A Reconfigurable HMM for Activity Recognition

Md. Kamrul Hasan[1], Husne Ara Rubaiyeat[2], Yong-Koo Lee[1], Sungyoung Lee[1]
[1]Ubiquitous Computing Lab, Department of Computer Engineering
[2]Bio Imaging Lab, Department of Biomedical Engineering
Kyung Hee University, 1 Socheon, Giheung, Yongin, Gyeonggi, 446-701, South Korea
E-mail: kamrul@oslab.khu.ac.kr, rubaiyeat@yahoo.com, yklee@khu.ac.kr, sylee@oslab.khu.ac.kr

***Abstract* — This paper presents a reconfigurable HMM for activity recognition in a home setting where sensors are deployed in the environment. Sensors activated in a sequence are captured and modeled using the HMM like construct. The model traces out the most likely sensor activation sequence responsible for the activity being performed. We recognize several Activities of Daily Living (ADLs) for Ubiquitous Healthcare.**

***Keywords* — HMM, Reconfigurable HMM, Activity Recognition.**

1. Introduction

In ubiquitous healthcare environment, it is essential to recognize users' Activity of Daily Living (ADLs) like, preparing meal, bathing etc. Monitoring and recognizing the daily activities can provide proactive healthcare at home [3]. Monitoring can reduce the number of low risk cases that come to hospitals-thereby reducing the financial burden on health systems and allowing more focus on patients at higher risk [2].

The lack of training sets that reflect daily activities puts us into a challenging state for learning ADLs. Existing deployment, such as PlaceLab [1] generate thousands of sensor data and only tens of activity instances. This is because users are reluctant to annotate the sensor values. Additionally, there are several ways of doing same activity. There can be noise in the data also, as users' movements are not always purposeful. Figure 1 shows a deviation of the user from the default course of movement.

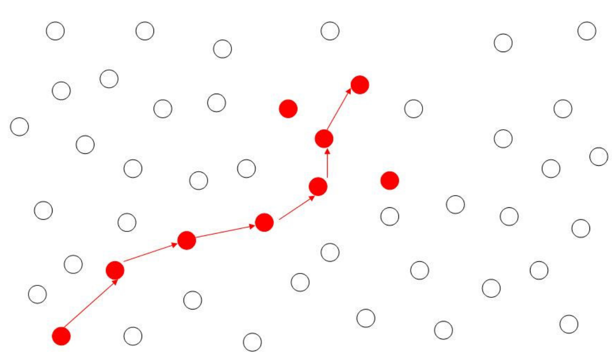

Figure 1: Example of sensor activations as user deviates from his default course of movement. Circles represent sensor and red circles represent activated sensors. Red circles connected through arrow are the default route.

So, it is difficult to produce an exact model. However, it might be possible to relate the sensor activation patterns with the activities. A sequence of sensor activations can indicate what the user is doing. With this view, we make an HMM with all the activated sensors in a sequence. We name it reconfigurable HMM as the number of states and their interconnection changes with time.

The rest of the paper is organized as follows: section 2 provides some related works, section 3 describes the Reconfigurable HMM construct, section 4 is the experimental results and section 5 is the conclusion and future works.

2. Related Works

Activity Recognition for a setup where sensors are deployed in the environment was addressed in [4], [5]. They used 77 on/off state sensors in a house to capture the daily activities. They got very low recognition accuracy (27% maximum) on the data collected. Even though they tried to incorporate temporal relationships, the accuracy decreased while increasing the number of features. In our experiment, we use the same data and show that temporal relationships among features actually increase the accuracy up to a certain limit.

Mobility based activity detection is a simplified version of activity recognition problem, where only user's motion is considered. Several approaches have been found in mobility based activity detection, such as Hierarchical HMM [6], Bayes Filter [7], Dynamic Bayesian Network [8], Naïve Bayes [9] etc. They concentrate on special techniques to use the specific domain knowledge. Our work provides a general model and shows an example how to use the approach for activity detection in one of the setups. All those works use temporal features created explicitly or achieved through placing several time slices of the same network sequentially. Those approaches need huge computation and may become intractable, if not carefully designed [2], [10]. On the other hand, our approach relates features in time sequence naturally through the model proposed.

