Using Markov Logic Network for On-line Activity Recognition from Non-Visual Home Automation Sensors

Pedro Chahuara, Anthony Fleury, François Portet, Michel Vacher

Using Markov Logic Network for On-line Activity Recognition from Non-Visual Home Automation Sensors

Pedro Chahuara[1], Anthony Fleury[2], François Portet[1], and Michel Vacher[1] ⋆

[1] Laboratoire d'Informatique de Grenoble
Grenoble 1/Grenoble INP/CNRS UMR 5217, F-38041 Grenoble, FRANCE
{pedro.chahuara, francois.portet, michel.vacher}@imag.fr
[2] Univ. Lille Nord de France, F-59000 Lille, France.
EMDouai, IA, F-59508 Douai Cedex, France.
anthony.fleury@mines-douai.fr

Résumé This paper presents the application of Markov Logic Networks(MLN) for the the recognition of Activities of Daily Living (ADL) in a smart home. We describe a procedure that uses raw data from non visual and non wearable sensors in order to create a classification model leveraging logic formal representation and probabilistic inference. SVM and Naive Bayes methods were used as baselines to compare the performance of our implementation, as they have proved to be highly efficient in classification tasks. The evaluation was carried out on a real smart home where 21 participants performed ADLs. Results show not only the appreciable capacities of MLN as a classifier, but also its potential to be easily integrable into a formal knowledge representation framework.

Keywords: Activity Recognition, Markov Logic Network, Support Vector Machine, Smart Home, Ambient Assisted Living.

1 Introduction

In the Ambient Assisted Living domain there is a increased interest in automatic human activity recognition from sensors [1,2,3]. Recognition of human activity is recognised as one important variable for human behaviour monitoring but it is also extensively studied for the provision of context-aware services in smart-phones and other smart objects [4].

In this paper, we focus on the recognition of activity within the home which is a private space in which multiple sensors, actuators and home automation equipments coexist. This research is linked to the Sweet-Home project which aims at developing a complete framework to enable voice command in smart homes. In this framework, the interpretation of the orders and the decisions to be made depend on the context in which the user is. This context is composed,

⋆. This work is part of the Sweet-Home project founded by the French National Research Agency (Agence Nationale de la Recherche / ANR-09-VERS-011)

among other information, of the user's current activity which is essential for decision making. For instance, if the user utters "turn on the light", the action will be different if she is awaking in the middle of the night (in that case, the best action could be to provide low intensity light using the bedside lamp) or if she is dressing up in the morning (in that case, the best action could be to provide high intensity light using the ceiling lamp). Knowing whether the user is sleeping or dressing permits fine grain decision making and facilitate disambiguation of situations interpretation.

Most of the progresses made in the field of activity recognition come from the computer vision domain [5]. However, the installation of video cameras in the user's home is not only raising ethical questions [1], but is also rejected by some users [6]. Other approaches rely on information from RFID tags [7] and wearable devices [8]. In the first case, putting RFID tags on objects makes fastidious the maintenance of the smart home since any new object implies technical manipulation to attach the corresponding sensor. The case of wearable sensors is sometimes not applicable when inhabitants do not want (or forget) to wear sensors all the time. Moreover, cost and dissemination of assistive technology would be more efficient if it is built on standard home automation technology with minimal technical addition. This is why, the solution developed in the SWEET-HOME project is to complete conventional sensors for home automation (infrared presence detectors, switches, etc.) by microphones to allow the user to control her environment through voice recognition.

This type of environment imposes constraints on the sensors and the technology used for recognition. Indeed, information provided by the sensor for activity recognition is indirect (no worn sensors for localisation), heterogeneous (from numerical to categorical), transient (no continuous recording), noisy, and non-visual (no camera). This application setting calls for new methods for activity recognition which can deal with the poverty and unreliability of the provided information, process streams of data, and whose models can be checked by human and linked to domain knowledge. To this end, we present a method based on Markov Logic Network (MLN) to recognise activities of daily living in a perceptive environment. MLN, a statistical relational method, makes it possible to build logical models that can deal with uncertainty. This method is detailed in Section 4. Before this, a state of the art in MLN based activity recognition is given in Section 2 and the SWEET-HOME project is introduced in Section 3. The method was tested in an experiment in a real smart home involving more than 20 participants. The experiment and the results are described in Section 5. The paper ends with a discussion and gives a short outlook on further work.

