www.s2is.org
International Journal on Smart Sensing and Intelligent Systems
ISSN 1178-5608

A VERSATILE LATTICE BASED MODEL FOR SITUATION RECOGNITION FROM DYNAMIC AMBIENT SENSORS

Sangeeta Mittal[1], Krishna Gopal[2] and S.L. Maskara[3]

[1]Department of Computer Science &Engineering, JaypeeInstitute of Information Technology, Noida, Uttar Pradesh - 201307, India

[2]Dean Academics and Research, JaypeeInstitute of Information Technology, Noida, Uttar Pradesh – 201307, India

[3]G-2W, SouraNiloy Housing Complex, 1 - KailashGhosh Road, Kolkata - 700008. India

Emails: smittal.150210@gmail.com[1], krishna.gopal@jiit.ac.in[2], maskara.shankar@gmail.com[3]

Corresponding Author: smittal.150210@gmail.com

Submitted: Oct. 9, 2012 ***Accepted: Feb. 10, 2013*** ***Published: Feb. 20, 2013***

Abstract-

Recent advances in acquisition, storage, and transmission of data from sensors in digital format has increased the need of tools to support users effectively in retrieving, understanding, and mining the information contained in such data.Extraction of domain specific actionable information like occurrence of one of the predefined "situations" is desirable. Major difficulties in achieving this extraction are 1) Source of Data, that is, number and type of sensors deployed is highly variable even for one type of application, 2)Availability of domain specific labeled training data is critical for computation of situations.In this paper, we propose a versatile method based on formal concept analysis to overcome these difficulties in modelingsensor based situations.Our method, making use of contexts as intermediate form of sensors data, works on any number and type of sensors. It is alsoinstance-independent and eliminates need of training, when applied to various instances of similar

application.For illustration, we model and perform real time recognition of activity of a person in indoor home environment with ambientsensors.The embedded sensors captureusage and proximity of human beings to objects.We apply the model learnt from one house, foractivity recognition of new persons across different new houses. The recognition results obtained have high precision and recall.

Index terms*:*Wireless Sensor Networks, Ambient Intelligence, Formal Concept Analysis, Situation Modeling, Activity Recognition, Lattice based Classification.

I. INTRODUCTION

Automated Situation Awareness can provide lot of proactive societal applications. It can enable monitoring and tracking entity of interest in fields like healthcare, activity, climate, and border surveillance. Application specific sensing using relevant physical wireless sensors are a cost effective way of accumulating fine-grained dataabout the components that describe situation of these environments. The major computational challenge iscorrect &real time extraction of situations or any other useful form from this data.Towards development of application independent methods to transform sensor data to situations, we define an intermediate form "current contexts". The overlapping set of contexts then defines situations and provides situation aware services. The humane application that has motivated this work is proactive care for theaging.Here the situation of interest is "activity" being carried by the monitored person. In case of patient, on situ "health condition"using appropriate sensors can be monitored. Medical professionals opine that one of the best ways to detect emerging medical conditions before they become critical is to look for changes in the activities of daily living (ADLs) [1]. These activities include timely food consumption, regular personal hygiene (toileting & bathing), medicines intake, and proper sleep. Development of computational systems that recognize such activities can enableremote automatic detection of changes in patterns of behavior of people at home that indicate decline in health[2].Wireless Sensors are highly useful in enabling round the clock monitoring of a person. Unlike conventional methods of monitoring like CCTV cameras, microphones, or wearable devices, use of sensors is non-obtrusive and does not compromise privacy of the person being monitored. Video and audio output from conventional monitoring devices require time consuming & complex preprocessing methods [3].On the other hand, the data emanating from sensors is simpler, very detailed, lightweight, voluminous but easier to preprocess. This makes sensors embedded in ambient surroundings a convenient way of monitoring activities of a person.Activity monitoring requires online activity prediction from a learnt model. Prediction itself requires real time abstractions out of huge volumes of data.When number and type of source sensors is large, the extraction of activities is done in hierarchical manner. For example, a person X's raw individual proximity sensors data, first contextual information like location "in kitchen", "using microwave and freezer," accessing "utensils and cupboards," can be extracted. This high-level information can be mapped to activity prediction"preparing breakfast". In other applications like health condition monitoring,definition of situations and contexts in terms of individual sensors data is done with help of domain expert. The set of current contexts extracted from sensor firings partially/completely aid activity recognition. In this work, activity recognition from multiple raw wireless sensors data is dealt. We, particularly, address the following:

1) Design a framework to define relationship between sensor data, contextual information, and comprehension of situations.
2) Develop a method to map variable dimensional sensor data to fixed dimension contexts.
3) Develop a method to modelsituations in terms of Contextual Information
4) Demonstrate use of designed framework to infer situations from new available Contextual Information.
5) Transfer the learnt model to any new application instance without any learning phase.

Recognition of situations in general from sensor data is complex and domain specific. Task-based recognition, template based matching, event trees and fuzzy inference rules [4-6] have been used in literature for this purpose. Most of these methods suffer from lack of generality, when transferred from one instance of an application to another.A general method should be transferrable to any new instance of same application in possibly different settings.To overcome the disadvantages of conventional methods, use of Formal Concept Analysis(FCA) has been made to define activities. Formal Concept Analysis (FCA) is an efficient method for data analysis, knowledge representation and information retrieval [7]. FCA requires structuring of data as 'Concepts'. Formally, concepts are a pair of a set of objects (extent) sharing a set of attributes (intent). The concept lattice organises the whole set of concepts as partial ordered sets. The concept lattice is used to visualize the conceptual structure and access it for finding patterns, regularities and exceptions in data. Use of Concept lattice for activityrecognition in presence of contexts obtained from noisy sensors data is considered appropriate for many reasons. Due to inherent uncertainty in wireless transmitted data, some context elements may not be present for activity recognition. The implication set of Concept lattice is useful to predict such contexts. In this paper we make use of publicly available datasets ofsensor dataannotated with some common activities of daily living of threedifferent personsliving in three differenthomes[8]. The evaluation using standard methods demonstrate ability of activity prediction of person at any time instance.The methods also represent transferability of the learnt model to an entirely new instance by applying the model on a different house with different person.

The paperhas been organized into six sections. Section IIelaborates on significance of situation assessment from sensor data and related methods. General introduction to the Formal Concept Analysis (FCA) is followed by inference of situations using FCAin section III.Section IV presents theexample sensor datasetsof human activity, design of contextual information andmapping of activities. The experimental results and evaluation of proposed methods in predicting activities are discussed in section V. The paper has been concluded in section VI.

II. SITUATION ASSESSMENT FROM SENSOR DATA

Sensors enable smart environments by providing highly detailed and frequent monitoring/ tracking information. Theraw sensor data may not provide this high-level information. It is critical to do this abstraction in real time to do any required actuation like providing medical help and controlling equipment etc. The situation of interest here is some common activities of a user living in a sensor-enabled house. In this section, we elaborate on the significance and requirements of activity extraction from sensor data. We also review some methods proposed by researchers in this direction.Application of various artificial intelligence methods can help in extraction of desired useful information.

a. Significance of Activity Assessment from Dynamic Sensors' Data

Sensors embedded in objects used in the activity and at various indoor locations can provide data about current activity taking place. Besides this, time of the day, schedule of the day, universal or known facts serve as logical sensors to provide data inputs. We give examples of some such sensors in Table 1.Data from these sensors when combined and organized in time becomes multidimensional instantaneous data for current activity. The number of these dimensions will change with the number of sensors providing data. On the other hand,the number of activities of interest is common and static. To deal with heterogeneity in number and type of sensors, it is convenient to derive an intermediate form like context from raw data. By "*Context,* "herewe mean any common information about user himself or his environment that may be useful in determining the current activity of the person. Individual contexts themselves defined from variable number of related sensors. Thus, "*Contextual Information"*(CI) information about activity is obtained in terms of number of contexts. TheCI thus obtained is lesser in size and more concise in usefulness. As the individual CI is trivial and transient, it is abstracted at a higher level to get situation information. In this work, we abstract, number of contexts occurring within same periodtoactivity within that period.As we do these abstraction steps, the size of the data to be handled is reduced and usefulness towards automated monitoring is increased. In Figure 1, we represent the setup ofsensors placed in objects and nearby secondary computational device extracting CI and situations from raw sensor data transmitted to it.The model thus built is useful in answering activity related queries by any interested third party software/ person in end user devices like mobile phone and computers.

