The 2010 IEEE/RSJ International Conference on
Intelligent Robots and Systems
October 18-22, 2010, Taipei, Taiwan

Interaction Models for Multiple-Resident Activity Recognition in a Smart Home

Yi-Ting Chiang, Kuo-Chung Hsu, Ching-Hu Lu, *Member, IEEE*,
Li-Chen Fu, *Fellow, IEEE* and Jane Yung-Jen Hsu, *Member, IEEE*

***Abstract*—Multi-resident activity recognition is among a key enabler in many context-aware applications in a smart home. However, most of prior researches ignore the potential interactions among residents in order to simplify problem complexity. On the other hand, multiple-resident activities are usually recognized using cameras or wearable sensors. However, due to human-centric concerns, it is more preferable to avoid using obtrusive sensors. In this paper, we propose dynamic Bayesian networks which extend coupled hidden Markov models (CHMMs) by adding some vertices to model both individual and cooperative activities. In order to improve performance of the model, we categorize sensor observations based on data association and some domain knowledge to model multiple-resident activity patterns. We then validate the performance using a multi-resident dataset from WSU (Washington State University), which only includes non-obtrusive sensors. The experimental result shows that our model performs better than other baseline classifiers.**

I. Introduction

FOR the purpose of service prevision, activity recognition progressively becomes desiderative in a smart home. Numerous intelligent applications can be fostered if more robust outcomes of activity recognition are given to a smart-home system. We exemplify this requirement using a commonly used reminder system in a smart home. A smart reminder system, automatically reporting scheduled task or information, is important for human computer interaction in a smart space and usually requires activity profiling. For example, the system can remind a resident that a specific food is about to expire during meal preparation. In addition, activity recognition is also needed in other context-aware applications, such as home automation, security surveillance, and healthcare. Due to these benefits, many researchers have focused on how to robustly sense activities of residents.

Manuscript received March 10, 2010. This work is sponsored by the National Science Council, Taiwan, under project NSC95-2218-E-002-072

Yi-Ting Chiang is with Dept. of Computer Science & Information Eng., National Taiwan University, Taiwan, R.O.C (corresponding author to provide phone: +886-2-3366-4888#336, e-mail: yitingchiang@gmail.com).

Kuo-Chung Hsu and Ching-Hu Lu are with Dept. of Computer Science & Information Eng., National Taiwan University, Taiwan, R.O.C. (e-mail: b93201043@ntu.edu.tw, jhluh@ieee.org).

Li-Chen Fu is with both Dept. of Computer Science & Information Eng. and Dept. of Electrical Eng., National Taiwan University, Taiwan, R.O.C. (e-mail: lichen@ntu.edu.tw).

Jane Yung-Jen Hsu is with both Department of Computer Science & Information Engineering and The Graduate Institute of Networking and Multimedia, National Taiwan University, Taiwan, R.O.C (e-mail: yjhsu@csie.ntu.edu.tw).

Many researchers have investigated the mechanisms for recognizing interactions among humans. In a recognition procedure, firstly sensors are deployed to detect the signals from both the environment and residents, and then the sensed signals are analyzed by activity recognizers to output estimated activities. According to the above discussion, two sensor and model related issues should be concerned.

Regarding sensor deployment and selection, several prior researches have been conducted for multiple-resident activity recognition in a smart home. Generally speaking, these approaches include demanding residents to carry wearable sensors [1] (e.g., RFIDs, infra-red receivers), installing cameras [2-3], or deploying pervasive sensors [4-5] (e.g., pressure sensors, reed switches) to collect important clues for later activity reasoning. What often seem to be ignored, however, is the human-centric concerns, since a smart home system, unlike a public space, should take ergonomic concerns into consideration. For example, a camera-based solution may violate a resident's privacy concerns, and wearable sensors usually cause inconvenience, these two sensors may be not suitable for a smart-home environment targeting maximum comfort. Therefore, Lu *et al.* classifies deployed sensors into seamless and seamful categories to take residents' ergonomic concerns into account, and they suggest taking advantage of as many seamless devices as possible especially for a smart-home environment [6]. As a result, sensor selection is nontrivial and deploying non-obtrusive and pervasive sensors for sensing information is suggested in a smart-home environment.