3. Reconfigurable HMM

An HMM is defined by a set of states with interconnection between them. A set of prior probabilities and output or emission probabilities for the states and a set of state transition probabilities are defined.

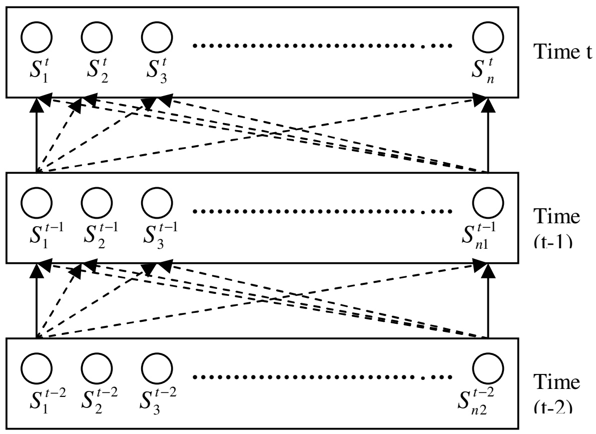

Figure 2: HMM constructed with the sensors in the last three time sequences for any particular activity. By box and arrows we indicate that all states in time (l-1) are connected with all the sensors in time l.

In our approach, each sensor is considered as a state in the HMM. For each activity, a separate HMM is defined. So, prior probability for each state in the HMM is actually the prior probability of the respective sensor being activated during the activity. A state transition probability is the conditional probability of one sensor being activated given another sensor activated within the particular activity period.

Sensors are activated and deactivated while the user is performing some Activity of Daily Life (ADL). If we take the samples of the sensor activations periodically, a sequence of activations can be obtained. The sequence of sensor activations may indicate what the user is actually doing. To capture the sequence information, we construct an HMM with the sensors activated in turns. We omit the state transition links among the sensors in the same sample period because those sensors infer concurrency, not sequence (see Figure 2).

After constructing the HMM, the maximum likelihood is calculated for each activity by evaluating their respective HMMs and the highest likely activity is taken. To overcome fluctuation, a decision is given only if the current activity is matched with the previously inferred activity.

The HMM is reconfigurable as the network changes based on the sensors present in the last few time sequences. The worst case is that all the sensors deployed stays activated. Keeping the worst case in mind, all the non-zero conditional probabilities are stored. Still the conditional probability table size is not big, unlike one may presume. This is because, ADLs are usually performed in a constrained physical area and the only the sensors of that physical area are activated. At this stage we are ready to introduce some terminologies:

$P_a(S_i)$ =Prior probability of sensor i for activity 'a'

$P_a(S_j \mid S_i)$ = conditional probability of S_j given S_i for activity 'a'

$P_a(O_{oi})$ =Probability of output O_o from sensor i for activity 'a'

S_i^t = HMM state constructed with sensor i at time t

$P_a(S_i^t)$ =Calculated probability (usually maximum likelihood) of S_i^t for activity 'a'

$P_a(S_j^t \mid S_i^{t-1})$ =State transition probability of S_j^t given S_i^{t-1} $= P_a(S_j \mid S_i)$

3.1 Calculating the Likelihood

After constructing the HMMs for each activity, we calculate the likelihood for them. We use a maximum likelihood algorithm on the HMM construct, a similar algorithm found in standard message passing literature [11], [12]. Our algorithm also takes the notion of Forward algorithm of HMM [13], where sum operation is replaced by max operation. It is to be noted that the maximum is not calculated for all the combination of states; because we know outputs and their respective hidden states. Each state calculates the maximum likelihood using the formula below:

$$P_a(S_i^l) = P_a(S_i)P_a(O_{oi}), \quad \text{l=last sequence}$$

$$P_a(S_i^l) = P_a(O_{oi})\max_c[P_a(S_i^l \mid S_c^{l-1})P_a(S_c^{l-1})], \quad \text{otherwise} \quad \text{……….} (1)$$