2 State of the art of Human Activity Recognition in the home

Automatic recognition of human activity (eating, talking, watching TV, etc..) can be defined as the identification of a sequence of atomic actions (taking a cooking utensil, lying, etc.). This involves abstracting the raw signals into symbols (proposals) temporally labelled (e.g. : door slamming at 11 :32), signatures of specific situations of atomic events are detected through a process of hierarchical abstraction. For example, the movements detected in the bedroom can be part of the activity "get up" which itself can be part of a plan of the day (e.g ; : Sunday morning). Automatic recognition of activity is one of the most active and most ambitious research areas because of the large amount of noise in the data and the difficulty of modelling situations ; for the same person, an activity can take place in many ways.

Approaches for activity recognition can be divided mainly into three categories : statistical, probabilistic and logic. In the former category, machine learning methods have been applied to classify activities from information related to pervasive environments. For instance, Fleury *et al.* [9] proposed Support Vector Machines (SVM) to implement a classifier using data from sensors in a real pervasive environment. If statistical methods such as SVM or Neural Networks can lead to good performance, they lack a formal base to represent uncertainty and the obtained models are not easily interpretable.

As information in pervasive environments is uncertain in most cases, probabilistic approaches are suitable candidates to be applied for activity recognition. For instance, Dynamic Bayesian networks were used by Van Kasteren *et al.* [3] to recognise elders' activities. Considering the temporal nature of activities as a succession of actions or sub activities, literature presenting a modelling of activity by Hidden Markov Models (HMM) is vast. For instance, Duong *et al.* [10] extended a conventional HMM to model the duration of an activity and Naaem *et al.* [11] defined activities as a composition of tasks modelled by hierarchical HMMs. However, despite improved expressiveness, these models require a large amount of training data. These training data are costly to obtain and are often not generalisable to other settings than the one in which they were acquired. Moreover, it remains difficult to integrate *a priori* high-level knowledge in these probabilistic models.

The logical approach offers an ideal framework to model explicit knowledge. Ontologies have been used for this task [12] since they feature readability and formal definitions while the activity recognition can be performed by an ontology reasoner as a problem of satisfiability. Moreover, under the logic approach, logic rules facilitate the implementation of expert knowledge within a model [2]. For instance, Augusto *et al.* [13] used logical models to represent the temporal relations among events to recognise activities. The main drawback of this methods is the lack of systematic handling of uncertainty.

Recently, Statistical Relational Learning (SRL) [14], a sub domain of machine learning, has gained much attention as it integrates elements of logic and probabilistic models. Under the SRL schema, models are defined in a formal logical language that makes them reusable and easy to verify, that systematically take uncertainty into account, and that allows easily inclusion of *a priori* knowledge. SRL has recently attracted attention in the domain of human activity modelling and recognition. For instance, Logic HMM and relational Markov networks are both SRL methods that were considered for activity recognition [15,16]. In our work we proposed to use Markov Logic Networks (MLN). To the best of our knowledge the closed approach to our is the one of Trans *et al.* [17] who detected activities in video streams. Other MLN-based activity recognition methods were defined and tested under different settings from the Sweet-Home project. For instance, [7] assumed RFID tags but this implies tagging every objects involved in the model and the practicability of the approach can be questioned. In our project, which is described in the next section input data imposes little constraint on the daily file of the user.

3 Sweet-Home : an Audio-Based Smart Home System

The Sweet-Home project (`sweet-home.imag.fr`) aims at designing a new smart home system based on audio technology focusing on three main aspects : to provide assistance via *natural man-machine interaction* (voice and tactile command), to ease *social inclusion* and to provide *security reassurance* by detecting situations of distress. If these aims are achieved, then the person will be able to pilot their environment at any time in the most natural way possible.