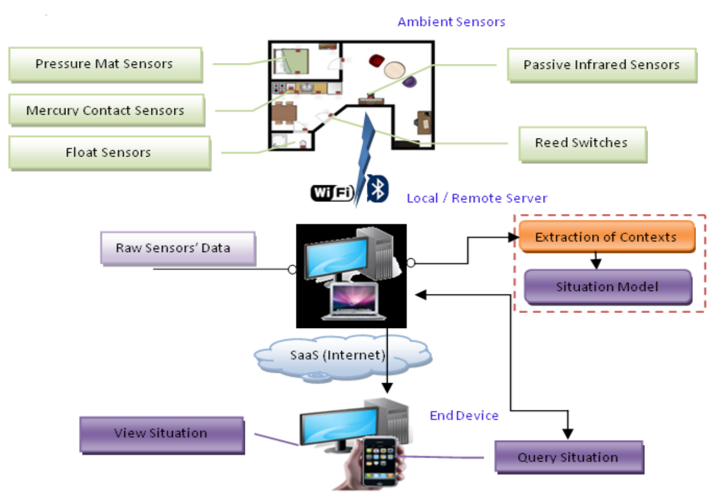


Figure 1 Setup of Ambient Sensor Based Data Collection, Processing and Distribution

For sake of commercialization, the situation model can also be provided on Software-as-a-service (SaaS) model and feedback and querying can be enabled.We show few other common types of sensors useful for abstraction of activities in Table 1.Due to uncertainties associated with wireless transmitted data from resource-constrained sensors, it is not straightforward to do conjunction to obtain context. Various probabilistic methods like Fuzzy Logic, Probabilistic Logic, Bayesian networks, Hidden Markov models, Dempster- Schafer theory of evidence, Rule Based Reasoning and Ontological Reasoning have been used to combine sensor data as relevant contexts [9][10].Categories of common contextual information useful in determining activity of a person may be Physical, Physiological & Environmental Knowledge.InTable 2, we describe, few relevant contexts abstracted from sensors of Table 1.In this work, we focus on abstraction of CI to activities. The set of common contexts are obtained from sensor data using one of the methods above or simply by majority voting. Theissue in designing methods for conversion of these contexts to activities is more of accuracy than of uncertainty and heterogeneity. Domain Knowledge is required for defining activities and characterizing contexts.One needs to answer questions like which activities are we interested in identifying and their characteristics in terms of contextual information. Activity determination is guided by set of rules that describe what contexts are present within a certain time frame, their order, and contexts that must negate.

Table 1: Various Sensors for Human Activity Recognition

Physical Environment Sensors	Physiological Sensors	Logical Sensors
Temperature	Body Temperature	Time of the Day
Humidity	Heart Beat Rate	Date
Pressure	Blood Pressure	Scheduled Event / Activity
Ambient Light	Accelerometers	Diary Entries
Ambient Sound	ECG	Age
Location	EEG	Gender
CO2 Sensors	RFId Tags for Identity	Information of co-residents
Camera	Passive Infrared Sensors	Visitor Information, if any
Microphone	Orientation	Any other habit related information

Table 2: Contexts and Respective Sensors

Contexts	Possible Context Values	Sensors Required
Spatial *(Where)*	Location: Office, Kitchen, Bedroom, Market/ Outdoor etc.	GPS, CO2, RFID, Light
Physical *(What)*	Ambient Climate, Light, Sound, Inner Health	Temp, Humidity, Pressure , light sensors, Microphone, SPO2,Blood Glucose
Temporal *(When)*	Current Time	Clock, Time of the Day
Identity *(Who)*	Person, Object	RFId Tags, State Change Sensors , Proximity Sensors
Physiological Context *(How)*	Body position	Orientation Sensors, Accelerometers
Logical Context*(Why)*	Habits, scheduled events,	Offline entries in temporal databases

The present work doesn't focus on methods of abstraction of contexts from sensors but on the methods of activity extraction from contexts.For this domain, the types of contexts as described in first column of Table 2 are considered relevant. A human activity can be described easily by getting contextual knowledge about these contexts. Apart from these"Why" context may be answered by correlating more than one activity.Given CI, recognition of activity needs to be accurate and inreal time. Most methods as will be described in next section are evaluated on these parameters. Besides this, one of the major challenges that we address here is eliminating the need to have costly labelled training data for every instance of activity recognition in similar scenarios. Particularly, here we demonstrate the feasibility of our approach in doing activity recognition for any new target house, given that the algorithm has been learnt from a labelled source house.

b. Related Work

Past decade witnessed development of a wealth of algorithms for creation and maintenance of a Wireless Sensor Networks. However, the question now has started to arise is, "what next"? Wireless Sensor Networks generate data on unprecedented scale and a lot of research is going on in harnessing this massive data to benefit underlying applications and ultimately human beings. Humans' quest is normally for qualitative descriptions of what is happening in the monitored environment. For example in tracking and monitoring applications, abstract information like good / alarming health condition, safe/ dangerous surroundings, normal/ abnormal activity are of more interest. The sensors cannot directly produce these descriptions. Lot of research [10-27] is focusing on automatically determining in which of these or any other "situation" the entity being monitored is in. The sensors can provide simulated detailed sensing of real world environments. To get information about desired situations, most works[11-12][14][17] do a hierarchical conversion of raw sensor values to qualitative actionable knowledge. They employ different machine learning methods and their variations. Cardelland Liu in [11] propose a layered architecture of situation representation and recognition along with sensor management and communication protocols. Situations are represented as events and composite events. Finite state machines and Distributed Commitment machines are worked out as recognition algorithms.

Oh et al in [12] demonstrate an indoor context management framework to monitor home environment with multiple habitants. They refer the Who, When, Why, Where and How model of instantaneous surrounding description. This model is used as reference for constituent contexts of situations defined in this work. Gu et al. [2] formulate activity recognition as a pattern-based classification problem, and propose an Emerging-Pattern based approach to recognize both simple and complex activities in a unified framework. They propose a segmentation algorithm based on feature relevance to segment the boundary oftwo adjacent activities. They report good classification performance on sequential and interleaved activities. Dahlbomet al. in [13] extract situations from video sensor contexts, regular grammars are used to match online abnormal behavior of any occupant of building as suspicious situation. Juan Ye [14], described defining and using situation lattices for situation description and inference. We apply a similar approach for expressing ADLs in hierarchical manner. The proposed activity lattices organize the activities and full or partial related contexts. We extend the work significantly by addressing the portability of the method to any new target application instance. Van Kasteren et al in [8] addressed similar problem using Hierarchical Hidden Markov Model and Conditional Random Fields. Thirunarayan et al. in [15] explained synthesis of high- level, reliable information for situation awareness by querying low-level sensor data. They make use of prolog type reasoning to convey confidence in sensor data interpretation and eliminate inconsistencies in situation descriptions. An integrated actual implementation framework requires modules for converting sensors data to semantic web notations like in XML. For reasoning, the inference engine is able to implement the rules in XML syntax or the data to be reasoned is to be brought in the form of PROLOG. Srivastav et al. in [16] define a framework for abstraction of situations from object and situations' information. The relational dependencies among objects are modeled as cross-machines called relational Probabilistic Finite State Automata (PFSA) using the xD-Markov machine construction. These PFSAs are mapped to situations. Clusters of sensor data are identified to summarize an individual event, the macro-cluster then integrate the information from multiple events in [17]. To facilitate scalable, flexible, and online analysis, the atypical cube is constructed, and a guided clustering algorithm is proposed to retrieve significant clusters in an efficient manner. The algorithm is fast as compared to other baseline solutions and hence suitable for online analysis. Fogarty and Hudson in [18] describe a toolkit for addressing issues in developing and deploying sensor-based statistical models. It caters to human mediation problems in earlier popular toolkits by eliminating mediation or using it as feedback. They improve HCI of applications by showing that using the model the interruptibility to human is reduced greatly.

Other different works [19-22][24]of enabling pervasive computing in general have also studied situations and their description mechanisms. Yau et al. [19] analyzed the semantics of situations and gave them formal representations. They consider Context as any instantaneous, detectable, and relevant property of the environment, the system, or users. An atomic situation is composed of contexts in terms of context operators, including function, arithmetic or comparison operators, and time constraints. A composite situation is composed of atomic or other composite situations in terms of logical operators and time constraints. This helps application designers to specify situations using formal expressions and manipulate them. Costa et al. [20] studied the classification of situations in terms of their composition. A situation can be an intrinsic, formal, or relational context situation, derived from a single, doubleand multiple pieces of context respectively. Loke [21] proposed representation of situations by decoupling the inference procedures of reasoning about context and situations from the acquisition procedure of sensor readings from context-aware systems. They apply a logic programming approach to characterizing situations, which helps the system designer in naturally individualizing situations descriptions in an application. Thomson et al provide reusable library of situation specifications [22]. They expressed different levels of granularity of a situation through specification inheritance. New specifications created as variations of existing ones enable interpretation of same situation at different levels of abstraction.

