Multioccupant Activity Recognition in Pervasive Smart Home Environments

ASMA BENMANSOUR, University of Tlemcen
ABDELHAMID BOUCHACHIA, Bournemouth University
MOHAMMED FEHAM, University of Tlemcen

Human activity recognition in ambient intelligent environments like homes, offices, and classrooms has been the center of a lot of research for many years now. The aim is to recognize the sequence of actions by a specific person using sensor readings. Most of the research has been devoted to activity recognition of single occupants in the environment. However, living environments are usually inhabited by more than one person and possibly with pets. Hence, human activity recognition in the context of multioccupancy is more general, but also more challenging. The difficulty comes from mainly two aspects: resident identification, known as data association, and diversity of human activities. The present survey article provides an overview of existing approaches and current practices for activity recognition in multioccupant smart homes. It presents the latest developments and highlights the open issues in this field.

Categories and Subject Descriptors: I. [**Computing Methodologies**]; I.2 [**Artificial Intelligence**]; I.2.4 [**Knowledge Representation and Reasoning**]: Machine Learning, Machine Learning Approaches, Learning in Probabilistic Graphical Models

General Terms: Theory, Design, Performance

Additional Key Words and Phrases: Human activity recognition, smart homes, pervasive environments, multiple occupancy

ACM Reference Format:
Asma Benmansour, Abdelhamid Bouchachia, and Mohammed Feham. 2015. Multioccupant activity recognition in pervasive smart home environments. ACM Comput. Surv. 48, 3, Article 34 (December 2015), 36 pages.
DOI: http://dx.doi.org/10.1145/2835372

1. INTRODUCTION

The greatest reason for the continued development of smart homes is to assist disabled and elderly people, especially those with chronic diseases, to accomplish their Activities of Daily Living (ADL) efficiently and consequently to enhance their well-being and independent living. Most of the research on smart homes has investigated the mono-occupant setting, where the assumption is that a living space is occupied by single individuals [Khan et al. 2012; Hu et al. 2009; Riboni et al. 2011; Kasteren et al. 2008; Sarkar et al. 2010; Kasteren et al. 2010, 2011; Nait Aicha et al. 2013; Gu et al. 2009a]. However, homes often have more than one occupant, referred to as multioccupancy. Even when old persons generally live alone, they could have pets and could receive guests like care professionals and family members. Therefore, developing solutions for multioccupancy is extremely vital.

Authors' addresses: A. Benmansour and M. Feham, STIC laboratory, University of Tlemcen, Tlemcen (13000), Algeria; emails: benmansour_asma@hotmail.fr, feham_m@yahoo.fr; A. Bouchachia, Department of Computing, Faculty of Science and Technology, Poole, BH12 5BB, UK; email: abouchachia@bournemouth.ac.uk.

ACM 0360-0300/2015/12-ART34 $15.00
DOI: http://dx.doi.org/10.1145/2835372

Multioccupancy has not been studied much so far, because the field is still young and because many outstanding challenges in single occupancy still need to be overcome before the research community can focus on multioccupancy [Prossegger and Bouchachia 2014]. Recently, an increasing interest has been witnessed acknowledging the prominence of multioccupancy as a research area in the context of smart homes and activity recognition.

There are two main technologies used to recognize human activities in smart environments including homes: computer vision [Nguyen et al. 2006; McCowan et al. 2005; Du et al. 2006, 2007; Natarajan and Nevatia 2007] and pervasive sensing [Prossegger and Bouchachia 2014; Crandall and Cook 2008a, 2008b, 2010; Hsu et al. 2010; Wilson and Atkeson 2005; Chiang et al. 2010; Cook et al. 2010; Singla et al. 2010; Alerndar et al. 2013; Wang et al. 2009, 2011; Gu et al. 2009a; Lin and Fu 2007; Chen and Tong 2014; Cook 2009; Crandall and Cook 2009]. In this survey, we focus only on the latter technology. Pervasive sensors are used to collect data related to the human physiology, the human activity, as well as the environment. Such data is processed in order to extract cues and patterns about various aspects such as the resident's profile, health status of the resident, the living environment, and the resident-environment interaction. Because of the very complex nature of human activities, the task of recognition in a pervasive context is very difficult, especially when pervasive data generated by sensors is noisy. Multioccupancy comes with specific scientific and technological challenges [Chen and Tong 2014] related to resident identification [Crandall and Cook 2008a, 2008b, 2010; Hsu et al. 2010; Wilson and Atkeson 2005; Cook et al. 2010; Alerndar et al. 2013; Chen and Tong 2014], activity tracking [Prossegger and Bouchachia 2014; Crandall and Cook 2009], behavior patterns of residents [Gu et al. 2009b], and conflict management [Hsu and Wang 2008].

While there have been several review papers published over the recent years devoted to activity recognition and to smart environments in general [Acampora et al. 2013; Chan et al. 2009; Sadri 2011], there is no study on multioccupancy that draws the picture of the current advances in this area; hence, the importance of this present survey. We will provide full coverage of techniques, methods, and open issues related to multioccupancy.

The remainder of this article is organized as follows. Section 1 describes the types of ADL. Section 2 presents the problems encountered in multioccupant activity recognition. Section 3 highlights the problem of multioccupancy focusing on two aspects: data association and interaction. Section 4 discusses sensor technology used in recent research related to pervasive multioccupant activity recognition. Section 5 presents the computational models used for modeling multioccupant activities. Section 6 presents some of the issues encountered in multioccupancy such as identification and interaction. Section 7 presents a sample of publicly available datasets. Section 8 provides examples of international research groups in the pervasive computing area for both single-occupant and multioccupant settings. Section 9 goes through a sample of open questions in this area. Section 10 concludes the article.

2. HUMAN ACTIVITY RECOGNITION

Activity recognition is the process of automatically identifying human actions from the data captured by various types of sensors. It is relevant to many real-world applications such as surveillance, assisted living, and healthcare. Modeling simple activities has been the focus of most of the activity recognition research, while complex activities have only recently started to attract attention from the ambient intelligence and pervasive computing communities [Kim et al. 2010]. Complex activities are common and can be performed by either single persons or by a group of people. We can therefore distinguish different types of activities:

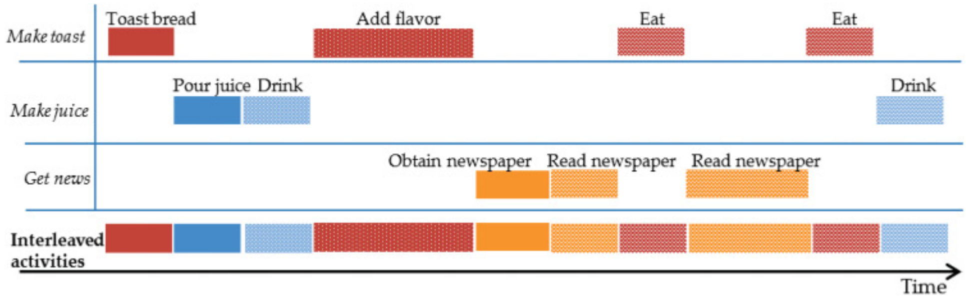

Fig. 1. Interleaved activities (a single occupant).

—*Complex activity* consists of many subactivities as fine-grained activities. For instance, the activity "*cooking soup*" could be modeled as a sequence of subactivities: *measure water, pour water into a pot, add contents of the bag, cook, and serve in a bowl*.
—*Simple activity* is usually an atomic activity that cannot consist of simpler activities, for instance, *pour water*.
—Moreover, we can distinguish two types of activities of daily life (ADLs):
—*Basic ADLs* refer to self-care tasks (e.g., eating, moving, dressing, bathing and showering, grooming, and toilet hygiene) [Roley et al. 2008].
—*Instrumental ADLs* are not essential for basic living, but they let an individual live independently in a community (e.g., doing housework, meeting with people, doing shopping, taking medicine, using technology, using transportation) [Bookman et al. 2007].

Most of the state-of-the-art research has investigated monitoring and assisting people in single-occupancy living spaces. Nevertheless, living spaces are usually inhabited by more than a single person; hence, designing solutions for handling multioccupancy is of prominent importance. In fact, recently, multioccupancy research has gained more attention. However, the pace of research is slow and many outstanding problems are still ahead. The reason for this is that there have been numerous other challenges with single occupancy to deal with before tackling multioccupancy.

The research work published on multioccupancy is mainly related to activity modeling and data association. The challenge is to find suitable models to address the problem of data association, to build activity recognizers that capture the various interactions between occupants. Data association is about the identification of the residents, by whom each sensor is triggered. That is about mapping sensed data to the occupant who actually caused the generation of the data. In activity modeling, we distinguish between five types of activities:

(1) *Sequential activities* where each activity is performed after another in a sequential fashion without any interweaving (e.g., make a phone call, washing hands, and then cooking).
(2) *Interleaving activities* where a single occupant switches between many activities (e.g., switching between chopping vegetables and stirring soup in the kitchen) as shown in Figure 1.
(3) *Concurrent activities* where a single occupant carries out more than one activity at a time (e.g., talking on phone, while cooking).
(4) *Parallel activities* where many occupants perform many activities at the same time (e.g., one occupant is watching TV in the living room, while the other is cooking in the kitchen).

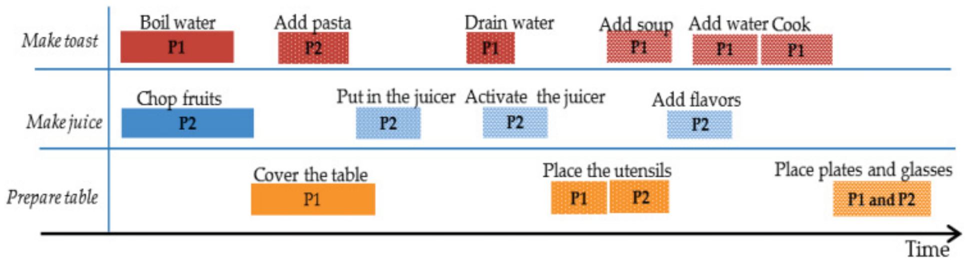

Fig. 2. Collaborative activities (two occupants: P1 and P2).

(5) *Collaborative activities* where many occupants work together in a cooperative manner such that each occupant performs certain actions of the same activity, either together (e.g., two persons moving a table by holding it by the ends) or in parallel (e.g., one person is chopping vegetables, while the other is boiling broth to make soup) as shown in Figure 2.

While the first three activity types are concerned with a single person (also termed as *exclusive activities*), but done in the presence of multiple occupants, the latter two are relevant for multioccupancy. Obviously, the complexity of ADLs increases as the number of occupants in the living environment increases and the activities tend to be cooperative (e.g., watching TV or play a board game). Also the cooperative activities tend to be generally instrumental.

The existing state-of-the-art literature on multioccupant smart homes indicates that these types of activities are not yet fully addressed. Many of the studies are done on simple scenarios like elementary activities [Wilson and Atkeson 2005] (e.g., whether a person moves or not) and sequential activities [Cook et al. 2010; Singla et al. 2010], although parallel exclusive and cooperative activities are the most frequent in nature. Almost, no work has addressed all types of activities. A more mature study in this area has been conducted by the computer vision community using normal cameras [Nguyen et al. 2006; McCowan et al. 2005; Du et al. 2006, 2007; Natarajan and Nevatia 2007]. Vision-based studies are nevertheless out of the scope of this survey.

3. MULTIOCCUPANCY PROBLEM

The challenge of multioccupant smart homes is to design a computational model to deal with the problem of data association (i.e., the identification of the resident) and to efficiently capture the interactions between the occupants.

3.1. Data Association

In a smart home environment shared by multiple residents, the identification of the resident is crucial. Recognizing who triggered the events allows efficient and accurate tracking of the residents' activities. The problem of data association consists of mapping the sensed data to the occupant causing its generation. Failing to do so, that data will not be useful and could even endanger the life of residents in telehealth/telecare context, if important actions are to be taken based on the assessment of such activity data. The data association problem is encountered either when using nonintrusive sensors that cannot directly identify residents in a smart home [Crandall and Cook 2008a, 2008b, 2010; Hsu et al. 2010; Wilson and Atkeson 2005; Cook et al. 2010; Alerndar et al. 2013; Chen and Tong 2014] or when using unlabeled data. All the studies in the literature show that data association is a fundamental problem when modeling activities in a multiple-occupant environment [Hsu et al. 2010].

3.2. Interaction

The main difference between a single-occupant environment that is characterized by exclusive activities and a multioccupant environment is the interaction between individuals to complete cooperative activities. Cooperative activities are usually interdependent activities. For example, a resident cannot do "toileting" because the bathroom is busy. Instead, he decides to do "dressing" [Hsu et al. 2010]. Clearly "dressing" takes place because "toileting" did not happen. But, "dressing" does not always occur when the bathroom is busy. Hence, these two activities are unrelated in the general case. Collective activities that involve many persons are usual in the daily life, such as watching TV, eating, gardening, etc.

People perform certain activities collectively because such activities require cooperation. In terms of interaction, we can distinguish two distinctive types of interdependence [Smith and Mackie 1999]: *social interdependence* and *task interdependence*. A task is socially interdependent if people rely on one another for its full completion, like playing monopoly. It would be more enjoyable to play monopoly in a group than alone. A task is interdependent if more than one person is required to accomplish the activity, like moving a table into another room of the house requires at least two persons.

Studies report that old people tend to isolate themselves [Anon. 2014], which may lead to dementia [Fratiglioni et al. 2000] or simply cause damage to their health [Cornwell and Waite 2009]. Recently, researchers have relied on wearable devices to study social behavior [Wang et al. 2009, 2011; Gu et al. 2009a; Gross 2007; Olguin et al. 2009; Eagle 2008; Choudhury 2004] in the context of multioccupant activity recognition. However, they have recognized the need to use nonintrusive sensors to monitor residents' behavior and develop real-world applications for older adults.

