Knowledge-Based Systems 121 (2017) 173–184

Contents lists available at ScienceDirect

Knowledge-Based Systems

journal homepage: www.elsevier.com/locate/knosys

Probabilistic ontology based activity recognition in smart homes using Markov Logic Network

K.S. Gayathri [a,*], K.S. Easwarakumar [b], Susan Elias [c]

[a] Department of Computer Science and Engineering, Sri Venkateswara College of Engineering, Sriperumbdur, India
[b] Department of Computer Science and Engineering, Anna University, Chennai, India
[c] School of Electronics Engineering, VIT University, Chennai, India

ARTICLE INFO

Article history:
Received 20 April 2016
Revised 15 November 2016
Accepted 19 January 2017
Available online 25 January 2017

Keywords:
Smart homes
Activity recognition
Ontology
Probabilistic approach
Markov Logic Network

ABSTRACT

Designing an activity recognition system that models various activities of an occupant is the fundamental task in creating a smart home. Activity Recognition (AR) modeling, has witnessed a comprehensive range of research, that focuses independently on probabilistic approaches and on ontology based models as well. The research presented in this paper introduces an innovative approach in AR system design that integrates probabilistic inference with the represented domain ontology. Data obtained from sensors are uncertain in nature and mapping uncertainty over ontology will not yield good accuracy in the context of AR. The proposed system augments ontology based activity recognition with probabilistic reasoning through Markov Logic Network (MLN) which is a statistical relational learning approach. The proposed system utilizes the model theoretic semantic property of description logic, to convert the represented ontology activity model to its corresponding first order rules. MLN is constructed by learning weighted first order rules that enable probabilistic reasoning within a knowledge representation framework. The experiments based on datasets obtained from smart home prototypes illustrate the effectiveness of integrating probabilistic reasoning over domain ontology and the result analysis shows enhanced recognition accuracy in comparison with existing approaches.

1. Introduction

Smart environments are enabled with intelligence, integrated into their surroundings for providing services that improve the quality of living [1]. In recent years, the research in smart environments have rapidly developed with the availability of low cost and low power sensors complemented with advances in wireless technologies and the emergence of research in fields such as Ambient Intelligence (AmI) and Ubiquitous computing. Disruptive technologies such as Internet of Things(IoT), augmented with Machine Learning paradigms revolutionized the design of smart environments [2,3].

Smart homes have found extensive applications in various socially relevant problems and as a result it is likely to have a huge impact on the future society[1]. Smart homes are designed to assist the occupant needing assistance to complete their daily routines independently without the help of a care taker [4,5]. The assistive smart home monitoring system recognizes and detects abnormality in occupant behavior and performs decision making to remotely alert the care taker during critical situations [6–8].

Activity recognition (AR) is the most important process in incorporating ambient intelligence into smart environments. It involves as a series of complex processes of activity monitoring, modeling, reasoning and decision making [2,3]. To illustrate the significance of activity recognition in the design of a smart home, let us consider a scenario - “Mary enters the Kitchen at 2:00 a.m and performs a sequence of actions using the objects kettle, stove and mug”. It is necessary for the smart home system to primarily recognize the ongoing activity in order to detect abnormality or to automate actions in the environment. In the above scenario, the ongoing activity is recognized as ‘cooking’ and it is associated with various contextual attributes (objects, space, and time). In order to integrate ambient intelligence in a smart home the fundamental task is activity recognition because, reasoning and inferences can be made only based on the ongoing activity [1,9].

The primary task in designing an activity recognition system involves the construction of an activity model that represents the occupant’s behavior and activity pattern for recognition [10]. The routine activities are the frequent activities that are repeated for a

* Corresponding author.
E-mail addresses: gayasuku@svce.ac.in, gayathrisivakumar80@gmail.com (K.S. Gayathri), easwarakumarks@gmail.com (K.S. Easwarakumar), susan.elias@vit.ac.in (S. Elias).

http://dx.doi.org/10.1016/j.knosys.2017.01.025
0950-7051/

few days. Such routine activities referred to as Activities of Daily Living(ADL) are preferred for activity modeling for the reason that these activities help in defining normal scenarios within the smart home [2]. It is therefore essential to study the characteristics of various ADLs of the occupant of the smart home so as to model a recognition system effectively[11].

ADLs executed by the occupant are generally described as complex patterns of events. An event in a smart environment is regarded to be atomic and indivisible in nature, for example 'lying on the bed', 'sitting on a chair' etc. A simple activity is characterized by an ordered sequence of events, for example, 'preparing meal' involves the sequence 'switching on stove', 'take pan' etc. Composite activity is two or more simple activities that appear within a time interval. Composite activities vary depending on the type of inter activity relationship connecting simple activities [12]. Sequential, interleaved and concurrent are the different ways in which composite activities are defined. Thus, activity recognition system needs to model both simple and composite activities.

Majority of the ADLs are common for most of the occupants, in terms of both functions and objects involved [13]. However, certain actions differ according to the individual's preference. The common functionalities can be grouped under different granularity. For example, 'Make hot drink' is a higher level granularity of 'Make coffee' and 'Make tea'. This activity granularity presents a hierarchical ordering among activities that enhances recognition accuracy and thus is preferred in activity modeling.

Moreover, the occupant's need and preferences may change over time, for example the occupant may take bath for 20 min during summer but he might prefer to take a bath only for 10 min during winter. The activity model needs to learn these preferences and update itself by representing the latest changes in occupant's interests [2]. Therefore the activity modeling mechanism employed must be flexible enough to incorporate these variations so as to handle activity diversity and dynamics.

Activities are executed under different context, for example specific locations, objects, time and situation [14]. Thus, the activity modeling approach should facilitate situation aware computing to handle spatio-temporal data through context modeling and recognition mechanisms. In smart homes, uncertainty materializes from situations that require recognizing incomplete sequences of events or those that emerge due error in sensor data [15]. The approach employed to model activity recognition should provide a means to handle uncertainty into the recognition system.

It is thus essential for any activity modeling approach to deal with the above mentioned challenges. The proposed activity recognition framework presented in this paper effectively tackles activity granularity, contextual knowledge, activity diversity through ontology model, while activity dynamics, data uncertainty are addressed through probabilistic reasoning over the represented domain ontology.

The remaining part of the paper is structured as follows: Section 2 reviews the related work on smart homes, presents various approaches for activity recognition and defines the motivation and scope of this paper. Section 3 describes the theoretical foundations of the proposed probabilistic ontology based activity modeling and recognition system. Section 4 describes the system prototype, experimental analysis and performance evaluation. Finally, Section 5 concludes the paper and outlines the future work.

2. Related work

An analysis and review of the existing Activity Recognition(AR) approaches are presented here, in-order to highlight the motivation of our proposed work. The challenges that need to be addressed in the design of a smart home were listed in the previous sections. The limitations of the existing activity recognition approaches are discussed and an appropriate solution strategy for AR modeling is proposed here.

Based on the type of sensors utilized for monitoring the occupants, the activity recognition is named as vision based or sensor based [15]. Vision based approaches make use of video cameras to monitor the occupant and the environment. Reasoning is performed over the collected video frames to recognize the ongoing activity. Though, vision based activity recognition has several advantages in modeling smart environment it has its own setbacks while being used in the design of the smart home. Privacy of occupant is a key concern in a home set up and thus cameras are not a suitable means to monitor the occupant for activity recognition.

Sensor based activity recognition employ wide varieties of sensors for activity monitoring. These sensors are either fixed on various objects in the environment or worn by the occupant [15]. The acquired data are then processed through statistical and knowledge engineering algorithms to facilitate activity recognition. The different types of sensors include wearable sensors like accelerometer, Global Positioning System (GPS) and Bio sensors [15,16]. The gesture of the occupant such as walking, running, and sitting is measured through accelerometer sensor. Global Positing System (GPS) is utilized for supervising the location of the occupant in an open mobile environment. The occupant's vital statistics like ECG, EEG, blood pressure, heart rate etc. are measured through bio sensors. The challenges in dealing with wearable sensors arises from the willingness of the occupant to put on such devices, battery existence, easiness of use and the size of device. Compared with other types of sensors, object sensors are advantageous as they can give an indirect indication of the occupant's activities [15].

Activity recognition has been broadly analyzed using Data driven and Knowledge driven approaches [15]. Data driven approach utilizes probabilistic and statistical machine learning strategies for the analysis and modeling sensor data. Knowledge driven paradigms utilize knowledge engineering and management techniques for activity modeling using domain information and artificial intelligence based reasoning techniques for inferences.

