Evaluation of Wearable Sensor Tag Data Segmentation Approaches for Real Time Activity Classification in Elderly

Roberto Luis Shinmoto Torres(✉), Damith C. Ranasinghe, and Qinfeng Shi

Auto-ID Lab, School of Computer Science,
The University of Adelaide South Australia, Adelaide 5005, Australia
{roberto.shinmototorres,damith.ranasinghe,javen.shi}@adelaide.edu.au

Abstract. The development of human activity monitoring has allowed the creation of multiple applications, among them is the recognition of high falls risk activities of older people for the mitigation of falls occurrences. In this study, we apply a graphical model based classification technique (conditional random field) to evaluate various sliding window based techniques for the real time prediction of activities in older subjects wearing a passive (batteryless) sensor enabled RFID tag. The system achieved maximum overall real time activity prediction accuracy of 95 % using a time weighted windowing technique to aggregate contextual information to input sensor data.

Keywords: Conditional random fields · RFID · Feature extraction

1 Introduction

The development of accurate human activity recognition methods is a growing field of study as many applications can be derived from this base. One application is the mitigation of high falls risk activities of older people in hospitals or age care facilities, as falls events occur especially in the bedroom [1]. A correct recognition of such high risk events can lead to an intervention to mitigate an event that can potentially cause further physical injury and mental distress [12]. For high falls risk mitigation the accurate recognition of real time activities is paramount as most falls occur during transfer activities, which are changes of static activities or locations (e.g. sit to stand, stand to sit or ambulating) [1,14]. In this article we consider activity recognition in the context of identifying high falls risk related activities.

Recent work on real time recognition of events have succeeded using body sensors [2,4,8,9,13,17], using tri-axial accelerometers, magnetometers and gyroscopes. These studies along with research using video images [11,18] or environment sensors [7,10] required the use of different time or data (number of

This research was supported by a grant from the Hospital Research Foundation (THRF) and the Australian Research Council (DP130104614).

I. Stojmenovic et al. (Eds.): MOBIQUITOUS 2013, LNICST 131, pp. 384–395, 2014.
DOI: 10.1007/978-3-319-11569-6_30

samples, pixels) based segmentation approaches to extract relevant information for data classification.

The difficulty for real time recognition of activities using sensors is that individual sensor readings are limited in time-space and by themselves provide little additional information about the related activity for the classifier to predict an activity correctly. In order to provide further contextual information to data collected [6], incoming sensor data stream is segmented for feature extraction prior to evaluation.

Using data segments is not precise as there is no pre-defined window size and sizes may differ depending on the application [4,8] or sensor platform used [2,8]. In addition, a passive sensor's data stream is not continuous and data collected can be incomplete or noise distorted. These factors can influence the information quality of the individual data segment. Moreover, the occurrence and information value of future readings are uncertain and the classifier relies exclusively on current and past information to emit a result. The classifier performance is also affected by the presence of data from unrelated activities (e.g. distant past activities' data or sensor readings from unrelated areas or activities) and data from unrelated activities may outweigh current activity information in a given data segment [7].

This paper describes several sliding window segmentation methods for a real time per sensor datum prediction of human activities from a sensor and ID data stream from a battery-less wearable radio frequency identification (RFID) platform, called W^{2}ISP [5]. Our main focus is the implementation and evaluation of the effectiveness of segmentation methods using a multi-class classifier to identify incoming activity data in real time. We used a conditional random field (CRF) classifier because of its desirable sequence dependency modelling capabilities. The main contributions of this study are: (i) development of a real time CRF based classifier for activity recognition of passive sensor and ID data streams; (ii) implementation and testing of multiple fixed and dynamic sized windowing methods for contextual information extraction based on data characteristics; and (iii) experimental demonstration of the accuracy of these methods using data gathered from a trial with older subjects (66–86 years old) in a clinical environment.

The rest of the paper is described as follows: Sect. 2 gives a brief overview of related work, Sect. 3 discusses the methodology for our windowing approaches. Section 4 shows the evaluation results from the various methods in the previous Section and finally conclusions are given in Sect. 5.

