

Research Article

Two-Layer Hidden Markov Model for Human Activity Recognition in Home Environments

M. Humayun Kabir, M. Robiul Hoque, Keshav Thapa, and Sung-Hyun Yang

Department of Electronic Engineering, Kwangwoon University, Seoul 139-701, Republic of Korea

Correspondence should be addressed to Sung-Hyun Yang; shyang@kw.ac.kr

Received 8 August 2015; Revised 3 December 2015; Accepted 7 December 2015

Academic Editor: Miguel A. Patricio

Copyright © 2016 M. Humayun Kabir et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Activities of Daily Livings (ADLs) refer to the activities that are carried out by an individual for everyday living. Recognition of ADLs is key element for building intelligent and pervasive environments. We propose a two-layer HMM to build a ADLs recognition model that can represent the mapping between low-level sensor data and high-level activity based on the binary sensor data. We used embedded sensor with appliances or object to get object used sequence data as well as object name, type, interaction time, and location. In the first layer, we use location data of object used sensor to predict the activity class and in the second layer object used sequence data to determine the exact activity. We perform comparison with other activity recognition models using three real datasets to validate the proposed model. The results show that the proposed model achieves significantly better recognition performance than other models.

1. Introduction

Activity recognition is an important task for many ubiquitous computing applications. In a home environment, activity recognition systems observe the resident's activities and can remind users to perform the missed activities or complete actions, helping them to provide some automated service. In a hospital environment, activity recognition system can remind a doctor or nurse to perform certain tests before operating. In a factory environment, it can monitor the activities of the workers and encourage them to act more safely. Activity recognition systems use several types of sensors, microphones and video cameras, RFID readers, wearable sensors, and embedded sensors with appliances (or object), to determine the state of the physical world. Based on the state of the physical world, activity recognition systems observe the behavior of persons and if necessary take actions in response. Microphones and video camera based activity recognition is complex to implement because it requires processing of multidimensional data and may violate issue of user's privacy [1–6]. It is effective to recognize primitive sequence of movements (walking, sitting, standing,

etc.) but difficult to recognize ADLs (showering, grooming, preparing meals, etc.) using wearable sensors [7–9]. Home users perform an activity by interacting with appliances (or objects) within nearby location at a given time. Appliances or object used information can be gotten by attaching embedded sensor with them [10–13]. Embedded sensors are inexpensive, invisible, and nonintrusive having no impact on regular life activities. Tapia et al. have presented how simple binary sensors have solid potential for solving the ADLs recognition problem in the home [11]. Binary sensors can also be applied in human-centric problems such as health and elder care [14–16]. Van Kasteren et al. use binary sensors and motion sensors for performing ADLs recognition in a house setting [17]. However sensors based activity recognition is challenging due to the inherent noisy nature of the input [18]. In this context the temporal probabilistic reasoning and machine learning approaches are very effective. Several object use based models have been used to recognize ADLs from binary sensors such as Bayesian Networks [11], Conditional Random Field (CRF) [19], Hidden Markov Model (HMM) [17], and Hidden semi-Markov Model [17]. In [20], the authors proposed Triplet Markov Chains (TMC) to deal with

nonstationary HMM by introducing an auxiliary process governing the regime switching of the hidden process.

The location of a person in an environment can provide important context information for activity recognition [21, 22]. Location based model is able to provide high-level abstraction for an activity. For example, bedroom is for sleeping, bathroom for showering, and kitchen for preparing meal. The same location can be used for different activities such as showering, brushing teeth, and toileting as bathroom activities and preparing meal, taking meals, and dishwashing as kitchen activities. So location information based modeling is not suitable for recognizing ADLs. Our aim is to build an activity recognition model that can represent the mapping between low-level sensor data and high-level activity based on the binary sensor data. We want to combine location of object used sensor with the object used data to overcome the limitation for each approach. Using embedded sensors with appliances or object we can get object used data as well as object name, type, interaction time, and location. We propose a two-layer Hidden Markov Model (HMM) for recognition of ADLs. In the first layer, we use location data of object used sensor to predict the activity class and in the second layer object used sequence data to determine the exact activity from the selected class. In our 2-layer HMM, first layer selects one activity class so second layer only considers the activities (variables) which are related to that class only. This approach reduces time complexity of this model. We evaluate and compare the activity recognition performance of the proposed models on three fully annotated real world datasets generated by Van Kasteren [23]. Based on our experiment, 2-layer HMM outperforms other activity recognition models. Using location data with object used data, this model can result in a significant increase in recognition performance.

The rest of the paper is organized as follows: Section 2 presents an overview of the data used in this study. Section 3 illustrates the details of the proposed model. Section 4 provides the experimental settings and results obtained followed by conclusion and future work in Section 5.

2. Dataset

We evaluate the recognition performance of the proposed model and compare it with naïve Bayes, Conditional Random Field (CRF), and Hidden Markov Model (HMM) on three fully annotated real world datasets generated by Van Kasteren et al. Wireless Sensor Networks are used to observe the behavior of inhabitants inside the houses by collecting binary temporal data. Wireless network nodes are equipped with various kinds of sensors: reed switches to measure whether doors and cupboards are open or closed, pressure mats to measure sitting on or lying in bed, mercury contacts to detect the movement of objects, passive infrared (PIR) to detect motion in a specific area, and float sensors to measure the toilet being flushed. Wireless network node sends an event to base station when the state of the digital input changes or

TABLE 1: Details of Kasteren datasets.

	House-A	House-B	House-C
Setting	Apartment	Apartment	House
Rooms	3	2	6
Duration	25 days	15 days	20 days
Sensors	14	23	21
Activities	10	13	16

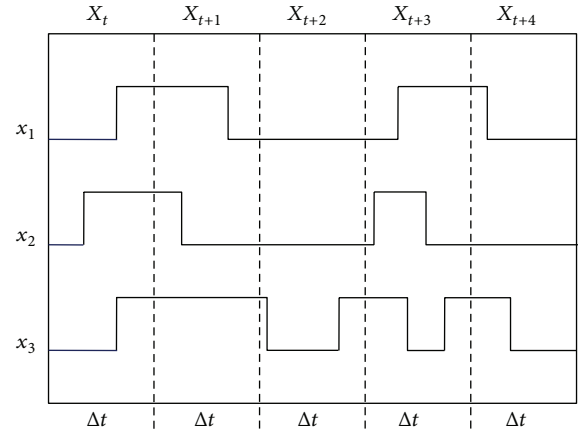


FIGURE 1: Temporal relation between sensor reading x_i and time intervals Δt .

when the analog input crosses some predefined threshold. An overview of the datasets is presented in Table 1.