2.1. Data driven activity modeling

Data driven approaches are modeled through generative approach or discriminative approaches [9,15,16]. Probabilistic models that are used in generative approaches give a comprehensive description about the input sensor data space. Classification model is employed in discriminative approach for constructing the activity model. Naive Bayes classifier (NBC), Hidden Markov Models (HMM) and Dynamic Bayesian Networks (DBM) are most commonly used generative approaches. Support Vector Machines (SVM) and Artificial Neural Networks (ANN) are most commonly used discriminative approaches [17,18]. Data driven approach needs a huge collection of sensor data to accurately model the occupant behavior. This introduces an overhead and is termed as 'cold start problem'. Furthermore re-usability is a concern where the complete task of data collection, analysis and modeling of activity needs to be done for building an activity model for a new occupant.

2.2. Knowledge driven activity modeling

Knowledge driven approaches construct activity models as a reusable contextual model that relate objects, space and time with occupant behavior. Knowledge is extracted from domain expert or mined from web using data mining techniques. The knowledge driven model is semantically clear, follows a uniform approach for representation, has no cold start problem, builds reusable activity model and thus scales to bigger environment thereby making it the most suitable approach for activity recognition [15]. Logic and ontology are most commonly used approaches for representing the domain knowledge [13]. Ontology based activity recognition is a recent approach and has gained enormous interest [13] in large scale adoption, application development and system prototyping, but has limitations in performing reasoning over temporal and uncertain data [19] making the modeling of composite activities a challenging task.

Thus, to model an activity recognition system for smart home, a hybrid activity modeling is required so as to integrate the essential characteristics of data driven and knowledge driven approach. To combine the advantages of ontology and probabilistic reasoning, design of a specific hybrid approach has been the focus of research presented in this paper.

2.3. Hybrid activity recognition: data driven and knowledge driven approach

This section focuses on modeling the activity recognition system through a combined data driven(probabilistic) and knowledge(ontology) driven approach.

Many existing activity recognition systems have employed hybrid approach for its design. Combined Ontological and Statistical Activity Recognition module(COSAR) is one such system where statistical approach is used to map the observations of the environment to an activity [20]. Using these inferences, an extended reasoning is done with ontology to perform context reasoning. Limitation of this approach is that the temporal information and duration of the activity are not effectively modeled and moreover the symbolic reasoning is done only to refine the statistical inference. Some of these have been addressed in the extended version [21] paper. The limitations of this approach in comparison with approach proposed in this paper are presented in the experimental section.

Another research work [22] builds a composite activity model using hybrid ontology and temporal approach where the Allen's temporal logic is modeled into the ontology and the Semantic Web Rule Language (SWRL) based inferences are used to make inference on occupants activity. But this approach is not modeled with sensor behavior data which reduces recognition accuracy.

The existing probabilistic ontology approach based on PR-OWL are modeled using UnBBayes [23]. UnBBayes is a Java plug-in that integrates probabilistic reasoning into ontology through Multi Entity Bayesian Network (MEBN) and supports the full integration of first order logic and probability theory. Though UnBBayes approach integrates probabilistic reasoning over ontology, its accuracy in recognition is low. It does not learn weights from the individuals of ontology through an efficient weight learning algorithm that models the training data at the best possible level. Moreover in UnBBayes, the local distribution of the random variables needs to be fixed while modeling and representing temporal data is difficult in Bayesian network.

To address the various issues of hybrid activity recognition system, a better modeling approach such as Markov Logic Network(MLN) [24] is required. MLN is an approach of statistical relational learning that merges probabilistic reasoning and inductive logic programming in a unified framework. Discriminative / generative machine learning mechanisms of MLN effectively models the activity recognition system with better accuracy. The following sections initially present details on independent modeling of activity recognition using Markov Logic and subsequently discusses its integration with ontology.

2.4. Probabilistic ontology based approach through Markov Logic Network

MLN has been applied to model several activity recognition systems in smart homes [25]. The MLN based activity model in [12,25] recognizes simple and composite activities and has been successful in the identification of normal and anomalous occupant behavior. Composite activity modeling described in [12] identifies concurrent, interleaved and sequential simple activities using MLN. The existing approach [26] of audio based activity recognition has employed MLN for recognizing activity from audio signals. This approach has limitations in dealing with contextual attributes to define the model, non inclusion of object based activity recognition, less efficient representation and reasoning of temporal data, inability in terms of composite activity recognition, lack of consistency check in MLN model and restricted fixed time interval segmentation.

Although, the precision of MLN based activity recognition has been satisfactory in certain context it has few specific limitations like in the representation of domain knowledge using first order logic and in the modeling of temporal and composite activities. The drawback with first order logic representation is that it does not automatically discover inconsistency among represented knowledge, lack of domain knowledge organization and in-feasibility in the hierarchical association of domain related concepts. Because of these limitations, the aspects of activity granularity and activity diversity are not well modeled in MLN.

Since ontology and Markov Logic Network is better in their own way for activity modeling, the proposed framework prefers to represent domain knowledge through ontology and facilitate probabilistic reasoning into ontology through novel utilization of MLN. Thus, the probabilistic ontology aims to offer consistent knowledge, enable uncertainty modeling, temporal modeling, contextual modeling and composite activity modeling in a unified framework. Furthermore, the challenge of activity diversity, activity dynamics, activity granularity and uncertainty modeling can be well addressed.

3. Proposed probabilistic ontology based activity modeling and recognition

The details of the proposed probabilistic ontology based activity recognition model are presented in this section. The novelty of the proposed design is in the integration of probabilistic reasoning into domain ontology using Markov Logic Network (MLN). The proposed Activity Recognition framework has two primary components namely Ontology based activity model and the Markov logic inference network (Fig. 1). Activity patterns mined using Event Pattern Activity Modeling (EPAM) framework [27] and domain knowledge are utilized in the proposed framework to construct the Terminology Box (TBox) and Assertion Box (ABox) of the ontology and transformed into a MLN to facilitate probabilistic context based reasoning. In the translation process, TBox of the ontology is transformed into first order logic based on the model theoretic semantics property [28] of description logic. Weight learning is then performed over the generated first order rules and individuals of the ABox, to generate the MLN activity model having the weighted first order rules. In real time applications, events are continuously captured and converted into discrete event sequences using the proposed algorithm that combines fixed interval based segmentation with location information. The generated event sequences are then processed by the MLN activity model to recognize the ongoing activity of the occupant of the smart home. The recognized activity is utilized by the decision making system to further process the relevant actions to be taken based on the requirements of the smart home application.