2 Related Work

There are several studies with as many approaches for the real time recognition of human activities using threshold and machine learning classification systems. Some real time results [4,8,13] imply the timely recognition of body positions or postures, however, these methods required a data buffer from which the classifier makes a prediction. Hence, results are available periodically rather than having

an output per individual sensor reading; some studies considered overlapping of data to provide faster output while having larger data buffers [8]. A study by Wang et al. [17] used a 1 s sliding window and a smartphone processing platform producing a recognition delay (time elapsed from ground truth to prediction of the ground truth) of $\sim$ 5.7 s. Other real time smartphone based studies relied on the sensors embedded in the device [9] or were used as a hub for other worn sensors [2]. Current smartphones are not imperceptible devices and their usage with older populations needs to be studied.

All these methods used different approaches to segmenting and windowing data. Most tried empirically different segment sizes to find that which maximized the resulting accuracy using the same set of features [8,17] or the window selection was arbitrary [9]; while others were limited by the underlying technology itself [4]. In addition, most of these studies produced periodical results only and used battery operated sensors which are bulky and not appropriate for older or frail subjects. Furthermore, none has been evaluated on an older population.

In the work of Krishnan et al. [7], several methods were applied to evaluate the best sliding window method for smart home data sets. Each method provided extra features for added information about the last received sample. The nature of irregular and incomplete data from environment sensors in the smart home is similar to that of passive worn sensors in an RFID platform, as is our case. This research study implements methods adapted from [3,7] to the ID and sensor data stream from sensor enabled passive RFID devices to evaluate time series data segmentation approaches for real time classification of scripted activities from older people.

3 Methodology

In this section, we present the developed windowing methods for feature extraction and describe the datasets and classification system used. These methods are based on the passive sensor platform (W^{2}ISP) constraints where signal collection is irregular, noisy and incomplete. Our RFID deployments used antennae location to obtain readings from targeted areas of high falls risk activities (in and around a bed and a chair) in two clinical rooms (Sect. 3.7).

Using a tri-axial acceleration (ADXL330) data stream from a W^{2}ISP we predicted the activity label (Sect. 3.7) that best represented every datum received. Our feature vector included: $V = [a_f, a_v, a_l, sin(\theta), RSSI, A_{ID}, \triangle T]$ for the recognition of bed exits, where a_f, a_v and a_l are frontal, vertical and lateral components of the tri-axial accelerometer sensor, $sin(\theta)$ refers to the body tilt angle, $RSSI$ is the strength (power) of the signal received from the W^{2}ISP by an RFID reader, $A_{ID} = \{aID_1, .., aID_A\}$ (where A is the number of antennae) is the identifier of the antenna that collected the datum and $\triangle T$ is the time difference between current and previous reading [15]. In [15] we obtained high accuracy for label detection using batch processing of activities sequences segmented by trial and by patient with a CRF classifier. In this study, we use CRF for real time activity recognition and apply different windowing techniques for feature extraction and predict the activity label of the last received sensor reading.

3.1 Activity Windowing (AW)

This approach considers that the system knows in advance the activity that is being performed and segments the data per each activity for both training and testing stages. Although knowing the activities performed beforehand in a real environment is implausible given that the separation of activities is unambiguous, maximum accuracy is expected from the predicted labels. Given this condition, we consider this technique our golden standard. However, this approach does not perform predictions in real-time but predicts samples in different sized pre-segmented batches where each batch represents a single activity. We use the generic set of features V for each observed sensor reading (Sect. 3) as input to the classifier.

3.2 Fixed Sample Windowing (FSW)

This approach considers a sliding window with fixed number of samples. The windowed sample sequence provides contextual information about the last sample in the window to enable the classifier to emit a more accurate prediction [7]. Different window sizes can better fit the duration of different activities (labels) as was the case in Sect. 3.1 when the activities are already known. To illustrate this case, consider resting positions such as lying or sitting on bed or the chair that usually last several minutes or hours when compared to dynamic activities such as walking. The lengths of such events are disproportionate and difficult to segment in real data; whereas a fixed sample segmentation is simple to produce.