Recognizing human activities from ambient and physiological sensors has attracted lot of research interest recently [23 -27]. One of the pioneer works in the area of activity recognition using ambient sensors was reported in [23]. The study was done by deploying large number of simple sensors in real environments. The work has become baseline for many further developments in this domain. Bicochhi et al.[24] describe situation recognition as classification problem, where input data source is a camera sensor and knowledge base of common sense of that field. The sensor data is segmented and classifications done on individual segments. The classified low-level labels are input to a domain specific "concept net". A shortest path algorithm is then used to recognize high-level situation. Experiments were done in classifying a) general types of location, (like road, square, park, shop, cinema, mall, restaurant, gym. (b) Domestic (kitchen, living room, bathroom, bedroom, garden) (c) working environments (meeting room, office, corridor, leisure room)(d) vehicles used (bike, car, bus, train). The database and concept net are user defined and hence may not be complete. Atallah et al. [25], do multi sensor fusion for health care monitoring. Distributed Inference is done using graphical models—that characterize the relationship between variables. They suggest three types of distributed inference at data, algorithm or decision level for activity prediction. Bayesian Framework for feature selection is used for feature reduction and outlier detection. Fischer et al [26] give a conceptual Framework for situation characterization, abstraction, recognition, and projection from sensor data. Snapshot of all sensor data at a point of time is considered as scene and over a period is considered as episode. Further, the data is quantitatively processed as an event. The quantitative relations are aggregated as qualitative relations, which are strongly connected to activities. Mahajan et al. [27] designed integrated systems with simple algorithms for events like noise, spike reporting instead of raw data. The decision about an event is taken within the sensor only. Before making a new physical sensor with these capabilities, a virtual sensor in a computer can be implemented to test the intelligence algorithms.

The method chosen here for analyzing and organizing situations is use of Formal Concept Analysis (FCA), which has a long history as a technique of data analysis conforming to the idea of Conceptual Knowledge Processing of symbolic data [28]. Before use in lattices the situations are represented as simple composition rules where each situation is specified to comprise of a set of contexts. FCA has already scaled well to web search within billions of web pages so these would be faster and scalable to retrieve situations. When an unprecedented situation occurs, probabilistic retrieval is possible using association rules; FCA can also help make the situation descriptions robust by indicating possible implications among attributes [29]. Human users can cross check descriptions using attribute exploration. The exploratory paradigm of FCA is better than decision trees as same tree may occur multiple times over the whole structure, but in concept lattice there is no duplication of information. The concept lattice designed as Activity Lattice will be useful fordiscovery, prediction and as a browsing space for activity retrieval. In next section, we introduce basic concepts of FCA and describe the construction of concept lattice using extracted contexts as features of situations.

III. CONSTRUCTION OF CONCEPT LATTICE OF SITUATIONS

In this section, the basics of concept lattices and its application to the organization of activities to study characteristics such as generalization and dependence is studied. The Concept data analysis methodrecognizes and generalizes structural similarities from data descriptions. It does not require any mathematical manipulations of probability distributions and still provide results even in uncertainty. In this section, the theoretical foundations of Formal Concept Analysis and their use for situation description are discussed.

a. Basic Definitions

Lattice is an algebraic structure represented as partial ordered set (poset) with binary joins and meet. Hasse Diagrams of Lattices represent elements in poset with relationship < =. One such diagram for a poset S with elements {a,s,d,f} , with relations a <s ; a<f ; s<d and s<f is shown in Figure 2 [7].

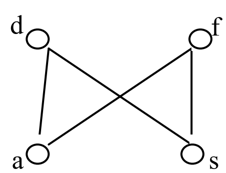


Figure 2 Line Diagram of Lattice of Set S

Some other terms useful in Formal Concept Analysis are defined as below:

Cross Table- A cross table is a triple k = <O,P,R>, where O is a finite set of elements called objects, P a finite set of elements called properties and R is a binary relation defined between O and P. The notations (g,m), or R(g,m)=1, mean that "formal object g verifies property m in relation R" . For example, cross table of 5 dummy objects with their 7 possible attributes is shown in Table 3, prepared in ConExp [30].

Formal Concept: A formal concept of a cross table<O,P,R> is a pair (A,B), where A ⊆ O, B ⊆P, such *f* (A) = B and *h* (B) = A. Sets A and B are called respectively the domain (extent) and range (intent) of the formal concept .

Table 3 Cross Table of 5 objects and their attributes

	Attr 1	Attr 2	Attr 3	Attr 4	Attr 5	Attr 6	Attr 7
Obj 1	X		X			X	X
Obj 2	X			X	X		X
Obj 3		X		X		X	X
Obj 4		X			X		X
Obj 5		X		X			

Extension: It consists of all objects belonging to the concept.

Intension: This set has all attributes common to all objects in extension.

Concept Lattice: From a cross table<O,P,R>, we can extract all possible formal concepts. The set of all concepts may be organized as a lattice, by defining the following partial order relation << between two concepts, (A1,B1) << (A2,B2) ⇔(A1 ⊆A2) and (B2 ⊆B1). The concepts (A1, B1) and (A2,B2) are called nodes in the lattice. The concept lattice for the cross table in Table 3 with total 15 concepts, is shown in figure 3.

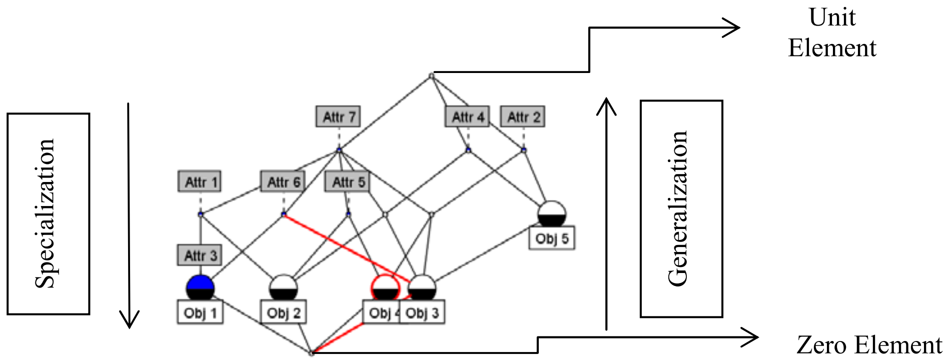


Figure 3Concept Lattice of Cross Table 3

The figure also represents the generic and specific concepts relations in the lattice. A concept lattice has two-way rootedness, with ‘top’ (unit element)&‘bottom’ (zero element) being two roots. Traversal towards top of the lattice corresponds to generalization and towards bottom specialization of retrieved results. Thus, it establishes a dual hierarchy between concepts. For example for dummy objects and attributes in Figure 3, creation of lattice converts the flat data into a more useful hierarchical one.Obj5 is a generalized form of rest of obj4 and obj3.Calculation of the concepts lattice from a cross table is a computation intensive task. Many algorithms have been proposed for generating the Galois lattice from a binary relation. Some algorithms generate complete lattice while others do it partially for large lattices. To improve efficiency of algorithms the options have been to either first generate only the concept list and then the diagram or both simultaneously. Few works, like one by Godin et al. [31] propose an incremental building of concepts lattices starting from given cross table. Such algorithm will be useful to obtain concepts lattices to our problem of research as user queries users can be inserted in the lattice and then it is possible to determine the most relevant situations given the deduced contexts. For implementation in software multi list, hashing or a hybrid method is employed[32]. The objects will replace activities and attributes will replacecontexts in the concept lattice we define in next sections.