On the other hand, model design will directly influence the performance of activity recognition. Unlike single-resident activity recognition, multiple-resident activity recognition should additionally consider data association and interactions among residents. Data association, which associates a triggered signal to its corresponding resident, is used to improve the performance of multiple resident activity recognition [4]. Moreover, there exists positive correlation between the accuracy of data association and that of activity estimation [7-8]. These proven suggestions turn our attention to modeling interactions among residents.

Before we introduce interaction modeling, some inherent properties of human interactions at home should be addressed. Human activities can be dependent or independent, and dependent activities can be exclusive or cooperative to one another. In formulating the relationship among human activities using probabilistic models, independent activities are probabilistically independent. However, dependent activities may not be necessarily probabilistically dependent, especially when they are mutually exclusive. For example, a person cannot perform “using computer” if the computer is occupied. If this person decides to watch TV instead because of the unavailability of the computer, there is causality between these two activities in this case since “watching TV” happens because of “using computer.” However, “watching TV” does not always occur when there is another person using the computer. Both activities are independent in most cases. If using computer is not a routine, it will be difficult for a probabilistic model to detect the above-noted causality between “using computing” and “watching TV.” In contrast, cooperative activities like playing chess or preparing meals together are usually probabilistically dependent.

Interaction modeling, which detects the occurrences of interactions and then classifies them into appropriate categories, has also been explored in some prior works[8-9]. However, in a smart home, considerable researches are devoted to either camera based approach or wearable-sensor one; little attention is paid to the suggested approach using pervasive-sensors. The purpose of this paper is to model the interaction by non-obtrusive pervasive sensor in a multi-resident smart home. The interactive activities we deal with in this work are those cooperative, or specifically, probabilistically dependent ones. Moreover, it is suggested that different types of features might be required for interaction modeling [10], thus motivating us to introduce usage of extra auxiliary nodes to describe interaction features.

The remainder of the paper is constructed as follows. After we describe the related work, section III introduces some properties of the dataset we use to evaluate the proposed approach. In the section IV, the methods of data processing and model design are given. The experiment results are given in the section V, and some discussion and the conclusion are presented in the last section.

II. Related Work

Numerous approaches focus on the interaction modeling in the literature, but to the best of our knowledge, few have addressed multiple-resident activity recognition using non-obtrusive sensors. Two exceptions are [4-5], and they both exploit the data association and apply hidden Markov models to infer activities. However, these works do not model the interactions among residents, thus motivating us to do more in-depth literature survey on interaction modeling.

As for interaction modeling using obtrusive sensors, a few approaches can be found. Briefly speaking, these approaches include coupled hidden Markov models (CHMMs [1-2, 11]), dynamic Bayesian networks (DBNs [12]), conditional random fields (CRFs [8]), and emerging patterns (EPs [9]). In addition to above-noted models, there are some other researches focusing on the assistance of hierarchical nodes [3] or duration nodes [13]. These approaches are beyond the scope of this paper, which concentrates on the interaction modeling.

Regarding CHMM related approaches, sensor observations for each person are used to learn a Markov chain and to describe the relations among residents. These approaches usually need to incorporate cameras along with large set of features to get a satisfactory accuracy. Regarding DBN related approaches, domain knowledge is often used to build their graphical structures. Moreover, scale decomposition could be exploited to model residents' interactions using cameras, but it causes sophisticated graphical structures. In the CRF approach [8], a single linear CRF with iterative inference is leaned to classify an activity of interest, and the CRF with decomposition inference uses one CRF to model the activity of one resident, resulting in multiple CRFs in a multiple-resident environment. Both of the two mechanisms do not model the human interaction. In the EP approach [9], the authors apply EPs, which describes significant changes between classes of activities, and design a confidence measure to decide whether two residents have interactions. However, due the current formulation of the EPs, they are intrinsically more prone to cooperative activities even if residents perform activities independently.

In this paper, we exploit auxiliary nodes to incorporate domain knowledge and extend CHMMs to model the activities sensed by non-obtrusive sensors.