The selection of window size is an empirical process, where the best result corresponds to the segment length that best fits all activities. For this analysis additional features are extracted and added to our generic feature vector V as contextual information. In contrast to [7] the set of extra features corresponds to the count of events reported by each antenna in the window, as we can differentiate sensor reading origin by the antenna used. Hence, the number of extra features is equal to the number of antennae present. Moreover, these summed amounts are further normalized to four levels of importance (0: unused, 1: low, 2: medium and 3: high importance) computed as follows:

$$F(i,k) = \left\lceil \frac{3 * c_{i,k}}{\sum_{j=1}^{A} c_{j,k}} \right\rceil \tag{1}$$

where A is the number of antennae and $c_{i,k}$ corresponds to the count of antenna i events in the window for the last reading k.

Two issues are present in this method. First, the window duration can span a long time and readings from distant past activities can affect the classifier decision. Second, a large volume of readings from previous activities unrelated to the current activity or spurious readings from distant antennae covering a distinct area present in the window can also alter the classifier decision. In order to meet these issues, weighted features are considered to balance the influence of unrelated data [7] and are described in Sects. 3.3 and 3.4.

3.3 Time Weighted Windowing (TWW)

This method is based on FSW in Sect. 3.2, as it uses a fixed sample window. However, more importance is given to events closer to the last sample to reduce the effect of historical events on the classifier. This technique gives each sample i in a window a distinct weight $T(i,k)$ which is a function of the time difference $\triangle_{i,k}$ between the last received sample k and sample i in the fixed time window. The evaluation of the weights is defined by an exponential function:

$$T(i,k) = \begin{cases} 0 & \text{if } \triangle_{i,k} >= \frac{1}{D} \\ \exp\left(-D * \triangle_{i,k}\right) & \text{else} \end{cases} \tag{2}$$

where D is the rate of decay, if $D > 1$ elements very close ($\triangle_{i,k}$ <1 s) are given priority as the exponential function decays quickly; smaller values of D allows the function to consider a wider time range of sensor readings. The method also considers a limited amount of time, bounded by $\triangle_{i,k} = \frac{1}{D}$, as larger values of $\triangle_{i,k}$ provide less weight. The extra features now consist of the sum of the weights for the readings corresponding to each antenna in the window and replaces the extra features from FSW method. Hence, the extra features are defined by the vector: $W = [\sum T(i,k)\delta(a_i, aID_1), \ldots, \sum T(i,k)\delta(a_i, aID_A)]$, where $a_i \in A_{ID}$ is the antenna corresponding to the i^{th} sample and the function $\delta(a_m, a_n)$ is defined as:

$$\delta(a_m, a_n) = \begin{cases} 0 & \text{if } a_m \neq a_n \\ 1 & \text{if } a_m = a_n \end{cases} \tag{3}$$

In addition, we normalize vector W to levels of importance using (1).

3.4 Mutual Information Windowing (MI)

Similarly to TWW (see Sect. 3.3), this method uses a fixed sample window (see Sect. 3.2). However, this approach intends to reduce the influence of readings from antennae focused on areas unrelated to the current activity. In general, samples of different activities are collected from antennae depending on the activity location. Nevertheless, readings from distant antennae occur in real data with low received energy ($RSSI$), although these readings are rare. We consider two types of mutual information between the i^{th} sample and the last sample k in a segmented window: (i) MI1: mutual information (MI) of two consecutive readings occurring from any pair of antennae as defined in [7] and given in (4); and (ii) MI2: the MI of two consecutive readings occurring from a given pair of antennae at any time (i.e. disregards the order of antennae occurrence while focusing only on antenna relationships), defined in (5), where N is the number of elements in the training sequence, $a_i \in A_{ID}$ and function $\delta(.)$ is as defined in (3).

$$MI1(m,n) = \frac{1}{N} \sum_{j=1}^{N-1} \delta(a_j, a_m)\delta(a_{j+1}, a_n) \tag{4}$$

$$MI2(m,n) = \frac{1}{N} \sum_{j=1}^{N-1} \left(\delta(a_j, a_m)\delta(a_{j+1}, a_n) + \delta(a_j, a_n)\delta(a_{j+1}, a_m)\right) \tag{5}$$