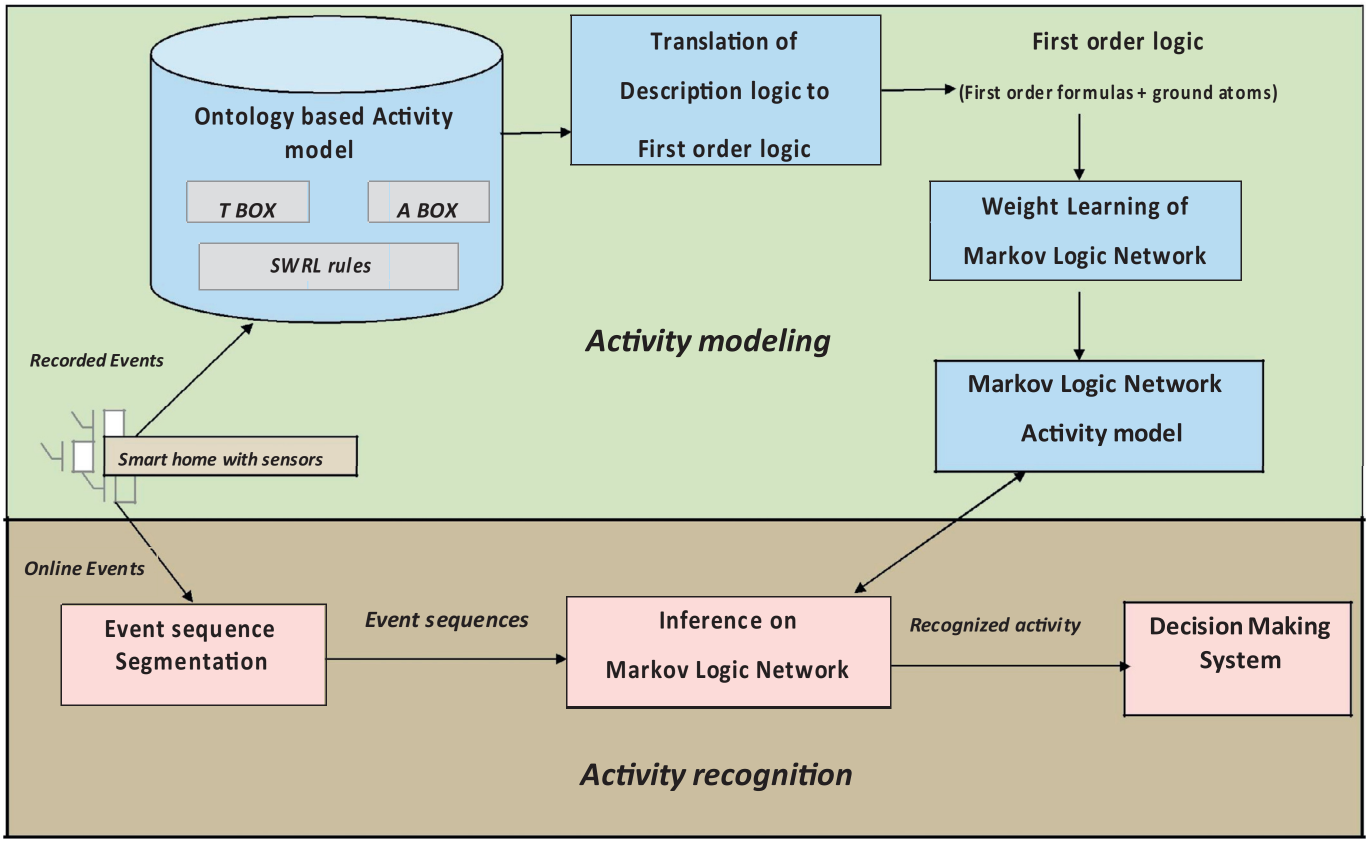

Fig. 1. Proposed probabilistic ontology based activity recognition system in smart home.

In the following subsections details of the various subsystems of the proposed framework is presented.

3.1. Ontology activity modeling

Semantic Web Ontology Language(OWL) [29] is preferred for knowledge modeling because of its expressive ability to effectively represent domain knowledge and perform reasoning [13]. Several smart home ontologies for activity modeling exists [19,30]. The ontology model presented in [21] is utilized in the proposed design as it models most of the essential attributes required in the design of smart home activity recognition system. This ontology is customized to complement the proposed design and is extended with temporal entities so as enable the essential functionalities for temporal modeling and composite activity modeling.

3.1.1. Composite activity modeling

Modeling composite activities (CA) are quintessential as the occupant habitually performs activities in complex patterns and these in turn are temporally related simple activities. Thus description logic that is used for constructing ontology activity model needs to incorporate and handle temporal data. OWL-Time is a temporal logic that defines classes to handle temporal data [31] and is utilized in the proposed ontology based modeling to represent composite activities. Extensive literature on description logic shows that the temporal inferences in ontology are handled effectively by defining rules. Semantic Web Rule Language (SWRL) [32] has the flexibility to define rules with temporal classes that enable temporal inferences over the represented ontology. SWRL requires description of temporal relationships between the time instances or between interval classes. Allen's temporal relations are effectual in bringing out a variety of possible temporal relations [33]. Composite activities are sequential, interleaved and concurrent in nature [34] and are represented in the model through SWRL rules with Allen's temporal relations.

The following are the concepts and properties to define composite activity model:

- ***CompositeADL*** - This concept represents composite ADL is the sub concept of ‘ADL’ concept
- ***SequentialCA, InterleavedCA, Concurrent CA*** - subconcepts of ‘CompositeADL’ for modeling sequential, interleaved and concurrent composite activities
  - ***hasSimpleactivity*** property - this property is used to relate the ‘SimpleADL’ with ‘Composite activity’

3.2. Allen's temporal logic based activity modeling

In addition to ontological modeling of relationships between the activities and entities as described above, the temporal relations among simple activities are modeled using Allen's temporal logic.

- *Sequential composite activities*
  Sequential composite activities are modeled through ‘before/after’ or ‘meets/metby’ Allen's temporal relations. The relations ‘before/after’ indicates that there is a gap between two time intervals whereas ‘meets/metby’ indicates that no gap between the two time intervals.

- *Interleaved and concurrent composite activity*
  The time interval between two simple activities that overlap completely or partially defines interleaved or concurrent activities. Nine temporal relations such as Overlaps / Overlapped by, Contains / during, Starts / Started by and Equals [33] are required to define interleaved and concurrent activities.

All the 13 Allen's temporal relations are first modeled by using the SWRL rules. These rules are then extended to design a composite activity recognition system. The Sequential CA is modeled using four SWRL rules whereas the Interleaved and Concurrent activities are modeled with eighteen SWRL rules. The following are samples of Allen's temporal relation modeled by SWRL rules for composite activity inference:

'After' Allen's relation

ProperInterval(?x), ProperInterval(?y), after(?c, ?b), hasBeginning(?y, ?c), hasEnd(?x, ?b) → intervalAfter(?y, ?x)

'Overlap' Allen's relation

ProperInterval(?x), ProperInterval(?y), before(?a, ?c), before(?c, ?b), before(?b, ?d), hasBeginning(?x, ?a), hasBeginning(?y, ?c), hasEnd(?x, ?b), hasEnd(?y, ?d) → intervalOverlaps(?x, ?y)

Sequential composite activity rule

CompositeADL(?c), hasSimpleactivity(?c,?ax), hasSimpleactivity(?c,?ay), SimpleADL(?ax), SimpleADL(?ay), hasTime(?tax, ?ax), hasTime(?tay, ?ay), hasTimeInterval(?tax, ?x), hasTimeInterval(?tay, ?y), ProperInterval(?x), ProperInterval(?y), intervalAfter(?y, ?x) → SequentailCA(?c)

The recognition of composite activity is illustrated with an example where, a simple activity 'clean' can follow a simple activity 'prepare breakfast', where the time interval of 'clean' activity satisfies 'intervalAfter' relation with the time interval of 'prepare breakfast' and hence the 'clean' and 'prepare breakfast' are concluded to be a sequential composite activity. The above ontology activity model can thus recognize both simple and composite activities. The inference mechanism on the represented ontology based activity model cannot handle data uncertainty as there is no methodology to incorporate probabilistic reasoning. The following sections describe an innovative approach that integrates probabilistic reasoning into the represented ontology based activity model.

3.3. Markov Logic Network for activity recognition

Markov Logic Network (MLN) is characterized as weighted first order rules, where weight is a real valued number reflecting the truth of the rule with that of the dataset [24]. Based on the weight allocation, first order rules in MLN is categorized into soft and hard [35]. The first order rules outline the structure of MLN and weights reflect the support of the rule with respect to real world facts. Weighted first order rule in MLN is represented as F_i, weight of the rule is represented as w_i and kth predicate of the rule F_i is represented as f_{ik} . During run time, inference is achieved by grounding the variables in the predicates of first order rules that aid in constructing a Markov network. Probabilistic inference using first order rules is achieved using equation 1 that calculates the joint probability of all the grounded atoms appearing together in the Markov network, where 'Z' is the normalization factor, 'X' is the set of 'n' contextual attributes and these attributes elements are represented by 'x', where x_j (where j takes a value 1 to n) assigns the value of the attributes to the predicate f_{ik}(x) [36].

$$P(X = x) = \frac{1}{Z} \exp\left(\sum_i w_i f_{ik}(x_j)\right) \tag{1}$$

3.4. Translation of ontology activity model into MLN activity model

A novel approach to integrate probabilistic inference into represented ontology activity model is proposed by converting ontology into its corresponding MLN activity model. The Algorithm 1 shows the steps involved in conversion of the ontology activity model into Markov Logic Network activity model. Lines 2–7 of Algorithm 1 generates the first order equivalent of TBox and SWRL rules of the ontology as described in Sections 3.1 and 3.1.1. Whereas, lines 8–10 of Algorithm 1 learns the weight of the first order logic using the ABox of the ontology. The Fig. 2 shows a visual representation of various steps in modeling MLN from the ontology based activity model.

Algorithm 1: Translation of ontology activity model into MLN activity model.

```
1  Procedure Ontology to MLN (Ontology activity model)
   Result: a Markov Logic Network (MLN) activity model
                                // Generation of first order rules
2  foreach concept in TBox of Ontology activity model do
3  |  MLN structure ← First order logic equivalent of concept
4  end
5  foreach SWRL rule in Ontology activity model do
6  |  MLN structure ← First order logic equivalent of SWRL
   |  rule
7  end
                          // Weight learning of MLN activity model
8  foreach Individuals in ABox of Ontology activity model do
9  |  Weighted MLN ← Weight_learning(MLN)
10 end
```

Ontology is modeled using OWL language. OWL is based on the Description Logic, and therefore (in most of the cases) can be deduced as formulae in first order logic as described in this work [28,37]. Thus, an expressive ontology model involving complex classes and relations can be translated to its corresponding FOL [28]. The possible DL to FOL is shown in Table 1. The key initiative following this interpretation is that concepts correspond to unary predicates, roles to binary predicates, and individuals correspond to constants and hence the translation of DL to FOL is sound and complete. Weight learning algorithms are employed to learn weights for the first order rules of MLN based activity model. The research work carried out in [38] shows that the learning algorithm used for modeling MLN is also sound and complete.