3. Experimental Setup

We are experimenting with the data gathered in MIT PlaceLab [2], [14]. The sensors deployed are 77 on/off state sensors. So, in this case, $P_a(O_{oi})$ is actually the activation probability of sensor i which is equivalent to the prior probability of sensor i, $P_a(S_i)$. So, the likelihood equation reduces to

$$Likelihood(Activity \quad a)$$

$$= \max_i \left(P_a(S_i^t)\max_c\left(P_a(S_i^t \mid S_c^{t-1})P_a(S_c^{t-1})\right) \right) \text{…..} \quad (2)$$

The data were collected in a central place and were labeled using Experience Sampling Method (ESM) [2]. The training set is small with noises as seen from their recognition results [2]. However, the result could further be improved through the use of temporal features constructed from sensor activation sequence, as indicated in [2].

The purpose of using MIT data in our work is to show the effectiveness of our model in a classification problem, where temporal relationship among the features is necessary.

We also assume that all the sensors are time synchronized and broadcast their states periodically all at the same time. We assume this ideal environment to show the improvement of recognition accuracy of ADLs.

We use the text data from [14], take samples of sensor values at every 5 seconds and feed through our simulator; our simulation program actually emulates the user activities recorded in the data file and provides a continuous sensor activation sequence needed for our algorithm. Our program runs in two phrases. First, we learn the prior and conditional probabilities for the sensors based on this data by frequency counting. In the second phrase, we emulate the ADLs, reproduce the sensor sequences and recognize the activities using algorithm given by equation 1 and modified in equation 2.

We also filter the data set to reduce the noise. For any activity recognition problem, it is desirable that the classifier recognizes the activity while it is being performed. In our algorithm, the classifier reaches such a state after passing T =5 samples, each taken in 5 seconds interval. However, doing this reduces the number of samples for large T. So, we relax this constraint and pass a decision for a sample whose previous two samples also have the same labeling.

We have found that, increasing the sequence length, T increases the recognition accuracy up to a certain limit. Using maximum likelihood belief propagation algorithm as given in equation 1, we have found average recognition accuracy as given in table 1. Here we can see that T=5 is the optimal sequence length.

Table 1: Average recognition accuracy for ADL data from MIT [14], using maximum likelihood algorithm in equation 1. The accuracy is shown against sensor sequence length.

Seq Len / Acc-uracy	2	3	4	***5***	6	7	8
	40	48	52	***54***	51	48	45

Table 2 presents the average accuracy results for all the activities given in [14] with a sequence length 5 and sensor values sampled at an interval of 5 seconds. The accuracy (54%) is double of the highest recognition accuracy (27%) found in [2].

Table 2: Activity wise recognition accuracy of our algorithm (sequence length 5 and sampling interval 5 seconds)

Activity	Number of Samples	Accuracy in Percentage (%)
Bathing	213	69.95
Toileting	232	56.03
Going_out_to_work	1	100.0
Preparing_lunch	362	35.63
Preparing_dinner	98	31.63
Preparing_breakfast	58	44.96
Dressing	129	69.76
Grooming	260	58.07
Preparing_a_snack	63	58.73
Preparing_a_beverage	60	70.0
Washing_dishes	60	76.66
Doing_laundry	150	61.33
Cleaning	140	18.57
Putting_away_dishes	48	81.25
Washing_hands	4	100.0
Putting_away_groceries	31	74.19
Other	16	93.75
Watching TV	35	48.57
Going_out_for_entertainment	5	80
Lawnwork	39	97.43
Putting_away_laundry	2	0
Average Accuracy		**54.02%**

However, the result is still not high enough to motivate some one. We believe that the medium accuracy of our result is because of the small training set and noise, which is also indicated in [2], [5].