The mutual information is built from the entire training data set prior to testing where a square matrix $(A \times A)$ with all possible antennae pairs and a triangular matrix is obtained for MI1 and MI2 respectively. These MI weights are used to build the extra features, where all sensor readings in a window are weighted in relation to the last reading and summed in relation to their corresponding antenna, obtaining the vector $W = [\sum MIr(i,k)\delta(a_i, aID_1), \ldots, \sum MIr(i,k)\ \delta(a_i, aID_A)]$ where $r = \{1, 2\}$. Vector W is normalized to levels of importance using (1).

3.5 Dynamic Windowing (DW)

This method considers a time based window of varying size using statistical properties of the data to continually adapt the window size. This method, first devised by Jeffery et al. [3] for cleaning of RFID data streams, was applied because the algorithm balances the need to provide contextual information by increasing the window size and reducing the window size when sensor readings become more sporadic. This method considers a stepped window size increments (0.5 s per sample) but reduction is rapid (halving the window size) when required [3]. We assume a standard epoch[1] duration of 0.25 s and sensor observation probability of 90 % [3]. For this method we use the extra features of the FSW method (see Sect. 3.2).

3.6 Fixed Time Windowing (FTW)

This method considers a sliding window of fixed time duration T^*, as opposed to a dynamic changing time window size as in DW. All readings within the time interval T^* from the last received reading k are considered in a window. Given the irregular collection of data (due to the nature of the passive device), the number of samples per segment will differ. For this method we use the extra features of the FSW method (see Sect. 3.2).

3.7 Datasets

We used data from two clinical room deployments as described in [15]; where both datasets (*RoomSet1* and *RoomSet2*) used four and three antennae respectively. In *RoomSet1* one antenna is placed on top of the bed (ceiling) and three on the walls focusing on the chair and around the bed providing a wide area of coverage. In *RoomSet2* two antennae were placed on top of the bed (ceiling) and one antenna focused on the chair. Fourteen older subjects were trialled (age:74.6 ± 4.9), wearing the W^2ISP on top of their garments. They performed a series of scripted activities which included: (i) lying (in bed); (ii) sitting in bed; (iii) sitting in chair; and (iv) ambulating. The CRF classifier was used to predict these four activity labels. In order to collect the ground truth, activities were annotated in real time by an observer. The same basic features were extracted

[1] Epoch refers to a group of RFID interrogation cycles.

for each dataset; however, the total number of features differed for both datasets as the aggregated extra features are based on their antennae deployment.

3.8 Classifier

In this research, we used a linear chain CRF, a model for structured classification (prediction) [16], as in a previous study [15]. We have preferred this method as it models the dependencies of activities in a sequence. Due to this advantage the system is trained using the complete sequence of activities of each training trial for parameter estimation with the exception of the first method (AW) as it assumes activities are previously known and independent from each other. In CRF, training and testing stages require the probabilistic inference of the target labels (hidden variables). In general, the inference process allows us to obtain: (i) marginal probabilities of the labels (using sum-product algorithm); and (ii) the most probable global state of all our hidden variables i.e. maximum a posteriori (MAP) assignment (using max-product algorithm). Our application requires real time response, thus we use the sum product algorithm (which only propagates the messages forwards) to find the marginal probability of the current hidden variable given the past and current observation efficiently. The prediction is done by maximizing the marginal probability.

The sum product algorithm propagates messages for every edge (i, j) connecting nodes i and j in a graph. In Fig. 1, circles represent the state variables $Y = \{y_1, \ldots, y_k\}$ and squares represent factors (node and edge potentials). The message updating and marginal probability are computed as follows:

$$m_{i,k}(s_k) = \sum_{s_i} (\psi(s_i)\psi(s_i, s_k) \prod_{t \in N_i \backslash \{k\}} m_{t,i}(s_i)) \tag{6}$$