The Table 2 shows some of weighted first order rules generated from the weight learning process of MLN. The *SR1* weighted rule represents an atomic event 'opengrocerycupboard' with its contextual attributes. The *SR2* describes the simple ADL 'preparebreakfast' with sequence of events. It can be observed that, during weight learning process, MLN substitutes all possible values for the attributes but any irrelevant change in the event reduces the weight of the rule as observed in *SR3* and *SR4*. This weight change reflects the absence of corresponding associations of atomic events for simple ADL in the dataset used for modeling. This is illustrated in *SR3* rule which indicates a change in an event from 'opengrocerycupboard' to 'toiletflush'.

The process of designing activity model of AR system prototype is completed and the following subsections describe the processes that need to be carried out during runtime.

3.5. Segmentation module to generate event sequences

The sensor events generated during run time are continuous and needs to be converted to event sequences by a segmentation process so as to enable activity recognition. In real time, it is complex to segment events to obtain complete event sequences of an activity. The precision in activity recognition depends on the generated event. The existing work [18] employs either fixed interval or location based segmentation and the events sequences generated are generally incomplete. Thus the proposed system introduces a hybrid of location and fixed interval based segmentation that reduces the generation of incomplete event sequences thereby increasing the accuracy of recognition process. The proposed segmentation approach Algorithm 2 at first employs location based

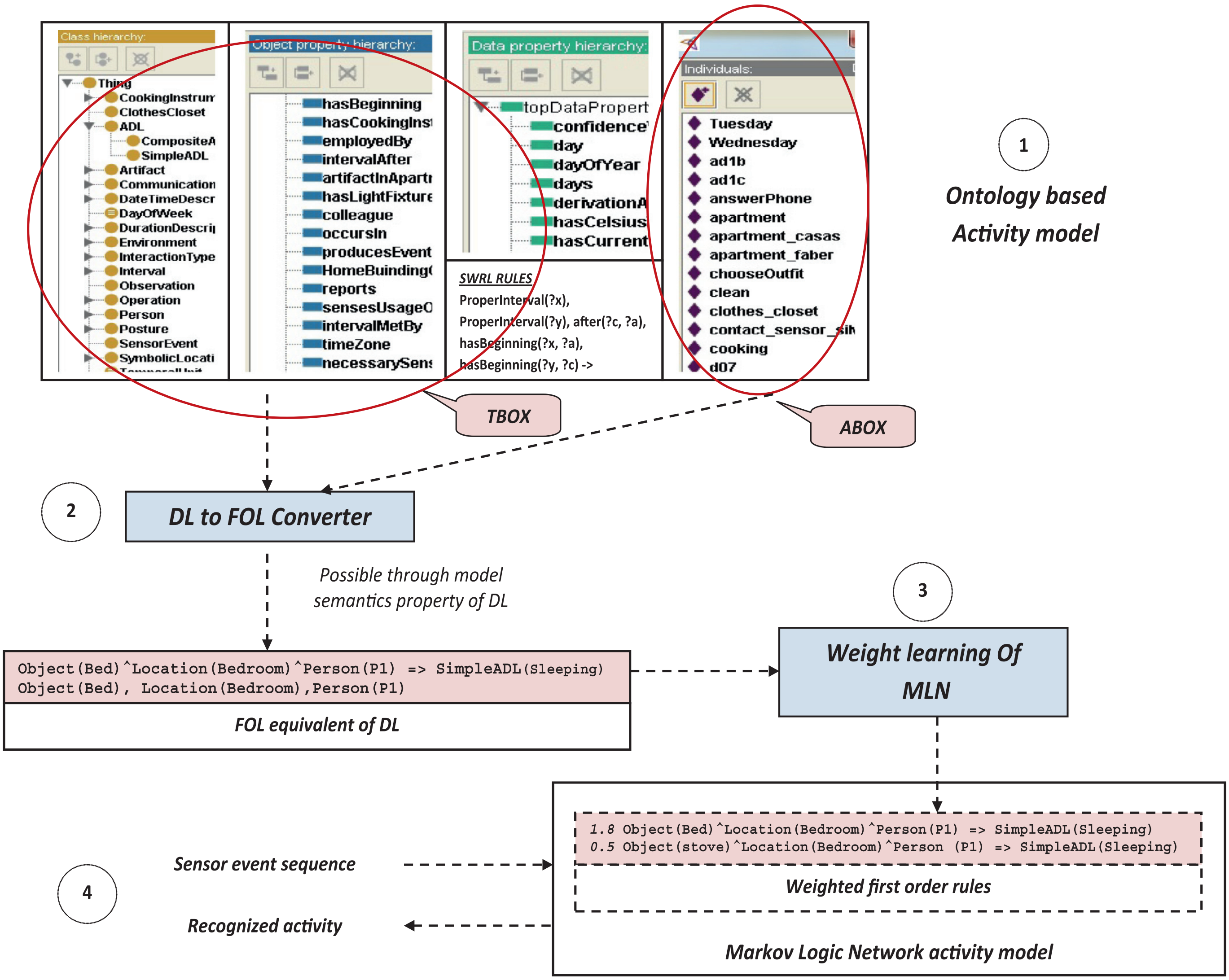

Fig. 2. Steps involved in modeling MLN based activity model from ontology activity model.

Table 1
First order logic equivalent for description logic.

Description logic representation	First order logic representation
SubClassOf(class1,class2)	∀ x, y: class1(x) → class2(x)
Transitive property(p)	∀ x, y, z: p(x,y) ∧ p(y,z) → p(x,z)
ProperInterval(?x),ProperInterval(?y), after(?c, ?a), hasBeginning(?x, ?a), hasBeginning(?y, ?c) -> intervalAfter(?y, ?x) SWRL rule, not a DL	∀ x, y ∊ time interval, ∀ a, c ∊ time instant: ProperInterval(x) ∧ ProperInterval(y) ∧ after(c, a) ∧ hasBeginning(x, a) ∧ hasBeginning(y, c) → intervalAfter(y, x)
ClassAssertion(class1, Anny)	Class1 (Anny)

Table 2
Sample soft rules in MLN activity model.

SR1	**2.39** $AtomicEvent(e) \wedge haslocation(e, kitchen) \wedge hasObject(e, cupboard) \wedge hasTimeInstance(e, 2008-07-299:30:28) \wedge hasTimeGranularity(e, morning) \rightarrow opengrocerycupboard(e)$
SR2	**1.9** $SimpleADL(a) \wedge hasEvent(a, opengrocerycupboard) \wedge hasEvent(a, OpenFridge) \wedge hasEvent(a, CookMicrowave) \rightarrow preparebreakfast(a)$
SR3	**0.8** $SimpleADL(a) \wedge hasEvent(a, toiletflush) \wedge hasEvent(a, OpenFridge) \wedge hasEvent(a, CookMicrowave) \rightarrow preparebreakfast(a)$
SR4	**0.1** $SimpleADL(a) \wedge hasEvent(a, toiletflush) \wedge hasEvent(a, openbathroomdoor) \wedge hasEvent(a, CookMicrowave) \rightarrow preparebreakfast(a)$

Algorithm 2: Segmentation over sensor data to generate event sequences.

```
1  Procedure Segmenting Sensor data(Sensordata)
     Data: Sensordata is a continuous sequence of event 'E'
     Data: sp.spatio is a table to maintain event sequences specific
           to every location, tp is duration of an event sequence
     Result: an event sequence 'A'
2  Intialise sp.spatio, tp as null
3  foreach event E_i in Sensor data do
4      tp = tp + (E_i.time - E_{i-1}.time)
5      if E_{i-1}.spatio == E_i.spatio then
6          Append E_i into sp.spatio
7          if tp ≥ threshold then
8              A ← sp.spatio
9              return A
10         end
11     end
12     if E_{i-1}.spatio != E_i.spatio then
13         A ← sp.spatio
14         return A
15     else
16     end
17 end
```

segmentation to produce event sequences and transfers to fixed time interval based, if the duration spent by the occupant in a location is below a minimum threshold or above a maximum threshold. The event sequences produced are then given to the recognition system to identify the activities of the occupant.