Time series data is discretized into a set of time slices of constant length Δt . A sensor event for time t is denoted by x_i , indicating whether sensor i fired at least once between time t and time $t + \Delta t$, with $x_i \in \{0, 1\}$. If N sensors are installed in a home, then the observation vector for each time slice will be $X_t = [x_1, x_2, \dots, x_{N-1}, x_N]^T$. Each time slice corresponds to a single data instance during data representation. The class of each data instances is defined by the activity label of the corresponding time slices. The activity at time slice t is denoted by y_t . It is the task to a classifier to find a mapping between a sequence of observations $\Theta = \{X_1, X_2, \dots, X_{T-1}, X_T\}$ and a sequence of labels $Y = \{y_1, y_2, \dots, y_{T-1}, y_T\}$ for total T time intervals as shown in Figure 1.

The activities or labels considered were not the same for all datasets. In the three datasets nineteen different ADLs were included for labels: “Brushing teeth”, “Breakfast”, “Dinner”, “Drink”, “Dressing”, “Dishwashing”, “Eating”, “Leaving house”, “Lunch”, “Medication”, “Preparing Breakfast”, “Preparing Lunch”, “Preparing Dinner”, “Relax”, “Showering”, “Sleeping”, “Shaving”, “Snack”, and “Toileting”. If the activity of a time interval does not match with labeled activity, then activity of the corresponding time interval is referred to as “others”. Table 2 shows percentage (%) of instances in each activity class of the used datasets.

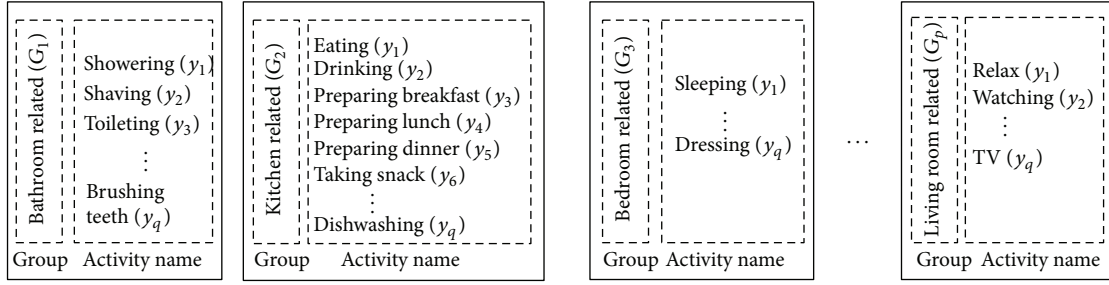


FIGURE 2: Grouping of the activities according to the location of sensor installed with different objects.

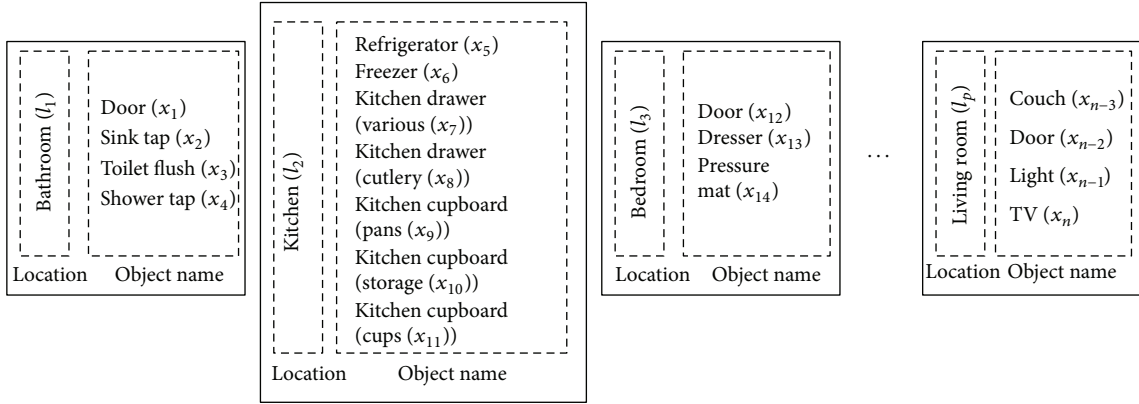


FIGURE 3: Classification of the location of sensor installed with different objects for each activity group.

TABLE 2: Instances of each activity classes in Kasteren dataset measured in percentage (%).

Activity	Kasteren-A	Kasteren-B	Kasteren-C
Breakfast	0.3	0.3	—
Brushing teeth	0.3	0.5	0.3
Dinner	1	1	—
Drink	0.1	0.1	0.1
Dressing	—	0.4	0.4
Dishwashing	—	0.5	—
Eating	—	—	3
Leaving house	50.2	51	46.5
Lunch	—	—	—
Medication	—	—	0.1
Others	13	7	5.1
Preparing Breakfast	—	2	2
Preparing Lunch	—	—	3
Preparing Dinner	—	4	3
Relax	—	—	8
Showering	0.8	0.6	0.6
Sleeping	33.5	32	27
Shaving	—	—	0.1
Snack	0.1	—	0.1
Toileting	0.7	0.6	0.7

3. Proposed Model

We have proposed a 2-layer HMM for recognition of ADLs. We used the locations information of the used object in the first layer and select a group which satisfied maximum joint probability. In the second layer we considered the data of used object and select one activity among the classes. Activities are grouped according to the location of objects used to perform similar activity. Figure 2 shows an example of the grouping of the activity. Sensors are also classified based on the location of the objects which are involved with each activity group. Figure 3 presents an example of classification of the location of sensor installed with different object for each activity group.