TABLE I
THE 15 ACTIVITIES ASKED TO PERFORM BY RESIDENTS IN CASAS PROJECT

Individual	***Cooperative***
Filling medication dispenser	*Moving furniture*
Hanging up clothes	*Playing checkers*
Reading magazine	*Paying bills*
Sweeping floor	*Gathering and packing picnic food*
Setting the table	
Watering plants	
Preparing dinner	

III. Dataset and Data Preprocessing

A. CASAS Dataset

The CASAS dataset, “WSU Apartment Testbed, ADL Multi-resident Activities”, was collected in the CASAS project in the WSU smart workplace [14]. There are totally 26 files of different participant pairs containing sensor events and their corresponding annotations whereby we utilize to validate our proposed models. There are approximately 400 to 800 sensor events in each file, totally 17,238 events in the whole dataset. The data represent the observation of two residents being asked to perform 15 activities of daily living (ADLs), which are listed in Table I. Of these activities, some are individual activities, which are performed by one resident and which are independent to other activities performed by the other residents. On the other hand, cooperative activities are simultaneously performed by one or more residents, and these activities are dependent. The sensors in the dataset include motion sensors, item sensors and cabinet sensors (see Fig. 1.(a)). The output of a sensor event and its annotation is in the format of (Date, Time, SensorID, Value, ResidentID, TaskID). The detail description of the dataset can be found in [5].

B. Data Preprocessing

In order to focus more on human interaction modeling, the data association is assumed to be given. Under this assumption, the events are separated according to their resident IDs and update each resident's state using a procedure as follows: the two residents (with their IDs A and B) in the environment are assigned a null activity and a null *feature vector* as their initial states. Each time getting an event of resident A, his/her current state is updated to a new state and the state of resident B remains unchanged. The same procedure is applied while an event of resident B occurs. Using the above-noted procedure, the original event sequences in the dataset will be transformed into two corresponding feature-vector sequences.

We apply three data preprocessing methods to generate three types of feature vectors in this paper. The three feature vectors are raw feature vector, loc-obj feature vector, and loc-obj with locoff feature vector. The details of these feature vector are described as follows.

The raw feature vector is represented as (*event*, *interaction*), where "*event*" is an integer representing a sensor and its state. There are 37 different sensors in the dataset, and each sensor has two states, on and off. The cardinality of possible integer of "*event*" is thus 75, including a null value.

The loc-obj and the loc-obj with locoff feature are used in the other two data preprocessing method that the *interaction* feature is unchanged, but the information of "*event*" in the raw feature vector are separated according to whether the event is generated by object sensors (item sensors and cabinet sensors) or location sensors (motion sensors) since the information implied in the motion sensor event is essentially different to that in the object sensor event. In the dataset, 10 object sensors are deployed, thus causing 20 possible cases for these object sensor events because each object sensor has two states. With a null value being added, the cardinality is 21. For the location sensors in loc-obj and loc-obj with locoff feature vector, a room-level granularity is choosed. That is, unlike the raw feature vector, the location sensors are embedded into a measure space. Specifically, we divide the environment into six regions as shown in Fig.1(b). Motion sensors in the same region are aggregated and mapped to the same index. The cardinality of this feature is seven because of the six regions and one null value.

The loc-obj with locoff feature vector adds an extra location related feature, locoff, into loc-obj feature vector. This feature records the "off" event of the motion sensor. Because we also take a room-level granularity in this feature, the cardinality of this feature is also seven. The locoff feature is used to indicate a recently occupied region that is now empty and this design provides extra information or correction to the loc-obj feature.

Finally, in all the three types of feature vectors, the dimension "*interaction*" is defined to be a binary feature which captures the information of activities between each resident. This feature becomes 1 if and only if the two residents are in the same region of the environment. We extract this feature because the "interactive activities" we are interested in are cooperative activities. As shown in Table I, these cooperative activities are usually performed by the residents in the same room in the dataset. This feature contains the information about the human interaction by observing if two residents are in the same region.