4. SENSING FOR MULTIOCCUPANCY ACTIVITY RECOGNITION

In terms of sensor deployment and selection, Table I presents different types of sensors used for multiple-occupant activity recognition. We can clearly distinguish two major classes of approaches: those relying on wearable sensors [Wang et al. 2009, 2011; Gu et al. 2009a] (e.g., Radio Frequency Identification (RFID)) and those based on infrastructure sensors (e.g., passive infrared sensors). Generally, researchers working on multioccupancy problems tend to use wearable sensors to reduce the problem complexity as these types of sensors can address the data association problem. However, it is often the case that smart home systems ignore ergonomic requirements. Wearable sensors offer the possibility of capturing fine-grained observations but cause inconvenience and are not appropriate for smart homes requiring privacy and comfort. Furthermore, this type of sensor is inappropriate for some people, especially for elderly people, who are not willing to wear them, tend to forget to wear them, or let the device's power source die.

Pervasive infrastructure sensors offer the advantage of being nonintrusive to the residents, but the data association problem is difficult to solve. The use of this type of sensor implies designing specific solutions for data association as in Crandall and Cook [2008a, 2008b, 2010], Hsu et al. [2010], Wilson and Atkeson [2005], Cook et al. [2010], Alerndar et al. [2013], and Chen and Tong [2014]. For instance, the authors in Wilson and Atkeson [2005] used infrastructure and nonintrusive sensors to monitor the residents at home. They studied the effect of sensor settings on the accuracy of resident identification. Three types of configurations were defined: *normal*, *extra*, and *fewer* configurations. The *normal configuration* contains one motion detector, one contact switch, and one pressure mat for each room. The *extra* configuration contains three motion detectors, three contact switches, and three pressure mats per room. The *fewer* configurations contain only one motion detector per room. They found that the

Table I. Summary of Selected Studies

Ref.	Dataset	Sensors	Data association	Activities covered[1]	Approach	Test & validation[2]	Evaluation metric[3]	Results
Crandall and Cook [2008a]	CASAS Student laboratory	Motion, Light, Door switch	Yes	-	NBC	Hold-out method	-Accuracy -False positive	-92% -7%
Crandall and Cook [2008b]	Same as Crandall and Cook [2008a]	Same as Crandall and Cook [2008a]	Yes	-	NBC and HMM	Same as in Crandall and Cook [2008a]	-Accuracy - False positive	-NBC: same as in Crandall and Cook [2008a]. -HMM: 84%; NBC: 76%
Crandall and Cook [2010]	CASAS B&B and TwoR	Motion, Door, Cabinet, Water flow, Power, light, Power	Yes	-	NBC and HMM on datasets B&B and TwoR	Threefold CV	-Accuracy rate -Average lag -Error rate	-NBC: 93.3% (B&B), 89.3% (TwoR) -HMM: 94% (B&B) 90.2% (TwoR)
Hsu et al. [2010]	CASAS Multiresi-dent ADLs	Motion, Item sensors, Cabinet, Water, Burner, Phone, Temperature	Yes	Sequential	2 CRFs (data association, activity recognition)	Leave-one-out CV	Average accuracy	50% (for raw representation of data = best representation)
			No	Sequential Parallel	CRF with decomposition inference			59% (R1), 64% (R2)
Wilson and Atkeson [2005]	Simulated data	Motion, Contact switches, Pressure mat, Beam	Yes	-	HMMs (one HMM for each resident)	One day of data	Time-slice accuracy	-Off-line learning: 100% (1 R) to 82% for 4Rs -Online learning: 100% (1 R) to 67% (4 Rs) -Combined learning: 100% (1 R) to 74% (4 Rs)
	Real-world data	RFID, Notion, Contact						-85.3% for 1R -82.1% for 2Rs -86.4% for 3Rs
Chiang et al. [2010]	Same as Hsu et al. [2010]	Same as Hsu et al. [2010]	No	Sequential Parallel Coopera-tive	-PHMM -CHMM -DBNs extended	Leave-one-out CV	- Accuracy (R1) -Accuracy (R2) -Joint accuracy	-78.85%, 81.62%, 84.72% -75.92%, 84.03%, and 86.44% -61.78%, 74.9%, 78.28%

(Continued)

Table I. Continued

Ref.	Dataset	Sensors	Data association	Activities covered[1]	Approach	Test & validation[2]	Evaluation metric[3]	Results
Cook et al. [2010]	Same as Hsu et al. [2010]	Same as Hsu et al. [2010]	Yes	Sequential	-Bayesian update -HMMs (1 for data association, 1HMM for recognition)	Threefold CV	-Average accuracy -Average precision -Average recall -Average f-score	-HMM: 90%, Bayes: 57% -HMM: 93% -HMM: 96% -HMM: 94%
[Singla et al. 2010]	Same as [Hsu et al. 2010]	Same as Hsu et al. [2010]	No	Sequential	1 HMM for all activities of both residents	Threefold CV	Average accuracy	60.60%
				Sequential Parallel	1 HMM for each resident			73.15%
Alerndar et al. [2013]	ARAS	Photocell, Infrared receiver, Force, Proximity, Sonar distance, Temp., Contact, Pressure mat	Yes	Sequential Parallel Interweaved Cooperative	1 HMM (combined label for multioccupant activities)	Leave-one-out CV	Average accuracy	-61.5% (House A) -76.2% (House B)
[Wang et al. 2009]	Real-world Data	Audio recorder, iMote2 with ITS400, RFID wristband reader	no	Sequential Parallel Interweaved Cooperative	CHMM	10-fold CV	-Time-slice overall accuracy -Time-slice single-occupant ADLs -Time-slice multi-occupant ADLs	-82.22% (R1) and 88.71% (R2). -74.79% and 85.11. -96.91% and 95.91%.
[Wang et al. 2011]	Same as [Wang et al. 2009]	Same as [Wang et al. 2009]	no	Sequential Parallel Interweaved Cooperative	CHMM	10-fold CV	Same as in [Wang et al. 2009]	Same results as in [Wang et al. 2009]
					FCRF			-86.7% and 86.37% -85.75% and 82.56% -87.02% and 88.84%
[Gu et al. 2009a]	Same as [Wang et al. 2009]	Same as [Wang et al. 2009]	no	Sequential Parallel Interweaved Cooperative	EPs (EPs for single-user ADLs and EPs for multi-occupant ADLs)	10-fold CV	-Time-slice single-occupant ADLs -Time-slice multi-occupant ADLs	-86.69% (R1) and 85.57% (R2) -95.06% (R1) and 95.71% (R2)

(Continued)

Table I. Continued

Ref.	Dataset	Sensors	Data association	Activities covered[1]	Approach	Test & validation[2]	Evaluation metric[3]	Results
[Lin and Fu 2007]	Real-world data	Motion, Thermometers, Humidity, light sensors, Smoke, RFIDs, Cameras	no	Sequential Parallel Inter-weaved Coopera-tive	Three-layer model Ontology + DBNs + BN	Leave-one-out CV	-Accuracy rate	-88.89%
[Chen and Tong 2014]	Same as [Hsu et al. 2010]	Same as [Hsu et al. 2010]	yes	Sequential Parallel Coopera-tive	HMM and CRF on multi-occupant activity recognition	3-fold CV	Average accuracy	-HMM: 75,77%, CRF: 75,38% (recognition) -HMM: 84,19%, CRF: 82,88% (data association)
					Same models but in the context of multi-label classification measurement	Hold-out method	-Average accuracy -Average error rate -Average precision -Average recall -Average f-score	-HMM:97.40%, CRF:97.25% -HMM:2.6%, CRF:2.75% -HMM:80.03%, CRF:80.05% -HMM:81.92%, CRF:79.91% -HMM:40.48%, CRF:39.99%

[1]"-" means that the authors didn't perform activity recognition, only data association was considered.
[2]R stands for "resident", R1 indicates Resident1.
[3]CV stands for "cross validation".

extra configuration achieves better accuracy regardless of the number of occupants. In contrast, with the fewer configurations, sensor observations do not provide enough information to the model to clearly identify the individuals and the model can be confused for a long period of time before it becomes able to distinguish between residents.

The authors in Lu et al. [2008] classify sensors into seamless and seamed ones following ergonomic criteria. They suggest taking advantage of many seamless sensors in the living space. In fact, by decreasing the number of seamed sensors, the behavior of people will not be much impacted. The authors also claim that developing a passive solution secures a clean design that separates technology from the smart space and consequently makes the space as natural as possible.

5. COMPUTATIONAL MODELS FOR MULTIOCCUPANCY ACTIVITY RECOGNITION

Diverse computational models have been applied in the context of single-occupancy ADLs ranging from probabilistic models to standard data mining and machine learning models like neural networks, decision trees, ontologies, etc. In the case of multioccupancy, however, no such diversity of models exists. Almost all of the proposed models are essentially probabilistic based on graphical models. This conclusion can easily be observed in the following Table I that illustrates also a summary of a set of representative research studies covering the sensors used, the type of activities covered, the models used, and the evaluation metrics as well the results obtained.

In the following sections we describe two classes of models applied in the context of multioccupancy: graphical models and association rule mining.

5.1. Graphical Models

Graphical models are the most popular computational models used in activity recognition in general. As their names indicate, graphical models are probabilistic models having the structure of graphs that represent conditional dependence between nodes, which are random variables.

Graphical models are defined as probability distributions that factorize according to a graph [Sutton and McCallum 2006]. The goal is to infer a matching sequence of hidden states that maximizes the probability of the activities given some sensor readings. We distinguish two classes of graphical models: *generative* and *discriminative*, which are explained next.

5.1.1. Generative Models. Generative models define the joint probability distribution and can be used to generate (sample) data from such distribution or to perform inference given a novel sequence of observations [Kasteren 2011]. We focus here only on models that have been applied in the context of multioccupancy.

5.1.1.1. Naive Bayes Classifier (NBC). The naive Bayes model can be considered as one of the most simplistic probabilistic models. It is a restricted version of the Bayesian Network (BN). In fact, the naive Bayes model assumes all data points (e.g., the events of sensors in the case of activity recognition) are independently and identically distributed. The class nodes have no parents and the attribute nodes are not connected. Moreover, NBC does not take into account any temporal relations between data points. The joint probability of observations and labels can be factorized as

$$p(X, Y) = \prod_{t=1}^{T} p(x_t \,|y_t)p(y_t), \tag{1}$$

where $p(y_t)$ is a prior probability over activities. To compute the conditional probability of labeled data (X,Y) in a straightforward way, we assume independence between input

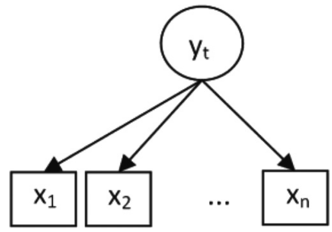

Fig. 3. Naive Bayesian representation.

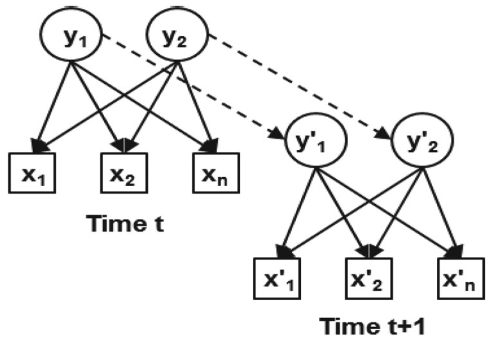

Fig. 4. Dynamic Bayesian network.

features given the input labels. The probability can then be written as follows:

$$p(x_t \mid y_t) = \prod_{i=1}^{N} p(x_t^i \mid y_t). \tag{2}$$

In our setting, the set X represents the sensor data, while Y represents the set of activities as shown in Figure 3.

Activity recognition can be considered as a classification problem where activities are regarded as classes [Van Laerhoven et al. 2003; Liao and Ji 2009]. NBC and, in general, conventional BNs are not suitable for modeling temporal processes because directed arcs of the network do not give any information about the time. In order to overcome this limitation, Dynamic Bayesian Networks (DBNs) were proposed as an upgrade of BNs.

5.1.1.2. Dynamic Bayesian Network (DBN). DBNs are designed to deal with temporal processes (time series). A DBN results from extending a BN by sequencing interlinked time-sliced instances of the BN as shown in Figure 4. When a DBN is applied to activity modeling, the observables, X_t, correspond to the sensor readings, while the unobservable variables, Y_t, correspond to the activities. A state at a specific time t depends on the previous states.

Formally, a DBN [Sanghai et al. 2005] is defined as a pair of BNs(B_1, $B\rightarrow$), where B_1 is prior, which defines the initial distribution $p(Z_1)$, and $B\rightarrow$ is a two-time-slice BN defining the transition distribution $p(Z_t|Z_{t-1})$ via a directed acyclic graph:

$$\begin{aligned} p(Z_t|Z_{t-1}) &= \prod_{i=1}^{N} p(Z_t^i|\text{Parents}(Z_t^i)), \\ p(Z_t) &= \prod_{t=1}^{T}\prod_{i=1}^{N} p(Z_t^i|\text{Parents}(Z_t^i)). \end{aligned} \tag{3}$$

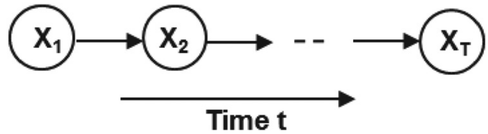

Fig. 5. Representation of a Markov model.

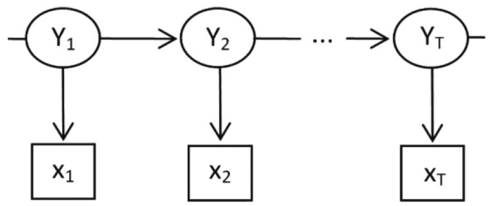

Fig. 6. HMM structure.