$$p(s_k) = \tfrac{1}{Z}\psi(s_k) \prod_{i \in N_k} m_{i,k}(s_k) \tag{7}$$

where $\psi(s_i)$ and $\psi(s_i, s_k)$ are node and edge potential respectively, s_i represents node i, $N_i \backslash \{k\}$ represents the set of neighbours of node i with the exception of node k, p is the marginal probability and Z is a normalization factor. In the case of our real time application, we are only interested in the marginal probability of the last variable (y_k) given the input observation as shown in Fig. 1, where the marginal probability of variable y_k reduces to:

$$p(y_k \mid x_{1:k}, \lambda) = \tfrac{1}{Z(\lambda, x_{1:k})}\, \psi(y_k; x_{1:k}, \lambda) m_{k-1,k}(y_k; x_{1:k}, \lambda) \tag{8}$$

where messages $m_{i,j}$ are calculated using (6) and $\lambda = \{\lambda_t\}$ represents the parameters estimated in training. Moreover, the message $m_{k-1,k}(y_k; x_{1:k}, \lambda)$ is recursive as it depends on previous messages as in the expression $m_{k-1,k}(y_k; x_{1:k}, \lambda) = \sum_{y_{k-1}} \psi(y_{k-1}; x_{1:k}, \lambda)\psi(y_{k-1}, y_k; x_{1:k}, \lambda) m_{k-2,k-1}(y_{k-1}; x_{1:k-1}, \lambda)$.

This derivation for the sum product is appropriate for real time streaming data as the sequence of information is always increasing and an activity label prediction is required for each datum. Therefore, we apply the expression in (8) for inference during testing.

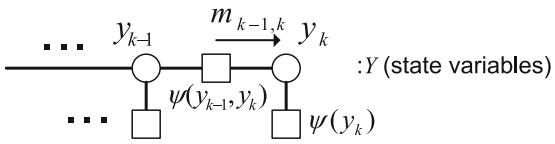

Fig. 1. Message propagation for the probability distribution of y_k

3.9 Statistical Analysis

The analysis of results was obtained using a 10-fold cross validation where general performance was measured using overall accuracy (referred as accuracy and given in (9)), and individual label performance using geometric mean (GM) and $Fscore$ as defined in (9), where N is the number of readings, TP is true positives and recall, specificity and precision are determined as per standard definitions. Results are shown as mean ± standard deviation (SD). Statistical significance is measured using a two-tailed two-sample t-test at 5 % significance level.

$$Accu = \frac{TP}{N} \quad GM = \sqrt{recall.specificity} \quad Fscore = \frac{2.recall.precision}{recall + precision} \tag{9}$$

4 Evaluation

The first evaluation corresponds to labelling a pre-segmented sequence in the AW method. High accuracies ($> 97.7\,\%$) and high $Fscore$ and GM values ($> 87\,\%$ and $> 92\,\%$) are obtained for *RoomSet1* and *RoomSet2* datasets (see Table 2). In *RoomSet*2, metrics for sitting-in-chair are affected by one test fold where only 17 samples were present for that activity which was missed (false negative), affecting all metrics. Accuracies for the FSW method are shown in Table 1(a), which tested sliding windows of 5 to 60 samples. Highest accuracy for *RoomSet*1 is achieved between $N = 10$ and 20; the largest source of error is caused by false negatives (FN) of sitting-in-bed and false positives (FP) in ambulating labels. These errors are minimal in AW approach. *RoomSet*2 is affected by one fold where samples collected during lying and sitting-in-bed caused mutual errors, which also affected the rest of the methods. Given that *RoomSet*2 metrics do not vary in this set of window sizes, we consider the window sample size of 15 as the best parameter for this method.

The TWW method was tested using a fixed window size of 15 samples as found in FSW method. For this study we tested decay rate values of $D = 2^{-7}$ to 2^0 as shown in Table 1(b). Highest accuracy for *RoomSet*1 occurs for $D = 2^{-2}$ and then drops slightly. In contrast, best performance for *RoomSet*2 occurs between $D = 2^{-1}$ to 2^{-3} with some folds affected by mutual error between lying and sitting-in-bed as in FSW method. However, accuracy is not significantly different across *RoomSet*2 ($p > 0.34$). Hence, we consider a decay rate of $D = 2^{-2}$ as the best parameter for this windowing case.