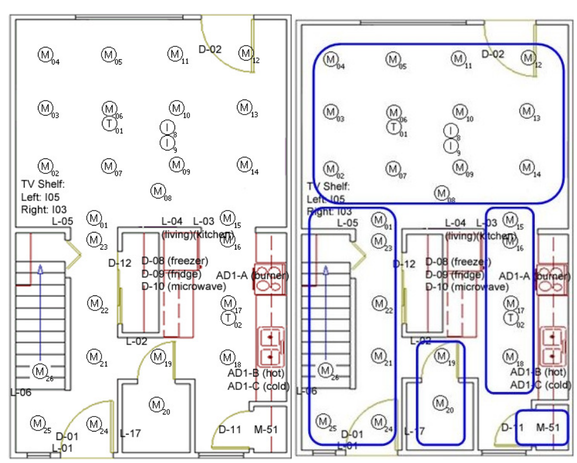
To sum up, we define three types of feature vectors, and each feature vector is generated by one corresponding data preprocessing method. The three feature vectors are raw, loc-obj, and loc-obj with locoff feature vector. These feature vectors are represented as (*event*, *interaction*), (*object*, *location*, *interaction*), and (*object*, *location*, *off-location*, *interaction*) respectively. We generate three types of data according to these feature vectors in the data preprocessing step and evaluate their effectiveness in our experiments.

IV. Human Interaction Modeling in Multi-resident Activity Recognition

In order to compare the performance among different models, we adopt the following three models in this paper, which are all based on the typical hidden Markov model. These three models are parallel hidden Markov model (PHMM), coupled hidden Markov model (CHMM) [10], and a dynamic Bayesian network which extends CHMM by adding some informative vertices.

A. Parallel Hidden Markov Model

PHMM is essentially a combination of two hidden Markov models without adding any edge. That is, this kind of model combines two HMMs without considering any relationship between them. In the CASAS dataset, there are two residents in the environment. Under the assumption that we know the generator of each sensor event, we can create one hidden Markov model for each resident. Let $A^{(1)}=\{a_{\{1:T\}}^{(1)}\}=\{a_1^{(1)},\ldots,a_T^{(1)}\}$ and $A^{(2)}=\{a_{\{1:T\}}^{(2)}\}=\{a_1^{(2)},\ldots,a_T^{(2)}\}$ be the sequences of performed activities from resident 1 and 2 respectively within the internal from time slice 1 to T. Let $O^{(1)}=\{o_{\{1:T\}}^{(1)}\}=\{o_1^{(1)},\ldots,o_T^{(1)}\}$ and $O^{(2)}=\{o_{\{1:T\}}^{(2)}\}=\{o_1^{(2)},\ldots,o_T^{(2)}\}$ be the sequences of sensor events corresponding to $A^{(1)}$ and $A^{(2)}$. We can construct two HMMs separately by $\{A^{(1)}, O^{(1)}\}$ and $\{A^{(2)}, O^{(2)}\}$, and combine



(a) (b)

Fig. 1. (a) Sensor Deployment (M: motion sensors, I: item sensors, D: cabinet sensors) (b) The room-level division

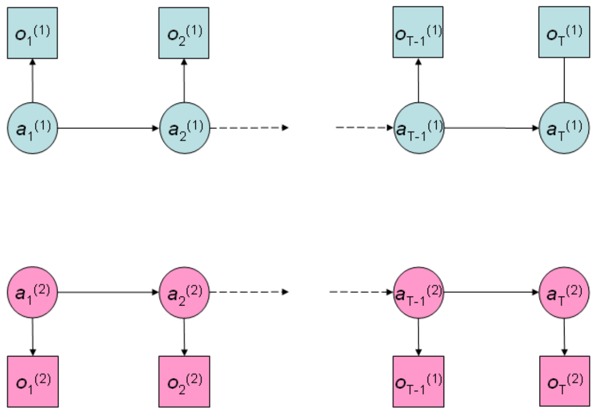


Fig. 2. PHMM consists of two independent HMMs

them independently to obtain a PHMM as shown in Fig. 2. Note that there are edges from $a_t^{(m)}$ to $a_{t+1}^{(n)}$ only when m equals to n. Moreover, we do not use the interaction feature we described in section III-B in this PHMM.