We consider $Y = \{y_1, y_2, \dots, y_m\}$ as the set of activities, where m is the total number of activities and these activities are divided into p groups; $G = \{G_1, G_2, \dots, G_p\}$ for all $G_i \subset Y$. Let $X = \{x_1, x_2, \dots, x_n\}$ be the set of sensors installed and $L = \{l_1, l_2, \dots, l_p\}$ the set of locations in an environment on deploying n number of sensors in p number of locations; then the location of sensors is $L_X = \{l_{x_1}, l_{x_2}, \dots, l_{x_n}\}$ for $l_{x_1} \in L$. Therefore $\Theta_L = \{L_{X_1}, L_{X_2}, \dots, L_{X_T}\}$ represents the set of locations of observed sensor sequences associated with the set of sensor observations sequence $\Theta = \{X_1, X_2, \dots, X_T\}$ at a given time T . Each group corresponds to a location.

3.1. First Layer of HMM. In the first layer, the classifier uses the location information of the observed sensors to classify

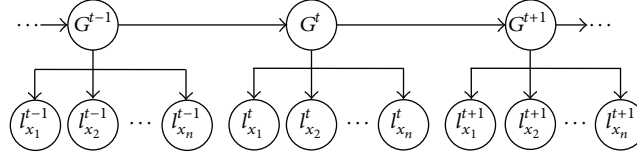


FIGURE 4: The graphical representation of the HMM used in the first layer of the activity model.

the group of activities. The first layer of this HMM is characterized by $\lambda_{s1} = \{\pi, A, B\}$, where π is the initial state distribution, A is the transition probability matrix, and B is the observation probability matrix for this layer. We assume each group of activities G_t is a hidden state and location information of the observed sensor l_{x_k} is the observations for first layer of HMM. The graphical representation of the first layer of HMM is shown in Figure 4. Initial state distribution is denoted by

$$\pi_i = P(G_1 = i), \quad 1 \leq i \leq p, \quad (1)$$

where p is the number of probability distributions. The state transition probability distribution represents the probability of transition from state i to state j and is represented by

$$a_{i,j} = P(G_t = j | G_{t-1} = i), \quad 1 \leq i, j \leq p. \quad (2)$$

The observation probability distribution indicates the probability that the state j would generate observation l_{x_k} :

$$b_j(k) = P(l_{x_k} | G_t = j), \quad 1 \leq j \leq p, \quad 1 \leq k \leq n. \quad (3)$$

The hidden G_t state at time t depends only on the previous hidden state G_{t-1} . The observation variable l_{x_t} at time t , namely, depends only on the hidden variable G_t . The goal is to find the joint probability distribution:

$$P(G, \Theta_L) = \prod_{t=1}^{|\Theta_L|} P(G_t | G_{t-1}) P(l_{x_t} | G_t). \quad (4)$$

The observation distribution $P(l_{x_t} | G_t)$ represents the probability that the activity group G_t would generate n distinct observation symbols. We apply naïve Bayes assumption for the observation distribution as

$$P(l_{x_t} | G_t = j) = \prod_{n=1}^{|L_X|} P(l_{x_n} | G_t = j). \quad (5)$$

3.2. Second Layer of HMM. In the second layer, the classifier uses the sensor data to classify the individual activity. For each activity group G_k , a separate HMM is used. Let $\lambda_{s2} = \{\lambda_{s2}^1, \lambda_{s2}^2, \dots, \lambda_{s2}^p\}$ be the set of all HMMs in the second layer in which HMM is characterized by $\lambda_{s2}^\theta = \{\pi_i^\theta, A^\theta, B^\theta\}$, where θ is the index that represents the HMM to be used and for all $\theta = |G|$. First layer calculates the joint probability distribution and selects the most probable groups for this index as an individual activity classification. The graphical representation of the second layer of HMM is shown in Figure 5. We assume all the activities in a group $G_k = \{y_1^k, y_2^k, \dots, y_q^k\}$ are a hidden

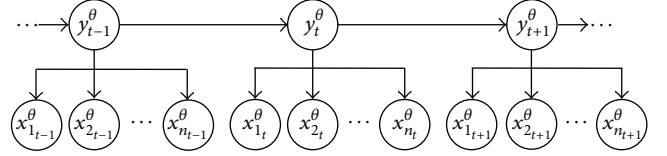


FIGURE 5: The graphical representation of the HMM used in the second layer of the activity model.

state where $1 \leq q \leq m$ and sensor data X_t at a time t are the observations for second-layer HMM. $\Theta^\theta = \{X_1^\theta, X_2^\theta, \dots, X_T^\theta\}$ is the observed sequence for sensor where θ represents the index of second-layer HMMs and $X_1^\theta \in X$. The initial state distribution (π_i^θ) is denoted by

$$\pi_i^\theta = P(y_1^\theta = i), \quad 1 \leq i \leq q, \quad (6)$$

where q is the number of probability distributions. The state transition probability distribution represents the probability of transition from state i to state j and is represented by

$$a_{i,j}^\theta = P(y_t^\theta = j | y_{t-1}^\theta = i), \quad 1 \leq i, j \leq q. \quad (7)$$

The observation probability distribution indicates the probability that the state j would generate observation x_k :

$$b_j^\theta(k) = P(x_k^\theta | y_t^\theta = j), \quad 1 \leq j \leq q, \quad 1 \leq k \leq n. \quad (8)$$

The hidden state y_t^θ at time t depends only on the previous hidden state y_{t-1}^θ . The observation variable x_t^θ at time t depends only on the hidden variable y_t^θ . The goal is to find the joint probability distribution:

$$P(Y^\theta, \Theta^\theta) = \prod_{t=1}^{|\Theta^\theta|} P(y_t^\theta | y_{t-1}^\theta) P(x_t^\theta | y_t^\theta). \quad (9)$$

The observation distribution $P(x_t^\theta | y_t^\theta)$ represents the probability that the activity y_t^θ would generate n distinct observation symbols. We apply naïve Bayes assumption for the calculation of observation distribution as

$$P(x_t^\theta | y_t^\theta = j) = \prod_{n=1}^{|X^\theta|} P(x_n^\theta | y_t^\theta = j). \quad (10)$$

3.3. Learning the Model Parameters. The model parameters are apparently learnt from learning samples. To estimate

the parameters from the sequences of observation Baum-Welch (BW) based algorithm is used [24, 25]. In BW method, for a given sequence of observations, an HMM with n states and an initial model λ^0 , the forward-backward (FB) algorithm computes the expected number of state transitions and state emissions based on the current model (E-step) and then reestimates the model parameters (M-step) using estimation formula during each iteration. After each iteration of the E-step and M-step, the likelihood of the observations increases until a convergence to a stationary point occurs. The reestimation formulas are as follows.