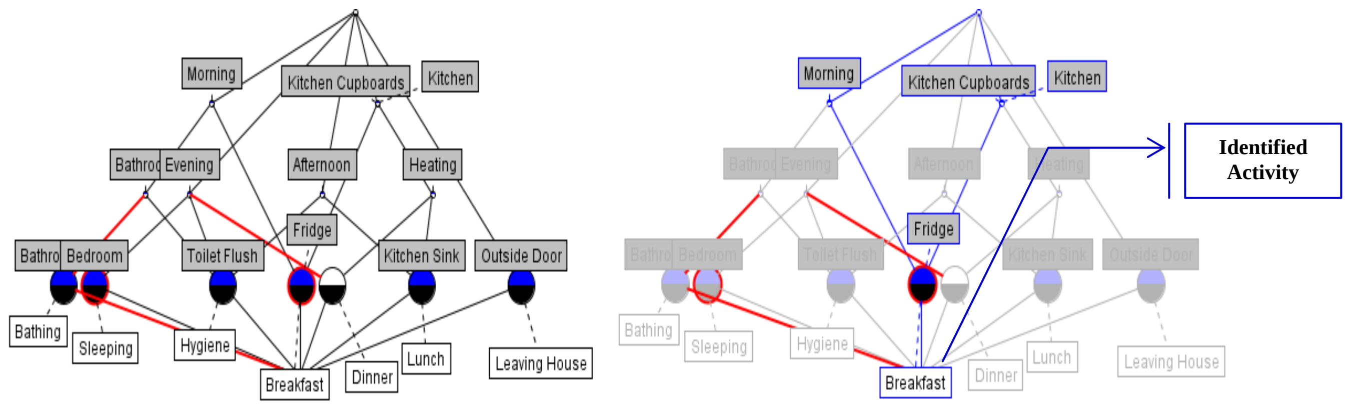
b. Lattice Navigation for Situation Recognition

The lattice can be used for recognizingsituationsby submitting to the lattice $Con_{present}$, the set of current context values. A simplified algorithm for doing so is described in figure 4.

```
Algorithm Predict_Situation
Inputs: Current Context Set: Con_present
      Concept Lattice: AL
Output: Possible Situations: Situs_R
Begin
Situs_R=Φ
M = all possible concepts of AL
For each (X,Y) ∈ M
      If(Y⊆{Con_present})
      Situs_R=X
If Situs_R=Φ
      Return "Unclassified"
End
```

Figure 4: Algorithm to navigate lattice foractivity query

In Figure 5(a), an example concept lattice as “Activity Lattice” of some common activities and related contexts is created. The activity lattice provides a visualization of semantic relationships between activities and contexts that are implicit in data. The lattice being a graphical structure is implemented as multi-list in memory. For traversal across all nodes as per algorithm in Figure 4, backtracked depth first search starting from top element is done. The lattice traversal is shown for given contexts set {Time: Morning, Using: fridge, kitchen cupboard, Location: Kitchen}.As shown in Figure 5(b) the activity breakfast is deduced from the lattice.



(a) (b)

Figure 5(a) Activity Lattice of some Activities(b) Lattice Navigation for Activity Recognition using present contexts

Using algorithm of Figure 4, only complete matching activities are inferred. In case, no intension completely matches the set of input contexts the input remains unclassified. Due to loss of data in wireless transmission, this can occur frequently in our domain.

IV. CONSTRUCTION OF ACTIVITY LATTICE FROM SENSOR DATA

The algorithms for construction and navigation of concept lattices as discussed in last section are applied to activity lattice. The information for creating cross table is unknown initially.

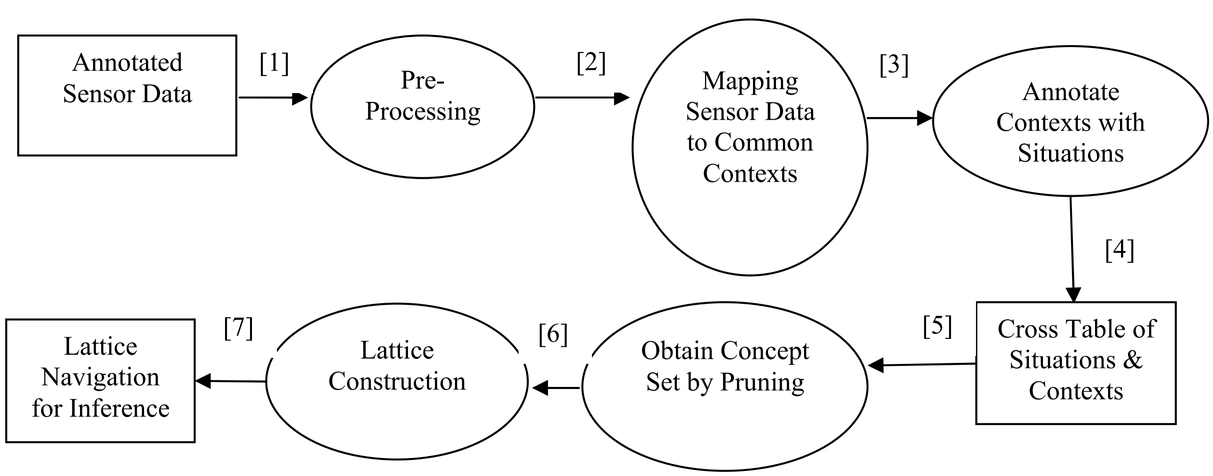


Figure 6 From Sensor Data to Situation Inference using Concept Lattice

Therefore, a training dataset is required which is used to create the cross table first and then the lattice of concepts from this table. The steps to do so are elaborated in figure 6. The construction of concept lattice here is done in steps shown in Figure 6 constructing a matrix of Situation-Context relations from sensor data followed by concepts and Hasse diagram. The resulting concept lattice is maintained as hashed list in memory or stored in a data file from where it may be later used for situation assessment. The datasets used are described in next section.

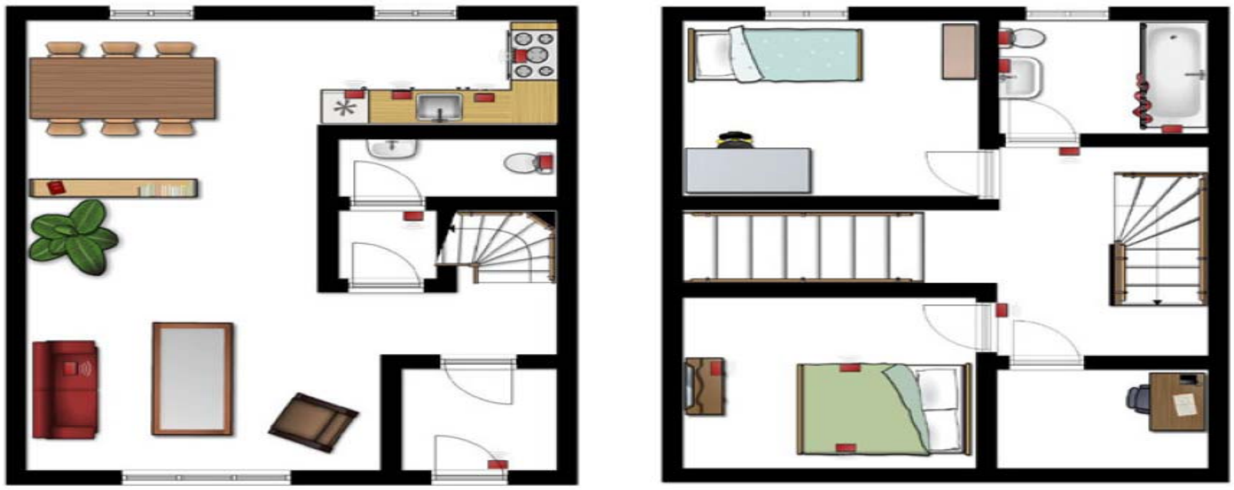
a. Annotated Sensor Datasets

The activity lattice is built to recognize some common ADLs of a single user in an indoor home environment. Homes and their furnishings have highly variable layouts, and individuals perform activities in many different ways. The same activity (e.g. brushing teeth) may result in a significantly different sensor activation profile based upon the habits, or routines of the home occupant and the layout and organization of the particular home. One approach to handling such variability is to use supervised learning with an explicit training phase.We have taken an example clustered third party real data set where the activities of interest are common and all other things are different in three houses occupied by single inhabitants. These common activities that were annotated by each inhabitant were “leave house”, “use toilet”, “take shower”, “go to bed”, “prepare breakfast”, “prepare dinner”, and “get drink”. These activities were chosen based on the Katz ADL index, a commonly used tool in healthcare to assess cognitive and physical capabilities of an elderly person [1].

(a) House A

(b) House B

Figure 7: Floor Plan of the house red rectangle boxes indicate sensor nodes in two houses [8]



(a) House C: First floor (b) House C: Second floor

Figure 8 The Floor Plan of House C is entirely different from other two houses

Times at which no activity is annotated is referred to as 'Idle'.The layouts of the houses as given in [8], are shown in Figure 7 and 8.All three houses are installed with different number and type of sensors, shown as red rectangles in the floor plan. Description of types of sensors, their place and purpose of installation is given in Table 4.

Table 4 Type, Location and Purpose of Sensors Used in Three Houses

Sensor Type	Location	Purpose
Passive Infra Red	Bathroom, Kitchen, Bedroom	Measures changes in infrared radiation and is generally used topick up the movement of a human being in a specific area
Reed switches	Installed on doors of rooms, cupboards, appliances like refrigerator, freezer and microwave	Measure whether doors are open or closed
Pressure Mats	On Bed, couch and chairs	Presence of someone on the objects
Mercury Contacts	Attached to cupboards and doors	Detect movement of objects
Float Sensors	Installed in toiletbasins, bathroom and kitchen sinks	Measure the fluid level in a basin

Besides difference in number of sensors being installed, there are other differences also like, there are two toilets in house C, the toilet in house B is in the same room as the shower, while the toilet and shower in house A are in separate rooms. Furthermore, the inhabitants differ as well, house A was occupied by a 26 year old male, house B by a 28 year old male and house C by a 57 year old male. There's difference even in the behaviour of all three persons. It is evident from the activity profile of two typical days of each house shown in Figure 9. The activities differ both in occurrence time and duration across the three houses.