Since the two HMMs in the PHMM are independent, the posterior of activities given the observation is just the multiplication of the two HMMs, which is:

$$
\begin{aligned}
&\prod_{m=\{1,2\}} P(A^{(m)} \mid O^{(m)}) \\
&= \prod_{m=\{1,2\}} \frac{P(A^{(m)})P(O^{(m)} \mid A^{(m)})}{P(O^{(m)})} \\
&\propto P(A^{(m)})P(O^{(m)} \mid A^{(m)}) \\
&= \prod_{m=\{1,2\}} P(a_1^{(m)}) \left(\prod_{t=2}^{T} P(a_t^{(m)} \mid a_{t-1}^{(m)}) \right) \left(\prod_{t=1}^{T} P(o_t^{(m)} \mid a_t^{(m)}) \right)
\end{aligned} \tag{1}
$$

B. Coupled Hidden Markov Model

Unlike the independent setting in a PHMM, the CHMM is, in contrast, a combination of two HMMs which are not independent. In a CHMM, there are directed edges from each $a_t^{(m)}$ to $a_{t+1}^{(n)}$ even if m is not equal to n. When two residents in a smart home act cooperatively, their activities will influence each other. Adding edges across two HMMs means that the activity of one resident at time t can affect not only the activity of his/herself but also that of the other, which is what we can observe when two people interact. However, like a PHMM, the interaction feature we extract from the dataset is ignored as

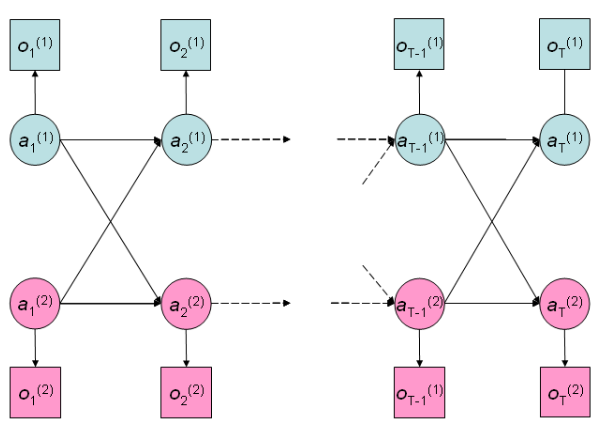


Fig.3. CHMM models human interaction by the cross HMM edges

well in the CHMM model. The graphical topology of a CHMM is shown in Fig.3.

Since the two HMMs are no longer independent in a CHMM, the posterior of the activity sequences given all the observation becomes:

$$
\begin{aligned}
&P(A^{(1)}, A^{(2)} \mid O^{(1)}, O^{(2)}) \\
&= \frac{P(O^{(1)}, O^{(2)} \mid A^{(1)}, A^{(2)})P(A^{(1)}, A^{(2)})}{P(O^{(1)}, O^{(2)})} \\
&\propto P(O^{(1)}, O^{(2)} \mid A^{(1)}, A^{(2)})P(A^{(1)}, A^{(2)})
\end{aligned} \tag{2}
$$

Given the condition independence in the graphical structure of Fig.3, we can factorize $P(O^{(1)}, O^{(2)} | A^{(1)}, A^{(2)})$ and $P(A^{(1)},A^{(2)})$ as follows:

$$
\begin{aligned}
&P(O^{(1)}, O^{(2)} \mid A^{(1)}, A^{(2)}) \\
&= P(o_1^{(1)}, o_1^{(2)} \mid a_1^{(1)}, a_1^{(1)}) \cdots P(o_T^{(1)}, o_T^{(2)} \mid a_T^{(1)}, a_T^{(1)}) \\
&= \prod_{t=1}^{T} P(o_t^{(1)} \mid a_t^{(1)}) P(o_t^{(2)} \mid a_t^{(2)}),
\end{aligned} \tag{3}
$$

and

$$
\begin{aligned}
&P(A^{(1)}, A^{(2)}) \\
&= P(a_1^{(1)}, a_1^{(2)}) \prod_{t=2}^{T} P(a_t^{(1)}, a_t^{(2)} \mid a_{t-1}^{(1)}, a_{t-1}^{(2)}) \\
&= \prod_{m=\{1,2\}} P(a_1^{(t)}) \prod_{t=2}^{T} P(a_t^{(m)} \mid a_{t-1}^{(1)}, a_{t-1}^{(2)})
\end{aligned} \tag{4}
$$