The input of the Sweet-Home system is composed of the information from the home automation system transmitted via a local network and information from the microphones transmitted through radio frequency channels. Rather than building communication buses and purpose designed material from scratch, the project tries to make use of already standardised technologies and applications. As emphasised in [18] the interoperability of ubiquitous computing elements is a well known challenge to address. Thus, the use of home automation standards ensure compatibility between devices, ease the maintenance and orient the smart home design toward cheaper solutions. While the home automation system provides symbolic information, raw audio signals must be processed to extract information from speech and sound. The extracted information is analysed and either the system reacts to an order given by the user or the system acts pro-actively by modifying the environment without an order (e.g. turns off the light when nobody is in the room). ù Output of the system thus includes home automation orders but also interaction with the user when a vocal order has not been well understood for example, or in case of alert messages (e.g. turn off the gas, remind the person of an appointment).

The Sweet-Home system will be piloted by an intelligent controller which will capture all streams of data, interpret them and execute the required actions. The diagram of this intelligent controller is depicted in Figure 1. All the knowledge of the controller is defined using two semantic layers : the *low-level* and the *high-level* ontologies. The former ontology is devoted to the representation of raw data and network information description while the high level ontology represents concepts being used at the reasoning level such as : Actions that can be performed in a home and the context in which a home can be (e.g., making coffee, being late). This separation between low and high levels makes possible a higher re-usability of the reasoning layer when the sensor network and the home must be adapted [19]. The estimation of the current context is carried

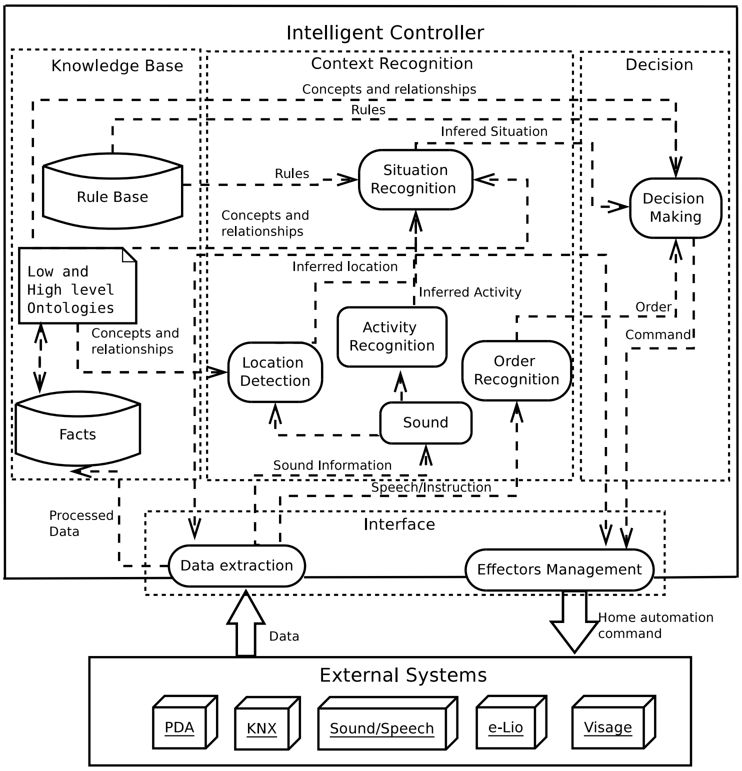

Figure 1. The Intelligent Controller Diagram

out through the collaboration of several processors, each one being specialised in a certain context aspect, such as location detection or activity recognition. All processors share the knowledge specified in both ontologies and use the same repository of facts. Furthermore, the access to the knowledge base is executed under a service oriented approach that allows any processor being registered to be notified only about particular events and saving any inferred information to be available to other processors. This data and knowledge centred approach permits to ensure that all the processors are using the same data structure and that the meaning of each piece of information is clearly defined among all of them. In addition, the chosen architecture is more flexible than a classical pipeline of processors, making possible the easy insertion of new processors. Once the current context has been determined, the controller evaluates if an action must be taken. What supports the process of decision is a set of logic rules which are part of the knowledge base.