The probability of being in state G_i at time t and state G_j at time $t + 1$, given the model λ_{s1} and the observation sequence Θ_L , can be estimated using forward (α) and backward (β) variables as follows:

$$\xi_t(i, j) = \frac{\alpha_t(i) a_{ij} b_j(O_{t+1}) \beta_{t+1}(j)}{\sum_{i=1}^N \sum_{j=1}^N \alpha_t(i) a_{ij} b_j(O_{t+1}) \beta_{t+1}(j)}. \quad (11)$$

The probability of being in state G_i at time t , given the observation sequence Θ_L and the model λ_{s1} , can be estimated as follows:

$$\gamma_t(i) = \sum_{j=1}^N \xi_t(i, j). \quad (12)$$

First layer of HMM ($\lambda_{s1} = \{\bar{\pi}_i, \bar{A}, \bar{B}\}$) is

$$\begin{aligned} \bar{\pi}_i &= \gamma_1(i) = \text{expected \# of times state } G_i \text{ at } t \\ &= 1, \\ \bar{a}_{i,j} &= \frac{\sum_{t=1}^{T-1} \xi_t(i, j)}{\sum_{t=1}^{T-1} \gamma_t(i)} \\ &= \frac{\text{Expected \# of transition from } G_i \rightarrow G_j}{\text{Expected \# of transition from } G_i}, \quad (13) \\ \bar{b}_j(k) &= \frac{\sum_{t=1}^T \gamma_t(j) \delta_{O_t} l_{x_k}}{\sum_{t=1}^T \gamma_t(j)} \\ &= \frac{\text{Expected \# of times in } G_j \text{ observing } l_{x_k}}{\text{Expected \# of times in } G_j}. \end{aligned}$$

Similarly the probability of being in state y_i^θ at time t and state y_j^θ at time $t + 1$, given the model λ_{s2}^θ and the observation sequence Θ , can be estimated using (11). Moreover the probability of being in state y_i^θ at time t , given the observation sequence Θ and the model λ_{s2}^θ , can be estimate using (12).

Second layer of HMM ($\lambda_{s2}^\theta = \{\bar{\pi}_i^\theta, \bar{A}^\theta, \bar{B}^\theta\}$, $1 \leq \theta \leq p$)

$$\begin{aligned} \bar{\pi}_i^\theta &= \gamma_1^\theta(i) = \text{expected \# of times state } y_i^\theta \text{ at } t \\ &= 1, \\ \bar{a}_{i,j}^\theta &= \frac{\sum_{t=1}^{T-1} \xi_t^\theta(i, j)}{\sum_{t=1}^{T-1} \gamma_t^\theta(i)} \\ &= \frac{\text{Expected \# of transition from } y_i^\theta \rightarrow y_j^\theta}{\text{Expected \# of transition from } y_i^\theta}, \quad (14) \\ \bar{b}_j^\theta(k) &= \frac{\sum_{t=1}^T \gamma_t^\theta(j) \delta_{O_t} x_k^\theta}{\sum_{t=1}^T \gamma_t^\theta(j)} \\ &= \frac{\text{Expected \# of times in } y_j^\theta \text{ observing } x_k^\theta}{\text{Expected \# of times in } y_j^\theta}. \end{aligned}$$

3.4. Inferring the Best State Sequence. To infer the best sequence of groups as well as activities we have used Viterbi algorithm which has been successfully applied with HMM to solve many activity recognition problems [24]. Viterbi algorithm can find out the best sequence efficiently using dynamic programming and discard a number of paths at each time step. The computational complexity of Viterbi algorithm is less than other algorithms (direct calculation, forward-backward). Inference is done in simultaneous manner: the groups and the activities are estimated simultaneously based on Viterbi algorithm.

4. Experimental Setup and Results

We evaluate the performance of the proposed 2-layer HMM in recognizing activities of ADLs and compare it with three well-known classifiers: naïve Bayes, CRF, and HMM using three real datasets. Sensor data were segmented in time slices of length $\Delta t = 60$ seconds where there were a total of 36,000 time slices for dataset of “Kasteren-A”, 21,600 time slices for “Kasteren-B”, and 28,800 time slices for dataset of “Kasteren-C”. The activities and sensors are grouped based on the location and tabulated in Table 3. The sensor networks deployed in the home environment generate raw data stream. The sensor data streams have been presented in three different feature representations [10].

Raw. This feature uses the sensor data directly as it was collected from the sensor network. The value is 1 when the sensor fires and 0 otherwise.

Change Point (CP). This feature indicates when a sensor changes value. The value is 1 when a sensor state goes from zero to one or vice versa and 0 otherwise.

TABLE 3: Details of group information for Kasteren datasets: activity classification based on location data.

Group information	House-A	House-B	House-C
Number of location groups	5	4	6
Name of location group	Bedroom, bathroom, front door, kitchen, and toilet	Bedroom, bathroom, front door, and kitchen	Bedroom, bathroom, front door, kitchen, living room, and toilet
Bedroom group activity list	Sleeping	Sleeping, dressing	Sleeping, dressing
Bathroom group activity list	Brushing teeth, showering	Brushing teeth, showering, and toileting	Brushing teeth, showering, and shaving
Front door group activity list	Leaving house	Leaving house	Leaving house
Kitchen group activity list	Breakfast, dinner, drinking, and snack	Breakfast, dinner, drinking, dishwashing, preparing breakfast, and preparing dinner	Drinking, dishwashing, eating, preparing breakfast, preparing lunch, preparing dinner, and snack
Living room group activity list	—	—	Relax
Toilet group activity list	Toileting	—	Toileting
Bedroom group sensor list	Door	Door, dresser, pressure mat	Door, dresser, and pressure mat
Bathroom group sensor list	Door	Door, toilet flush	Door, shower tap, and sink
Front door group sensor list	Front door	Front door	Front door
Kitchen group sensor list	Microwave, refrigerator, freezer, kitchen cupboard (cup), kitchen cupboard (plates), kitchen cupboard (various), and dishwasher	Microwave, refrigerator, freezer, kitchen sink, kitchen cupboard (various), kitchen cupboard (plates), kitchen drawer (cutlery), and dishwasher	Microwave, refrigerator, freezer, kitchen cupboard (cup), kitchen cupboard (plates), kitchen cupboard (various), and dishwasher
Living room group sensor list	—	—	Couch pressure mat
Toilet group sensor list	Door, toilet flush	—	Door, toilet flush