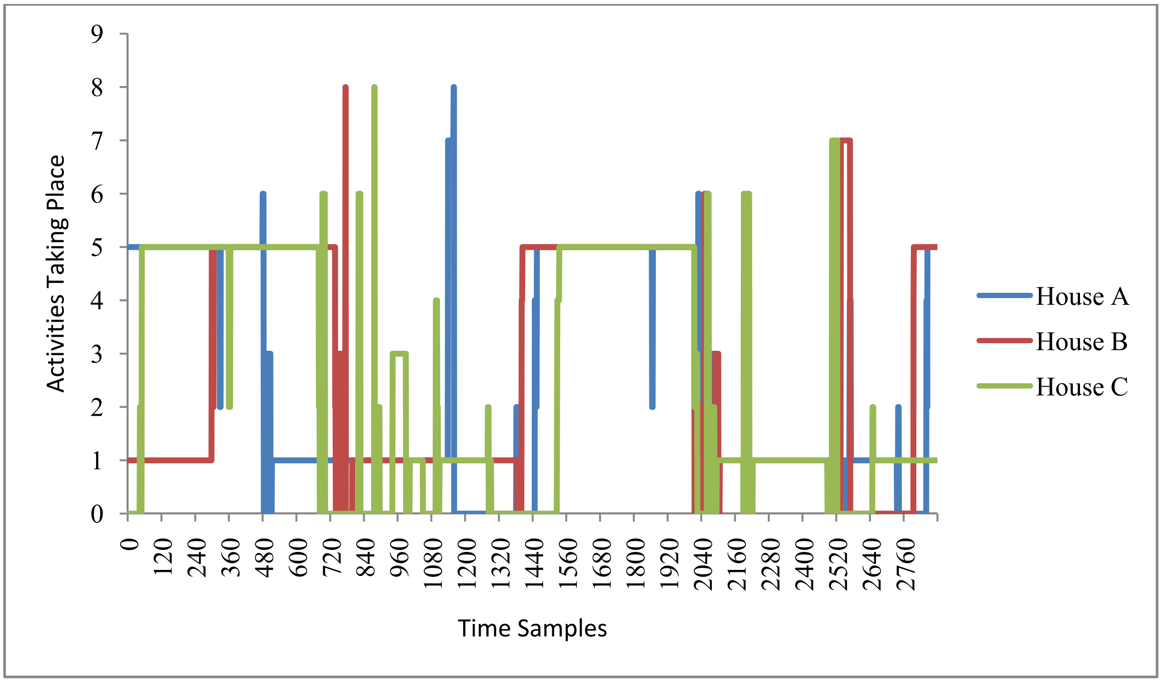


Figure 9: Activities in all three houses for two typical days

Table 5 shows the number of separate instances of activities and the percentage of time each activity takes up in the data set[8]. It can be seen that some activities occur very frequently (e.g. personal hygiene), while others that occur less frequently have a longer duration and therefore take up more time (e.g. leaving and sleeping).The frequency of each activity has large deviations ranging from about 60% to 0.1%. The activity recognition algorithms have to work well in both cases. There also has to be room for accounting sensor errors like for instance, the "grocery cupboard" reed switch kept on firing for all the time in dataset of house B. Such obvious errors are removed before using the dataset. The provided datasets are in two parts, one is the sensor data tables with timestamps and the other is the manually annotated activities with time stamps. These two sets of information are merged and a dataset of sensor firing status along with activity of each minute is prepared. The Matlab scripts provided by [8] are used for this purpose.

a. Sensor Data Sets - Mapping to Common Contexts (Steps [1]-[3])

All houses differ in the sensor feature space where each sensor of a house represents a feature. This is the case in the datasets used here and also may be obvious in any other real life scenario. The modelling of common set of activities has to take into account this variance. This is donehere by mapping the minute wise data of raw sensors of each house to presence or absence of

Table 5 Number of instances and percentage of time activities occur in the dataset

		House A		**House B**		**House C**	
Activity ID	Activity	No. of Occurrences	%age of Time spent on Activity	No. of Occurrences	%age of Time spent on Activity	No. of Occurrences	%age of Time spent on Activity
A1	Not in House	33	50.5	24	59.6	47	45.7
A2	Personal Hygiene	114	1	27	0.4	89	1
A3	Taking Shower	23	0.8	11	0.6	14	0.8
A4	Brushing Teeth	16	0.1	13	0.2	26	0.4
A5	Sleeping	24	33.2	14	29.4	19	29.2
A6	Preparing Breakfast	20	0.3	9	0.5	18	0.6
A7	Preparing Dinner	9	0.9	6	0.5	11	1.1
A8	Preparing Drink	20	0.2	8	0.1	10	0.1
A0	Idle	-	13	-	8.7	-	21.1

common set of contexts that define the common activities of a person living alone inside a house. The variable number of sensors is mapped to fixed number of domain specific contexts. The definition of common set of contexts is inspired by the W31H model of activity descriptiondiscussed in second section [12]. According to the model definition, we take location within house and the objects specific to the location defined as "Objects Used" as first two contexts. Each context is in itself multi-valued. For example, the location in an apartment can be any of the bedroom, kitchen, washroom, living room and bathroomand so are the objects used. A logical context as "current time"is also considered. Time is mapped to one of the 24 hourly time slices of a day. Time Slices are counted from 0 to 23 where 0^{th} hour is midnight. Start time of each firing is taken. Besides these, the objects used and location of the person in previous time slice is also taken as context. The individual sensors of every house were then mapped to common set of contexts as per Table 6.

Table 6: Mapping of Sensors to Common Set of Contexts

Context Category	Context Values	Sensors of House A	Sensors of House B	Sensors of House C
Objects	Bedroom Objects		'press bed links' , 'PIR bedroom' , kwik dresser	'bed right and left pressure mat' ,'dresser, pir '
	Kitchen Heating Objects	'Microwave'	'kwik stove lid' , ' toaster', 'microwave'	'microwave, reed '
	Kitchen Storage Objects	'Cups cupboard' , Fridge, 'Plates cupboard' , 'Freezer' ,'Pans Cupboard' ,'Groceries Cupboard'	' fridge',' cupboard groceries', ' cupboard plates'	'fridge, reed ','freezer, reed', 'cutlary drawer, mercury switch' , 'cupboard pots and pans, reed ' , 'cupboard storage bins, reed ' , 'cupboard herbs and plates, reed ', 'cupboard bowl and cups'
	Cleaning Objects	'Dishwasher' , 'Washing Machine'		
	Bathroom Objects	'ToiletFlush'	' toilet flush',' PIR badkamer' , ' PIR bathroom' , ' sink float'	,'toilet flush upstairs, flush ', 'toilet flush downstairs. flush ' , 'bathtub, pir ','sink upstairs, flush'

Location	Toilet	'Hall-Toilet door'	' toilet door'	'toilet door downstairs'
	Bathroom	'Hall-Bathroom door'		'bathroom swingdoor left and right'
	Bedroom	'Hall-Bedroom door'	'Bedroom door'	'bedroom door'
	Outside	'Frontdoor'	' frontdoor'	'frontdoor, reed '
	Kitchen		' PIR kitchen'	'drawer with keys to backdoor'
Time	24 Slices of 1 Hour Each			

After discussion with several experts, Table 6 was arrived at for mapping. To give statistical basis to this categorization an unsupervised agglomerative clustering was applied on the individual houses to find intuitive correlations among the sensors and identifying the common ones across all houses [33]. This was done for all three houses, and is shown for House A and C in Figures 10 a&b. After analysing the clustering and heatmaps of cross correlation matrix within various sensors of a house, the mapping as in Table 6was found suitablefor finding contexts of person to in turn define his activity.