4 Method

As shown in the previous section, activities is an important element of the context. In this study, the activities under consideration are daily living tasks such as sleeping, dressing, eating, communicating, etc. [20] that a person performs during a temporal interval. Given that the data streams provide quite poor, indirect (no wearable sensors) and sporadic (no continuous measurement of some variables) information as they are only composed of classical home automation sensors and microphones data ; fine grain activity recognition, such as screwdriver usage, is impracticable. However, each instance of activity is composed of a set of events that generate observations from the set of home automation sensors. Our hypothesis is that these set of observations are signatures of the activities and that they can be described by statistics of predefined variables computed over temporal windows shorter that the minimal activity duration. Although activities captured in this manner might be gross, we argue that they can be sufficient to provide contextual information to the decision module.

The method to recognise activities from the streams of raw sensor data goes through different levels of abstraction as depicted Figure 2. The raw data are composed of symbolic timestamped values (e.g., infra-red sensors), state values (e.g., switches), time series (e.g., temperature) and signals (e.g., microphones). A pre-processing stage extracts higher-level information such as speech, sounds, location and agitation of the inhabitant. To represent the stream of data, all the raw and abstracted data are summarised as attribute vectors, each of which corresponding to a temporal windows of size W. So, at current time t, every segment of the data of interval $]t - W, t]$ seconds is represented by a vector $v(t)$ which is used as input to a classification model M which gives the main activity a of the person during $]t - W, t]$. All of the classification models were acquired using supervised machine learning techniques. This section summarises the pre-processing stage, and details the attributes and the classifier model.

4.1 Preprocessing

The raw data captured within the smart home (see bottom of Fig. 2) contains information that must be extracted for enhanced activity recognition. Three types of abstraction are considered : localisation of the inhabitant — which is of primary importance for activity recognition —, speech and sound recognition — which is important for activities of communication — and activity level — which specifies how active the person was during the temporal window.

Speech/sound detection In this approach, sound events are detected in real-time by the AuditHIS system [21]. Briefly, the audio events are detected in real-time by an adaptive threshold algorithm based on a wavelet analysis and an SNR (Signal-to-Noise Ratio) estimation. The events are then classified into speech or everyday life sound by a Gaussian Mixture Model. The microphones being omnidirectional, a sound can be recorded simultaneously by multiple microphones ; AuditHIS identifies these simultaneous events.

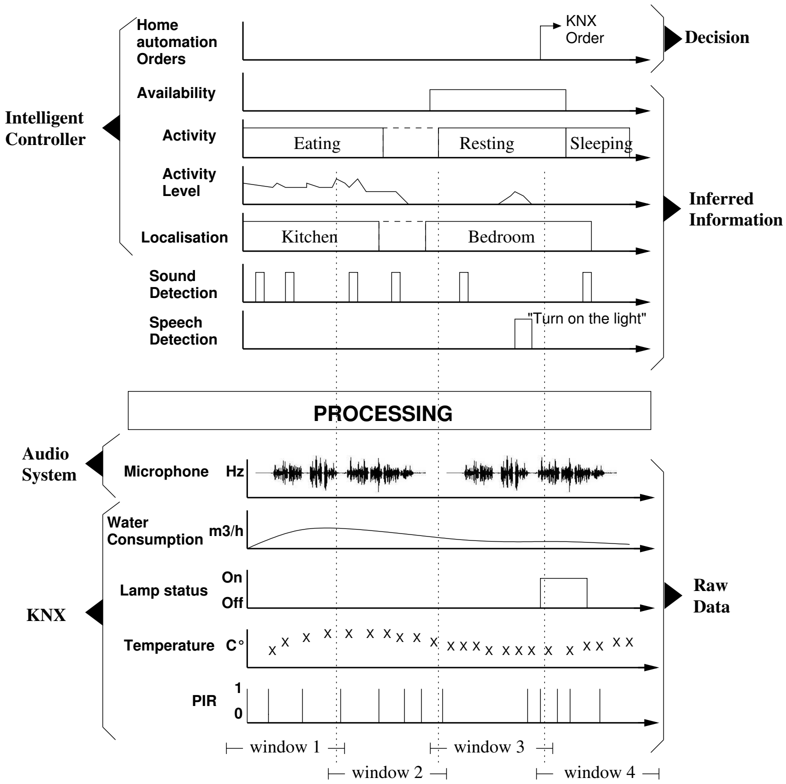

Figure 2. Temporal windowing for computing the vectors from the sensors data.