Z_t^i is a node at time slice t; it can be a hidden node, an observation node, or a control node (optional), while Parents(Z_t^i) are parent nodes of Z_t^i and can be at either time slice t or $t-1$.

5.1.1.3. Markov Model. A Markov Model (MM) is a simplification of a DBNs that models the temporal aspect of processes as shown in Figure 5. The first order Markov assumption was proposed to simplify the dependence relationship between consecutive states. It stipulates that the present state at time t depends only on the previous one: $p\,(y_t|y_1, y_2, y_3, \ldots, y_{t-1}) = p\,(y_t|y_{t-1})$. Thus, the future state depends only on the current state, not on past states [Sutton and McCallum 2006]; that is, y_t depends only on y_{t-1}.

A specific case of MMs is called the Markov Chain (MC) and corresponds to the case where the states are all observable. A MC is a sequence of random variables X_1, X_2, $X_3, \ldots, X_T$ with the Markov property. Formally,

$$p(X_{t+1} = x\,|X_1 = x_1, X_2 = x_2, \ldots, X_t = x_t) = p(X_{t+1} = x\,\big|X_t = x_t)\,. \tag{4}$$

The possible values of X_i form a countable set S called the state space of the chain.

MCs are not popular in human activity modeling since we cannot always directly recognize the activities from sensory data. In general, only simple activities can be modeled using MCs [Kim et al. 2010]. Interestingly, MCs were applied in Crandall and Cook [2008a] to model the data association problem in order to identify the residents.

5.1.1.4. Hidden Markov Model (HMM). The most popular generative temporal probabilistic model is the HMM. In contrast to MCs, HMM consists of hidden and observable states. The data $(x_1, x_2, \ldots, x_T)$ is therefore assumed to be generated by a temporal process whose states are hidden, $(y_1, y_2, \ldots, y_T)$, as shown in Figure 6.

HMM relies on two assumptions: the first order Markov assumption in relation to the independence of hidden states and the conditional independence of observation parameters stipulating that $p\,(x_t|y_t, x_1, x_2 \ldots x_{t-1}, y_1, y_2 \ldots y_{t-1}) = p\,(x_t|y_t)$. The observable state at time t, x_t, depends only on the current hidden state y_t. That is, the probability of observing x_t while being at y_t is independent of all other observable and hidden variables [Sutton and McCallum 2006].

The joint probability $p\,(x, y)$ of the observations and hidden states can be factorized as follows:

$$p(x, y) = \prod_{t=1}^{T} p(y_t\,|y_{t-1})p(x_t\,\big|y_t), \tag{5}$$

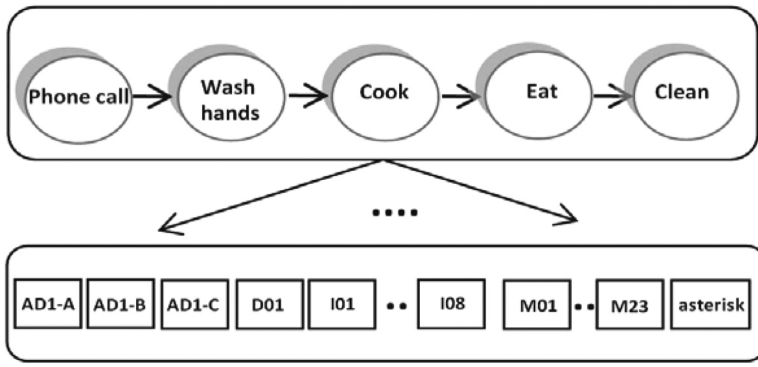

Fig. 7. Representation of a global HMM.

where $p(y_t|y_{t-1})$ and $p(x_t|y_t)$ indicate the probability of transition between the two consecutive hidden states y_{t-1} and y_t and the probability of observing x_t at state y_t, respectively [Sutton and McCallum 2006]. Given the sequence of observables, the maximum of the joint probability corresponds to the highly probable sequence of hidden states.

In the context of activity recognition, similar modeling is adopted like with the previous computational models. That is, the hidden states are the activities and observations are the sensed data as shown in Figure 7. The activities are represented as ellipses, while the observables (sensors) are represented as rectangles. The links between the hidden states are labeled with the transition probabilities and those between the hidden states and the observables are labeled with the emission probabilities.

Usually when activities are sequential, it is possible to separate activity data and then create one HMM for each activity. However, if the activities are interleaved, this way of modeling is not suitable because the interlacement of the activities will be disregarded. In addition, finding the optimum number of hidden states for each HMM corresponding to an activity is another issue. Creating an HMM for each activity would lead to having the same sensor model for each activity, but the number of hidden states for each activity is unknown. In fact, the authors in Khan et al. [2012] used the accuracy to find the optimum number of hidden states and suggested using techniques applied for Hierarchical Dirichlet Process HMM (HDP-HMM) [Hu et al. 2009] and infinite HMMs [Pruteanu-Malinici and Carin 2008].

In some cases, even when the complex activity is decomposed, its subactivities also form complex activities that are not directly observable (hidden). For instance, the activity "Prepare a dinner" can be decomposed into the activity "prepare a drink" and the activity "cook" and each of them also includes subactivities. Then individually trained HMMs on the activities can be combined to build a global HMM. Thus, hierarchical graphical models (e.g., Hierarchical HMM or Abstract HMM) look more suitable in this case.

5.1.1.5. Parallel Hidden Markov Model (PHMM). A parallel hidden Markov model consists of a set of independent HMMs. In other terms, PHMMs are standard HMMs that are used in parallel under the assumption that the corresponding individual processes being modeled evolve independently from one another with independent output. Therefore, when applied for activity recognition in a k-occupant environment [Chiang et al. 2010], PHMM will consist of k independent HMMs, one for each occupant, where the hidden states correspond to the activities and the observations correspond to the sensor values.

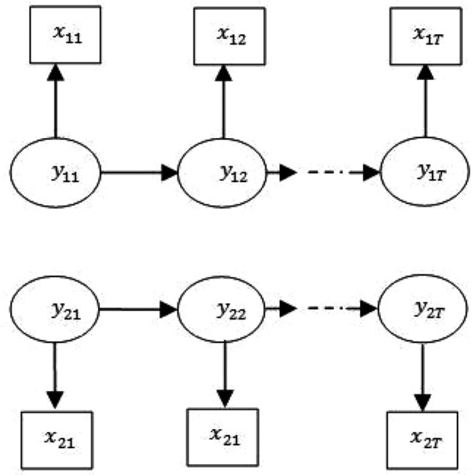

Fig. 8. Representation of PHMM consisting of two HMMs.

Formally, let $Y_{m\{1,2,\ldots,k)} = \{y_{m1}, y_{m2} \ldots y_{mT}\}$ be the sequence of activities by the resident m over the time $t = 1 \ldots T$ and the corresponding sequence of observations $X_{m\{1,2,\ldots,k)} = \{x_{m1}, x_{m2} \ldots x_{mT}\}$. Then, given the observations, the posterior probability of the activities is computed as the product of the k HMMs thanks to their independence [Chiang et al. 2010]:

$$\prod_{m=\{1,2,\ldots,k\}} p\,(Y_m|\,X_m) = \prod_{m=\{1,2,\ldots,k\}} \frac{P\,(Y_m)\,P\,(X_m|\,Y_m)}{P\,(X_m)} \propto P\,(Y_m)\,P\,(X_m|\,Y_m)$$
$$= \prod_{m=\{1,2,\ldots,k\}} p(y_{m1}) \left[\prod_{t=2}^{T} p(y_{mt}\,|\,y_{mt-1})\right] \left[\prod_{t=1}^{T} p(x_{mt}\,|\,y_{mt})\right]. \quad (6)$$

Figure 8 shows the case of $k = 2$.

While the application of PHMMs is easy and straightforward, their capabilities are limited in the context of multioccupant activity recognition due to the lack of interaction between HMMs and data independence between the models. This is the main reason why PHMMs have been mainly used for modeling the occupants rather than modeling the interaction.

5.1.1.6. Coupled Hidden Markov Model (CHMM). A CHMM is a combination of a set of HMMs that interact with one another in different ways. In a CHMM, a state in an HMM depends on the previous states from the other HMMs. Figure 9 illustrates the general scheme of coupling, where the hidden states of two chains HMM are fully connected.

For activity modeling, coupled HMMs can be very useful in collaborative activities by many inhabitants. For instance, in Cook [2009] and Hsu and Wang [2008] CHMMs are used to model two interacting agents performing three different activities. Each agent can be in one of five states: *walk at normal speed*, *walk slowly*, *run (walk fast)*, *stand (chat)*, and *change direction*. Three interaction cases between the agents were devised: *follow*, *approach+talk+continue separately*, and *approach+talk+continue together*. Clearly, this type of scenario requires the communication between the two models to coordinate the actions by the two agents. In Chiang et al. [2010], CHMM

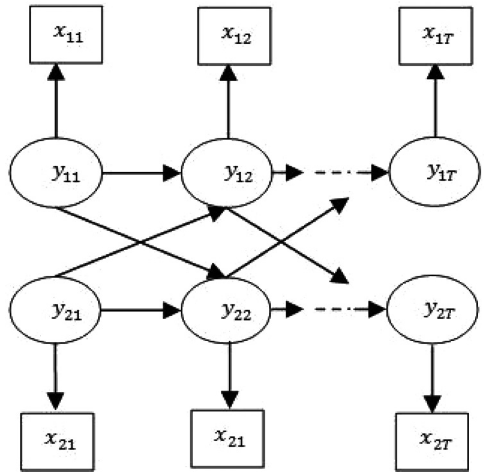

Fig. 9. Two chains HMM fully coupled.

is applied to recognize the multioccupant collaborative activities by two occupants showing that it outperforms the regular HMMs.

The recognition of the most likely activity is computed using the posterior of the activity sequences as follows:

$$\begin{aligned} p(Y_1,\ldots,Y_k|X_1,\ldots,X_k) &= \frac{p(X_1,\ldots,X_k|Y_1,\ldots,Y_k)p(Y_1,\ldots,Y_k)}{p(X_1,\ldots,X_k)} \\ &\propto p(X_1,\ldots,X_k|Y_1,\ldots,Y_k)p(Y_1,\ldots,Y_k) \qquad (7) \\ &= \left(\prod_{m=1}^{k}\prod_{t=1}^{T} p(x_{mt}|y_{mt})\right)\left(\prod_{m=1}^{k} y_{m1}\prod_{t=2}^{T} p(y_{mt}|y_{1t-1},\ldots,y_{kt-1})\right). \end{aligned}$$

5.1.2. Discriminative Models. In contrast to generative models where we attempt to model the joint probability distribution of paired observations and activity sequences $p(Y, X)$, in discriminative models we rather attempt to directly model the conditional probabilities of the activities given the sequence of observations $p(Y|X)$. Moreover, generative models assume that the observations are independent, which is not always satisfied.

In the following, we will present some of the discriminative models used in multioccupant activity recognition. These are undirected graphical models; hence, the same activity modeling like with generative models is adopted as will be explained in the following.

5.1.2.1. Conditional Random Field (CRF). The linear-chain CRF model is one of the most popular discriminative models for dealing with sequential data. It is more flexible compared to HMM, because it does not assume any independence among the observation sequences. Like HMM, CRF is applied to determine the most likely sequence of states given the sequence of observations.

As shown in Figure 10, a linear-chain CRF is an undirected acyclic graph where the hidden sate y_t depends only on the previous state y_{t-1} and the observation x_t depends only on the hidden state y_t. The conditional probability distribution is defined as a

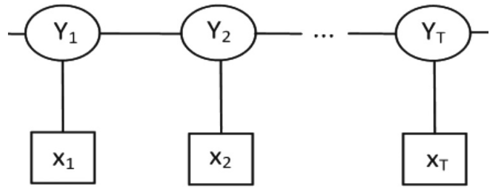

Fig. 10. The linear-chain CRF model.

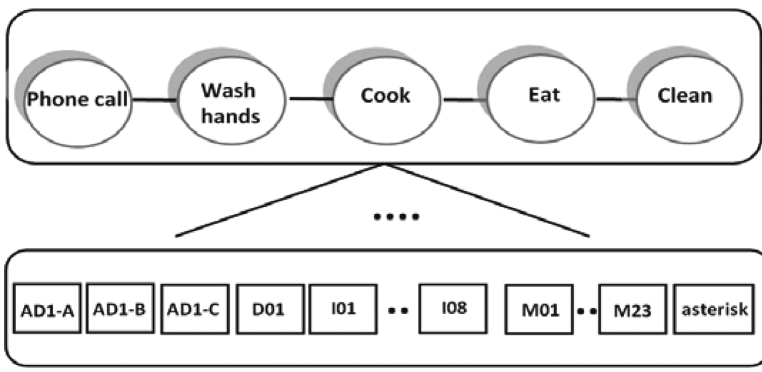

Fig. 11. Representation of a global linear-chain CRF.

multiplication of feature functions exponents:

$$p(Y|X) = \frac{1}{Z(X)} \prod_{t=1}^{T} \exp \sum_{k=1}^{K} \lambda_k f_k(y_t, y_{t-1}, x_t). \tag{8}$$

Here, T is the number of observations and K is the number of feature functions used to approximate the probability distribution, and $\lambda_K(k = 1...K)$ are learning weights associated with the feature functions $f_k(y_t, y_{t-1}, x_t)$ that are estimated by training. The expression $\lambda_K f_K(y_t, y_{t-1}, x_t)$ is known as the energy function, while the exponential of the energy function is known as the potential function [Bishop 2006]. The quantity $Z(X)$ is a normalization term so that the probability distribution adds up to 1 resulting in a proper conditional probability as follows:

$$Z(X) = \sum_{y} \left\{ \prod_{t=1}^{T} \exp \sum_{k=1}^{K} \lambda_K f_K(y_t, y_{t-1}, x_t) \right\}. \tag{9}$$

Figure 11 shows how a CRF is applied for activity recognition. Activities are represented as hidden states and the sensor readings correspond to the observables.