Results for MI windowing indicates that both MI1 and MI2 approaches (see Sect. 3.4) obtained similar results as shown in Table 2, where results are almost

identical. From the MI1 matrices shown below, we can see that for *RoomSet*1 there is little interaction between antennae, with higher values for self transition of antennae as found in [7]; in *RoomSet*2 there are strong transition values for antennae aID_2 and aID_3 both of which are located on top of the bed and report most in-bed and around-bed sensor readings.

MI1*RoomSet1* (%)

	aID_1	aID_2	aID_3	aID_4
aID_1	35.12	0.25	2.18	3.28
aID_2	0.22	8.77	0.04	0.06
aID_3	2.23	0.03	16.48	4.33
aID_4	3.24	0.04	4.36	19.26

MI1*RoomSet2* (%)

	aID_1	aID_2	aID_3
aID_1	3.19	0.16	0.31
aID_2	0.17	21.92	18.15
aID_3	0.4	18.14	37.42

The DW method produced high accuracy (94.6 %) results for *RoomSet*2 but for *RoomSet*1 results are comparable to those of previous windowing methods as shown in Table 2. However, $Fscore$ and GM are significantly different compared to MI ($p < 0.01$). Partial results for FTW are shown in Table 1(c), where time durations of 1 to 128 s were tested. The highest accuracies were achieved with a 4 s time window for both datasets. In *RoomSet*2, there is little variation among different time windows. In contrast, for *RoomSet*1 this method achieves the highest accuracy for all tested methods as seen in Table 2 but this result is still significantly different from that of the golden standard ($p \sim 10^{-7}$).

Finally, we combined time and space based segment contextual information extraction with expected performance improvement. Combination of FTW and MI1, introduced mutual information rather than counting events in FTW. The results for *RoomSet*1 lie between the amalgamated methods; however, the overall performance for *RoomSet*2 is lower than FTW and MI1. The combination case of TWW and MI1 introduced extra features from both methods, achieving the highest accuracy for *RoomSet*2 which is comparable with our golden standard ($p = 0.14$). Results for *RoomSet*1 are slightly lower than the best performance with FTW method (see Table 2). This is because TWW+MI1 had a rich extended information. However, in FTW+MI1 the fixed time window did not bring enough mutual information as the number of samples in a segment can be as low as one, performing lower than counting per antenna samples.

Further analysis of Table 2 indicates low $Fscore$ and GM values for *RoomSet*1 created mostly by FNs and FPs (false positives) for sitting-in-bed and ambulating respectively. These errors were reduced in the FTW case but not greatly. Metrics for *RoomSet*2 were affected by two main causes, one fold in particular produced large FN and FP for lying and sitting-in-bed respectively and few readings for sitting-in-chair label which in some folds are ignored produced recall, Fscore and GM values of zero and affected these averages. In both datasets, the low composition of ambulating activity data (not shown) compared to the other states affected individual and average metrics as only a couple of seconds of data are retrieved as a subject walked away from the reading range of antennae near the bed or chair.

These results indicate the real time inference algorithm (marginal inference) found most difficulty predicting ambulating and sitting-in-bed samples

for *RoomSet*1. In general, no method was able to produce maximum results for both datasets, although results for FTW and TWW+MI1 methods are comparable ($p > 0.48$) in both datasets. *RoomSet*2 disposition using two antennae focusing on the bed was able to avoid classification error in *RoomSet*1, however there were relatively smaller number of sensor readings for sitting-in-chair as only one antenna powers the tag in *RoomSet*2. In terms of real time analysis, test inference calculation time for both datasets is of 5.12 µs and 7.31 µs per sample respectively and this time is minimal compared with the observed minimum inter-sample duration time of 25 ms. These results were obtained using algorithms implemented in MATLAB scripts and mex code but these algorithms will run faster if developed in a low level language as C/C++ and therefore our results also demonstrate that the classifier is capable of real time sample prediction.