Localisation In smart homes, cheap localisation can be performed by using infra-red sensors detecting movement but these sensors can lack sensitivity. To improve localisation, our approach fuses information coming from different data sources namely, infra-red sensors, door contacts and microphones. The data fusion model is composed of a two-level dynamic network [22] whose nodes represent the different location hypotheses and whose edges represent the strength (i.e., certainty) of the relation between nodes. This method has demonstrated a correct localisation rate from 63% to 84% using several uncertain sources.

Activity Level In smart homes, the level of activity is a measure of the frequency of actions of the person in the house. This makes it possible to distinguish from activities that ask for many actions (e.g., cleaning the house) from quieter ones (e.g., sleeping). In our approach, level of activity is estimated not only from infra-red sensors but also from sound information as well as door contacts. These sources of data are fused using a linear model. Level of activity — which is a numerical measure— must not be confused with activity — which is the category of daily living activities being performed.

4.2 Generation of the Vectors of Attributes

The traces generated from human activities are difficult to generalise due to the high inter and intra-person variability of realisation of a task. This is why statistic attributes and inferred information were chosen to summarise the information present in each window. In total 69 attributes were extracted for each temporal window. They are summarised in Table 1. In addition to the set of attributes, the seven activities considered in the study are given at the end of the Table. It can be noticed that an unknown class is present. This class is composed of all windows during which none of the seven activities is being performed.

4.3 Markov Logic Network (MLN)

MLN [23] combines first-order logic and Markov Networks. A MLN is composed of a set of first-order formulae each one associated to a weight that expresses a degree of truth. This approach soften the assumption that a logic formula can only be true or false. A formula that does not contain variables is

Table 1. Attributes used for activity recognition (all the attributes are numerical)

Attributes	Number	comments
PourcentageLoc_x	4	ratio of time spent at room x
PredominantLoc	1	most occupied room during the temporal window
LastRoomBeforeWindow	1	last room in which the person was just before the current window
TimeSinceInThisRoom	1	Time elapsed since the person entered the room
ActivationDeactivationWindow_y	6	number of state changes of the door contact of the window y
ActivationDeactivationDoor_w	4	number of state changes of the door contact of the door w
ActivationDeactivationLight_z	6	number of state changes of the light z
ActivationDeactivationCommDoor_f	5	number of state changes of the door contact of the furniture f
DetectionPIR_x	2	number of time the movement detector fired in room x
AmbientSensor	13	difference of value between temporal windows of : CO2, temperature, humidity, brightness, water and electricity
Power_LastUse	3	Id of the last used sockets or switches.
PercentageTime_Sound	1	ratio of time occupied by sounds during the temporal window
PercentageTime_Speech	1	ratio of time occupied by speech during the temporal window
sound_m	7	number of sound occurrences detected by microphone m
speech_m	7	number of speech occurrences detected by microphone m
PercentageAgitation_	6	number of events per windows for the category : room, doors, electricity, water, sounds and speech
TotalAgitation	1	sum of the PercentageAgitation_ in the temporal window
Class	1	eating, tidying up, eliminating, communicating, dressing up, sleeping, resting, unknown

ground and is a *ground atom* if it consists of a single predicate. A set of ground atoms is a *possible world*. All possible worlds in a MLN are true with a certain probability which depends on the number of formulae they satisfy and the weights of these formulae. A MLN, however, can also have hard constraints by giving a infinite weight to some formulae, so that worlds violating these formulae have zero probability. Let's consider F a set of first-order logic formulae, $w_i \in \mathbf{R}$ the weight of the formula $f_i \in F$, and C a set of constants. During the inference process, every MLN predicated is grounded and Markov network $M_{F,C}$ is constructed where each random variable corresponds to a ground atom. The obtained Markov network allows to estimate the probability of a possible world $P(X = x)$ by the equation 1 :

$$P(X = x) = \frac{1}{Z} exp\left(\sum_{f_i \in F} w_i n_i(x)\right) \tag{1}$$

where $Z = \sum_{x' \in \chi} exp\left(\sum_{f_i \in F} w_i n_i(x')\right)$ is a normalisation factor, χ the set of possible worlds, and $n_i(x)$ is the number of true groundings of the i-th clause in the possible world x. Exact inference in MLN is intractable in most cases, so Markov Chain Monte Carlo methods are applied [23].