5.1.2.2. Factorial Conditional Random Field (FCRF). A FCRF combines many linear-chain CRFs (called chains) by linking not only the hidden states of each chain to input, but also linking the hidden states of the chains to result in cotemporal connections [Sutton et al. 2007] (see Figure 12). The cotemporal connections allow an efficient representation of the interactions between the chains.

The application of a FCRF for modeling multioccupant activities is straightforward. We can think of using one chain for each occupant, where the hidden states of the

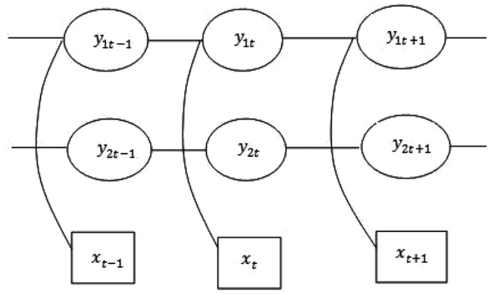

Fig. 12. Representation of a FCRF consisting of two CRF chains.

chains representing the activities are cotemporally connected to model the interaction [Wang et al. 2011]. Figure 12 illustrates an FCRF model as a combination of two chains, each representing an occupant. The two sequences $\{y_{1t-1}, y_{1t}, y_{1t+1}\}$, $\{y_{2t-1}, y_{2t}, y_{2t+1}\}$ are the activities at time $t-1$, t, $t+1$ of resident 1 and resident 2, respectively. The corresponding sequence of observations $\{x_{t-1}, x_t, x_{t+1}\}$ represent sensor readings at the timesteps $t-1$, t, $t+1$.

Following the notations used in Section 5.1.2.1, a FCRF is given by the following posterior probability:

$$p(Y\,|X) = \frac{1}{Z(X)}\left(\prod_{t=1}^{T}\prod_{i,j}^{A}\exp\left(\sum_{k=1}^{K}\lambda_k f_k(y_{it}, y_{jt-1}, X)\right)\right)\left(\prod_{t=1}^{T-1}\prod_{i}^{A}\exp\left(\sum_{k=1}^{K}\lambda_k f_k(y_{it}, y_{it+1}, X)\right)\right)$$
$$\left(\prod_{t=1}^{T-1}\prod_{i}^{A}\exp\left(\sum_{k=1}^{K}\lambda_k f_k(y_{it}, y_{it}, X)\right)\right), \qquad (10)$$

where A is the set of activities and Y and X indicate the set of hidden and observable state sequences. That is, $Y = \{Y_1, Y_2, \ldots, Y_T\}$ where $Y_t = \{y_{1t}, y_{2t}, \ldots, y_{At}\}$ and Y_{it} represents the state of the i^{th} activity at time t. The observable state sequence X is defined in a similar way. $Z(X)$ is a normalization factor obtained over X.

Given an observation sequence X, to find the most likely sequence of activities states, the Maximum-A-Priori (MAP) algorithm is applied once the marginal probability of all node pairs is computed. Actually there exist many inference algorithms like the forward-backward algorithm, loopy belief propagation, mean field free energy, and junction tree [Wang et al. 2011; Sutton et al. 2007].

5.2. Emerging Patterns (EPs)

Association rule mining is about finding interesting relations between features in data. Such relations are rules whose right-hand side and left-hand side are frequent itemsets (i.e., set of features). Itemsets can also be used to distinguish between datasets, and in such case they are called Emerging Patterns (EPs). Thus, EPs can be considered as itemsets with support that changes significantly between datasets.

In the context of activity recognition, EPs are applied to model the activities using the discriminating features. An EP of an activity is the set of features that are the most discriminating for that activity. The set of EPs of an activity form the corresponding *activity model*. The set of features of an activity is selected as an EP if the frequency count of such features changes from that activity's instances to the rest of other activities' instances.

A data instance refers to all observations that are part of an activity during a continuous period of time. The support of an itemset V, in a dataset X, is given as

$$\mathrm{S}_X(V) = \sigma_X(V)/|X|, \tag{11}$$

where σ_X is the number of instances in X that include V. $|X|$ is the total number of instances in X. Using the notion of support, we can compute the growth measure to identify EPs as follows. Given two activities A and B, the Growth Rate (GR) of an itemset V, denoted as $GR(\mathrm{V}, A, B)$, is given as follows:

$$GR(V, A, B) = \begin{cases} 0 & \text{if } S_A(V) = 0 \text{ and } S_A(V) = 0 \\ \infty & \text{if} S_B(V) = 0 \ \text{ and } S_B(V) > 0 \\ \dfrac{S_A(V)}{S_B(V)} & \text{otherwise} \end{cases}. \tag{12}$$

An itemset V is an EP of an activity B if and only if its GR exceeds a given threshold ρ (i.e., $GR(V, A, B) > \rho$); that is, the change from A to B is significant.

EPs are thoroughly discussed in Gu et al. [2009a]. An example presented therein assumes the activity "cleaning a dining table" and the following itemset "object@cleanser, object@plate, object@wash_cloth, and location@kitchen" is an EP. The authors in Gu et al. [2009a] apply an efficient algorithm described in Li et al. [2007] to discover EPs from sequential activity data. Such EPs are used to construct the activity model in a single-resident setting. Using the epSICAR algorithm described in Gu al. [2009a], not only sequential activities but also concurrent and interleaved activities can be identified. Going a step further, the authors apply EPs in a multioccupant setting by mining EPs for each activity and for each resident.

6. FACETS OF MULTIOCCUPANT ACTIVITY RECOGNITION

As shown in Table I, some studies have focused on solving the data association problem [Crandall and Cook 2008a, 2008b, 2010; Hsu et al. 2010; Wilson and Atkeson 2005; Cook et al. 2010; Alerndar et al. 2013; Chen and Tong 2014]. Some other studies have considered that the data association problem had been already resolved and consequently focused on modeling the activities [Hsu et al. 2010; Cook et al. 2010; Chiang et al. 2010; Singla et al. 2010; Wang et al. 2009, 2011; Gu et al. 2009a; Lin and Fu 2007]. As the two issues of data association and activity recognition tend to be treated separately, we will discuss in Section 6.1 the data preprocessing approaches designed for the two problems separately. Section 6.2 presents the studies dealing with data association. For activity recognition, two methodologies will be discussed. According to the first methodology, activity recognition models for residents are independent, ignoring the interactions between the occupants (Section 6.3). In the second methodology, the models take the interaction into account (Section 6.4). Section 6.5 discusses the application of knowledge-based approaches for multioccupant activity recognition. Scalability of all the studies is discussed in Section 6.6. Limitations of the different approaches are summarized in Section 6.7. Evaluation metrics for assessing the performance of algorithms for both problems are discussed in Section 6.8, respectively.

6.1. Preprocessing Methods

All multioccupant approaches presented in this survey are data-driven approaches that rely on data to construct the activity model. As a result, these approaches may be sensible to the representation of the activity data. This later is often incomplete, inconsistent, and prone to errors. Hence, generally it is preprocessed with the intention of making the activity recognition problem easier to solve so that (1) data meets the computational model applied for developing the activity recognition algorithms and (2) efficient use of the raw data is guaranteed through a new representation.

Data preprocessing is an important step in the data mining process. It can consist of several tasks such as cleansing, transformation, normalization, feature extraction, and selection. For instance, feature extraction is used in Crandall and Cook [2008a, 2008b] to generate new features from the raw data. In particular, the date and the time information stamps are used to extract different features like "hour of day," "part of day," "day of week," and "hour of day," which are applied to handle the problem of data association. The studies in Crandall and Cook [2008a, 2008b] discuss the impact of the best temporal feature in capturing the differences in behavior between individuals showing that "hour of day" significantly enhances the classifier performance in resident identification. They also show that depending on the facets of the dataset (e.g., the habits of residents, type of environment, student laboratory or real home), different kinds of features can lead to different classification results. For example, hour-of-the-day is the most discriminating feature, because the dataset was collected from a student laboratory. Furthermore, in comparing the performance of a NBC and a HMM for data association, the studies conclude that feature extraction is not valuable for all types of classifiers and in this case, it is valuable for a NBC but not for a HMM. As a result, in Crandall and Cook [2010], the authors applied a HMM but without feature extraction to deal with data association.

To check the effect of preprocessing, the authors in Hsu et al. [2010] investigate three configurations: raw data, environment data, and room-level data. The raw data is obtained by removing the date and time from the observations and is represented using the sensor ID combined with its reading value. The environment data consist of all data captured in the house. For the *room-level* data, a preprocessing method is applied to represent each room by a feature. However, the environment data does not help in discriminating the residents. This later is "on" if and only if one of the motion sensors in the room is "on." This feature also does not help in discriminating between the residents either. The experiments show that the raw data allows obtaining the best recognition results compared to the other two datasets.

In terms of preprocessing methods for activity recognition models, the authors in Chiang et al. [2010] apply three data preprocessing methods to obtain raw feature, loc-obj feature, and loc-obj with locoff feature vectors. The three types of vector are represented as a tuple (event, interaction), where "event" in raw data is an integer indicating a sensor and its state. "Event" in loc-obj and loc-obj with locoff indicates whether it was captured by object sensors (e.g., item and cabinet sensors) or by location sensors (e.g., motion sensors). "Interaction" is only used in loc-obj with locoff to indicate whether the residents were in the same room or not. The results show that better results are obtained with raw features, which is consistent with the previous work in Hsu et al. [2010]. The low performance in the two preprocessing methods may be attributed to only the issue of representing a location sensor by the corresponding index of the room. The model confuses in the case of many activities sharing the same room. Adding all historical data of triggering of the events would better discriminate between activities as reported by the raw feature vector.

On the other hand, the authors in Chen and Tong [2014] use the same dataset as in Hsu et al. [2010] and Chiang et al. [2010], but preprocess it in a different way. An observation is represented as a binary vector whose length corresponds to the number of sensors. At time t, a position in the vector is set to 1 if the i^{th} sensor changed state. However, such a representation ignores the date and time as features.

The authors in Wang et al. [2009, 2011] and Gu et al. [2009a] investigated a dataset obtained by means of wearable sensors (e.g., three-axis acceleration, audio, location, and tagged objects). New features are extracted from the raw data like the mean, variance, energy, frequency-domain entropy, correlation, location name, and object name.

Likewise, in Lin and Fu [2007] the light and motion data serve to derive new features like bright, dim, dark, no light, triggered, and nontriggered.

However, some authors like in Wang et al. [2009, 2011], Gu et al. [2009a], and Lin and Fu [2007] do not compare their preprocessing methods against raw representation to show the effectiveness.

6.2. Data Association

To avoid the problem of data association in multioccupant activity recognition, some studies relied on wearable sensors [Wang et al. 2009, 2011; Gu et al. 2009a]. Because of the inconvenience of the wearable sensors in some situations, the use of infrastructure sensors has also been investigated. The challenge there is that infrastructure sensors cannot directly identify individuals. In the following, we will discuss the computational approaches used in the context of data association.

The authors in Crandall and Cook [2008a] apply a NBC on raw data for data association, but obtain low performance. In fact, the NBC tends to assign activities to the resident who produced most of the sensor events in the training data, presumably due to imbalance of the training data. But after adding some feature like the temporal feature "hour-of-the-day," the NBC shows better discrimination between the residents. In another study by the same authors [Crandall and Cook 2008b], the HMM is found to outperform the NBC when using only raw data. Feature extraction is valuable for the NBC, but does not affect the HMM. In a third study by these authors [Crandall and Cook 2010], the results show that both algorithms perform well on other real-world datasets B&B and TwoR using the same experimental setting described in Crandall and Cook [2010] with a slightly better performance for HMM.

To investigate the correlation between data association and activity recognition, CRF is applied in Hsu et al. [2010]. As expected, the quality of data association impacts activity recognition if both are integrated in one system. In this study, a two-layer cascade is proposed. Each layer consists of a CRF model. The first layer is designed for data association such that the CRF's hidden states represented the residents, while the observables correspond to the sensor events and activity labels. The second CRF in the second layer is dedicated to activity recognition. Thus, the hidden states in the CRF correspond to the activities, while the observables are sensor reading and resident labels resulting from the previous layer. Likewise, the authors in Cook et al. [2010] construct one HMM to recognize the residents followed by another HMM to recognize the activities. The disadvantage of these cascades is that the recognition accuracy depends on the performance of the data associator.

On the other hand, the authors in Wilson and Atkeson [2005] propose one motion model for each resident using a particle filter based approach in order to identify the optimal assignment of sensors to the residents. They study the impact of varying both the number of residents and the number of particle filters to accurately identify a resident on simulated data. The number of residents varies between one and five and the number of particle filters varies between one and 20 [Wilson and Atkeson 2005]. An insignificant improvement of accuracy is observed after 20 particle filters. Also, more particles are usually required to recognize multiple occupants. Moreover, the accuracy decreases as the number of occupants increases.

6.3. Independent Models for Residents

Many studies address the problem of multioccupant activity recognition but they neither model interaction among residents [Hsu et al. 2010; Wilson and Atkeson 2005; Cook et al. 2010; Singla et al. 2010], nor do they consider real situations where residents perform separate, interleaved, parallel, or cooperative activities [Cook et al. 2010; Singla et al. 2010]. Often interaction is modeled only in a noncomplex setting.

The authors in Hsu et al. [2010], Wilson and Atkeson [2005], Singla et al. [2010], and Lin and Fu [2007] claimed that multioccupant activities can be better recognized if individual models for the residents are learned. In the following, we will discuss the approaches used to create independent models for residents.