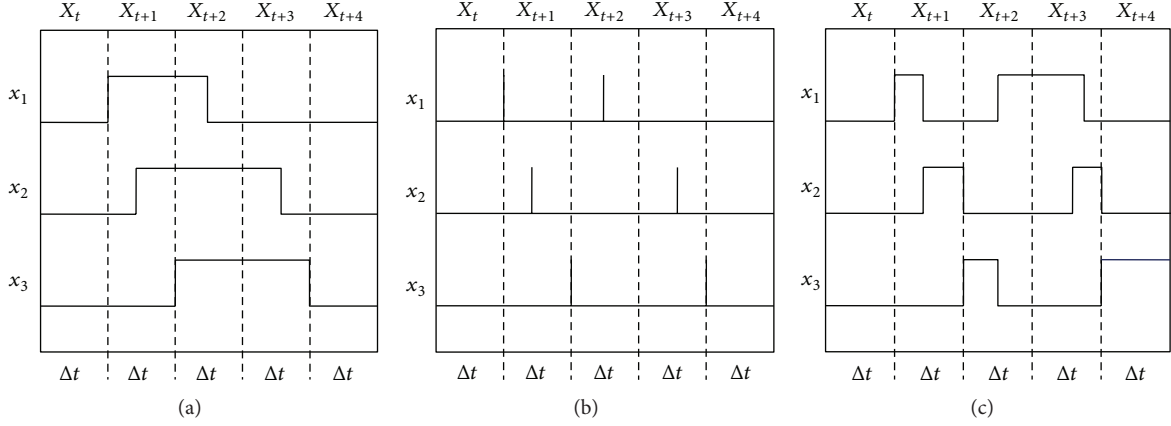


FIGURE 6: Feature representations of (a) Raw, (b) Change Point, and (c) Last-Fired.

TABLE 4: Confusion matrix showing the true positives (TP), total of true labels (TT), and total of inferred labels (TI) for each class of activity.

True	Inferred					
	1	2	3	...	N	
1	TP ₁	ε ₁₂	ε ₁₃	...	ε _{1N}	TT ₁
2	ε ₂₁	TP ₂	ε ₂₃	...	ε _{2N}	TT ₂
3	ε ₃₁	ε ₃₂	TP ₃	...	ε _{3N}	TT ₃
⋮	⋮	⋮	⋮	⋮	⋮	⋮
N	ε _{N1}	ε _{N2}	ε _{N3}	...	TP _N	TT _N
	TI ₁	TI ₂	TI ₃	...	TI _N	Total

Last-Fired (LF). This feature indicates which sensor fired last. The sensor that changed state last continues to value 1 and changes to 0 when another sensor changes state.

We accommodate three distinct features, Raw, Change Point (CP), and Last-Fired (LF), and four combined features: Raw + CP, Raw + LF, CP + LF, and Raw + CP + LF. The combined feature representation is a concatenation of the feature matrices. Figure 6 shows the three distinct features.

The data were split into a test and training set using a “leave one day out” approach in which one day of sensor data is used for testing and the remaining days are used for training. The process is repeated for each day and the average performance is measured. The performance of the proposed 2-layer HMM is evaluated using precision, recall, and *F*-measure. *F*-measure data were expressed as a mean ± standard deviation and significance was analyzed using Student’s *t*-tests. Statistical significance was considered as $P < 0.05$. *F*-measure, precision, recall, and accuracy can be calculated using the confusion matrix shown in Table 4:

$$F\text{-Measure} = \frac{2 \cdot \text{precision} \cdot \text{recall}}{\text{precision} + \text{recall}}. \quad (15)$$

The precision and recall are defined as

$$\begin{aligned} \text{Precision} &= \frac{1}{N} \sum_{i=1}^N \frac{TP_i}{TI_i}, \\ \text{Recall} &= \frac{1}{N} \sum_{i=1}^N \frac{TP_i}{TT_i}, \\ \text{Accuracy} &= \frac{\sum_{i=1}^N TP_i}{\text{Total}}. \end{aligned} \quad (16)$$

The recognition accuracy of each activity for three datasets is shown in the confusion matrix separately. In the case of the “Change Point + Last-Fired” feature the highest accuracy is achieved for each activity. Tables 5–7 show confusion matrix for 2-layer HMM using the “Change Point + Last-Fired” feature of three datasets.

The average *F*-measure values for the three datasets are tabulated in Tables 8–10. Rows in the tables correspond to the different distinct and combined features whereas columns represent the experimental results for each activity recognition model. In case of the “Change point + Last-Fired (CP + LF)” feature the highest *F*-measure value is achieved. *F*-measure data are expressed as a mean ± standard deviation and significance is analyzed using Student’s *t*-test. Statistical significance is considered as $P < 0.05$.

The experimental results for House-A of Kasteren datasets are similar to House-B of Kasteren datasets. The 2-layer HMM achieves the best *F*-measure value for all datasets using distinct feature “Change Point (CP)” and combined feature “Change Point + Last-Fired (CP + LF).” The average performance of each activity model for all datasets is shown in Figure 7. From Figure 7 we can conclude that the 2-layer activity model outperforms the other activity models.

Time complexity arises while calculating the joint probability of each state sequence with the observed series of events. For example, an HMM having N states will need N^N state transition probabilities, 2^N output probabilities (assuming

TABLE 5: Confusion matrix of two-layer HMM using “Change Point + Last-Fired” feature for House-A of Kasteren dataset.