Learning an MLN consists of two independent tasks : structure learning and weight learning. Structure can be obtained by applying machine learning methods, such as Inductive Logic Programming, or rules written by human experts. Weight learning is an optimisation problem that requires learning data. The most applied algorithm in the literature is *Scaled Conjugate Gradient* [24].

The approach we implemented uses a list or rules modelling the capacity of each feature value to discriminate an activity. Formally, this rules have the following structure $feature_i(X, V_i) \rightarrow class(X, V_c)$ where the variables X, V_i, and V_c represent : the temporal window to be classified, the value of the i^{th} feature, and the value of the target class. Before creating the model, all numerical variables were discretised. In addition, the temporal relation between instances were modelled by means of the following rule $previous(X, V_c) \rightarrow class(X, V_c)$. The adequacy of including this rule comes from the fact that daily activities use to follow a certain pattern within the routine of an inhabitant. The weights for this model were learnt by *Scaled Conjugate Gradient*. Inference on test data is performed as though it was an on-line system. Thus, the classification of a temporal window not only considers the features describing it but also the results of the previous window classification.

4.4 Support Vector Machine

Support Vector Machine (SVM) [25] is a classification and regression method. In essence, a SVM is a mathematical entity for maximizing a particular mathematical function with respect to a given collection of data (i.e. samples, measurements, records, patterns or observations). Similar to neural networks, SVMs possess the well-known ability of being universal approximators of any multivariate function. Consequently, they are of particular interest for modeling highly nonlinear complex systems.

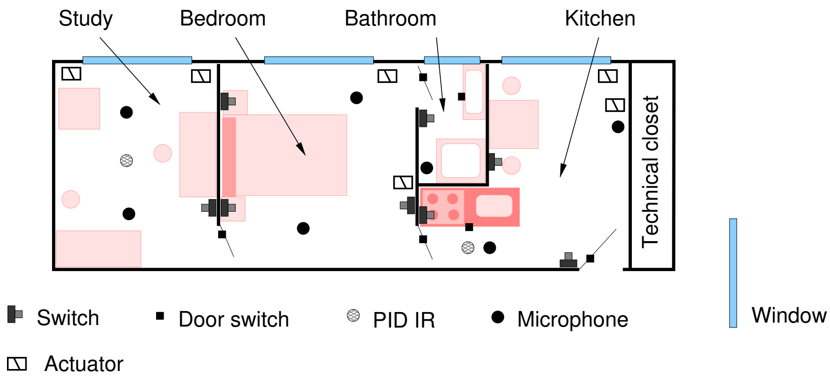

Figure 3. The DOMUS Smart Home

The simplest model of SVM is the so-called maximal margin classifier, which works only for data which are linearly separable in the feature space. However, SVM can be extended to treat nonlinear casesby the use of Kernel functions[26].

This method has already been used for activity recognition[9] achieving acceptable rates of accuracy. That is why, we have chosen SVM as a baseline for our proposed model.

5 Experiments and results

The method was applied on data collected in a smart home during an experiment involving 21 persons. This section describes the pervasive environment in which the tests were performed (sec. 5.1), the data set (sec. 5.2) that was acquired and the results of the activity recognition (sec. 5.3).

5.1 Pervasive Environment

In the SWEET-HOME project, the main pervasive environment considered is the DOMUS smart home depicted Figure 3. It was adapted to acquire realistic corpus and test the developed techniques. This smart home was set up by the Multicom team of the Laboratory of Informatics of Grenoble, partner of the project. It is a thirty square meters suite flat including a bathroom, a kitchen, a bedroom and a study, all equipped with sensors and effectors such as infrared presence detectors, contact sensors, video cameras (used only for annotation purpose), etc. In addition, seven microphones were set in the ceiling. The technical architecture of DOMUS is based on the KNX bus system[3], a worldwide ISO standard (ISO/IEC 14543) for home and building control. Bus devices can either be sensors or actuators needed for the control of building equipments such as : lighting, shutters, security systems, energy management, heating, ventilation