Specifically, some studies [Wilson and Atkeson 2005; Singla et al. 2010] show that motion models can be useful for disambiguation of activities, because people usually tend to follow regular habits. One HMM is used to model each resident. Likewise, one CRF per resident is proposed in Hsu et al. [2010]. In this latter study, the accuracy reported is greater than the accuracy in Singla et al. [2010] using the same benchmark. This seems to confirm that CRF performs better than the HMM in handling complex situations. The modeling of activities separately is good when there is less collaboration among the residents. Thus, if the data contains cooperative activities, the accuracy will be low.

In Lin and Fu [2007] separated models for residents are applied using a layered Bayes network-based architecture that models the interaction between the residents. Each layer in the model received the results from the previous layer. In the first layer, the input consists of the sensor readings along with the location data related to each resident. In the second layer, one DBN for each resident is used to model the activities. In the third layer a BN is used to model the interaction between the residents.

The authors in Kasteren et al. [2011] investigate the use of Hierarchical HMM (HHMM) in a single-occupancy setting showing higher accuracy compared to the HMM and the Hidden Semi Markov Model (HSMM). Furthermore, the authors in Nguyen et al. [2006] apply this model in a multioccupant setting by constructing a separate HHMM for each resident and reported a high accuracy of the model when tested on video data.

To the best of our knowledge, hierarchical models have not yet been investigated for multiresident activity recognition in the context of pervasive sensing. It would be interesting to apply HHMMs for multioccupancy to check their ability to infer high level behavior and to deal with parallel and cooperative activities.

6.4. Interaction Modeling

In contrast to the pervasive setting, much work on interaction modeling has been done in computer vision [McCowan et al. 2005; Du et al. 2006, 2007; Natarajan and Nevatia 2007]. In the following, we summarize the approaches discussed in the literature.

Recently, a number of studies on modeling resident's interaction in pervasive environment have been conducted [Chiang et al. 2010; Alerndar et al. 2013; Wang et al. 2009, 2011; Gu et al. 2009a; Lin and Fu 2007; Chen and Tong 2014]. Existing approaches include both supervised [Chiang et al. 2010; Alerndar et al. 2013; Wang et al. 2009, 2011; Lin and Fu 2007; Chen and Tong 2014] and unsupervised approaches [Gu et al. 2009a]. In supervised approaches, we can enumerate the HMM [Alerndar et al. 2013; Chen and Tong 2014], CHMM [Chiang et al. 2010; Wang et al. 2009, 2011], PHMM [Chiang et al. 2010], the BNs [Lin and Fu 2007], and FCRFs [Wang et al. 2011].

The authors in Chiang et al. [2010] investigate close-proximity interaction using an *interaction feature*, that is, a binary feature, which is set to 1 if the two residents are in the same region of the environment and to 0 otherwise. The study shows that the presence of residents in the same room does not imply that the residents are involved simultaneously in cooperative activities. Although the contribution of this interaction feature is not significant, the model is more accurate than without it. Using the dataset of Chiang et al. [2010], the authors in Cook et al. [2010] investigate the detection of close-proximity interaction using a Bayesian approach. It is found that the number of events generated during interaction is more important compared to the number of interactions detected. More interestingly, it is found that the physical proximity does not imply interaction.

In Gu et al. [2009a], the authors apply EPs, which describe important changes from one activity to the other. A confidence measure is proposed to determine if the residents had interacted. EPs are mined for exclusive activities and for cooperative activities. However, EPs tend to recognize the activities as cooperative activities even when they are not.

To study the effect of interaction modeling on the efficiency of multioccupant activity recognition, the authors in Chiang et al. [2010] compare the performance of three models: PHMM, CHMM, and CHMM extended with auxiliary nodes. The results show that the extended CHMM performs the best, while CHMM outperforms PHMM. Considering the same context, the authors in Wang et al. [2011] find that FCRF and CHMM perform similarly but CHMM outperforms FCRF in the case of multioccupant activities. For example, the accuracy of the cooperative activity *watching TV is* higher with CHMM (100%) than with FCRF (70.5%). Finally, as the models in Wang et al. [2009, 2011] and Gu et al. [2009a] use the same dataset, comparing their recognition performance shows that the EPs approach performs the best for activity recognition in terms of accuracy, scalability, and robustness.

In Natarajan and Nevatia [2007] Coupled HSMM (CHSMM) and CHMM are used to model multiresident activities using a dataset related to simultaneous hand gesture obtained by camera in the context of sign language. CHSMM outperforms CHMM by a difference of 20%–30% accuracy rate. Considering multioccupancy, transfer learning can be applied by substituting the two hands by two residents and test the ability of the model to deal with parallel and cooperative activities in a pervasive setting.

6.5. Knowledge-Driven vs Data-Driven Approaches

We point out that all work previously presented is data driven using mainly probabilistic algorithms to build activity models. Knowledge-driven approaches, on the other hand, use ontology and symbolic representation (i.e., logic) to specify the semantic relations of activities as in the ontology *snapshot* approach used in Riboni et al. [2011]. Although data-driven techniques exploit temporal information, which is a very important aspect in activity recognition, the knowledge-driven techniques have been proven to be effective in single-resident activity recognition [Riboni et al. 2011]. When extended with simple forms of temporal reasoning, knowledge-based methods are comparable to the state-of-the-art techniques based on HMMs [Riboni et al. 2011].

Interestingly enough and to the best of our knowledge, ontology modeling has not yet been fully investigated in the context of multioccupant activity recognition. An exception to this is the work described in Lin and Fu [2007] (already mentioned in Section 6.2) where a combination of data-driven and knowledge-driven methods is proposed resulting in a layered approach. In the first layer, ontology is used to interpret raw data from sensors by exploiting knowledge about the residents and their relationships. In the second layer a DBN is applied to learn single-occupant preferences and in the third layer a BN is used to learn multioccupant preferences.

This study, however, focuses on learning user preferences, not on recognizing the activities. Its merit lies in the fact that it provides a unified framework for recognizing both individual preferences and cooperative preferences. Considering multioccupancy, such a layered model can be applied by substituting the preferences by the activities and test the ability of the model to deal with parallel and cooperative activities in the pervasive setting.

Another interesting study was presented in Alerndar et al. [2013] and Chen and Tong [2014] where knowledge-driven and data-driven approaches for multiresident activity recognition are combined. In fact, the authors of that study exploited some simple knowledge of multiresident activities by defining "*combined labels*." Specifically, each observation in the dataset is represented by a label pair (activity label of resident 1, activity label of resident 2). The pair is then converted into a scalar to result in a combined label that represents the two activities of the two residents. Using the dataset, all possible combinations are collected. After mapping each pair of multiresident activities labels in the training dataset to their combined label, an HMM is applied to construct the activity model.

In Chen and Tong [2014], HMM and CRF were applied to construct the activity model. In both HMM and CRF models, hidden states represent the combined labels and the observations represent the sensor readings. The authors apply a two-stage method in the inference step. In the first stage of the method, CRF and HMM are applied to infer the combined label state. In the second stage, the combined label states are inversely mapped onto the corresponding residents' activity labels. The results show that this approach increases the average accuracy by approximately 10% in comparison with the approaches described in Hsu et al. [2010] and Singla et al. [2010] using the same dataset. The results are slightly better for HMM in comparison with CRF. To the best of our knowledge, the approach used by the authors in Alerndar et al. [2013] and Chen and Tong [2014] is the only one that allowed one to solve both data association and activity recognition at the same time.

6.6. Applicability, Adaptability, and Scalability of Multioccupancy Models

The applicability and the adaptability of all models used in the context of multioccupancy have not yet been investigated. Existing multiresident activity recognition systems are trained on private datasets [Wilson and Atkeson 2005; Wang et al. 2009, 2011; Gu et al. 2009a; Lin and Fu 2007] or on publicly available datasets [Prossegger and Bouchachia 2014; Crandall and Cook 2008a, 2008b; Hsu et al. 2010; Chiang et al. 2010; Cook et al. 2010; Singla et al. 2010; Alerndar et al. 2013; Chen and Tong 2014]. Thus, the models are closely adjusted to the living space and the person, to the training data, and to the types of activities monitored in the home. Then, such recognition models would only be applicable to that environment be it a mono-occupant setting or multioccupant setting.

To overcome the preceding limitations, the authors in Sarkar et al. [2010] suggest the use of an alternative source of activity data that is a web data. Although their approach would work for almost any environment, the web data is clean and therefore cannot be used for real-world systems. The proposed activity model is developed for a single-resident setting and thus the authors only discuss its scalability in terms of adding new activities. However, dealing with scalability in a multiresidential setting should not only consider new activities, but also new occupants. Clearly, the scalability of the models in terms of the number of residents is the most important issue.

Referring to the scalability of data association algorithms, all studies in the literature have considered only a two-resident situation [Crandall and Cook 2008a, 2008b, 2010; Hsu et al. 2010; Cook et al. 2010], except Wilson and Atkeson [2005]. In fact, the authors in Wilson and Atkeson [2005] study the impact of varying the number of residents from one to five on simulated data and from one to three on real-world data using HMM. It is found that there is no difference in accuracy when varying the number of residents on a real-world dataset in comparison with simulated data. This is a good sign that the model can be applied in real-world environments. However, the accuracy drops, because the complexity, which depends on the number of occupants, increases [Wilson and Atkeson 2005]. On simulated data, an accuracy of 100% is obtained for one occupant and only 67% for four occupants.

To the best of our knowledge, the scalability of activity recognition algorithms has never been considered. Nevertheless, as the authors in Gu et al. [2009a] point out, EP-based models are scalable to additional occupants. Adding a new occupant would only imply mining the set of EPs for each activity monitored in the environment for this occupant. The scalability of the model presented in Alerndar et al. [2013] and Chen and Tong [2014] seems to be feasible too. In fact, the authors use a combined label to represent the two activities of two residents at the same time. Adding another resident implies combining three labels and hence, implies only increasing the number of label combinations. The authors in Chiang et al. [2010] noted that the scalability of the activity model would be more difficult to achieve and that training and inference will be computationally highly demanding, especially for CHMM and FCRF. In a nutshell, the scalability problem is a challenging research avenue that is about the general issue of learning more generalized multioccupant activity models.

6.7. Limitations of Multioccupant Activity Recognition Systems

As we mentioned earlier, a number of studies on multioccupant activity recognition have been carried out using the pervasive computing technology. Some of them investigate activity recognition ignoring data association. In this context, the studies in the second approach of Hsu et al. [2010] and the second approach of Singla et al. [2010] applied an individual model for each resident (see Section 6.3 for additional details). The advantage of this approach lies in the fact that it is easily scalable to new residents. It only requires learning a new chain for the new resident. Furthermore, this approach is suitable in case residents follow their regular routines and do not much interact with each other. Thus, in the case of more interactions taking place between residents, this approach may not be suitable.

CHMM and FCRF (see Section 5.1.1.6 and Section 5.1.2.2, respectively) used in Chiang et al. [2010] and Wang et al. [2009, 2011], respectively, offer the advantage of modeling both parallel and cooperative activities. In contrast to CHMM, FCRF does not require using a data association variable in order to construct the activity model, since all residents' activities at time t depend on all occupants' data. Another, major limitation of these models lies in their scalability to new residents in the setting. In contrast to the independent models for residents, adding a new resident in the environment implies relearning the model again on all occupants' sensory data, that is, both old resident's sensory data and new resident's sensory data.

Studies in the first approach of Hsu et al. [2010] and Cook et al. [2010], Alerndar et al. [2013], and Chen and Tong [2014] present the advantage of solving both data association and multiresident activity recognition. However, the methodology differs from Hsu et al. [2010] and Cook et al. [2010] (see Section 6.2) to Alerndar et al. [2013] and Chen and Tong [2014] (see Section 6.5). The solution suggested in Cook et al. [2010] requires solving the data association problem before activity recognition. The disadvantage of this approach is that the misclassification of the resident by the data associator strongly impacts the recognition of the activity. Moreover, adding a new resident in the environment implies retraining both the data association model (adding a hidden state representing the new resident) and the activity recognition model (adding the hidden states corresponding to the activities of the new resident).

Furthermore, the activity recognizer in the first approach of Hsu et al. [2010] nd Cook et al. [2010] and the first approach of Singla et al. [2010] consists of a single-chain CRF and a single-chain HMM, respectively. In the inference step, one activity label is inferred representing either the activity label of resident 1 or resident 2. Tracking the activity of each resident requires recognizing the activities of all residents at each timestep. Using a single-chain HMM or CRF in which the hidden states represent the activities of all residents is not suitable for activity recognition.

A similar approach presented in Alerndar et al. [2013] and Chen and Tong [2014] (see Section 6.5) also use a single-chain HMM and a single-chain CRF, but the hidden states refer to combinations of activity labels that are obtained by aggregating pairs of activities from the occupants. In the inference step, the combined label is converted back into the individual activity labels. By doing so, conventional graphical models, like HMM and CRF, can be applied to multioccupancy. This approach presents the advantage of solving both data association and recognizing both parallel and cooperative activities simultaneously. Moreover, although the concept is applicable regardless of the number of occupants, the process needs to be repeated again on all occupants' data if new occupants are added.

Comparing HMM and CRF, the study in Kasteren et al. [2008] reports that HMM is more appropriate than CRF for imbalanced activity data that contains dominant activities. In the "Ubicomp dataset" described in Kasteren et al. [2008], we can encounter more events related to the activity "going to bed" than those related to the activity "toileting."