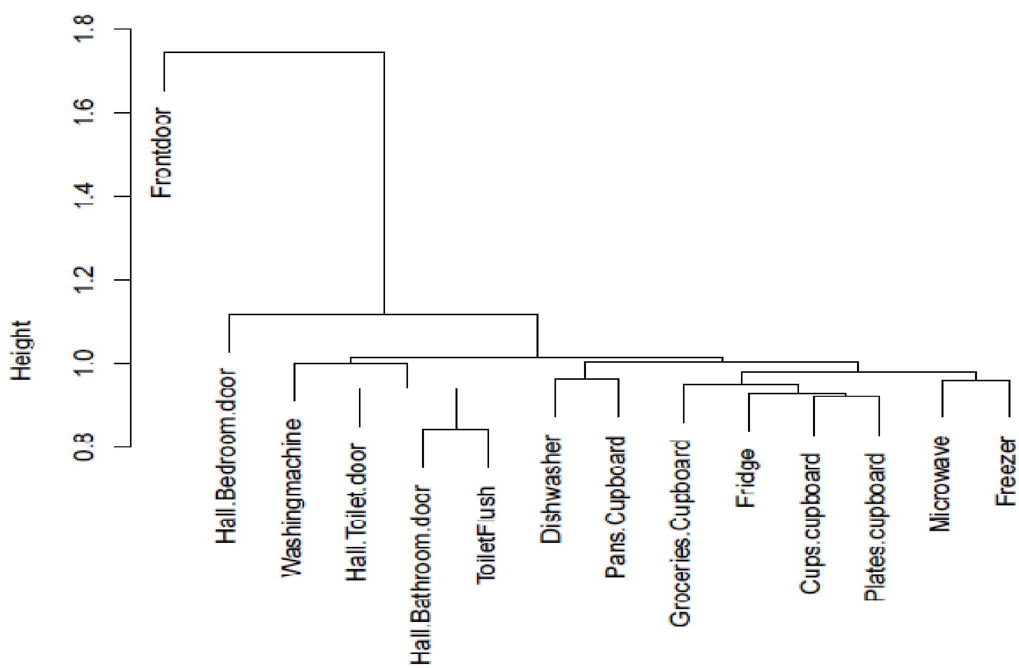


Figure 10(a) Correlation Clusters in House A

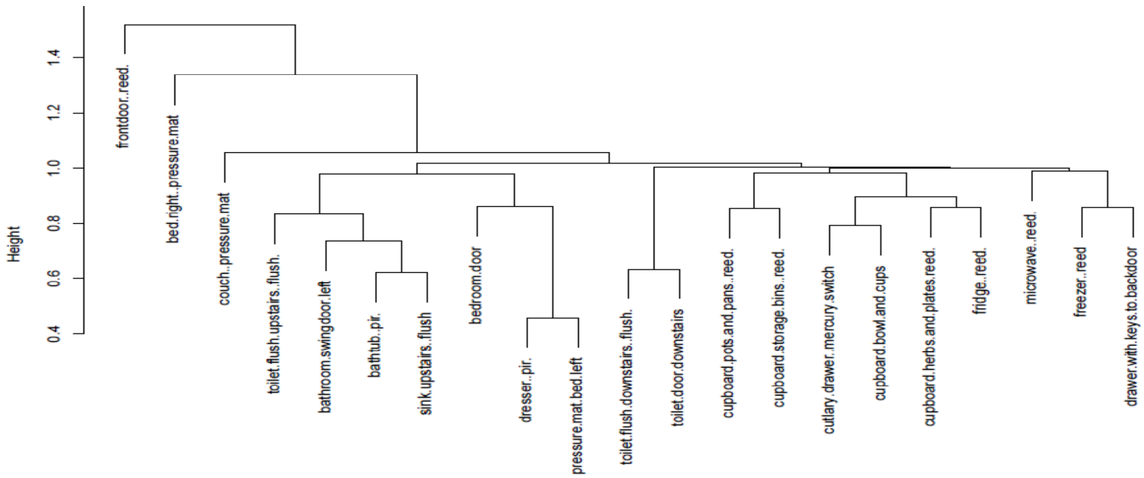


Figure 10(b) Correlation Clusters in House C

We can see that sources for some categories of context values are not available in certain houses. For example, in House A no sensors are present to map the context of bedroom objects used. This context will alwaysconsiderbeing absent in House A. Thus after obtaining the above table, it is deduced that any newsensor will be mapped to the existing contexts. If required, new category of contexts can easily be added with description of constituting sensors.

b. Creation of Activity Lattice (Steps [4]-[6])

The model is built in two phases: Firstly out of all possible nodes of the lattice, the nodes are annotated with frequency of occurrence in training data. The nodes that don’t occur in the data are pruned. In the second phase, the nodes obtained from first phase are annotated with situation occurrence vectors. After this phase, if there are nodes that have no situations occurring are further pruned. A step-by-step methodology of building activity lattices goes as follows:

1) Abstract sensor data into context categories – each unique context combination is defined as preliminary node

2)Annotate the preliminary nodes with activityoccurrence vectors calculated from training data

3)Connect the nodes with semantic edges using Formal Concept Analysis methods of section 3.

4) Update activity occurrence vector of every compound node while moving towards top of lattice.

For each node, the activity occurrence ratio of each node is calculated and the activity having maximum ratio is labeled asoccurring activity with that node. Other activities with non-zero occurrence ratio are considered as possibly occurring activities. With leave one day out approach of training, the number of nodes in activity lattices of House A, B and C came out to be 773, 1156 and 914 respectively which is far lesser than total possible nodes in subsets of each lattice.

c. Inferring Activities from the Concept Lattice

For inferring activities of unknown sensor datasets, we use depth first search methods with backtracking. When given the input context vector, inference starts by evaluating the predicates on all nodes from the top of the lattice. It stops finally at the node exactly matching the input or at the bottommost node giving output that the inputcouldn't be classified.With these inference procedures, we used standard leave one day out method on individual houses to prepare training and test data. Inference experiments were also done to check the validity of model against a new house. Alternately, we used one house as source, while others for testing. The details of results of these experiments and evaluation criteria are discussed in the next section in more detail.

V. RESULTS AND PERFORMANCE ANALYSIS

For analysis of the performance of our proposed method, we first define the appropriate evaluation metrics and then apply these to experiments conducted.

a. Metrics used for Performance evaluation

The activity recognition is a classification problem of determining the current class of activity. As the datasets considered here are unbalanced, that is, some classes appear much more frequent than others do; we use a number of metrics for evaluating classification.

Confusion Matrix: A confusion matrix has actual classes as rows and classified as columns. The size of matrix is square of number of possible classes. The elements of matrix in i^{th} row define how many times class "I" was classified as it and all other classes. An ideal classifier gives non-zero entries only on the left diagonal of this matrix. Any other values are misclassifications. The diagonal of the matrix contains the true positives (TP), while the sum of a row gives us the total of ground truth labels (TT) and the sum of a column gives us the total of inferred labels (TI).

Precision -- The precision gives probability of correctness of activity classification. Precision is more about overall class accuracy and is calculated as per equation (1).

$$\text{Precision} = \frac{1}{N}\sum_{i=1}^{N}\frac{TP_i}{TI_i} \qquad (1)$$

***Recall* –** Recall rate gives how many times a particular activity is correctly recognized. Use equation (2) for finding recall. A less frequent activity has more chances of having better recall rate.

$$\text{Recall} = \frac{1}{N}\sum_{i=1}^{N}\frac{TP_i}{TT_i} \qquad (2)$$

***F-Measure* –** is a metric that balances out the bias of above two metrics, hence it is used to measure overall appropriateness of a classification method.

$$\text{F} - \text{Measure} = \frac{2.\text{Precision}.\text{Recall}}{\text{Precision}+\text{Recall}} (3)$$

Accuracy -- The accuracy is one of the most intuitive measures and represents the percentage of correctly classified time slices.

$$\text{Accuracy} = \frac{\sum_{i=1}^{N}\text{TP}_\text{i}}{\text{Total}} \qquad (4)$$

The calculation of these metrics for a particular case is done as in equations below.

To calculate the over all effect, precision and recall for each class is computed separately and average taken over all classes.

b. Effect of Choosing Meta Features

As discussed in section 5, all three data sets differ in dimension, that is, the actual number of features sensed differs in each house. The lattice-basedmodel on individualhouses takes actual sensor data as source. We useleave one day out approach where all except one random day is chosen for training the model. The performance obtained on various evaluation metrics of such models of all three houses isshown inTable7. Similarly the common set of meta features extracted from each house as per section 5 are used for modeling individual houses with data used in the same way as above.

Table7: Performance Summary of all three houses on using actual sensors data and extracted Meta features

House	Feature Space	Accuracy	Precision	Recall	F-Measure
A	Sensor Data	96.28	62.3	71.5	66.6
	Contextual Information	90.9	60.9	58.3	59.5
B	Sensor Data	88.6	71.3	69.5	70.3
	Contextual Information	79	56.8	64.3	60.3
C	Sensor Data	96.4	56.0	59.8	57.8
	Contextual Information	86.85	51.6	57.1	54.2

The table represents that performance is slightly degraded when derived contextual information are used. This is due to loss of sensor specific information in mapping. The performance is good enough

c. Effect of amount of training data

In this experiment, the performance of common contextual information is explored from the perspective of amount of training data available in a new house. We learnt activity lattices from partial data of individual houses and tested on remaining data of the same house. For this purpose data of 25, 40 and 50 and so on up to 90% of random days are chosen from the entire annotated dataset available.