Input to the Algorithm 2 is 'sensor data', a sequence of events say *E1, E2, E3....* A data structure 'sp.spatio' is maintained for each spatial location that stores the event sequences generated with respect to a particular location. In segmentation process, the duration of the event sequence is updated in variable 'tp'. Both 'tp' and 'sp.spatio' are initialized to NULL in Line 2 of the Algorithm 2. Line 3 to Line 17 of the Algorithm 2, carries out the process of segmentation where the steps are repeated for each event in the event sequences. Line 4, updates the duration 'tp', the duration between two consecutive events as $tp = tp + (E_i.time - E_{i-1}.time)$. Line 5 ensures that two consecutive events have same spatial location, if so it is appended into the corresponding spatial table in line 6. The fixed duration threshold of a spatial location is compared with the duration 'tp' in Line 7 to ensure that the recognition process is not postponed beyond the threshold of the spatial location. This threshold used in the experimentation fluctuates according to the spatial location and in practice is set to the minimum average duration of an activity of a particular spatial location. The threshold thus fixed ensures that complete information regarding event sequences are obtained for an activity. In case, duration 'tp' exceeds the threshold, the Algorithm 2 as shown in line 8 returns the event sequence presented in spatial table instantly for activity recognition. If in-case, two events do not belong to identical spatial location as in line 12 of Algorithm 2, returns the event sequences available in spatio table for recognition. And, hence Algorithm 2 follows a hybrid of location and fixed interval for segmentation process.

3.6. Inference using MLN activity model generated from ontology activity model

Probabilistic inferences for the event sequences are performed by a Maximum-A-Posterior (MAP) query over MLN to predict the most probable activity of the occupant and explained in Section 3.3. The event sequences generated by segmentation modules are employed to position the variables in first order rules of MLN activity model [36]. Simple activity 'SA' recognized using MLN obtains the time interval from the time instant of its atomic events 'E' as shown in the Fig. 3. Later, composite activities are inferred from simple activities using Allen's temporal rules modeled within MLN using SWRL rules of ontology as explained in Section 3.1.1. The recognized simple and composite activity is given to the decision making system for response activation. The response could be a simple alert, or a call to caretaker or voice alert message, based on the smart home environmental set up.

Table 3
Description of simple ADLs.

Activity label	Activity description
ADL 1	Fill medication dispenser
ADL 2	Wash DVD
ADL 3	Water plants
ADL 4	Answer the phone
ADL 5	Prepare birthday card
ADL 6	Prepare soup
ADL 7	Clean
ADL 8	Choose outfit

4. Experimental analysis

The objectives of the experimental analysis of the proposed probabilistic ontology based activity modeling and recognition system is to evaluate the proposed probabilistic ontology model for activity recognition in comparison with existing approaches.

4.1. Dataset description

The dataset for the experimental analysis is derived from WSU CASAS smart home project [39]. The dataset represents activities of 21 occupants performing eight different simple ADLs individually and in addition the dataset also provides information of occupants executing activities in sequential, interleaved and concurrent manner. Details regarding wireless sensor networks and type of sensors related to the dataset is presented in this work [39]. The description of eight simple ADLs considered for modeling the proposed system is given in the Table 3. Moreover, the dataset contains event description of ADLs executed in simple and interwoven manner making it suitable to model the proposed simple and composite activity recognition system.

4.2. Experimental setup

Experiments were conducted with WSU CASAS smart home dataset, where the dataset is divided into training and test dataset. The activity recognition system is modeled with the training dataset and evaluation is done with test dataset. 'Leave one day out' approach is preferred to split the given smart home dataset into training and test data as it facilitates cross validation of the recognition system [40]. WSU CASAS dataset considered here contains data of the occupant performing various ADLs on thirteen different days (spanned over a period of two and half months) and hence the experiment is subjected to a thirteen fold cross validation. The WSU CASAS experimental dataset contains equal number of instances (twenty one instances) for all eight ADLs and hence does not suffer from class imbalance problem. The proposed probabilistic ontology based activity recognition system is evaluated with three well known metrics *Precision*, *Recall* and *F-measure*.

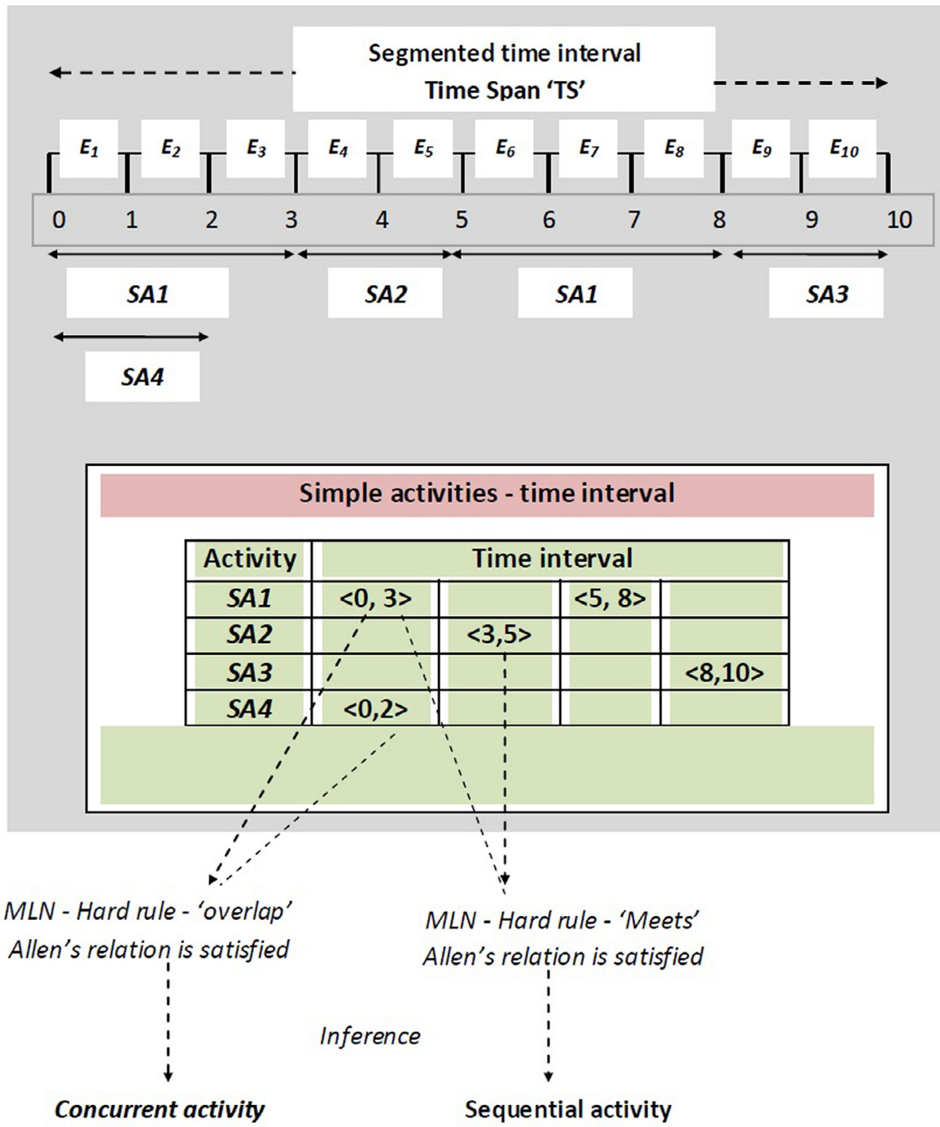

Fig. 3. Illustration of temporal reasoning in proposed system.

4.3. Implementation of ontology model

The ontology based activity model is constructed using Protege 4.3, an ontology editor [41]. *HermiT* reasoner [42] of Protege is preferred for reasoning over the modeled ontology because of its ability to perform consistency check and to make inferences on SWRL rules modeled within the system. The experiments were conducted with ontology based model to infer the ongoing activities. The ontology is constructed with various concepts and properties to design simple and composite activity model as described in Section 3.1.

4.4. Implementation of probabilistic reasoning using MLN

Transformation of description logic to first order logic is done as per Algorithm 1 and is implemented by using Incerto, a tool for probabilistic reasoning for semantic web that uses the model theoretic semantic property to convert description logic into corresponding first order logic [43]. Incerto further integrates Alchemy [44], an open source tool for MLN to perform weight learning and carry out probabilistic reasoning over first order logic.

The test dataset may possibly contain both simple and composite activities and thus demands appropriate segmentation of test dataset into comprehensive event sequences so as to carry out effectual activity recognition. Experiments were conducted on both fixed interval segmentation and proposed hybrid location and fixed interval based segmentation. The time slices for fixed interval segmentation should be maintained in a way that it is long enough to be discriminative and small enough to provide good accuracy and hence after various experiments time slice was fixed to 150 s for WSU CASAS dataset.