	Actual activity	Predicted activity										Recall
		Breakfast	Brushing teeth	Dinner	Drinking	Leaving house	Others	Sleeping	Showering	Snack	Toileting	
	Breakfast	75	0.0	3.0	5.7	0.0	6.5	0.0	5.5	4.3	0.0	75.0
	Brushing teeth	0.0	75.5	0.0	0	0	10.2	0	7.4	0	6.9	75.5
	Dinner	6.6	0.0	78	7.6	0	7.8	0	0	0	0.0	78.0
	Drinking	4.5	0.0	5.3	78	0	11	0	0	1.2	0.0	78.0
	Leaving house	0.0	0.0	0.0	0	85	15	0	0	0	0.0	85.0
	Others	5.3	5.5	5.8	5.2	5.8	50	5.7	5.6	5.6	5.5	50.0
	Sleeping	0.0	0.0	0.0	0	0	19.5	80.5	0	0	0.0	80.5
	Showering	0.0	7.6	0.0	0	0	7.2	0	77	0	8.2	77.0
	Snack	6.5	0.0	5.7	5.6	0	10.2	0	0	72	0.0	72.0
	Toileting	0.0	6.0	0.0	0	0	9.7	0	7.3	0	77	77.0
	Precision	76.61	79.80	79.75	76.39	93.61	33.99	93.38	74.90	86.64	78.89	

TABLE 6: Confusion matrix of two-layer HMM using “Change Point + Last-Fired” feature for House-B of Kasteren dataset.

Actual activity	Predicted activity												
	Breakfast	Brushing teeth	Dinner	Drinking	Dressing	Leaving house	Others	Preparing Breakfast	Preparing Dinner	Sleeping	Showering	Toileting	Using dishwasher
Breakfast	75	0.0	4.3	4.7	0.0	0.0	6.5	3.5	3.0	0.0	0.0	0.0	3.0
Brushing teeth	0.0	75.5	0.0	0.0	0.0	0.0	10.2	0.0	0.0	0.0	7.4	6.9	0.0
Dinner	6.6	0.0	78	7.6	0.0	0.0	7.8	0.0	0.0	0.0	0.0	0.0	0.0
Drinking	4.5	0.0	3.3	78	0.0	0.0	8.2	2.0	2.0	0.0	0.0	0.0	2.0
Dressing	0.0	0.0	0.0	0.0	80	0.0	10	0.0	0.0	10.0	0.0	0.0	0.0
Leaving house	0.0	0.0	0.0	0.0	0.0	85	15	0.0	0.0	0.0	0.0	0.0	0.0
Others	5.3	5.5	5.4	5.2	0.0	4.8	60	0.0	0.0	4.7	4.6	4.5	0.0
Preparing Breakfast	5.0	0.0	4.0	3.0	0.0	0.0	8.0	65	10	0.0	0.0	0.0	5.0
Preparing Dinner	3.0	0.0	5.0	3.0	0.0	0.0	8.0	10	66	0.0	0.0	0.0	5.0
Sleeping	0.0	0.0	0.0	0.0	9.5	0.0	10.0	0.0	0.0	80.5	0.0	0.0	0.0
Showering	0.0	7.6	0.0	0.0	0.0	0.0	7.2	0.0	0.0	0.0	77	8.2	0.0
Toileting	0.0	6.0	0.0	0.0	0.0	0.0	9.7	0.0	0.0	0.0	7.3	77	0.0
Using dishwasher	3.0	0.0	4.0	2.0	0.0	0.0	5.0	5.0	5.0	0.0	0.0	0.0	76.0
Precision	73.24	79.80	75.00	75.36	89.38	94.65	36.23	76.02	76.74	84.55	79.96	79.71	83.51

TABLE 7: Confusion matrix of two-layer HMM using “Change Point + Last-Fired” feature for House-C of Kasteren dataset.

	Brushing teeth	Predicted activity										Recall				
		Drinking	Dressing	Eating	Leaving house	Medication	Others	Preparing Breakfast	Preparing Lunch	Preparing Dinner	Relax	Sleeping	Showering	Snack	Shaving	Toileting
Actual activity	Brushing teeth	72	0.0	0.0	0.0	0.0	12	0.0	0.0	0.0	0.0	0.0	5.0	0.0	6.0	5.0
	Drinking	0.0	80	0.0	0.0	0.0	9.0	2.0	0.0	2.0	0.0	0.0	0.0	0.0	0.0	2.0
	Dressing	0.0	0.0	75	0.0	0.0	10	0.0	0.0	0.0	8.0	7.0	0.0	0.0	0.0	0.0
	Eating	0.0	0.0	0.0	75	0.0	6.0	5.0	5.0	5.0	0.0	0.0	0.0	4.0	0.0	0.0
	Leaving house	0.0	0.0	0.0	0.0	85	15	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	Medication	0.0	0.0	0.0	0.0	0.0	85	10	0.0	0.0	0.0	5.0	0.0	0.0	0.0	0.0
	Others	3.0	3.0	3.0	3.0	4.8	50	3.0	3.0	3.0	2.0	4.7	4.5	3.0	4.5	2.5
	Preparing Breakfast	0.0	5.0	0.0	5.0	0.0	8.0	65	6.0	6.0	0.0	0.0	0.0	5.0	0.0	0.0
	Preparing Lunch	0.0	5.0	0.0	5.0	0.0	8.0	6.0	65	6.0	0.0	0.0	0.0	5.0	0.0	0.0
	Preparing Dinner	0.0	5.0	0.0	5.0	0.0	8.0	6.0	6.0	65	0.0	0.0	0.0	5.0	0.0	0.0
	Relax	0.0	0.0	0.0	0.0	0.0	5.0	0.0	0.0	0.0	85	10	0.0	0.0	0.0	0.0
	Sleeping	0.0	0.0	5.0	0.0	0.0	9.5	0.0	0.0	0.0	5	80.5	0.0	0.0	0.0	0.0
	Showering	6.0	0.0	0.0	0.0	0.0	8.0	0.0	0.0	0.0	0.0	76	0.0	0.0	5.0	5.0
	Snack	0.0	5.0	0.0	5.0	0.0	10	5.0	5.0	5.0	0.0	0.0	0.0	65	0.0	0.0
	Shaving	6.0	0.0	0.0	0.0	0.0	10	0.0	0.0	0.0	0.0	0.0	6.0	0.0	72	6.0
	Toileting	4.0	0.0	0.0	0.0	0.0	9.0	0.0	0.0	0.0	0.0	0.0	5.0	0.0	5.0	77
Precision	79.12	77.67	90.36	72.81	96.59	94.65	26.67	70.65	72.22	70.65	85	75.09	78.75	74.71	77.83	78.97

TABLE 8: *F*-measure expressed in percentage (%) for House-A of Kasteren datasets.