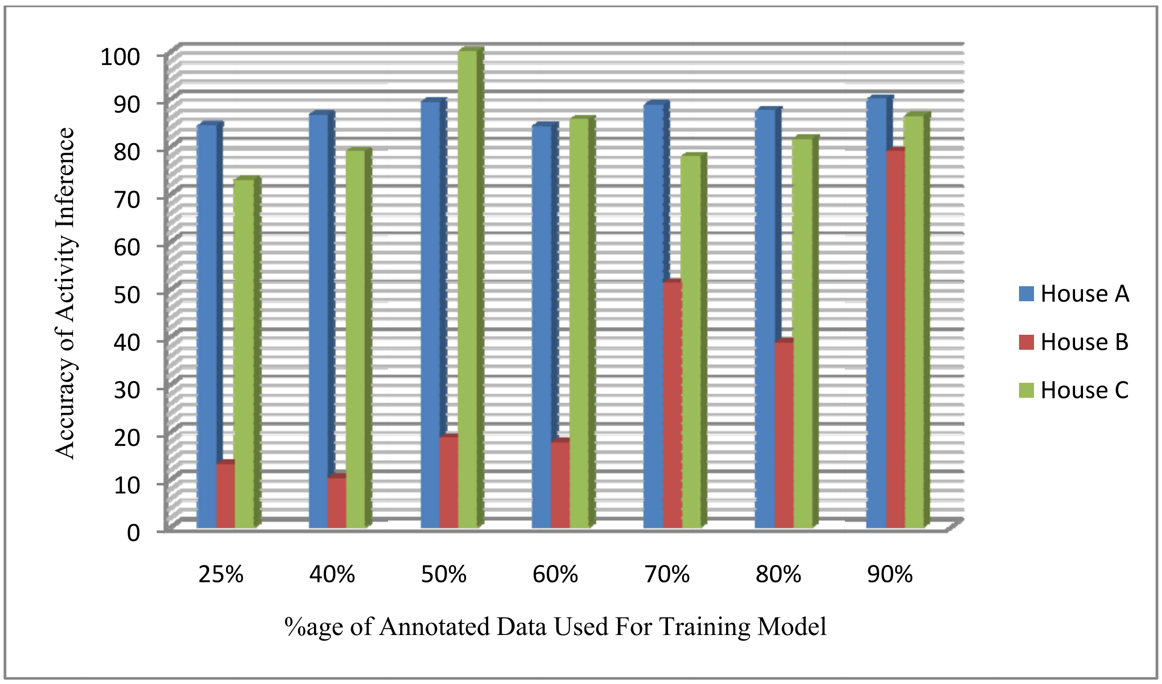


Figure 12Accuracy of House A learnt on varying number of training days

Results inFigure 12 overall suggest that the amount of training data does not affect the accuracy much. We can reach this conclusion by observing prediction accuracy in houses A and C. The results are very different in house B where we detect a much lower accuracy. There are two reasons for this. Firstly, the annotated number of days of house B is least, that is, 11 days. Secondly, on closely observing data we find that the routine of resident in House B is most irregular. Therefore, one typical day does not match much with others. This leads to reduced accuracy in case of House B.

Table 8: Confusion Matrix of House A& C activities

Actual/ Inferred	**A0**		**A1**		**A2**		**A3**		**A4**		**A5**		**A6**		**A7**		**A8**	
	HA	**HC**	**HA**	**HC**	**HA**	**HC**	**HA**	**HC**	**HA**	**HC**	**HA**	**HC**	**HA**	**HC**	**HA**	**HC**	**HA**	**HC**
A0	271	668	125	24	2	3	14	2	0	3	77	5	31	98	0	38	156	2
A1	62	292	2480	1471	0	0	2	0	0	0	76	34	0	0	0	0	0	0
A2	7	14	5	1	6	7	0	2	1	1	6	3	0	0	1	0	1	0
A3	1	1	1	0	1	1	33	2	0	8	0	0	0	0	0	0	0	0
A4	0	9	0	0	2	0	0	0	2	4	2	2	0	0	0	0	0	0
A5	1	1	84	45	1	1	0	0	0	0	2239	1264	0	0	0	0	0	0
A6	0	1	0	0	0	0	0	0	0	0	1	0	8	22	0	0	0	0
A7	36	12	0	1	0	1	0	0	0	0	0	0	0	0	8	91	1	0
A8	5	0	0	0	0	0	0	0	0	0	0	0	0	8	1	0	3	0

The confusion matrix of Table 8 shows the misclassifications occurring in the lattice based model of houses A and C. The model separates the classeswell and most of the confusion is due tothe un-annotated time.Those time slices are classified into any other activity, which may actually be correct but we do not have the ground truth available to compare with. Other major errors are due to misclassifications in sleeping and left house activities. Both these activities result in no sensor firings. Moreover, activities on same time slices are different during weekdays and weekends.Such differences, if many, can be handled by considering type of the day as another logical sensor.

d. Performance of transferred model across houses:

The lattice-based models for individual houses are supervised models and require annotated datasets for learning. Collection of training data for doing activity monitoring in every new house on every new person is a costly practice. This also limits the feasibility of sensor-based solution for activity monitoring. In this experiment, we evaluate the performance of lattice-based model learnt from source house for inferring activities in target houses. The accuracy of activity inference in target houses, when one of the houses is chosen as source for learning model is shown in Figure 13.

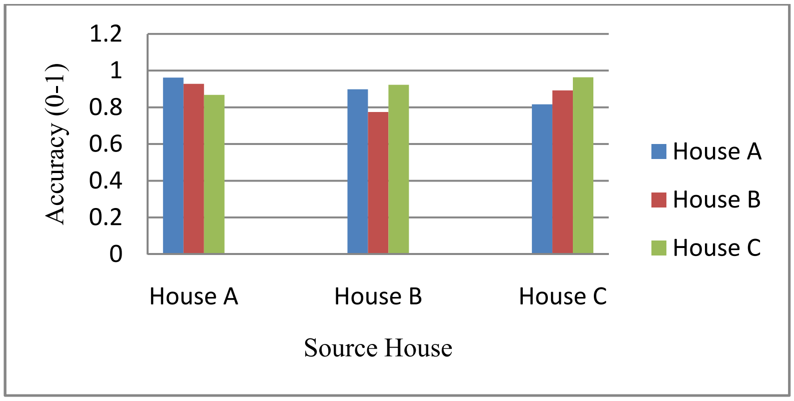


Figure13. Precision of average activity classification on transferring lattice learnt on Source House to target Houses

From the figure, it is clear that despite differences in layouts of the house and habits of the residents, the generic information on activity classification is transferred to the target house. Given the fact that the method is unsupervised for targets, precision achieved isacceptable in classified time slices.The main problem in porting this model is of unclassified instances as shown in Figure 14.

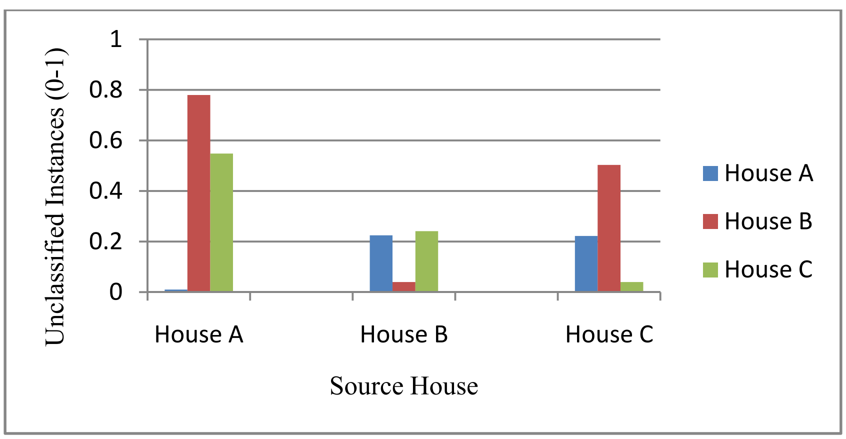


Figure 14: Unclassified instances all possible Source Houses

As mentioned earlier the routine of resident in house B is most unpredictable.There are instances of him having breakfast at noon and going to bed early morning etc. Such irregularities are lesser in other two datasets. Hence, when models learnt from any of the other two houses are applied to house B, lot of time slices are not recognized and are left unclassified. A separate model can be prepared for these unclassified instances. One of the mechanisms is to take feedback from the resident in case of only unclassified activity instance and update the learnt lattice. Another possibility is to use fuzzy formal concept analysis to eliminate non-classifications.The proposed method is tested only on single resident houses. In case of multiple residents, redesigning of sensors to context may have to be done to segregate users [34]. We are also considering single activity being done by the user at a time here. In practice, people frequently interleave one activity with another e.g. like preparing breakfast and showering may be done parallel.The model still needs to be evaluated in these circumstances and refined if required.