The proposed hybrid (location and fixed interval) segmentation as explained by the Algorithm 2 sets the threshold specific to spatial location. After analyzing the results of various experiments, threshold is set to a minimum of average duration of all activities performed in that location. For the WSU CASAS dataset, in the kitchen location there are two ADLs namely 'Prepare Soup' and 'Clean' and the average duration of these activities are 313 s and 198 s respectively. Hence the threshold for hybrid segmentation approach in kitchen location is set to 198 s. The Fig. 4 compares fixed interval based segmentation with the hybrid location and fixed interval based segmentation using proposed probabilistic ontology based approach. F-measures of fixed interval based seg-

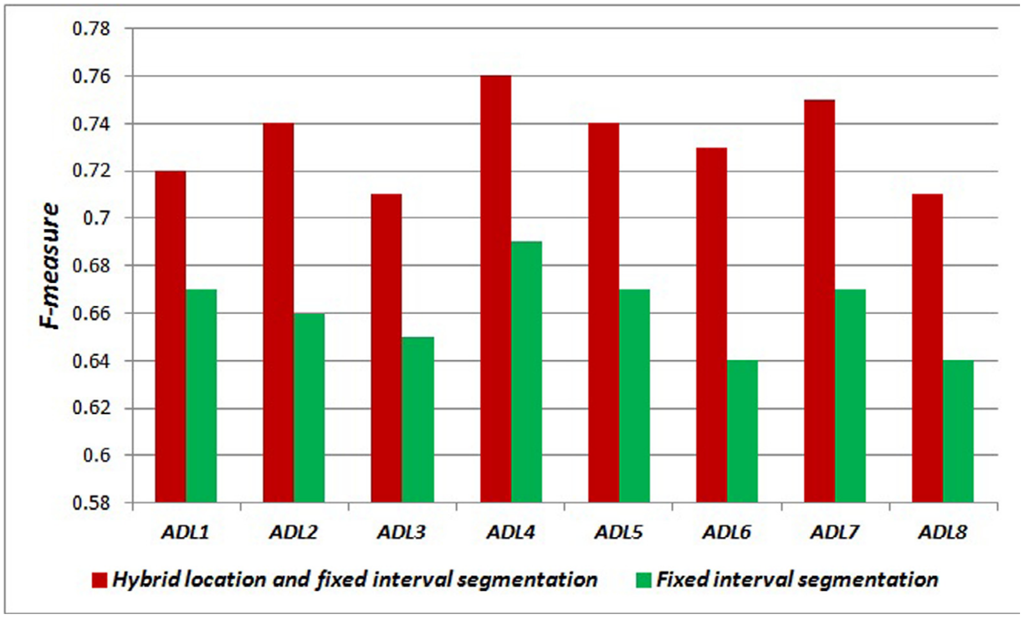

Fig. 4. Performance evaluation of the proposed hybrid (location and fixed interval) segmentation with existing fixed interval segmentation.

Table 4
Comparision of proposed system with EPAM.

	Existing (EPAM)			Proposed (Probabilistic ontology)		
	Precision	*Recall*	*F-measure*	*Precision*	*Recall*	*F-measure*
ADL1	0.79	0.65	0.72	0.98	0.92	0.95
ADL2	0.78	0.69	0.74	0.96	0.97	0.97
ADL3	0.75	0.67	0.71	0.94	0.93	0.94
ADL4	0.73	0.78	0.76	0.93	0.97	0.95
ADL5	0.78	0.69	0.74	0.97	0.98	0.98
ADL6	0.72	0.74	0.73	0.95	0.98	0.97
ADL7	0.78	0.72	0.75	0.98	0.94	0.96
ADL8	0.73	0.69	0.71	0.94	0.96	0.95
sequential	–	–	–	0.95	0.93	0.94
concurrent	–	–	–	0.91	0.94	0.93
interleaved	–	–	–	0.93	0.91	0.92

mentation are low for the reason that incompleteness in generated event sequences in most cases is incredibly high. Hybrid approach of segmentation in most cases is location specific and in certain cases is fixed interval therefore the likelihood of attaining complete event sequences are high in this approach rather than fixed interval based approach. Since complete event sequences are generated by hybrid location and fixed interval based segmentation, the experiments with proposed probabilistic ontology based activity modeling employs this approach to segment the test dataset into event sequences.

4.5. Experimental results

The effectiveness of the proposed probabilistic ontology based activity recognition system in classifying the simple and composite activity is compared with various existing approaches and its details are presented in this section.

4.5.1. Performance analysis of proposed system and comparison with Ontology based model

The focus of Event Pattern Activity Modeling (EPAM) [27] framework is to model ontology for simple activity recognition by extracting patterns from the dataset. From the experimental results as shown in Table 4, it is observed that proposed approach outperforms EPAM in terms of F-measure. The uncertainty in sensor data is not modeled in EPAM because of which its F-measures are low for all ADL activities. The proposed probabilistic ontology based on MLN is capable of carrying out probabilistic inference to recognize an activity even with uncertain data and produces a most probable activity for the given uncertain event sequence. The last three lines in the Table 4 indicates that composite activities are not modeled in EPAM and considered in the proposed system through modeling Allen's temporal relations using SWRL rules because of which the F-measure is high.

4.5.2. Performance analysis of proposed system and comparison with existing classification approaches

Artificial Neural Network (ANN) [45], Support Vector Machine [46], Bayesian Network (BN) [47] and Hidden Markov Model [48] are the commonly used statistical approaches to classify the sensor events to recognize the ongoing activity of the occupant. Weka [49], open source software for data mining was used in modeling the existing systems.

The Fig. 5 compares the F-measures of all of these approaches with that of the proposed system. The F-measures are low for ANN as it has limitation in performing inferences over the incomplete and uncertain inputs. Moreover ANN does not have any provision to incorporate domain knowledge because of which com-

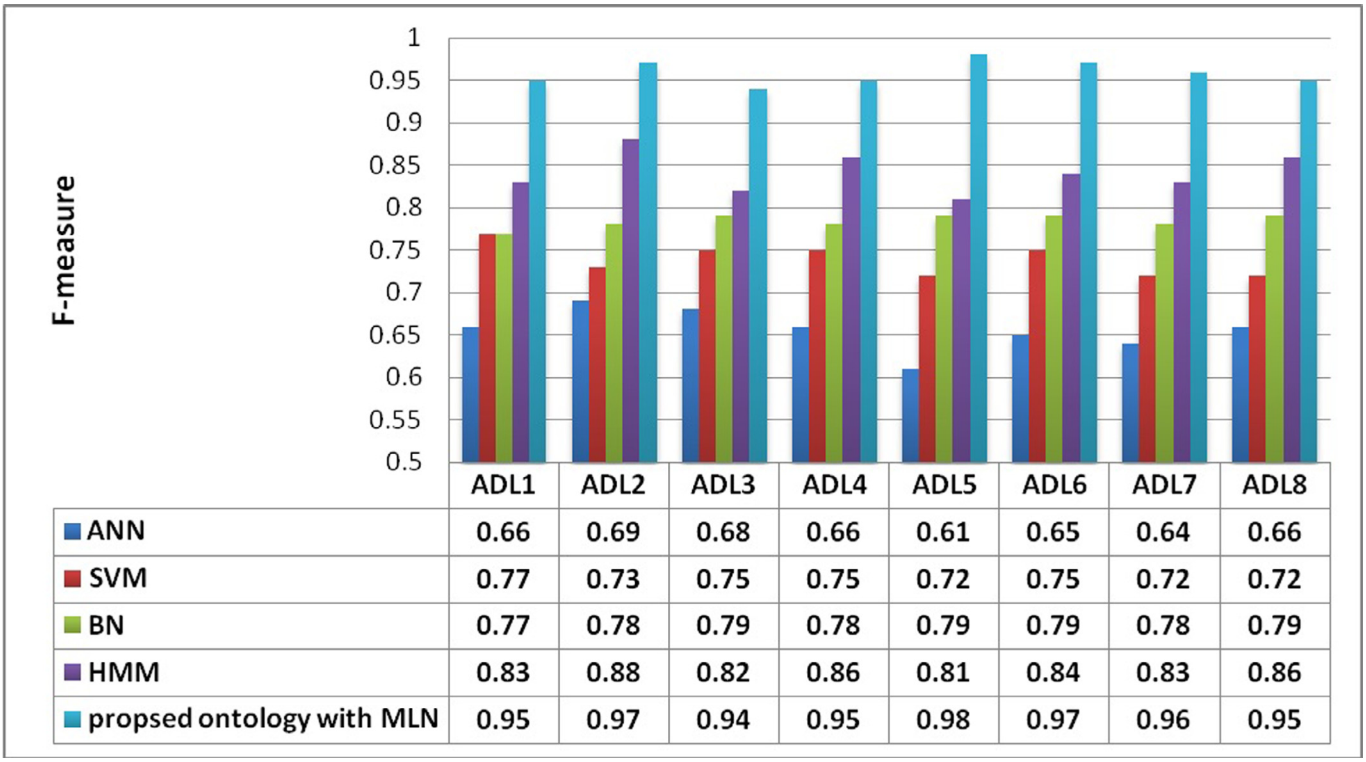

Fig. 5. Performance evaluation of the proposed system with existing statistical approaches.

posite activities could not be modeled. Whereas, SVM is a classification model that is built using marginal vectors and the kernel function considered that ably models the high dimensional feature space and hence its F-measures are high compared to ANN. Both ANN and SVM do not handle uncertainty and incomplete data and hence a probabilistic statistical approach is required to pay attention to uncertain data. BN handles uncertainty through probabilistic reasoning and executes a MAP query to make inferences on ongoing activity of the occupant. The limitation of BN approach is that temporal data and context level modeling are not well addressed. HMM, on the other hand handles uncertainty, temporal modeling much better than BN but domain knowledge modeling is difficult.