6.8. Evaluation Issues

To evaluate the performance of computational models for both data association and activity recognition models, the evaluation method should describe how the data is to be used, how training is to be carried out, and how validation and testing are to be conducted. The performance metrics, used for evaluating the model, are very important in the validation of any model. Selecting the adequate metrics strongly depends on the specific problem (e.g., classification, regression) at hand. In the following, we will display the different criteria used in the literature references mentioned in this survey to assess the performance of multioccupant activity recognizers.

$$Accuracy = \frac{1}{E}\sum_{i=1}^{E}[\inf erred(i) = true(i)], \tag{13}$$

$$FalsePositiveRate = \sum_{i=1}^{l} fp_i, \tag{14}$$

$$ErrorRate = \sum_{i=1}^{l} \frac{fn_i + fp_i}{tp_i + fn_i + fp_i + tn_i}, \tag{15}$$

$$TimeSliceAccuracy = \frac{1}{N}\sum_{i=1}^{N}[\inf erred(i) = true(i)], \tag{16}$$

$$Accuracy_i = \frac{tp_i + tn_i}{tp_i + fn_i + fp_i + tn_i}, i = 1 \ldots l, \tag{17}$$

$$AverageAccuracy = \frac{1}{l}\sum_{i=1}^{l} Accuracy_i, \tag{18}$$

$$Average\,\mathrm{Pr}ecision = \frac{1}{l}\sum_{i=1}^{l}\frac{tp_i}{tp_i + fp_i}, \tag{19}$$

$$Average\mathrm{Rec}all = \frac{1}{l}\sum_{i=1}^{l}\frac{tp_i}{tp_i + fn_i}, \tag{20}$$

$$AverageFscore = \frac{\Pr ecision \times \text{Rec}all}{\Pr ecision + \text{Rec}all}, \tag{21}$$

$$AverageErrorRate = \frac{1}{l}\sum_{i=1}^{l}\frac{fn_i + fp_i}{tp_i + fn_i + fp_i + tn_i}, \tag{22}$$

where l is the total number of classes/activities and N is the total number of time slices when sensory data is discretized using a constant length. E is the total number of sensor events. $[a = b]$ is a binary indicator having the value 1 if true and 0 otherwise. tp_i is the number of true positives (instances from the i^{th} class that are correctly recognized as being from the i^{th} class), tn_i is the number of instances recognized as not part of class i and indeed they are not (true negatives), fp_i is the number of instances that are incorrectly recognized as part of the i^{th} class (false positives), and fn_i is the number of instances recognized as part of the i^{th} class, while they are not (false negatives). The quantities tp_i, tn_i, fp_i, and fn_i for each class are computed from the confusion matrix.

In the context of data association, classes are the residents, whereas in activity recognition classes represent the activities. Authors in Crandall and Cook [2008a, 2008b] study the impact of adding new features (hour of day, day of week, part of day, part of week) to the data and used NBC and HMM in order to deal with data association. Two measures are used: the *accuracy rate* (Equation (14)) and the *false positive rate* (Equation (15)). They compute each of the two metrics for each feature type to select the feature that reports the best results and to evaluate the effect of features on the efficiency of resident identification. For instance, in a two-occupant home, a person would spend much more time at home than the other one and, hence, the probability that an event would be generated by a person will be attributed to the person that caused most of the events resulting in a high false positive rate. A good resident classification would result in high accuracy and a low false positive rate.

When comparing NBC and HMM for resident identification, the authors in Crandall and Cook [2010] use the *average lag* to assess the performance of their HMM model. The average lag is defined as the average number of events after a transition in sensor data before HMM correctly classifies the resident. An average lag of 1 indicates that HMM improperly classified one event in each transition from one resident to the other one in activity data before correctly recognizing the resident causing the events. The authors also use the *error rate* (Equation (16)), which represents the ratio of errors made when classifying a number of instances. The authors in Wilson and Atkeson [2005] use *time-slice accuracy* (Equation (17)) to evaluate the effectiveness of their HMM based resident identification problem. Time-slice accuracy is the ratio of correctly classified time slices when data is discretized using a time length. On the other hand, for evaluating activity recognition, the authors in Chiang et al. [2010] compute the accuracy for each resident separately. They also compute the joint accuracy, which is counted when the activity recognized for both resident 1 and resident 2 is correct.

Existing datasets in activity recognition such as Activity Recognition with Ambient Sensing (ARAS) [Alerndar et al. 2013] and CASAS "Multiresident ADLs" [Singla et al. 2010] are imbalanced, which means that some classes have more instances in the dataset than do other classes. Hence, because of the class imbalance the correct classification of each class is equally important for activity recognition; many studies in the field tend to apply the *average accuracy* measure (Equation (19)). As a result, the authors in Hsu et al. [2010], Cook et al. [2010], Singla et al. [2010], Alerndar et al. [2013], and Chen and Tong [2014] use the *average accuracy* to assess the performance of their activity models. In Cook et al. [2010], and Chen and Tong [2014] many measures are applied: the average accuracy, the *average precision* (Equation (20)), the *average recall* (Equation (21)), and the *average f-score* (Equation (22)). In addition to these metrics [Chen and Tong 2014], the authors also use the *average error rate* (Equation (23)).

On the other hand, the authors in Wang et al. [2009, 2011], Gu et al. [2009a], and Lin and Fu [2007] discretize the activity data before modeling. The length of time slice is set to 1s in Wang et al. [2009, 2011] and to 15s in Gu et al. [2009a]. They use *time-slice accuracy* (Equation (16)) as an evaluation metric for the activity model. Furthermore, the authors in Wang et al. [2009, 2011] and Gu et al. [2009a] apply single-user ADLs time-slice accuracy and multioccupant ADLs time-slice accuracy. The reason these authors separate single-user results from multioccupant results is to show the ability of each classifier to recognize each of the two types of activities. As portrayed in the comparison study between CHMM and FCRF in Wang et al. [2011], CHMM is better than FCRF in recognizing multioccupant activities.

From all evaluation metrics used by research studies mentioned in this survey, we clearly distinguish two classes of measures: the standard *accuracy* (Equation (13)) [Crandall and Cook 2008a, 2008b, 2010; Chiang et al. 2010] or a variant of its named *time-slice accuracy* (Equation (16)) [Wilson and Atkeson 2005; Wang et al. 2009, 2011; Gu et al. 2009a; Lin and Fu 2007] and the *average accuracy* (Equation (18)) [Hsu et al. 2010; Singla et al. 2010; Alerndar et al. 2013; Chen and Tong 2014].

7. DATASETS FOR MULTIPLE-RESIDENT ACTIVITY RECOGNITION

Usually human activity recognition systems are developed and evaluated using datasets. Publicly available datasets are important for the research community to create standardized test beds that could be used for evaluating the performance of activity recognition algorithms and for comparison purposes. Among the benchmark datasets that are freely available, there exist many single-resident ones used in Riboni et al. [2011], Kasteren et al. [2008], Sarkar et al. [2010], and Kasteren et al. [2010, 2011].

However, there is a real need for datasets collected from houses with multiple occupants. The CASAS (Center for Advanced Studies in Adaptive Systems) group has collected several multioccupant activity datasets: “twor.2009,”[4] “twor.summer.2009,”[5] “twor.2010,”[6] “Tulum,”[7] “tulum2,”[8] “cairo,”[9] and “Multiresident ADLs.”[10] Likewise, the ARAS[11] group has collected a multioccupant dataset, named ARAS, which includes two datasets, House A and House B. To the best of our knowledge, these datasets are the only ones publicly available recorded from multiple residents using pervasive sensors.

Table II summarizes the characteristics of the multioccupant datasets, which will be described further in the following sections.

7.1. CASAS Multioccupant Datasets

Multioccupant datasets of CASAS were collected in the WSU smart apartment test bed. Multioccupant activities were obtained using clinical questionnaires [Reisberg et al. 2001]. Activities were annotated by recording the start and end times of the activities via a handwritten diary. We can distinguish two types of datasets: (i) *unscripted*

[4]http://ailab.wsu.edu/casas/datasets/twor.2009.zip.
[5]http://ailab.wsu.edu/casas/datasets/twor.summer.2009.zip.
[6]http://ailab.wsu.edu/casas/datasets/twor.2010.zip.
[7]http://ailab.wsu.edu/casas/datasets/tulum.zip.
[8]http://ailab.wsu.edu/casas/datasets/tulum2.zip.
[9]http://ailab.wsu.edu/casas/datasets/cairo.zip.
[10]http://ailab.wsu.edu/casas/datasetdlmr.zip.
[11]http://www.cmpe.boun.edu.tr/aras/.

Table II. Characteristics of ARAS and CASAS Multi-Occupant Datasets

Dataset	# of residents	Duration	# of Sensors	# of ADL	# of sensor events	Environment	Scripted	Annotation medium
House A of ARAS	1 pair	1 month (continuous)	20	27	2 592 000	Real house	No	GUI
House B of ARAS	1 pair	1 month (continuous)	30	27	2 592 000	Real house	No	GUI
"MultiResident ADLs"	26 pairs	Spread over 2 months	37	15	17 258	Lab.	Yes	diaries
"twor. 2009"	1 pair	Continuous period of 2 months	71	9	137 789	Lab.	No	diaries
"twor.summer. 2009"	1 pair	Continuous period of 2 months	86	8	772 544	Lab.	No	diaries
"twor. 2010"	1 pair	2009–2010 academic year	87	13	2 804 813	Lab.	No	diaries
"tulum"	1 pair	4 months (Several days are missing)	20	9	486 912	Lab.	No	diaries
"tulum2"	1 pair	2009–2010 academic year	36	15	1 085 902	Lab.	No	diaries
"cairo"	1 pair +1 pet	Continuous period of 2 months	32	11	726 534	Lab.	No	diaries

Table III. Activities of the Unscripted CASAS Multioccupant Datasets

"twor. 2009"	"twor.summer. 2009"	"twor. 2010"	"tulum"	"tulum2"	"cairo"
-Clean -Meal preparation -Bed to toilet -Personal hygiene -Sleep -Work -Study -Wash bathtub -Watch TV	-Bed to toilet -Cleaning -Cooking -Grooming -Shower -Sleep -Wake up -Work	-Bathing -Bed to toilet -Eating -Enter home -Housekeeping -Leave home -Meal preparation -Personal hygiene -Sleep -Not sleeping in bed -Wandering in room -Watch TV -Work	-Cook breakfast -Cook lunch -Enter home -Group meeting -Leave home -Eat breakfast -Snack -Wash dishes -Watch TV	-Bathing -Bed to toilet -Eating -Enter home -Leave home -Meal preparation -Personal hygiene -Sleeping in bed -Wash dishes -Watch TV -Work bedroom 1 -Work bedroom 2 -Work living room -Work table -Yoga	-Bed to toilet -Breakfast -Sleep -Wake -Work in office -Dinner -Laundry -Leave home -Lunch -Night wandering -Take medicine

activity datasets like "twor.2009," "twor.summer.2009," "twor.2010," "tulum," "tulum2," and "cairo" and (ii) *scripted* activity dataset like "Multiresident ADLs." Activities considered in the unscripted multioccupant datasets and the scripted multioccupant one are listed in Tables III and Table IV, respectively.

The WSU smart apartment test bed is equipped with many types of sensors: motion sensors, door sensors, temperature sensors, light switch sensors, water flow sensors, burner sensor, phone sensor, and item sensors. A summary of the type and the number of sensors used for recording activity is shown in Table V.

In the following, we give some details of the scripted and unscripted datasets.

Table IV. Activities Covered by the Scripted Multioccupant Dataset of CASAS

Exclusive	*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	

Table V. Sensors Used in CASAS Multioccupant Datasets

	"Multiresident ADLs"	"twor. 2009"	"twor. summer. 2009"	"twor. 2010"	"tulum"	"tulum2"	"cairo"
Motion sensors	27	51	51	51	18	31	27
Door sensors	8	9	15	15			
Light sensors		7	10	11			
Item sensors	2	1	4	4			
Temperature sensors			5	5	2	5	5
Electricity sensors			1	1			
Water flow sensors		2					
Burner sensors		1					

7.1.1. Unscripted Multioccupant Datasets. To the best of our knowledge, "twor.2009," "twor.summer.2009," "twor.2010," "tulum," "tulum2," and "cairo" have not yet been used in multioccupant activity recognition research. Each of these datasets was collected through a pair of residents who performed unscripted activities. Specifically, "tulum" and "tulum2" represent activity data of a married couple, whereas the "cairo" dataset consists of three types of data: the activity data of a volunteer adult couple, the motion data related to their dog, and the data related to their children who come sometimes to visit. All these datasets account for intrasubject variability.

Although these datasets stemmed from a laboratory on a voluntary basis, they were recorded continuously in time. The recording time for "twor.2009," "twor.summer.2009," and "cairo" was approximately 2 months, for "tulum" 4 months, and for both "twor.2010" and "tulum2" approximately 1 year. Multioccupant activities in these datasets are described by records of the form (Date, Time, SensorID, Value). Each activity is delimited by specific markers: (ResidentID_ActivityName Begin) and (ResidentID_ActivityName End).

7.1.2. Scripted Multioccupant Dataset. "Multiresident ADLs" collection has been used in many studies [Hsu et al. 2010; Chiang et al. 2010; Cook et al. 2010; Singla et al. 2010; Chen and Tong 2014]. It was generated in a laboratory setting and therefore it does not fully reflect on real-world scenarios. "Multiresident ADLs" was collected through 26 pairs of volunteers who performed scripted activities. Such activities are predetermined and were repeatedly performed. This collection accounts for intersubject variability, yet it is not sufficient for explaining real-world situations. This dataset was not recorded continuously in the time, and instead it was spread over 2 months. Multiresident activities in "Multiresident ADLs" come in the format of (Date, Time,

Table VI. Activities Simulated by ARAS

House A and House B	
-Other	-Having shower
-Going out	-Toileting
-Preparing breakfast	-Napping
-Having breakfast	-Using Internet
-Preparing lunch	-Reading book
-Having lunch	-Shaving
-Preparing dinner	-Brushing teeth
-Having dinner	-Talking on the phone
-Washing dishes	-Listening to music
-Having snack	-Cleaning
-Sleeping	-Having conversation
-Watching TV	-Having guest
-Studying	-Changing clothes
	-Laundry

Table VII. Sensor Infrastructure Used by ARAS

	House A		*House B*
1	Wardrobe photocell	2	Kitchen cupboards CSs
1	Convertible couch photocell (Resident 2's bed)	1	House DCS
		2	Wardrobe DCSs
1	TV infrared receiver	1	Shower cabinet DCS
2	Couch force sensors	1	Tap distance sensor
2	Chair proximity sensors	3	Chair force sensors
1	Fridge photocell	11	Fridge photocell
1	Kitchen drawer photocell	2	Kitchen drawer photocell
1	Wardrobe photocell	2	Couch pressure mat
1	Bathroom cabinet photocell	1	Bed pressure mat
1	House DCS	1	Armchair pressure mat
1	Bathroom DCS	1	Bathroom door sonar distance
1	Shower cabinet DCS	1	Kitchen sonar distance
1	Hall sonar distance	1	Closet sonar distance
1	Kitchen sonar distance		
1	Tap proximity sensor		
1	Water closet proximity sensor		
1	Kitchen temperature sensor		
1	Bed force sensor		

SensorID, Value, ResidentID, TaskID). A full description of this dataset can be found in Singla et al. [2010].