VI. CONCLUSIONS

In this work, we extracted the activities of a person living independently in his house using data from sensors embedded in his surroundings. Lattice based model is created by learning profile of

normal activities a person does within a day. Such model will be useful in unobtrusive remote monitoring of person's health as any deviation in routine can be recognized. An alarm can be raised in such case and hencetimely action be taken to assess the real health condition of the elderly. Availability of various sensors with diminishing costs will witness more sensor-based environments. We proposed to learn activities of user from sensors via context mapping. Algorithm based on concept lattices using FCA was then used to derive and recognize activities based on context. Sensor Datasets of common activities done by three different people living in three different sentient houses are used for evaluation. The results represent feasibility of the method in inferring activities of person in future time slices. More importantly, the models also work well across the houses and persons. This method following similar steps can be useful for situation recognition in more complex situations like remote border monitoring and livestock tracking.

REFERENCES

[1] S.Katz et al, "Progress in development of the index of ADL",*The Gerontologist,* pp 20-30Vol. 10, Issue 1,1970.

[2] T.Gu, L.Wang, W. Zhanqing,T. Xianping and L. Jian,"A pattern mining approach to sensor-based human activity recognition", IEEE Transactions on Knowledge and Data Engineering, pp 1359-1372, Vol. 23, Issue 9, September 2011.

[3] D.Guan , T.Ma,W.Yuan, Y. Lee , AM JehadSarkar, “Review of Sensor-based Activity Recognition Systems”, pp. 418-33 Vol. 28, Issue 5, IETE Tech Rev,Oct 2011.

[4] A. Bauer and Y. Fischer, “Task-oriented Situation Recognition”, In Proc. of Cyber Security, Situation Management and Impact Assessment II; and Visual Analytics for Homeland Defense and Security II, SPIE Vol. 7709,2010.

[5] J. Ye, S. Dobson and S. McKeever, “Situation Identification Techniques in Pervasive Computing: A Review”, Pervasive Mobile Computing, pp 36-66, Vol. 8, Issue 1, February 2012.

[6] A. Dahlbom, L. Niklasson,G. Falkman and A. Loutfi, "Towards Template-Based Situation Recognition," in Proc. of Intelligent Sensing, Situation Management, Impact Assessment and Cyber-Sensing, SPIE Vol. 7352, 2009.

[7] C Carpineto, G Romano, “Concept Data Analysis: Theory and Applications”, John Wiley, 2004.

[8] T VKasteren, G Englebienne and B Kröse, "Transferring Knowledge of Activity Recognition Across Sensor Networks", In Proceedings of the 8th international conference on Pervasive Computing, pp 283-300, 2010.

[9] S Mittal and S L Maskara, “A Review of Bayesian Belief Network Structure Learning Algorithms”, In Proceedings of 8th International Conference on Information, Communications and Signal Processing (ICICS) , 13 -16 Dec, Singapore.

[10] S Mittal, AAggarwal, S L Maskara, “Application of Bayesian Belief Networks for Context Extraction from Wireless Sensors Data”, 14th International Conference on Advanced Communication Technology (ICACT2012), pp 410 -415, Korea, 2012.

[11] R Cardell-Oliver and Liu Wei, "Representation and Recognition of Situations in Sensor Networks,” IEEE *Communications Magazine*, Vol.48, Issue 3, pp.112-117, March 2010.

[12] Y Oh, J Han and W Woo, “A context management architecture for large-scale smart environments,” IEEE Communications Magazine, Vol. 48, Issue 3, March 2010

[13] A. Dahlbom, L. Niklasson, G. Falkman and A. Loutfi, "Towards template-based situation recognition," in Proceedings of SPIE on Intelligent Sensing, Situation Management, Impact Assessment, and Cyber-Sensing, Vol. 7352. 2009.

[14] J Ye, L Coyle, S Dobson and P Nixon, "Representing and manipulating situation hierarchies using situation lattices,"*Revue d'intelligenceartificielle*, Vol. 22, Issue 5, pp 647-667, 2008.

[15]K. Thirunarayan, C. Henson and A. Sheth, "Situation awareness via Abductive Reasoning from Semantic Sensor data: A preliminary report," International Symposium on Collaborative Technologies and Systems, pp.111-118, May 2009.

[16] A Srivastav, W Yicheng, E Hendrick,I Chattopadhyay, A Ray and S Phoha,"Information Fusion for Object &Situation Assessment in Sensor Networks," In Proceedings ofAmerican Control Conference (ACC), pp.1274-1279, 2011.

[17] Lu-An Tang, Yu Xiao, Kim Sangkyum et al., "Multidimensional Sensor Data Analysis in Cyber-Physical System: An Atypical Cube Approach," International Journal of Distributed Sensor Networks, Vol. 2012, Article ID 724846, 19 pages, 2012.

[18] J Fogarty and S E. Hudson, "Toolkit Support for Developing and Deploying Sensor-Based Statistical Models of Human Situations", In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp 135-144, 2007 , ACM, New York,USA.

[19] S. Yau, D. Huang, H. Gong andY. Yao, "Support for Situation Awareness in Trustworthy UbiquitousComputing Application Software," Software: Practice and Experience, pp 893-921, Vol. 36, No. 9,2006.

[20] P. Costa, G. Guizzardi, J. Almeida, L. Pires and M. Van Sinderen, "Situations inConceptual Modeling of Context,"In Proceedings of the 10^{th} IEEE International Enterprise Distributed Object Computing Conference Workshops, pp 6-16, October, 2006.Hong Kong.

[21]S. W. Loke, "Representing and Reasoning with Situations for Context-Aware Pervasive Computing: ALogic Programming Perspective", Knowledge Engineering Review, pp 213-233, Vol. 19 Issue 3, 2004.

[22]G. Thomson, S.Terzis and P. Nixon, "Situation Determination with Reusable Situation Specifications," InProceedings of 4th IEEE International Conference on Pervasive Computing and Communications, pp 620-623, March 2006.

[23] T. Emmanuel, S. Intille and K. Larson, "Activity Recognition in the Home Using Simple and Ubiquitous Sensors", In Proceedings of 2nd International Conference on Pervasive Computing in LNCS, Springer, Vol. 3001, pp 158-175,2004.

[24] N. Bicocchi, M. Lasagni andF. Zambonelli, "Bridging Vision and Commonsense forMultimodal Situation Recognition in Pervasive Systems," 12thIEEE International Conference on Pervasive Computing and Communications,pp48-56, March 2012.

[25]L. Atallah et al., “Distributed Inferencing with Ambient and Wearable Sensors”, Wireless Communications and Mobile Computing, pp 117-131, Vol. 12, Issue 1, 2010.

[26] Y. Fischer, A. Bauer and J.Beyerer, "A Conceptual Framework for Automatic Situation Assessment," in Proceedings of First International Multi-Disciplinary Conference on Cognitive Methods in Situation Awareness and Decision Support (CogSIMA), pp.234-239, 2011.

[27]A. Mahajan, C. Oesch, H. Padmanaban, L. Utterback, S. Chitikeshi and F. Figueroa, “Physical and Virtual Intelligent Sensors for Integrated Health Management Systems”, International Journal on Smart Sensing and Intelligent Systems, pp 559 – 575, Vol. 5, No. 3, September 2012.

[28] B.Ganter and R.Wille, “Formal Concept Analysis: Mathematical Foundations”, Springer Verlag, 1999.

[29] K. Dalkir, “Knowledge Management in Theory and Practice”, The MITPress, 2011.

[30] http://conexp.sourceforge.net/

[31]J. Poelmans et al., "Formal Concept Analysis in Knowledge Discovery: ASurvey", In Proceedings of 18th International Conference on Conceptual Structures: From Information to Intelligence,pp 139-153, 2010.

[32]J. Herethet al.,"Conceptual Knowledge Discovery and Data Analysis", In Proceedings of 8th International Conference onConceptual Structures: Logical, Linguistic, and Computational Issues, pp 421-437, 2000.

[33]Y. Zhao, K. George and U. Fayyad, “Hierarchical Clustering Algorithms for Document Datasets”, Data Mining and Knowledge Discovery, pp 141-168, Vol.10, Issue 2, March 2005.

[34] A. Mannini and A. Sabatini, “Machine Learning Methods for Classifying Human Physical Activity from On-Body Accelerometers”, Sensors,pp. 1154-1175,Vol. 10,Issue 2, February 2010.