From the Fig. 5, it is observed that the F-measure of the proposed approach is better than the existing approaches. The reasons for high F-measure are that the proposed approach provides an excellent framework to integrate probabilistic reasoning into ontology approach through MLN, integrates rich temporal context within domain knowledge modeling, and constructs MLN through weight learning of first order rules (ontology) that reflects the truth both in the data and domain knowledge.

4.6. Performance analysis of proposed system and comparison with UnBBayes

The proposed approach of integrating ontology with probabilistic reasoning is compared with UnBBayes [23]. The results of comparison is shown in the Fig. 6 and it is understood that though UnBBayes approach integrates probabilistic reasoning over ontology, its F-measures are low when compared to MLN based probabilistic integration with ontology. This is because; the MLN learns weights from the individuals of ontology through an efficient weight learning algorithm that models the training data at the best possible level. Moreover, the proposed approach modes temporal data much better than UnBBayes.

Table 5
Comparison of F-measures of proposed approach with existing probabilistic ontology activity recognition system.

	Probabilistic ontology AR	
	Exisiting	*Proposed*
ADL1	0.85	0.95
ADL2	0.81	0.97
ADL3	0.72	0.94
ADL4	0.72	0.95
ADL5	0.81	0.98
ADL6	0.88	0.97
ADL7	0.57	0.96
ADL8	0.88	0.95

4.7. Performance analysis of proposed system with existing probabilistic ontology activity recognition framework

The comparison of the proposed approach with a probabilistic ontology activity recognition framework [21] is presented in this subsection. The framework formulates initial hypotheses about simple activities by matching object contextual information of observed events with an activity. These hypotheses are eventually refined with probabilistic reasoning over semantic constraints derived from the ontology. The comparative F-measure value in Table 5 highlights the ability of the proposed system in recognizing the diverse composite event patterns. This is accomplished in the proposed design through the integration of weight learning of MLN, mapping of contextual attributes (object, spatial location, time instant) of an event collectively to an activity, modeling Allens temporal relation and hybrid segmentation of event sequences.

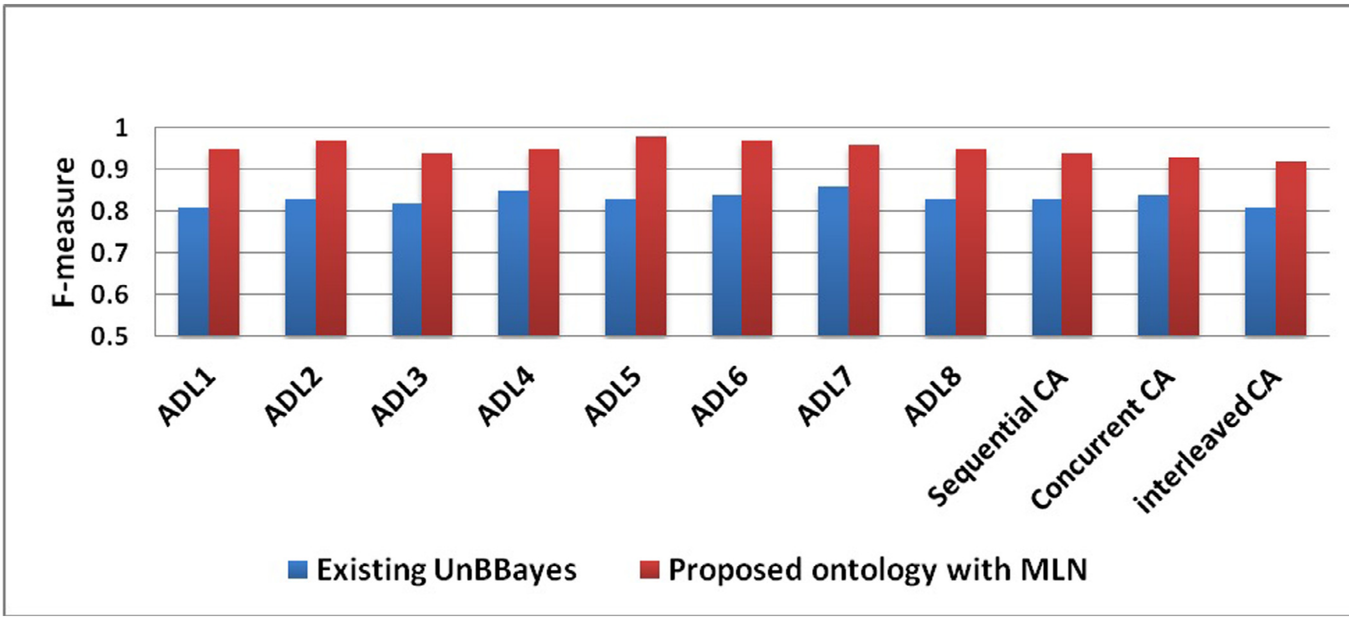

Fig. 6. Performance evaluation of the proposed system with existing UnBBayes approaches.

5. Conclusion

Activities of Daily Living (ADL) in a smart home were modeled using a hybrid approach, integrating data and the knowledge driven approach for activity recognition. Novelty is introduced in the design of the proposed activity recognition model through the incorporation of ontology and Markov Logic Network (MLN) thereby leveraging the strengths of ontological modeling and probabilistic reasoning. The domain knowledge is represented through ontology based activity model that represents simple and composite activities. The constraint in handling uncertainty in an ontology based model is resolved through a competent Markov Logic Network that provides probabilistic reasoning. MLN's are represented by means of weighted first order rules and thus the ontology activity model is converted to its equivalent first order logic facilitated through model semantic property of description logic. The proposed hybrid segmentation approach can automatically segment continuous sensor events into discrete sequences ensuring comprehensiveness of input data fed to the AR system. The experimental analysis of the proposed system shows improvements in the accuracy of activity recognition. Though the proposed system is advantageous in terms of providing additional knowledge for reasoning, there is a need for further exploration of knowledge engineering to enhance performance. The future work would be to extend the activity recognition system to characterize the occupant behavior from similar temporal activities, and to model multiple occupants in a single home and multiple smart home scenarios too.