7.2. ARAS Collection

ARAS data was collected from two pairs of residents performing a large variety of activities [Alerndar et al. 2013]. The first pair consists of two males, while the second is a couple. This collection of two-home dataset offers a better opportunity to study and compare activity recognition algorithms more realistically. ARAS data accounts for intrasubject variability and do not account for the intersubject one. It reflects on the natural behavior of the residents during 2 months. An important feature of ARAS data is that it contains a large variety of human activities and a large number of activity occurrences. Activities and sensors considered by ARAS are presented in Table VI and Table VII, respectively. Annotation of the activities was achieved by the residents themselves using a simple Graphical User Interface (GUI). Several instances of GUI applications were placed in the most convenient places in the houses. This way of doing annotations is more accurate than using a diary.

ARAS collection offers the advantage of being ready to use. Each day of recordings consists of a $22 \times 86{,}400$ matrix that is stored in a file. In all, the collection consists of

30 files for each dataset. The first 20 columns are the sensor binary values, fired/not fired; columns 21 and 22 contain the activity labels for resident 1 and resident 2, respectively. Using a constant time interval to discretize the data allows representing the sensor readings in time slices. This representation leads to a better discrimination between activities as shown in Kasteren et al. [2008, 2010, 2011] on other datasets. For instance, in Kasteren et al. [2010] two experiments were run. In the first experiment different lengths of the time slice were tested including (1s, 10s, 30s, 60s, 300s, and 600s) finding that short time slices produced better recognition results. 30s and 60s produced the best results. In the second experiment, NB, HMM, HSMM, and CRF models were compared using a number of feature representations that consisted of raw, change point, and last representation. The change point representation produced the best results. It was found that the recognition performance of the activity model is strongly influenced by the time-slice length and the feature representation.

8. INTERNATIONAL RESEARCH GROUPS

Several research groups have equipped experimental living spaces with pervasive sensors for human activity recognition research like GeorgiaTech Aware Home Research Initiative (AHRI),[12] Intel research laboratory in Seattle,[13] DOMUS (Domotics and Mobile computing Research)[14] at Sherbrooke University (Canada), and the Place Lab at Massachusetts Institute of Technology (MIT).[15]

However, only few research groups have been working on multioccupant activity recognition. The NTUWisdom Family (Attentive Home)[16] targets the family environment as shown in recent work [Hsu et al. 2010; Chiang et al. 2010]; whereas others [Lin and Fu 2007] looked at the problem of multiuser preference modeling. Members of the Institute of Computer Software (ICS) at Nanjing University worked on multiresident activity recognition from wearable sensors [Wang et al. 2009, 2011; Gu et al. 2009a]. Although the ARAS group project published a multioccupant dataset (see Section 7.2), which includes a variety of pervasive sensors as well as a variety of activities, they did not work on multioccupant activity recognition except in Alerndar et al. [2013]. The CASAS group seems to be a driving force in the area of human activity recognition, be it for a single-occupant or a multioccupant setting. One of their major contributions in this area is making around 24 datasets publicly available.

9. OPEN RESEARCH QUESTIONS

In order to bring the multioccupant activity recognition systems to a more mature stage, some research avenues require further investigation. Next, a list of open research questions is discussed.

9.1. Complex Activity Recognition in Multiresident Setting

In real-world situations, human activities are often carried out in a complex way. The existing research literature dealing with multioccupancy has not fully addressed the problem of cooperative activities in a way to cope with different situations like these:

1. Interleaved or concurrent activities performed in parallel by multiple residents: Each resident performs his/her activities in a concurrent or an interleaved manner

[12]http://www.awarehome.gatech.edu/drupal/.
[13]http://www.intel.com/research/network/seattle_human_activity_recognition.htm.
[14]http://www.domus.usherbrooke.ca/.
[15]http://web.mit.edu/cron/group/house_n/placelab.html.
[16]http://www.attentivehome.org/index.html.

and at the same time another resident performs his/her activities in a concurrent or an interleaved manner.

2. There exist more complex situations in which a resident switches between an activity and a collaborative activity or performs both in a concurrent manner.
3. Ambiguity of interpretation: The interpretation of similar activities may differ depending on the context; for example, an activity *turn the water tap* can be part of many activities like *cooking* and *drinking* and the model should be able to handle these situations, which appear in both single and multioccupant settings.

9.2. Scalability of the Activity Model

All the studies discussed in this survey use datasets that are related to only two occupants and do not investigate the scalability of the models proposed therein. Evaluating such models with more than two residents is an important aspect for real-world situations.

In Section 6.6, we discussed the scalability of data association models and activity recognition models separately, as researchers working on multioccupant activity recognition tend to focus on one of the two latter problems. First, we pointed out that dealing with scalability in a multiresidential setting should not only consider new activities, but also new occupants. Second, the scalability of the models in terms of the number of residents is the most important issue. Although all studies presented in multioccupant activity recognition have only considered a two-resident situation [Chiang et al. 2010; Cook et al. 2010; Singla et al. 2010; Alerndar et al. 2013; Wang et al. 2009, 2011; Gu et al. 2009a; Lin and Fu 2007; Chen and Tong 2014], some of them [Alerndar et al. 2013; Gu et al. 2009a; Chen and Tong 2014] would be easily scalable to additional occupants than others [Chiang et al. 2010; Cook et al. 2010; Singla et al. 2010; Wang et al. 2009, 2011].

9.3. Resolution of Conflicts

People using the same resources may have different preferences. For instance, one prefers watching TV while the light is on, and another prefers it to be off. This is valid for most of the shared but parallel activities.

In Davidoff et al. [2006], a fieldwork with 12 families is presented. The study reports on six social characteristics of home life including *multiple users' conflicts at home*. The study points out that an understanding of these characteristics should be more tightly coupled with what services should ultimately be developed for, and how these services should be implemented. It also concludes that smart home systems need to participate in value decisions and in negotiating a group goal setting.

In this context, the authors in Hsu and Wang [2008] propose a resource management system for a multioccupant smart home. The system relies on the strategy of agent conceding negotiation to manage the smart home resources. The system consists of three components named as home ontology, device controller, and resource allocator. *The home ontology* describes the spatial organization of the smart home as well as information on the devices equipping the living space. *The device controller* is responsible for collecting information about the residents. It applies case-based reasoning to predict the resources a resident may need. The controller finds matching cases in the case base to determine the potential resource conflicts. *The resource allocator* relies on Belief Desire Intention (BDI) agents, a communication blackboard, and conceding negotiation mechanisms to manage conflicts over resources. To implement the smart home system, a BDI agent is assigned to each resident. The blackboard enables the BDI agents to communicate and facilitates the management of the resource conflicts. In terms of conceding negotiation, each agent is assigned a computed conceding risk and in case there is conflict the one with the lowest risk is chosen as conceder of the resource that looks for another resource. The negotiation cycle continues until a common resource use plan is obtained.

9.4. Pervasive Multioccupant Activity Datasets

The quantitative comparison of multioccupant activity recognition methods is not straightforward because studies use different datasets. The lack of standard benchmarks makes it difficult to exhaustively and fairly evaluate multioccupant activity recognition models. The studies presented in Wang et al. [2009, 2011] and Gu et al. [2009a] relied on a private dataset as can be seen in Table I. Studies described in Hsu et al. [2010], Chiang et al. [2010], Cook et al. [2010], Singla et al. [2010], and Chen and Tong [2014] and in Prossegger and Bouchachia [2014] and Alerndar et al. [2013] used CASAS "Multiresident ADLs" and ARAS datasets, respectively. Selecting the best approach from all these studies is not possible as they do not rely on the same activity data.

As the process of collecting activity data requires financial resources that are not within the reach of all research laboratories, researchers tend to use publicly available datasets. Hence, the motivation for presenting eight datasets publicly available may serve as benchmarks for future research on multioccupancy (see Section 7). It must be emphasized that there is a real lack of pervasive multioccupant activity recognition datasets; in particular, datasets that include activity data of more than two individuals and covering the various types of cooperative activities. This will allow researchers to experiment the scalability of their activity models proposed.

9.5. Online Learning and Inference for Real-Time Multioccupancy

In comparison to offline activity recognition, online activity recognition has not been much investigated. Indeed, most of the work presented in this survey is based on offline supervised learning. A few works based on online learning have been, however, presented in Kasteren et al. [2008] and Bouchachia and Vanaret [2014]. Online learning and online inference are required for some situations to adapt the models incrementally as new data becomes available or to make decisions in pseudo real time, respectively.

Often a monitoring system needs to make inference instantly in situations that may render an elderly person at risk; for example, forgetting to take medication. In these situations we need to detect these unusual events at time in order to intervene. The latter problem is of more importance in a multioccupant setting as it may put not only the person who caused the unusual events at risk and vulnerable but also all the residents at home; for example, forgetting the stove is on. Online inference is also important in situations in which the smart home system would have temporary occupants, such as guests. In such cases, the system needs to recognize the new guests and to distinguish between them and the residents. Because of the relevance of online learning and online inference in this context of activity monitoring, it is important that more effort should be devoted to it.

10. CONCLUSION

So far, research related to multioccupant smart homes has devoted significant attention to the application of graphical probabilistic algorithms to model and recognize activities. This survey emphasizes the importance of the various technology aspects to fully realize the multioccupancy paradigm. While there has been much effort invested on the single-occupancy paradigm, multioccupancy has recently started to be the central focus of many studies. Clearly, there were and are still many outstanding scientific questions related to single occupancy to be dealt with before dealing with those specific to multioccupancy.

In this survey article, we pointed out the major issues pertaining to activity recognition in the context of multioccupancy taking data association and interaction into account. We discussed in detail the state-of-art computational models used for modeling collaborative activities, the existing benchmark datasets, and the evaluation metrics. Toward the end of this contribution, we highlighted some of the open questions for future investigation by the research community.

REFERENCES

Giovanni Acampora, Diane J. Cook, Parisa Rashidi, and Athanasios V. Vasilakos. 2013. A survey on ambient intelligence in healthcare. *Proc. IEEE* 101, 12 (Dec. 2013), 2470–2494. DOI:http://dx.doi.org/10.1109/JPROC.2013.2262913

Hande Alerndar, Halil Ertan, Ozlern D. Incel, and Cem Ersoy. 2013. ARAS human activity datasets in multiple homes with multiple residents. In *Proceedings of the 7th International Conference on Pervasive Computing Technologies for Healthcare (PervasiveHealth'13)*. 232–235. DOI:http://dx.doi.org/10.4108/icst.pervasivehealth.2013.252120

Anon. 2014. Centers for Disease Control and Prevention. Retrieved October 13, 2014 from http://www.cdc.gov/.

Christopher M. Bishop. 2006. *Pattern Recognition and Machine Learning (Information Science and Statistics)*. Springer-Verlag, New York.

Ann Bookman, M. Harrington, L. Pass, and E. Reisner. 2007. Family Caregiver Handbook: Finding Elder Care Resources in Massachusetts. (2007). Retrieved October 15, 2014 from http://web.mit.edu/workplacecenter/hndbk/.

