that activities may overlap in two levels.

Another point worth mentioning is about our trial to benefit from the knowledge of the previous activities with the purpose of enhancement in our proposed method's performance. However, the results suggest a 12% accuracy drop for HomeA and 15% drop for HomeB. The primary cause of the mentioned accuracy drops can be due to misclassification error propagation. The proposed model does not have access to the ground truth of the previous activity. Therefore, in case of erroneous prediction of the previous activity, this error will propagate through the system.

As the results provided in Table 7 suggest, our proposed method outperforms the others in the matter of accuracy. F1-measure metric has not followed the same trend as accuracy, hence methods with the better accuracy did not necessarily obtain greater F1-measure.

Generally, classes of activity that have the same local domain such as *Napping-Eating Meal*, *Housekeeping-Meal preparation*, *Enter Home-Leave Home* are tough to discriminate, due to their locally similar patterns. Furthermore, no specific place or description has been considered as a tutorial to do activities. Therefore in some cases, activities have been done in different ways or places. For instance, resident Eats his Meal on the couch which causes conflict between Eating and Napping activity, while he was supposed to eat his meal on the eating table. It is worth noting that, taking into consideration time of the day or other statistical features can compensate local overlapping as it does for *Personal Hygiene* and *Bed to Toilet transition*.

## 6 Conclusions and future work

Robust online activity recognition in the domain of smart environments is considered as one of the most dominant challenges. While most of the previous approaches suffer from window size parameter tuning, this paper has highlighted a new activity recognition approach on sensor data stream which determines the beginning and end of the activities. In addition, our proposed model is capable of recognizing interrupted activities. The experimental results of this study indicate the efficiency of the proposed model compared to the existing ones.

In general, the accuracy of the activity recognition methods cannot exceed a certain amount due to the uncertainty of human behavior. Nevertheless, uncertain predictions can be offered through a recommendation context in smart environments.

Further work needs to be performed to achieve improvement on recognition of *Other* class as the most problematic activity class. Current recognition systems require a fair amount of labeled data to reach a certain satisfactory accuracy while data acquisition and labeling in this domain is impractical. Future studies on the current topic are therefore recommended in the domain of Knowledge Transfer to reduce the crucial need of labeled data.

## References

- Al Machot F, Mayr HC, Ranasinghe S (2016) A windowing approach for activity recognition in sensor data streams. In: 2016 Eighth international conference on ubiquitous and future networks (ICUFN). IEEE, pp 951–953
- Al Machot F, Mosa AH, Ali M, Kyamakya K (2018) Activity recognition in sensor data streams for active and assisted living environments. *IEEE Trans Circuits Syst Video Technol* 28(10):2933–2945
- Cook D, Schmitter-Edgecombe M, Crandall A, Sanders C, Thomas B (2009) Collecting and disseminating smart home sensor data in the CASAS project. In: Proceedings of the CHI workshop on developing shared home behavior datasets to advance HCI and ubiquitous computing research, pp 1–7
- Cook DJ (2012) Learning setting-generalized activity models for smart spaces. *IEEE Intell Syst* 27(1):32–38
- Cook DJ, Youngblood M, Heierman EO, Gopalratnam K, Rao S, Litvin A, Khawaja F (2003) Mavhome: an agent-based smart home. In: Proceedings of the first IEEE international conference on pervasive computing and communications, 2003. (PerCom 2003). IEEE, pp 521–524
- Cook DJ, Krishnan NC, Rashidi P (2013) Activity discovery and activity recognition: a new partnership. *IEEE Trans Cybern* 43(3):820–828