Feature	Naïve Bayes	CRF	HMM	2-layer HMM
Raw	48 ± 16	50 ± 18	50 ± 16	60 ± 13
Change Point (CP)	50 ± 17	65 ± 15	71 ± 14	72 ± 10
Last-Fired (LF)	58 ± 15	60 ± 14	62 ± 13	67 ± 12
Raw + CP	55 ± 16	62 ± 16	55 ± 15	71 ± 11
Raw + LF	54 ± 14	63 ± 14	69 ± 13	72 ± 12
CP + LF	65 ± 16	70 ± 16	72 ± 15	76 ± 13
Raw + CP + LF	60 ± 18	68 ± 14	68 ± 14	70 ± 12
Average	56 ± 16	63 ± 15	64 ± 16	70 ± 12

TABLE 9: *F*-measure expressed in percentage (%) for House-B of Kasteren datasets.

Feature	Naïve Bayes	CRF	HMM	2-layer HMM
Raw	49 ± 10	52 ± 13	50 ± 13	60 ± 12
Change point (CP)	52 ± 8	60 ± 17	60 ± 11	71 ± 10
Last-Fired (LF)	55 ± 11	55 ± 12	55 ± 12	68 ± 13
Raw + CP	59 ± 12	62 ± 18	57 ± 9	67 ± 14
Raw + LF	60 ± 8	65 ± 12	56 ± 12	66 ± 13
CP + LF	66 ± 12	68 ± 13	70 ± 15	76 ± 11
Raw + CP + LF	62 ± 9	67 ± 15	68 ± 8	72 ± 11
Average	58 ± 10	61 ± 14	59 ± 11	69 ± 12

TABLE 10: *F*-measure expressed in percentage (%) for House-C of Kasteren datasets.

Feature	Naïve Bayes	CRF	HMM	2-layer HMM
Raw	40 ± 10	50 ± 13	52 ± 10	62 ± 12
Change Point (CP)	58 ± 8	60 ± 14	70 ± 8	72 ± 13
Last-Fired (LF)	57 ± 12	58 ± 16	65 ± 11	70 ± 11
Raw + CP	55 ± 10	55 ± 17	60 ± 9	65 ± 12
Raw + LF	58 ± 9	57 ± 15	66 ± 8	68 ± 13
CP + LF	60 ± 13	65 ± 13	72 ± 12	74 ± 13
Raw + CP + LF	55 ± 11	63 ± 18	68 ± 11	71 ± 11
Average	55 ± 10	63 ± 15	65 ± 10	69 ± 12

all the outputs are binary), and N^2L time complexity to derive the probability of an output sequence of length L [26]. Time complexity is uncontrollable for realistic problems as the number of possible hidden node sequences typically is extremely high. To reduce time complexity of probabilistic model we proposed 2-layer HMM.

Using 2-layer HMM we can reduce the number of hidden states as well as the number of symbols in each hidden state. Comparison of time complexity for HMM and 2-layer HMM is tabulated in the following table. We calculate

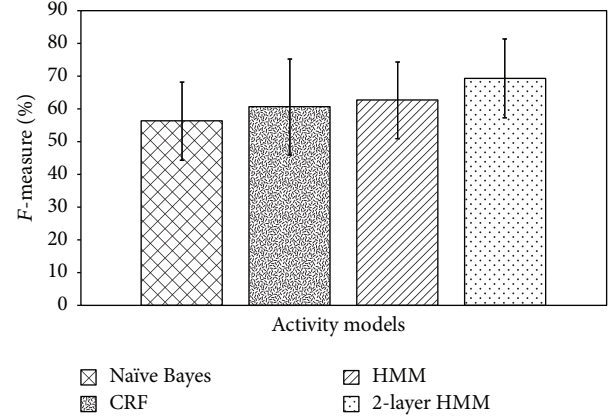


FIGURE 7: Performance comparison with the existing systems.

output sequence length L , multiplying the number of symbols emitted in corresponding layer by the number of time slices with time duration of 60 s. In House-A of Kasteren dataset, the number of hidden states and the number of observation symbols in each state for the HMM are 10 and 14, respectively. Then time complexity can be calculated as $N^2L = 10^2 \times 14 \times 1440 = 2,016,000$. For 2-layer HMM we consider 5 hidden states and 14 observation symbols in first layer. Second layer is composed of 5 sublayers and the numbers of hidden states and observation symbols for each sublayer are 1 and 1, 2 and 1, 1 and 1, 4 and 7, and 1 and 2, respectively. Time complexity for 2-layer HMM can be calculated as $N^2L = (5^2 \times 14 + 1^2 \times 1 + 2^2 \times 1 + 1^2 \times 1 + 4^2 \times 7 + 1^2 \times 2) \times 1440 = 676,800$. Similarly, time complexity of HMM and proposed 2-layer HMM can be calculated for House-B and House-C of Kasteren datasets and are tabulated in Table 11.

5. Conclusion

We have presented new approach to recognize ADLs using binary sensors in home environments. We proposed 2-layer HMM using location information in first layer and object used information in another layer. We used three real datasets to evaluate the recognition performance of the proposed model and then make a comparison with the other activity recognition models. For evaluation we considered *F*-measure, a measure that considers the correct classification of each class as equally important. *F*-measure data were expressed as a mean ± standard deviation and significance was analyzed using Student's *t*-test. Statistical significance was considered as $P < 0.05$. Comparison results show that combination of location data with object used data can lead to significantly better performance in real world activity recognition. This work demonstrates that ADLs recognition can be done effectively by binary sensor network deployed in home environments. In the future we can apply machine learning schemes for learning the model parameters.

TABLE 11: Comparison of time complexity (N^2L) for different models.

House-A		House-B		House-C	
HMM	2-layer HMM	HMM	2-layer HMM	HMM	2-layer HMM
2,016,000	676,800	5,597,280	1,002,240	7,741,440	1,644,480

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgment

This work was supported by the Industrial Strategic Technology Development Program (10041788, Development of Smart Home Service based on Advanced Context-Awareness) funded by the Ministry of Trade, Industry & Energy (MI, Korea).