Abdelhamid Bouchachia and Charlie Vanaret. 2014. GT2FC: An online growing interval type-2 self-learning fuzzy classifier. *IEEE Trans. Fuzzy Syst.* 22, 4 (Aug. 2014), 999–1018. DOI:http://dx.doi.org/10.1109/TFUZZ.2013.2279554

Marie Chan, Eric Campo, Daniel Estève, and Jean-Yves Fourniols. 2009. Smart homes—Current features and future perspectives. *Maturitas* 64, 2 (Oct. 2009), 90–97. DOI:http://dx.doi.org/10.1016/j.maturitas.2009.07.014

Rong Chen and Yu Tong. 2014. A two-stage method for solving multi-resident activity recognition in smart environments. *Entropy* 16, 4 (April 2014), 2184–2203. DOI:http://dx.doi.org/10.3390/e16042184

Yi-ting Chiang, J. Y. J. Hsu, Ching-Hu Lu, and Li-Chen Fu. 2010. Interaction models for multiple-resident activity recognition in a smart home. In *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS'10)*. 3753–3758. DOI:http://dx.doi.org/10.1109/IROS.2010.5650340

Diane J. Cook. 2009. Multi-agent smart environments. *J. Ambient Intell. Smart Env.* 1, 1 (Jan. 2009), 51–55. DOI:http://dx.doi.org/10.3233/AIS-2009-0007

Diane J. Cook, Aaron Crandall, Geetika Singla, and Brian Thomas. 2010. Detection of social interaction in smart spaces. *Cybern. Syst.* 41, 2 (Feb. 2010), 90–104. DOI:http://dx.doi.org/10.1080/01969720903584183

Erin Y. Cornwell and Linda J. Waite. 2009. Social disconnectedness, perceived isolation, and health among older adults. *J. Health Soc. Behav.* 50, 1 (March 2009), 31–48. DOI:http://dx.doi.org/10.1177/002214650905000103

Aaron S. Crandall and Diane J. Cook. 2009. Coping with multiple residents in a smart environment. *J. Ambient Intell. Smart Env.* 1, 4 (Dec. 2009), 323–334. DOI:http://dx.doi.org/10.3233/AIS-2009-0041

Aaron S. Crandall and Diane J. Cook. 2008a. Resident and caregiver: Handling multiple people in a smart care facility. In *Proceedings of the AAAI Fall Symposium—AI in Eldercare: New solutions to Old Problems* (2008). 39–47. DOI:http://dx.doi.org/10.1.1.329.4737

Aaron S. Crandall and Diane J. Cook. 2008b. Attributing events to individuals in multi-inhabitant environments. In *Proceedings of the IET 4th International Conference on Intelligent Environments* (August 2008). 1–8. DOI:http://dx.doi.org/10.1049/cp:20081164

Aaron S. Crandall and Diane J. Cook. 2010. Using a hidden Markov model for resident identification. In *Proceedings of th 6th International Conference on Intelligent Environments (IE'10)*. 74–79. DOI:http://dx.doi.org/10.1109/IE.2010.21

Scott Davidoff, Min Kyung Lee, John Zimmerman, and Anind Dey. 2006. Socially-aware requirements for a smart home. In *Proceedings of the International Symposium on Intelligent Environments*. 41–44. DOI:http://dx.doi.org/10.1.1.104.5771

Youtian Du, Feng Chen, and Wenli Xu. 2007. Human interaction representation and recognition through motion decomposition. *IEEE Signal Process. Lett.* 14, 12 (Dec. 2007), 952–955. DOI:http://dx.doi.org/10.1109/LSP.2007.908035

Youtian Du, Feng Chen, Wenli Xu, and Yongbin Li. 2006. Recognizing interaction activities using dynamic Bayesian network. In *Proceedings of the 18th International Conference on Pattern Recognition (ICPR'06)*. 618–621. DOI:http://dx.doi.org/10.1109/ICPR.2006.977

Nathan Eagle. 2008. Behavioral inference across cultures: Using telephones as a cultural lens. *IEEE Intell. Syst.* 23, 4 (July 2008), 62–64. DOI:http://dx.doi.org/10.1109/MIS.2008.58

Laura Fratiglioni, Hui-Xin Wang, Kjerstin Ericsson, Margaret Maytan, and Bengt Winblad. 2000. Influence of social network on occurrence of dementia: A community-based longitudinal study. *Lancet* 355, 9212 (April 2000), 1315–1319. DOI:http://dx.doi.org/10.1016/S0140-6736(00)02113-9

Jane Gross. 2007. A Grass-Roots Effort to Grow Old at Home. *N. Y. Times* (August 2007).

Tao Gu, Zhanqing Wu, Xianping Tao, H. K. Pung, and Jian Lu. 2009a. epSICAR: An emerging patterns based approach to sequential, interleaved and concurrent activity recognition. In *Proceedings of the IEEE International Conference on Pervasive Computing and Communications (PerCom'09)*. 1–9. DOI:http://dx.doi.org/10.1109/PERCOM.2009.4912776

Tao Gu, Zhanqing Wu, Liang Wang, Xianping Tao, and Jian Lu. 2009b. Mining emerging patterns for recognizing activities of multiple users in pervasive computing. In *Proceedings of the 6th Annual International Conference on Mobile and Ubiquitous Systems: Networking Services (MobiQuitous'09).* 1–10. DOI:http://dx.doi.org/10.4108/ICST.MOBIQUITOUS2009.6818

Chien-Chang Hsu and Liang-Zong Wang. 2008. A smart home resource management system for multiple inhabitants by agent conceding negotiation. In *Proceedings of the IEEE International Conference on Systems, Man and Cybernetics (SMC'08)*. 607–612. DOI:http://dx.doi.org/10.1109/ICSMC. 2008.48 11344

Kuo-Chung Hsu, Yi-Ting Chiang, Gu-Yang Lin, Ching-Hu Lu, Jane Yung-Jen Hsu, and Li-Chen Fu. 2010. Strategies for inference mechanism of conditional random fields for multiple-resident activity recognition in a smart home. In *Trends in Applied Intelligent Systems*, Lecture Notes in Computer Science. Springer, Berlin, 417–426.

Derek Hao Hu, Xian-Xing Zhang, Jie Yin, Vincent Wenchen Zheng, and Qiang Yang. 2009. Abnormal activity recognition based on HDP-HMM models. In *Proceedings of the 21st International Joint Conference on Artifical Intelligence (IJCAI'09)*. Morgan Kaufmann Publishers Inc., 1715–1720.

Shehroz S. Khan, Michelle E. Karg, Jesse Hoey, and Dana Kulic. 2012. Towards the detection of unusual temporal events during activities using HMMs. In *Proceedings of the Conference on Ubiquitous Computing (UbiComp'12)*. ACM, New York, 1075–1084. DOI:http://dx.doi.org/10.1145/2370216.2370444

Eunju Kim, S. Helal, and D. Cook. 2010. Human activity recognition and pattern discovery. *IEEE Pervasive Comput.* 9, 1 (Jan. 2010), 48–53. DOI:http://dx.doi.org/10.1109/MPRV.2010.7

Jinyan Li, Guimei Liu, and Limsoon Wong. 2007. Mining statistically important equivalence classes and delta-discriminative emerging patterns. In *Proceedings of the 13th International Conference on Knowledge Discovery and Data Mining (SIGKDD'07). ACM*, New York, 430–439. DOI:http://dx.doi.org/10.1145/1281192.1281240

Wenhui Liao and Qiang Ji. 2009. Learning Bayesian network parameters under incomplete data with domain knowledge. *Pattern Recognit.* 42 (2009), 3046–3056.

Zhi-Hao Lin and Li-Chen Fu. 2007. Multi-user preference model and service provision in a smart home environment. In *Proceedings of the IEEE International Conference on Automation Science and Engineering (CASE'07)*. 759–764. DOI:http://dx.doi.org/10.1109/COASE.2007.4341737

Ching-Hu Lu, Yung-Ching Lin, and Li-Chen Fu. 2008. Hide and not easy to seek: A hybrid weaving strategy for context-aware service provision in a smart home. In *IEEE Asia-Pacific Services Computing Conference (APSCC'08)*. 595–600. DOI:http://dx.doi.org/10.1109/APSCC.2008.77

Iain McCowan, Daniel Gatica-Perez, Samy Bengio, Guillaume Lathoud, Mark Barnard, and Dong Zhang. 2005. Automatic analysis of multimodal group actions in meetings. *IEEE Trans. Pattern Anal. Mach. Intell.* 27, 3 (March 2005), 305–317. DOI:http://dx.doi.org/10.1109/TPAMI.2005.49

Ahmed Nait Aicha, Gwenn Englebienne, and Ben Kröse. 2013. How lonely is your grandma?: Detecting the visits to assisted living elderly from wireless sensor network data. In *Proceedings of the Conference on Pervasive and Ubiquitous Computing Adjunct Publication (UbiComp'13)*. ACM, New York, 1285–1294. DOI:http://dx.doi.org/10.1145/2494091.2497283

Pradeep Natarajan and Ramakant Nevatia. 2007. Coupled hidden semi Markov models for activity recognition. In *Proceedings of the Workshop on Motion and Video Computing (WMVC'07)*. IEEE, 10. DOI:http://dx.doi.org/10.1109/WMVC.2007.12

Nam Nguyen, Svetha Venkatesh, and Hung Bui. 2006. Recognising behaviours of multiple people with hierarchical probabilistic model and statistical data association. In *Proceedings of the 17th British*

Machine Vision Conference (BMVC'06). The British Machine Vision Association and Society for Pattern Recognition, 1239–1248.

Daniel O. Olguin, Benjamin N. Waber, Taemie Kim, Akshay Mohan, Koji Ara, and Alex Pentland. 2009. Sensible organizations: Technology and methodology for automatically measuring organizational behavior. *IEEE Trans. Syst. Man Cybern. Part B Cybern.* 39, 1 (Feb. 2009), 43–55. DOI:http://dx.doi.org/10.1109/TSMCB.2008.2006638

Markus Prossegger and Abdelhamid Bouchachia. 2014. Multi-resident activity recognition using incremental decision trees. In *Adaptive and Intelligent Systems*, Abdelhamid Bouchachia (Ed.). Lecture Notes in Computer Science. Springer International Publishing, 182–191.

Iulian Pruteanu-Malinici and Lawrence Carin. 2008. Infinite hidden Markov models for unusual-event detection in video. *IEEE Trans. Image Process.* 17, 5 (May 2008), 811–822. DOI:http://dx.doi.org/10.1109/TIP.2008.919359

Barry Reisberg, Sanford I. Finkel, John Overall, and Hellmut Erzigkeit. 2001. The Alzheimer's disease activities of daily living international scale (ADL-IS). *Int. Psychogeriatr. IPA* 13, 2 (June 2001), 163–181.

Daniele Riboni, Linda Pareschi, Laura Radaelli, and Claudio Bettini. 2011. Is ontology-based activity recognition really effective? In *Proceedings of the IEEE International Conference on Pervasive Computing and Communications Workshops (PERCOM Workshops'11)*. 427–431. DOI:http://dx.doi.org/10.1109/PERCOMW.2011.5766927

Susanne Smith Roley et al. 2008. Occupational therapy practice framework: Domain and practice 2nd ed. *Am. J. Occup. Ther. Off. Publ. Am. Occup. Ther. Assoc.* 62, 6 (Dec. 2008), 625–683.

Fariba Sadri. 2011. Ambient intelligence: A survey. *ACM Comput. Surv.* 43, 4 (Oct. 2011), 36:1–36:66. DOI:http://dx.doi.org/10.1145/1978802.1978815

Sumit Sanghai, Pedro Domingos, and Daniel Weld. 2005. Relational dynamic Bayesian networks. *J. Artif. Intell. Res.* 24, 2005 (2005), 759–797.

Jehad A. M. Sarkar, Young-Koo Lee, and Sungyoung Lee. 2010. ARHMAM: An activity recognition system based on hidden Markov minded activity model. In *Proceedings of the 4th International Conference on Uniquitous Information Management and Communication (ICUIMC'10)*. ACM, New York, 71:1–71:9. DOI:http://dx.doi.org/10.1145/2108616.2108702

Geetika Singla, Diane J. Cook, and Maureen Schmitter-Edgecombe. 2010. Recognizing independent and joint activities among multiple residents in smart environments. *J. Ambient Intell. Humaniz. Comput.* 1, 1 (March 2010), 57–63. DOI:http://dx.doi.org/10.1007/s12652-009-0007-1

Eliot R. Smith and Diane M. Mackie. 1999. *Social Psychology*. Routledge, London.

Charles Sutton and Andrew McCallum. 2006. An introduction to conditional random fields for relational learning. In *Introduction to Statistical Relational Learning*. MIT Press, 93–128.

Charles Sutton, Andrew McCallum, and Khashayar Rohanimanesh. 2007. Dynamic conditional random fields: Factorized probabilistic models for labeling and segmenting sequence data. *J. Mach. Learn. Res.* 8 (May 2007), 693–723.

Tanzeem Choudhury and Alex Pentland. 2004. Characterizing social networks using the sociometer. In *Proceedings of NAACOS 2004*.

Tim L. M. van Kasteren. 2011. *Activity Recognition for Health Monitoring Elderly Using Temporal Probabilistic Models*. Ph.D. dissertation. University of Amsterdam.

Tim L. M. van Kasteren, Gwenn Englebienne, and Ben J. A. Kröse. 2011. Hierarchical activity recognition using automatically clustered actions. In *Proceedings of the 2nd International Conference on Ambient Intelligence (AmI'11)*. 82–91.

Tim L. M. van Kasteren, Gwenn Englebienne, and Ben J. A. Kröse. 2010. Human activity recognition from wireless sensor network data: Benchmark and software. In *Activity Recognition in Pervasive Intelligent Environments*, Liming Chen, Chris D. Nugent, Jit Biswas, and Jesse Hoey (Eds.), Atlantis Ambient and Pervasive Intelligence. Atlantis Press, 165–186.

Tim L. M. van Kasteren, Athanasios Noulas, Gwenn Englebienne, and Ben Kröse. 2008. Accurate activity recognition in a home setting. In *Proceedings of the 10th International Conference on Ubiquitous Computing (UbiComp'08)*. ACM, 1–9. DOI:http://dx.doi.org/10.1145/1409635.1409637

Kristof Van Laerhoven, Nicky Kern, Hans W. Gellersen, and Bernt Schiele. 2003. Towards a wearable inertial sensor network. In *Proceedings of the IEE Eurowearable Conference (2003)*. 125–130. DOI:http://dx.doi.org/10.1049/ic:20030158

Liang Wang, Tao Gu, Xianping Tao, Hanhua Chen, and Jian Lu. 2011. Recognizing multi-user activities using wearable sensors in a smart home. *Pervasive Mob. Comput.* 7, 3 (June 2011), 287–298. DOI:http://dx.doi.org/10.1016/j.pmcj.2010.11.008

Liang Wang, Tao Gu, Xianping Tao, and Jian Lu. 2009. Sensor-based human activity recognition in a multi-user scenario. In *Ambient Intelligence*, Manfred Tscheligi et al. (Eds.). Lecture Notes in Computer Science. Springer, Berlin, 78–87.

Daniel H. Wilson and Chris Atkeson. 2005. Simultaneous tracking and activity recognition (STAR) using many anonymous, binary sensors. In *Pervasive Computing*, Hans-W. Gellersen, Roy Want, and Albrecht Schmidt (Eds.). Lecture Notes in Computer Science. Springer, Berlin, 62–79.

Received February 2015; revised June 2015; accepted September 2015