C. CHMM + Interaction vertices

Recall that we assume that the residents in an environment interact only when they are in the same region. We can extract one more binary feature, the interaction feature, accordingly from the dataset. We utilize the interaction feature in the third model as follows: t CHMM model is extended by adding one vertex i_t representing the interaction feature in each time slice t. Some activities are performed interactively, so there should be edges from $a_t^{(m)}$ to i_t. Moreover, people may use different objects when they perform activities solely or cooperatively, this leads to extra edges from i_t to $o_t^{(m)}$. From the aspect of modeling a dynamic Bayesian network, $a_t^{(1)}$ and $a_t^{(2)}$ are two hidden-state vertices at time frame t with causal links to the observation vertices i_t, $o_t^{(1)}$, and $o_t^{(2)}$. After unrolling the time frame and adding edges from $a_t^{(m)}$ to $a_{t+1}^{(n)}$ and i_t to i_{t+1} for all time interval, we get a resultant dynamic Bayesian network as shown in Fig.4.

Inferring the most likely activity sequences given the observation and interaction feature sequences becomes rather complex. By the Bayes rule:

$$
\begin{aligned}
&P(A^{(1)}, A^{(2)} \mid O^{(1)}, O^{(2)}, I) \\
&= \frac{P(O^{(1)}, O^{(2)}, I \mid A^{(1)}, A^{(2)})P(A^{(1)}, A^{(2)})}{P(O^{(1)}, O^{(2)}, I)} \\
&\propto P(O^{(1)}, O^{(2)}, I \mid A^{(1)}, A^{(2)})P(A^{(1)}, A^{(2)})
\end{aligned} \tag{5}
$$

In the extended DBN, the graph topology is no longer the same with the original CHMM. According to the d-separation

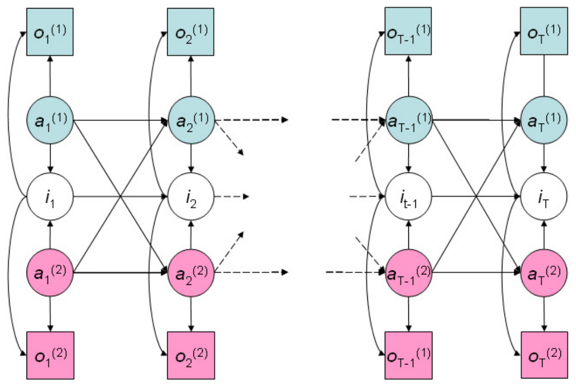


Fig.4. The DBN with additional vertices tries to capture more human Interaction information than a CHMM

criteria[15], however, despite of these newly added vertices that the CHMM does not define, we can still separate the joint probability of activity sequence as follows:

$$\begin{aligned} &P(A^{(1)}, A^{(2)}) \\ &= P(a_1^{(1)}, a_1^{(2)}) \prod_{t=2}^{T} P(a_t^{(1)}, a_t^{(2)} \mid a_{t-1}^{(1)}, a_{t-1}^{(2)}) \\ &= \prod_{m=\{1,2\}} P(a_1^{(t)}) \prod_{t=2}^{T} P(a_t^{(m)} \mid a_{t-1}^{(1)}, a_{t-1}^{(2)}), \end{aligned} \tag{6}$$

The "state transition" parameters in this model are identical to those in a CHMM. However, the observation and interaction feature sequences become

$$\begin{aligned} &P(O^{(1)}, O^{(1)}, I \mid A^{(1)}, A^{(2)}) \\ &= P(o_1^{(1)}, o_1^{(2)}, i_1 \mid a_1^{(1)}, a_1^{(2)}) \prod_{t=2}^{T} P(o_t^{(1)}, o_t^{(2)}, i_t \mid a_t^{(1)}, a_t^{(2)}, i_{t-1}) \\ &= P(i_1 \mid a_1^{(1)}, a_1^{(2)}) \left(\prod_{t=2}^{T} P(i_t \mid a_t^{(1)}, a_t^{(2)}, i_{t-1}) \right) \\ &\quad \left(\prod_{m=\{1,2\}} \prod_{t=2}^{T} P(o_t^{(m)} \mid a_t^{(m)}, i_t) \right) \end{aligned} \tag{7}$$

where the probability of observation here depends on not only the activity but also the interaction features.