- Espinilla M, Medina J, Hallberg J, Nugent C (2018) A new approach based on temporal sub-windows for online sensor-based activity recognition. *J Ambient Intell Humaniz Comput*. <https://doi.org/10.1007/s12652-018-0746-y>
- Fleury A, Noury N, Vacher M (2009) Supervised classification of activities of daily living in health smart homes using SVM. In: Annual international conference of the IEEE engineering in medicine and biology society, 2009. EMBC 2009. IEEE, pp 6099–6102
- Gayathri K, Elias S, Ravindran B (2015) Hierarchical activity recognition for dementia care using Markov logic network. *Pers Ubiquitous Comput* 19(2):271–285
- Gjoreski H, Roggen D (2017) Unsupervised online activity discovery using temporal behaviour assumption. In: Proceedings of the 2017 ACM international symposium on wearable computers. ACM, pp 42–49
- Hao J, Bouzouane A, Bouchard B, Gaboury S (2018) Activity inference engine for real-time cognitive assistance in smart environments. *J Ambient Intell Humaniz Comput* 9(3):679–698
- Helaoui R, Niepert M, Stuckenschmidt H (2011) Recognizing interleaved and concurrent activities using qualitative and quantitative temporal relationships. *Pervasive Mobile Comput* 7(6):660–670
- Hong X, Nugent CD (2013) Segmenting sensor data for activity monitoring in smart environments. *Pers Ubiquitous Comput* 17(3):545–559
- Hoque E, Stankovic J (2012) Aalo: Activity recognition in smart homes using active learning in the presence of overlapped activities. In: 2012 6th international conference on pervasive computing technologies for healthcare (PervasiveHealth). IEEE, pp 139–146
- Intille SS, Larson K, Beaudin J, Tapia EM, Kaushik P, Nawyn J, McLeish TJ (2005) The placelab: a live-in laboratory for pervasive computing research (video). In: Proceedings of PERVASIVE 2005 video program
- Jalal A, Kamal S (2018) Improved behavior monitoring and classification using cues parameters extraction from camera array images. *Int J Interact Multimed Artif Intell* 5(7):1–8
- Jurek A, Nugent C, Bi Y, Wu S (2014) Clustering-based ensemble learning for activity recognition in smart homes. *Sensors* 14(7):12285–12304
- Kabir MH, Hoque MR, Thapa K, Yang SH (2016) Two-layer hidden Markov model for human activity recognition in home environments. *Int J Distrib Sens Netw* 12(1):4560365
- Kasteren T, Englebienne G, Kröse B (2010) An activity monitoring system for elderly care using generative and discriminative models. *Pers Ubiquitous Comput* 14(6):489–498
- Katz S, Downs TD, Cash HR, Grotz RC (1970) Progress in development of the index of ADL. *The Gerontologist* 10(1–Part-1):20–30
- Kidd CD, Orr R, Abowd GD, Atkeson CG, Essa IA, MacIntyre B, Mynatt E, Starner TE, Newstetter W (1999) The aware home: a living laboratory for ubiquitous computing research. In: International workshop on cooperative buildings. Springer, pp 191–198
- Krishnan NC, Cook DJ (2014) Activity recognition on streaming sensor data. *Pervasive Mobile comput* 10:138–154
- Kröse B, Van Kasteren T, Gibson C, Van Den Dool T et al (2008) Care: context awareness in residences for elderly. In: International conference of the international society for gerontechnology. Pisa, Tuscany, Italy, pp 101–105
- Li M, O’Grady M, Gu X, Alawlaqi MA, O’Hare G, et al (2018) Time-bounded activity recognition for ambient assisted living. In: IEEE transactions on emerging topics in computing
- Malazi HT, Davari M (2018) Combining emerging patterns with random forest for complex activity recognition in smart homes. *Appl Intell* 48(2):315–330
- Mozer MC (1998) The neural network house: an environment that adapts to its inhabitants. In: Proceedings AAAI spring symposium intelligent environments, vol 58
- Nazerfard E (2018) Temporal features and relations discovery of activities from sensor data. *J Ambient Intell Humaniz Comput*. <https://doi.org/10.1007/s12652-018-0855-7>
- Nef T, Urwyler P, Büchler M, Tarnanas I, Stucki R, Cazzoli D, Müri R, Mosimann U (2015) Evaluation of three state-of-the-art classifiers for recognition of activities of daily living from smart home ambient data. *Sensors* 15(5):11725–11740
- Nguyen NT, Phung DQ, Venkatesh S, Bui H (2005) Learning and detecting activities from movement trajectories using the hierarchical hidden Markov model. In: IEEE computer society conference on computer vision and pattern recognition, 2005, CVPR 2005, vol 2. IEEE, pp 955–960
- Okeyo G, Chen L, Wang H, Sterritt R (2014) Dynamic sensor data segmentation for real-time knowledge-driven activity recognition. *Pervasive Mobile Comput* 10:155–172
- Quero J, Orr C, Zang S, Nugent C, Salguero A, Espinilla M (2018) Real-time recognition of interleaved activities based on ensemble classifier of long short-term memory with fuzzy temporal windows. In: Multidisciplinary digital publishing institute proceedings, vol 2. p 1225
- Riboni D, Szttyler T, Civitarese G, Stuckenschmidt H (2016) Unsupervised recognition of interleaved activities of daily living through ontological and probabilistic reasoning. In: Proceedings of the 2016 ACM international joint conference on pervasive and ubiquitous computing, ACM. pp 1–12
- Sánchez D, Tentori M, Favela J (2008) Activity recognition for the smart hospital. *IEEE Intell Syst* 23(2):50
- Singh D, Merdivan E, Psychoula I, Kropf J, Hanke S, Geist M, Holzinger A (2017) Human activity recognition using recurrent neural networks. In: International cross-domain conference for machine learning and knowledge extraction. Springer, pp 267–274
- Sun FT, Yeh YT, Cheng HT, Kuo C, Griss M, et al (2014) Nonparametric discovery of human routines from sensor data. In: 2014 IEEE international conference on pervasive computing and communications (PerCom). IEEE, pp 11–19
- Szewczyk S, Dwan K, Minor B, Swedlove B, Cook D (2009) Annotating smart environment sensor data for activity learning. *Technol Health Care* 17(3):161–169
- Wan J, O’Grady MJ, O’Hare GM (2015) Dynamic sensor event segmentation for real-time activity recognition in a smart home context. *Pers Ubiquitous Comput* 19(2):287–301
- Wen J, Zhong M, Wang Z (2015) Activity recognition with weighted frequent patterns mining in smart environments. *Expert Syst Appl* 42(17–18):6423–6432
- Yala N, Fergani B, Fleury A (2015) Feature extraction for human activity recognition on streaming data. In: 2015 international symposium on innovations in intelligent systems and applications (INISTA). IEEE, pp 1–6
- Yala N, Fergani B, Fleury A (2017) Towards improving feature extraction and classification for activity recognition on streaming data. *J Ambient Intell Humaniz Comput* 8(2):177–189
- Yan S, Liao Y, Feng X, Liu Y (2016) Real time activity recognition on streaming sensor data for smart environments. In: 2016 international conference on progress in informatics and computing (PIC). IEEE, pp 51–55
- Zhang T, Wong KBY, Aghajan H (2014) Concurrent activation events based trajectory propagation in smart environments. *J Ambient Intell Human Comput* 5(6):867–880
- Zhang T, Fu W, Ye J, Fischer M (2017) Learning movement patterns of the occupant in smart home environments: an unsupervised learning approach. *J Ambient Intell Humaniz Comput* 8(1):133–146