References

[1] M.R. Alam, M.B.I. Reaz, M.A.M. Ali, A review of smart homes - past, present, and future, IEEE Trans. Syst., Man, Cybern., Part C 42 (6) (2012) 1190–1203.
[2] J. Augusto, H. Nakashima, H. Aghajan, Ambient intelligence and smart environments: A state of the art, in: Handbook of Ambient Intelligence and Smart Environments, 2010, pp. 3–31, doi:10.1007/978-0-387-93808-0_1.
[3] D.J. Cook, J.C. Augusto, V.R. Jakkula, Ambient intelligence: technologies, applications, and opportunities, Pervasive Mob. Comput. 5 (4) (2009) 277–298.
[4] U. Cortés, C. Urdiales, R. Annicchiaric, Intelligent healthcare managing: An assistive technology approach, in: IWANN'07, 2007, pp. 1045–1051.
[5] N. Graham, W. J, Understanding Alzheimer's Disease and Other Dementias, British Medical Association, Family Doctor Books, Poole, Dorset, U.K, 2009.
[6] J.C. Augusto, M. Huch, A. Kameas, J. Maitland, Handbook of ambient assisted living - technology for healthcare, rehabilitation and well-being, Handbook of Ambient Assisted Living, 11, IOS press, 2012.
[7] S. P, G.K. S., An unsupervised pattern clustering approach for identifying abnormal user behaviors in smart homes, Int. J. Comput. Sci. Netw. 2 (2) (2013).
[8] A. Lotfi, C.S. Langensiepen, S.M. Mahmoud, M.J. Akhlaghinia, Smart homes for the elderly dementia sufferers: identification and prediction of abnormal behaviour, J. Ambient Intell. Hum. Comput. 3 (3) (2012) 205–218.
[9] A. Aztiria, A. Izaguirre, J.C. Augusto, Learning patterns in ambient intelligence environments: a survey, Artif. Intell. Rev. 34 (1) (2010) 35–51.
[10] D.J. Cook, Learning setting-generalized activity models for smart spaces, IEEE Intell. Syst. 27 (1) (2012) 32–38. http://doi.ieeecomputersociety.org/10.1109/MIS.2010.112.
[11] D.J. Cook, S.K. Das, How smart are our environments? an updated look at the state of the art, Pervasive Mob. Comput. 3 (2) (2007) 53–73, doi:10.1016/j.pmcj.2006.12.001.
[12] K.S. Gayathri, S. Elias, S. Shivashankar, Composite activity recognition in smart homes using markov logic network, in: Proceedings of the 2014 IEEE 11th Intl Conf on Ubiquitous Intelligence and Computing and 2014 IEEE 11th Intl Conf on Autonomic and Trusted Computing and 2014 IEEE 14th Intl Conf on Scalable Computing and Communications and Its Associated Workshops (UIC-ATC-ScalCom), in: UIC-ATC-SCALCOM '14, IEEE Computer Society, Washington, DC, USA, 2014, pp. 880–887, doi:10.1109/UIC-ATC-ScalCom.2014.98.
[13] L. Chen, C.D. Nugent, H. Wang, A knowledge-driven approach to activity recognition in smart homes, IEEE Trans. Knowl. Data Eng. 24 (6) (2012) 961–974.
[14] A.C. Tran, S. Marsland, J. Dietrich, H.W. Guesgen, P. Lyons, Use cases for abnormal behaviour detection in smart homes., in: Y. Lee, Z.Z. Bien, M. Mokhtari, J.T. Kim, M. Park, J. Kim, H. Lee, I. Khalil (Eds.), ICOST, Lecture Notes in Computer Science, vol. 6159, Springer, 2010, pp. 144–151.
[15] L. Chen, J. Hoey, C.D. Nugent, D.J. Cook, Z. Yu, Sensor-based activity recognition, IEEE Trans. Syst., Man, Cybern., Part C 42 (6) (2012) 790–808.
[16] L. Atallah, G.-Z. Yang, The use of pervasive sensing for behaviour profiling - a survey., Pervasive Mob. Comput. 5 (5) (2009) 447–464.
[17] P. Rashidi, D.J. Cook, L.B. Holder, M. Schmitter-Edgecombe, Discovering activities to recognize and track in a smart environment, IEEE Trans. Knowl. Data Eng. 23 (4) (2011) 527–539. http://doi.ieeecomputersociety.org/10.1109/TKDE.2010.148.
[18] P. Rashidi, D.J. Cook, Keeping the resident in the loop: adapting the smart home to the user., IEEE Trans. Syst., Man, Cybern., Part A 39 (5) (2009) 949–959.
[19] N.D. Rodríguez, M.P. Cuéllar, J. Lilius, M.D. Calvo-Flores, A survey on ontologies for human behavior recognition, ACM Comput. Surv. 46 (4) (2014) 43:1–43:33, doi:10.1145/2523819.
[20] D. Riboni, C. Bettini, Cosar: hybrid reasoning for context-aware activity recognition., Pers. Ubiquitous Comput. 15 (3) (2011) 271–289.
[21] D. Riboni, T. Sztyler, G. Civitarese, H. Stuckenschmidt, Unsupervised recognition of interleaved activities of daily living through ontological and probabilistic reasoning, in: Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing, in: UbiComp '16, ACM, New York, NY, USA, 2016, pp. 1–12, doi:10.1145/2971648.2971691.
[22] G. Okeyo, L. Chen, H. Wang, R. Sterritt, A hybrid ontological and temporal approach for composite activity modelling, in: TrustCom, 2012, pp. 1763–1770.
[23] S. Matsumoto, R.N. Carvalho, M. Ladeira, P.C.G. da Costa, L.L. Santos, D. Silva, M. Onishi, E. Machado, K. Cai, in: Unbbayes: A Java Framework for Probabilistic Models in AI, Java in Academia and Research, 2011, p. 34.

[24] M. Richardson, P. Domingos, Markov logic networks, Mach. Learn. 62 (1–2) (2006) 107–136, doi:10.1007/s10994-006-5833-1.
[25] K. Gayathri, S. Elias, B. Ravindran, Hierarchical activity recognition for dementia care using markov logic network, Pers. Ubiquitous Comput. (2014) 1–15, doi:10.1007/s00779-014-0827-7.
[26] P. Chahuara, A. Fleury, F. Portet, M. Vacher, Using markov logic network for on-line activity recognition from non-visual home automation sensors, in: Ambient Intelligence - Third International Joint Conference, AmI 2012, Pisa, Italy, November 13–15, 2012. Proceedings, 2012, pp. 177–192, doi:10.1007/978-3-642-34898-3_12.
[27] K.S. Gayathri, S. Elias, S. Shivashankar, An ontology and pattern clustering approach for activity recognition in smart environments, in: Advances in Intelligent Systems and Computing, 258, Springer, 2014, pp. 833–843.
[28] F. Baader, I. Horrocks, U. Sattler, Description logics, in: Handbook of Knowledge Representation, 2008, pp. 135–179, doi:10.1016/S1574-6526(07)03003-9.
[29] [link]. URL http://www.w3.org/TR/owl-features/.
[30] [link]. URL https://http://everywarelab.di.unimi.it/palspot.
[31] [link]. URL http://www.w3.org/TR/owl-time/.
[32] [link]. URL http://www.w3.org/Submission/SWRL/.
[33] V.R. Jakkula, D.J. Cook, Learning temporal relations in smart home data, in: Proceedings of the second International Conference on Technology and Aging. Canada, 2007.
[34] G. Okeyo, L. Chen, H. Wang, Combining ontological and temporal formalisms for composite activity modelling and recognition in smart homes, Future Gener. Comput. Syst. 39 (0) (2014) 29–43. http://dx.doi.org/10.1016/j.future.2014.02.014. Special Issue on Ubiquitous Computing and Future Communication Systems
[35] P. Chahuara, A. Fleury, F. Portet, M. Vacher, Using markov logic network for on-line activity recognition from non-visual home automation sensors., in: F. Paternò, B.E.R. de Ruyter, P. Markopoulos, C. Santoro, E. van Loenen, K. Luyten (Eds.), AmI, Lecture Notes in Computer Science, vol. 7683, Springer, 2012, pp. 177–192.
[36] [link]. URL http://guangchun.files.wordpress.com/2012/03/mlnprob1.pdf.
[37] P.C. de Oliveira, P.J. de Sousa Gomes, Probabilistic Reasoning in the Semantic Web Using Markov Logic, July 2009 (MSc Thesis).
[38] [link]. URL http://resources.mpi-inf.mpg.de/d5/teaching/ss12/sum/slides/06-markov-logic-handout.pdf.
[39] G. Singla, D. Cook, M. Schmitter-Edgecombe, Tracking activities in complex settings using smart environment technologies, Int. J. BioSci., Psychiatry Technol. 1 (1) (2009) 25–35.
[40] T. Van Kasteren, G. Englebienne, B.J. Kröse, Human activity recognition from wireless sensor network data: Benchmark and software, in: Activity recognition in pervasive intelligent environments, Atlantis Press, 2011, pp. 165–186.
[41] [link].URL http://protege.stanford.edu/.
[42] B. Glimm, I. Horrocks, B. Motik, G. Stoilos, Z. Wang, Hermit: an OWL 2 reasoner, J. Autom. Reasoning 53 (3) (2014) 245–269, doi:10.1007/s10817-014-9305-1.
[43] [link]. URL https://code.google.com/p/incerto/.
[44] [link]. URL http://alchemy.cs.washington.edu/.
[45] R. Begg, R. Hassan, Designing Smart Homes, Springer-Verlag, Berlin, Heidelberg, 2006, pp. 146–164.
[46] A. Fleury, M. Vacher, N. Noury, S. Member, Svm-based multimodal classification of activities of daily living in health smart homes: sensors, algorithms, and first experimental results, IEEE Trans. Inf. Technol. Biomed. (2010) 274–283.
[47] E. Nazerfard, D.J. Cook, Using bayesian networks for daily activity prediction., AAAI Workshop: Plan, Activity, and Intent Recognition, AAAI Workshops, WS-13-13, AAAI, 2013.
[48] G. Singla, D.J. Cook, M. Schmitter-Edgecombe, Recognizing independent and joint activities among multiple residents in smart environments, J. Ambient Intell. Human. Comput. 1 (1) (2010) 57–63.
[49] [link]. URL http://www.cs.waikato.ac.nz/ml/weka/.