We implement the three models using GMTK, the Graphical Models Toolkit[16]. The details about what inference and training algorithms are used by GMTK can be found in the GMTK document. Note that the difference between the above dynamic Bayesian network and the CHMM is whether the interaction feature is taken into consideration. The reason we select these three models in this paper is thus clear: the three models are similar with the exception of how much they take the human interaction into account. We want to verify if this difference affects the performance of these three models under a multi-resident environment. The difference will be more clarified when comparing among Fig.2, 3, and 4. We can see the independent property, the cross edges, and the additional interaction vertices in the three graphical models.

V. Experiments and Result

In order to examine whether taking human interaction into account can help a multi-resident model get better accuracy, we make use of GMTK[16] to implement the three models and run the experiments on the CASAS "WSU Apartment Testbed, ADL Multiresident Activities" dataset. There are 26 files in this dataset. Each file corresponds to one participant pair. We run leave-one-out cross validation experiments in this paper. We use the three different data pre-processing methods we describe in section III-B to separate the sensor events into two concurrent event sequences according to their resident IDs and the extracted features. The results of our experiments are shown in Table II. In each model, the accuracy using raw feature set is better than the others, which is consistent with our previous work[8]. This decline in accuracy for more manipulation may be attributed to the sensitivity to noises. Furthermore, in each data pre-processing method, the dynamic Bayesian network extended from CHMM outperforms the original CHMM and PHMM, and the CHMM performs better than the PHMM.

The result conforms to our expectation. Recall that a PHMM ignores the data dependency of the two hidden sequences in the two parallel HMMs. If this independency assumption does not hold, the PHMM may be biased. The CHMM with additional interaction vertices tries to capture more human interaction information than the original CHMM. The interaction feature we manually labeled may be error-prone. First, it is an overlapping feature, which means that it can be computed from the other features, and information subservient to the performance as well as the noises in the dataset will all be replicated. Second, the assumption about the distance between residents who have interactions is not always true. The residents being at the same room does not mean that they will perform cooperative activities. As a result, the information this interaction feature provides may be limited. However, the model with the assistance of this feature is still more accurate than that of the others. From this result, we will argue that human interaction is a necessary and important factor in designing multi-resident activity recognition models.

TABLE II

Experimental results of the three models runnig on the dataset with three data preprocessing methods. Sub.1 and Sub.2 are resident 1 and 2's activities, and joint is counted when both sub.1 and sub.2 are correct.

	PHMM			CHMM			CHMM+interaction vertices		
Feature set	Sub.1	Sub.2	Joint	Sub.1	Sub.2	Joint	Sub.1	Sub.2	Joint
loc_obj	0.6760	0.5487	0.3906	0.7132	0.7179	0.6217	0.7732	0.7678	0.7045
loc_obj with locoff	0.6990	0.5823	0.4254	0.7319	0.7305	0.6249	0.7666	0.7776	0.6947
raw	0.7885	0.7592	0.6178	0.8162	0.8403	0.7490	0.8472	0.8644	0.7826

VI. Conclusion and Future Work

In this paper, we adopted three different directed graphical models including PHMMs, CHMMs, and the DBNs extended from CHMM with interaction node, to identify the activities in a multi-resident environment. We extract two kinds of features, the object used and the locations of residents, from the dataset, and create the human interaction feature based on the assumption of the distance between two interacting persons. Note that in the CHMM, the human interaction has already been preliminarily identified, and it outperforms PHMM, a model ignoring the dependency between two activity (hidden) sequences. Moreover, adding auxiliary nodes of human interaction to CHMM can still improve the accuracy. The result confirms that human interaction plays an important role in building an intelligent system for smart home that often has more than one resident in it.