Table 1. Partial accuracies for FSW, TWW and FTW methods

(a) Accuracy for Fixed Sample Window method						
Datasets	N = 5	N = 10	N = 15	N = 20	N = 30	N = 60
RoomSet1	70.6±6.0 %	71.2±6.1 %	72.5±5.9 %	72.2±6.2 %	70.6±6.2 %	70.7±5.9 %
RoomSet2	94.9±4.3 %	91.9±11.5 %	91.8±11.2 %	91.7±11.3 %	91.2±11.6 %	93.5±6.9 %
(b) Accuracy for Time Weighted Window method						
Datasets	$D = 2^0$	$D = 2^{-1}$	$D = 2^{-2}$	$D = 2^{-3}$	$D = 2^{-4}$	$D = 2^{-7}$
RoomSet1	70.3±6.1 %	71.6±6.1 %	73.1±7.5 %	71.8±6.2 %	71.7±6.0 %	71.8±6.1 %
RoomSet2	91.3±11.2 %	94.3±5.0 %	91.7±11.2 %	93.5±4.6 %	92.1±11.1 %	90.2±10.8 %
(c) Accuracy for Fixed Time Window method						
Datasets	$T = 1$	$T = 2$	$T = 4$	$T = 8$	$T = 16$	$T = 128$
RoomSet1	70.5±6.2 %	73.9±7.1 %	78.2±7.3 %	73.6±6.8 %	71.9±6.7 %	70.2±5.7 %
RoomSet2	91.5±11.3 %	91.6±11.4 %	92.3±11.2 %	92.2±11.1 %	92.2±11.0 %	91.2±11.0 %

Table 2. Classification accuracy for all tested methods for both datasets, including average Fscore and GM for all activities.

	RoomSet1 (%)			*RoomSet2 (%)*		
Method	*Accuracy*	*Fscore*	*GM*	*Accuracy*	Fscore	GM
AW	98.1±1.8	93.5±5.5	96.1±3.7	97.7±3.6	87.0±21.0	92.8±16.6
FSW	72.5±6.0	57.9±6.5	76.6±5.5	91.8±11.2	67.5±23.2	82.5±17.7
TWW	73.1±7.5	59.0±9.2	77.1±8.0	91.7±11.2	69.3±21.4	84.4±18.0
MI1	70.8±6.0	55.2±5.0	74.0±4.3	94.4±5.2	68.6±20.5	84.3±17.2
MI2	70.8±6.1	55.3±5.4	74.1±4.9	93.8±5.2	67.5±20.4	83.3±16.8
DW	74.7±8.1	61.2±10.2	79.2±8.4	94.6±4.7	68.4±22.7	83.4±18.6
FTW	**78.2 ± 7.3**	**65.1 ± 11.5**	**82.1 ± 9.2**	92.3±11.2	68.5±23.9	82.8±18.7
FTW+MI1	71.5±6.0	56.6±5.6	75.2±4.8	91.7±11.1	67.1±22.9	82.9±17.8
TWW+MI1	77.1±7.8	63.8±11.5	81.0±9.1	**95.0 ± 4.2**	**71.6 ± 20.2**	**85.8 ± 16.8**

5 Conclusions

In this study we have developed a number of sliding window based data segmentation techniques for real time prediction of human activities where contextual information was introduced as extra features to the input observation to improve classification accuracy. Although no segmentation method exceeded the golden standard for both datasets, methods TWW+MI1 and FTW achieved high performance metrics in both datasets and had comparable results with the golden standard with *RoomSet2* dataset. In general *RoomSet2* achieved better results than *RoomSet1*, due to its more focused antennae disposition which caused less prediction errors as opposed to a wider area coverage as in *RoomSet1*. Moreover, the system can process real time streaming data using fixed or variable windowing approaches on a sample by sample basis with high accuracy as in the case of *RoomSet2*; and using a CRF classifier which learned the model using complete sequences of activities and applied into real time label prediction.

A limitation comes from the scripted nature of the activity datasets. Nonetheless, all related sensor worn research (see Sect. 2) were based on scripted set of activities, where execution order was random or sequential. Further analysis is required to determine whether these segmentation techniques based on scripted models can perform well with unscripted and undirected activities.