4. Conclusion

We propose a model for activity recognition model in a home setting with simple state sensors attached to everyday devices. We plan to use our model in a ubiquitous home environment for learning user's behavior for proactive healthcare. We are continuing our work to further improve the accuracy and optimize the algorithm.

ACKNOWLEDGEMENTS

This work was supported by the MIC (Ministry of Information and Communication), Korea, under the ITFSIP (IT Foreign Specialist Inviting Program) supervised by the IITA (Institute of Information Technology Advancement) and MIC (Ministry of Information and Communication), Korea, under the ITRC (Information Technology Research Center) support program supervised by the IITA (Institute of Information Technology Advancement) (IITA-2006-C1090-0602-0002).

REFERENCES

[1] S. Intille, K. Larson, E. Tapia, J. Beaudin, and J. Kaushik. Using a live-in laboratory for ubiquitous computing research, LNCS Springer-Verlag, 2006.

[2] E. Tapia, S. Intille, and K. Larson. Activity Recognition in the home setting using simple and ubiquitous sensors. Proceedings of PERVASIVE 2004, vol. LNCS 3001, A. Ferscha and F. Mattern, Eds. Berlin Heidelberg: Springer-Verlag, 2004, pp. 158-175.

[3] Matthai Philipose, Kenneth P. Fishkin, Mike Perkowitz, Donald J. Patterson, Dieter Fox, Henry Kautz, and Dirk Hähnel. Inferring Activities from Interactions with Objects. IEEE Pervasive Computing, October 2004, pp. 50-57.

[4] E. Tapia, S. Intille, and K. Larson. Activity Recognition in the home setting using simple and ubiquitous sensors. Proceedings of PERVASIVE 2004, vol. LNCS 3001, A. Ferscha and F. Mattern, Eds. Berlin Heidelberg: Springer-Verlag, 2004, pp. 158-175.

[5] E. Munguia Tapia. Activity Recognition in the Home Setting Using Simple and Ubiquitous Sensors. M.S. Thesis Media Arts and Sciences, Massachusetts Institute of Technology, September 2003.

[6] Sebastian Luhr, Hung H. Bui, Svetha Venkatesh, Geoff A. W. West. Recognition of Human Activity through Hierarchical Stochastic Learning. First IEEE International Conference on Pervasive Computing and Communications (PerCom'03).

[7] Donald J. Patterson, Lin Liao, Dieter Fox, and Henry Kautz. Inferring High-Level Behavior from Low-Level Sensors. Ubicomp, 2003.

[8] Donald J. Patterson, Dieter Fox, Henry Kautz, Matthai Philipose. Fine-Grained Activity Recognition by Aggregating Abstract Object Usage. In the proceedings of the 2005 Ninth IEEE International Symposium on Wearable Computers (ISWC'05).

[9] Christopher R. Wren and Emmanuel Munguia Tapia. Toward Scalable Activity Recognition for Sensor Networks. In Proceedings of The Second International Workshop in Location and Context-Awareness (LoCA '06), pp. 168-185, Springer vol. 3987 / 2006.

[10] Sangho Park, Jake K. Aggarwal: A hierarchical Bayesian network for event recognition of human actions and interactions. Multimedia Syst. 10(2): 164-179 (2004).

[11] Alexander T. Ihler. Inference in Sensor Networks: Graphical Models and Particle Models. Ph. D Thesis, MIT, 2005.

[12] Christopher M. Bishop. Chapter 8: Graphical Models, Pattern Recognition and Machine Learning, Springer 2006.

[13] Rabiner L. R., Juang B. H. An introduction to hidden Markov models. IEEE ASSP Magazine: 4-15, Jan 1986.

[14] MIT Activity Recognition data: http://courses.media.mit.edu/2004fall/mas622j/04.projects/home/