Data association and human interaction are two important factors to extend single-resident activity recognition models to multi-resident ones. In our previous work[8], we showed the importance of data association in multi-resident activity recognition. In this paper, we assume data association is given and focus on modeling the interaction between the residents. A model which can combine both two factors will be a better solution to the multi-resident activity recognition problem. Furthermore, scalability of the model in the number of residents is another important issue. In this paper, we use a dataset that has only two residents in it. In an environment containing more than two residents, this model may fail. Adding one sequence for each resident in graphical models will become intractable because of the number of vertices/edges in the graph. The resource for training and inference will be extremely demanding when the number of vertices/edges increases. How to learn a more generalized multi-resident activity model will be an interesting and challenging research.

Acknowledgment

Special thanks to the Washington State University CASAS project for the dataset and Professor Cook in CASAS project for helping us clarify some details about the dataset.

References

[1] L. Wang, *et al.*, "Sensor-Based Human Activity Recognition in a Multi-user Scenario " in *Ambient Intelligence*, ed, 2009, pp. 78-87.

[2] L. McCowan, *et al.*, "Automatic analysis of multimodal group actions in meetings," *IEEE Transactions on Pattern Analysis and Machine Intelligence,* vol. 27, pp. 305-317, 2005.

[3] N. Nguyen, *et al.*, "Recognising Behaviours of Multiple People with Hierarchical Probabilistic Model and Statistical Data Association," presented at the British Machine Vision Conference 2005.

[4] D. H. Wilson and C. G. Atkeson, "Simultaneous Tracking and Activity Recognition (STAR) Using Many Anonymous, Binary Sensors.," presented at the Pervasive, 2005.

[5] G. Singla, *et al.*, "Recognizing independent and joint activities among multiple residents in smart environments," *Ambient Intelligence and Humanized Computing Journal,* 2010.

[6] C.-H. Lu, *et al.*, "Hide and Not Easy to Seek: A Hybrid Weaving Strategy for Context-aware Service Provision in a Smart Home," presented at the IEEE Asia-Pacific Services Computing Conference, Yilan, Taiwan, 2008.

[7] D. J. Cook and S. K. Das, "How smart are our environments? an updated look at the state of the art," *Pervasive Mobile Computing,* vol. 3, pp. 53-73, 2007.

[8] K.-C. Hsu, *et al.*, "Strategies for Inference Mechanism of Conditional Random Fields for Multiple-Resident Activity Recognition in a Smart Home," presented at the The Twenty Third International Conference on Industrial, Engineering & Other Applications of Applied Intelligent Systems (IEA-AIE), Córdoba, Spain, 2010.

[9] T. Gu, *et al.*, "Mining Emerging Patterns for recognizing activities of multiple users in pervasive computing," presented at the Mobile and Ubiquitous Systems: Networking & Services, MobiQuitous, 2009.

[10] Y. Du, *et al.*, "Human Interaction Representation and Recognition Through Motion Decomposition," *Signal Processing Letters,* vol. 14, pp. 952-955, 2007.

[11] M. Brand, *et al.*, "Coupled hidden markov models for complex action recognition," in *Proceedings of IEEE CVPR97*, 1997.

[12] D. Youtian, *et al.*, "Recognizing Interaction Activities using Dynamic Bayesian Network," in *Pattern Recognition, 2006. ICPR 2006. 18th International Conference on*, 2006, pp. 618-621.

[13] P. Natarajan and R. Nevatia, "Coupled Hidden Semi Markov Models for Activity Recognition," presented at the Proceedings of the IEEE Workshop on Motion and Video Computing, 2007.

[14] W. S. University. *CASAS Smart Home Project* Available: http://ailab.eecs.wsu.edu/casas/

[15] J. Pearl, *Probabilistic reasoning in intelligent systems: networks of plausible inference*. San Francisco, CA, USA: Morgan Kaufmann Publishers Inc., 1988.

[16] J. Bilmes. *The Graphical Models Toolkit (GMTK).* Available: http://ssli.ee.washington.edu/~bilmes/gmtk/