Finally, this work sets the foundation for high level applications such as high falls risk activities (bed and chair ingress or exit, room exiting and bathroom access) recognition in real time.

References

1. Becker, C., Rapp, K.: Fall prevention in nursing homes. Clin. Geriatr. Med. **26**(4), 693–704 (2010)
2. Györbíró, N., Fábián, A., Hományi, G.: An activity recognition system for mobile phones. Mob. Netw. Appl. **14**(1), 82–91 (2009)
3. Jeffery, S., Franklin, M., Garofalakis, M.: An adaptive rfid middleware for supporting metaphysical data independence. VLDB J. **17**(2), 265–289 (2008)
4. Karantonis, D., Narayanan, M., Mathie, M., Lovell, N., Celler, B.: Implementation of a real-time human movement classifier using a triaxial accelerometer for ambulatory monitoring. IEEE Trans. Inf. Technol. Biomed. **10**(1), 156–167 (2006)
5. Kaufmann, T., Ranasinghe, D.C., Zhou, M., Fumeaux, C.: Wearable quarter-wave microstrip antenna for passive UHF RFID applications. Int. J. Antennas Propag. (2013)
6. Kern, N., Schiele, B., Schmidt, A.: Recognizing context for annotating a live life recording. Pers. Ubiquit. Comput. **11**(4), 251–263 (2007)
7. Krishnan, N.C., et al.: Activity recognition on streaming sensor data. Pervasive Mob. Comput. **10**, 138–154 (2014)
8. Lee, M.W., Khan, A., Kim, J.H., Cho, Y.S., Kim, T.S.: A single tri-axial accelerometer-based real-time personal life log system capable of activity classification and exercise information generation. In: 32nd Annual International Conference IEEE Engineering in Medicine and Biology Society, pp. 1390–1393 (2010)

9. Lee, Y.-S., Cho, S.-B.: Activity recognition using hierarchical hidden markov models on a smartphone with 3D accelerometer. In: Corchado, E., Kurzyński, M., Woźniak, M. (eds.) HAIS 2011, Part I. LNCS, vol. 6678, pp. 460–467. Springer, Heidelberg (2011)
10. Lu, C.H., Fu, L.C.: Robust location-aware activity recognition using wireless sensor network in an attentive home. IEEE Trans. Autom. Sci. Eng. **6**(4), 598–609 (2009)
11. Messing, R., Pal, C., Kautz, H.: Activity recognition using the velocity histories of tracked keypoints. In: IEEE 12th International Conference on Computer Vision, pp. 104–111 (2009)
12. Oliver, D.: Prevention of falls in hospital inpatients. agendas for research and practice. Age Ageing **33**(4), 328–330 (2004)
13. Ranasinghe, D.C., Shinmoto Torres, R.L., Hill, K.D., Visvanathan, R.: Low cost and batteryless sensor-enabled radio frequency identification tag based approaches to identify patient bed entry and exit posture transitions. Gait & Posture (2013)
14. Robinovitch, S.N., Feldman, F., Yang, Y., Schonnop, R., Lueng, P.M., Sarraf, T., Sims-Gould, J., Loughin, M.: Video capture of the circumstances of falls in elderly people residing in long-term care: an observational study. The Lancet **381**(9860), 47–54 (2012)
15. Shinmoto Torres, R.L., Ranasinghe, D.C., Shi, Q., Sample, A.P.: Sensor enabled wearable rfid technology for mitigating the risk of falls near beds. In: 2013 IEEE International Conference on RFID, pp. 191–198 (2013)
16. Sutton, C., McCallum, A.: An introduction to conditional random fields. Found. Trends Mach. Learn. **4**(4), 267–373 (2012)
17. Wang, L., Gu, T., Tao, X., Lu, J.: A hierarchical approach to real-time activity recognition in body sensor networks. Pervasive Mob. Comput. **8**(1), 115–130 (2012)
18. Xia, L., Chen, C.C., Aggarwal, J.: Human detection using depth information by kinect. In: 2011 IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops, pp. 15–22 (2011)