References

- [1] T. Duong, D. Phung, H. Bui, and S. Venkatesh, "Efficient duration and hierarchical modeling for human activity recognition," *Artificial Intelligence*, vol. 173, no. 7-8, pp. 830–856, 2009.
- [2] V. Kellokumpu, M. Pietikäinen, and J. Heikkilä, "Human activity recognition using sequences of postures," in *Proceedings of the IAPR Conference on Machine Vision Applications (MVA '05)*, pp. 570–573, Tsukuba, Japan, May 2005.
- [3] J. K. Aggarwal and Q. Cai, "Human motion analysis: a review," *Computer Vision and Image Understanding*, vol. 73, no. 3, pp. 428–440, 1999.
- [4] R. Polana and R. Nelson, "Detecting activities," *Journal of Visual Communication and Image Representation*, vol. 5, no. 2, pp. 172–180, 1994.
- [5] J. A. Ward, P. Lukowicz, G. Tröster, and T. E. Starner, "Activity recognition of assembly tasks using body-worn microphones and accelerometers," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 28, no. 10, pp. 1553–1567, 2006.
- [6] C. Sminchisescu, A. Kanaujia, Z. Li, and D. Metaxas, "Conditional models for contextual human motion recognition," in *Proceedings of the 10th IEEE International Conference on Computer Vision (ICCV '05)*, vol. 2, pp. 1808–1815, IEEE, Beijing, China, October 2005.
- [7] T. Huynh and B. Schiele, "Towards less supervision in activity recognition from wearable sensors," in *Proceedings of the 10th IEEE International Symposium on Wearable Computers (ISWC '06)*, pp. 3–10, Montreux, Switzerland, October 2006.
- [8] Ó. D. Lara and M. A. Labrador, "A survey on human activity recognition using wearable sensors," *IEEE Communications Surveys & Tutorials*, vol. 15, no. 3, pp. 1192–1209, 2013.
- [9] L. Bao and S. S. Intille, "Activity recognition from user-annotated acceleration data," in *Pervasive Computing*, A. Ferscha and F. Mattern, Eds., vol. 3001 of *Lecture Notes in Computer Science*, pp. 1–17, Springer, Berlin, Germany, 2004.
- [10] T. L. M. Van Kasteren, G. Englebienne, and B. J. A. Kröse, "Human activity recognition from wireless sensor network data: benchmark and software," in *Activity Recognition in Pervasive Intelligent Environments*, vol. 4 of *Atlantis Ambient and Pervasive Intelligence*, pp. 165–186, Atlantis Press, 2011.
- [11] E. M. Tapia, S. S. Intille, and K. Larson, "Activity recognition in the home using simple and ubiquitous sensors," in *Pervasive Computing*, vol. 3001 of *Lecture Notes in Computer Science*, pp. 158–175, Springer, Berlin, Germany, 2004.
- [12] M. Philipose, K. P. Fishkin, M. Perkowitz et al., "Inferring activities from interactions with objects," *IEEE Pervasive Computing*, vol. 3, no. 4, pp. 50–57, 2004.
- [13] I. Fatima, M. Fahim, Y.-K. Lee, and S. Lee, "Analysis and effects of smart home dataset characteristics for daily life activity recognition," *Journal of Supercomputing*, vol. 66, no. 2, pp. 760–780, 2013.
- [14] D. H. Wilson and C. Atkeson, "Simultaneous tracking and activity recognition (STAR) using many anonymous, binary sensors," in *Pervasive Computing*, vol. 3468 of *Lecture Notes in Computer Science*, pp. 62–79, Springer, Berlin, Germany, 2005.
- [15] D. J. Cook and M. Schmitter-Edgecombe, "Assessing the quality of activities in a smart environment," *Methods of Information in Medicine*, vol. 48, no. 5, pp. 480–485, 2009.
- [16] D. L. de Ipiña, I. D. de Sarralde, and J. García-Zubia, "An Ambient assisted living platform integrating RFID data-on-tag care annotations and twitter," *Journal of Universal Computer Science*, vol. 16, no. 12, pp. 1521–1538, 2010.
- [17] T. L. M. Van Kasteren, A. Noulas, G. Englebienne, and B. J. Kröse, "Accurate activity recognition in a home setting," in *Proceedings of the Conference on Autonomous Agents and Multiagent Systems (AAMAS '07)*, pp. 1–9, Honolulu, Hawaii, USA, May 2007.
- [18] P. Palmes, H. K. Pung, T. Gu, W. Xue, and S. Chen, "Object relevance weight pattern mining for activity recognition and segmentation," *Pervasive and Mobile Computing*, vol. 6, no. 1, pp. 43–57, 2010.
- [19] D. L. Vail, M. M. Veloso, and J. D. Lafferty, "Conditional random fields for activity recognition," in *Proceedings of the Conference on Autonomous Agents and Multiagent Systems (AAMAS '07)*, article 235, pp. 1–8, Honolulu, Hawaii, USA, May 2007.
- [20] P. Lanchantin, J. Lapuyade-Lahorgue, and W. Pieczynski, "Unsupervised segmentation of randomly switching data hidden with non-Gaussian correlated noise," *Signal Processing*, vol. 91, no. 2, pp. 163–175, 2011.
- [21] D. Fox, "Location-based activity recognition," in *Proceedings of the 30th Annual German Conference on Advances in Artificial Intelligence (KI '07)*, p. 51, Springer, Osnabruck, Germany, September 2007.
- [22] C. Zhu and W. Sheng, "Motion- and location-based online human daily activity recognition," *Pervasive and Mobile Computing*, vol. 7, no. 2, pp. 256–269, 2011.
- [23] T. L. M. Van Kasteren, "Datasets for Activity Recognition," <https://sites.google.com/site/tim0306/datasets>.
- [24] L. R. Rabiner, "A tutorial on hidden Markov models and selected applications in speech recognition," *Proceedings of the IEEE*, vol. 77, no. 2, pp. 257–286, 1989.
- [25] W. Khreich, E. Granger, A. Miri, and R. Sabourin, "A survey of techniques for incremental learning of HMM parameters," *Information Sciences*, vol. 197, pp. 105–130, 2012.
- [26] J. A. Bilmes, "What HMM can do," *IEICE Transactions on Information and Systems*, vol. E89-D, no. 3, pp. 1–24, 2006, Technical Report of University of Washington, Department of Electrical Engineering, Number UWEETR-2002-0003.