3. www.knx.org

and air-conditioning systems, interfaces to service and building control systems, remote control, metering, audio video control. . .Besides KNX, several field buses coexist in Domus, such as UPnP (Universal Plug and Play) for the audio video distribution, X2D for the opening detection (doors, windows, and cupboards), RFID for the interaction with tangible objects (not used in the Sweet-Home project). More than 150 sensors, actuators and information providers are managed in the flat. A residential gateway architecture has been designed, supported by a virtual KNX layer seen as an OSGI service (Open Services Gateway Initiative). This layer guarantees the interoperability of the data coming from the different field buses and allows the communication between them and towards virtual applications, such as activity tracking. More than 60 bundles delivering more than 60 services are running. Thanks to this gateway, all home automation elements as well as multimedia elements can be controlled and parametrised remotely.

The sensors that were used in the experiments were the following :
- 7 radio microphones set into the ceiling (2 per room except for the bathroom). These microphones are useful to detect communication activities, recognise speech or detect abnormal sounds ;
- switches and door contact detectors connected to the KNX network which are useful for device usage (fridge, cupboard), informing about movement (e.g., door used between to rooms), knowing whether a window is open or a socket activated.
- Two infra-red movement detectors connected to the KNX network ;
- ambient sensors, such as temperature, humidity, CO_2, etc. ;
- electricity and water meters ;
- video cameras only used to perform annotation of the data.

5.2 Experimental Data

An experiment was conducted to acquire a representative data set of activities performed at home. The Domus flat was designed to seem as normal as a standard flat (e.g., no visible wire, no prominent sensor) so that any participant's activity instance would be as close as possible to an instance performed in their usual manner. 21 persons (including 7 women) participated to the experiment to record all sensors data in a daily living context. The average age of the participants was 38.5 ± 13 years (22-63, min-max). To make sure that the data acquired would be as close as possible to real daily living, the participants were asked to perform several daily living activities in the smart home. A visit, before the experiment, was organised to make sure that the participants will find all the items necessary to perform the activities. No instruction was given to any participant about how they should behave. All the home automation data was recorded at the virtual KNX layer apart from video data which was recorded separately and audio data which was recorded in real-time thanks to a dedicated PC embedding an 8-channel input audio card [27].

The experiment consisted in following a scenario of activities without condition on the time spent and the manner of achieving them (having a breakfast, simulate a shower, get some sleep, clean up the flat using the vacuum, etc.). During this first phase, participants uttered 40 predefined casual sentences on the phone (e.g., “Allo”, “J’ai eu du mal à dormir”) but were also free to utter any sentence they wanted (some did speak to themselves aloud). At the end, more than 26 hours of data were recorded. All the activities as well as other information such as location, position, etc. were annotated using the advene software [4].

The stream of data was analysed window by window to generate 2122 instances described by the 69 attributes detailed in Section 4.2. The temporal window size was chosen empirically to be 1 minute. Moreover, given that some activities might intersect with several windows, a 25% overlap was applied. This gives an on-line system that recognises activities every 45 seconds.

5.3 Results

The MLN approach was applied to the experimental data collected in the smart home. In order to assess improvement, we also applied a Naive Bayes (NB) classifier known to be simplistic but often efficient and a more elaborate approach based on Support Vector Machine (SVM) which showed very good results on a similar task [9]. These two classifiers NB and SVM provide the baseline results. The learning and testing strategy used was leave-one-out which consisted in learning models successively on 20 participants and testing on the remaining one. The results on the 21 combinations are then averaged to give the overall performance.

Table 2 shows the precision and recall for the methods. The overall accuracy achieved with MLN, 85.3%, is significantly higher than the one obtained with SVM, 59.6% and Naive Bayes, 66.1% . The unknown class case, which corresponds to temporal windows that have not been labelled, exhibits the lowest performance due to the fact that none of the methods can characterise the class. These unlabelled windows can, indeed, take place in very different circumstances, mainly within the transition of two activities, what makes difficult to model them. When the Unknown class is excluded from the data set the overall performances significantly increase showing again the very good performance of MLN (acc. = 90.5%).

We believe that, in spite of its simplicity, our MLN model is well suited for activity modelling and classification because it covers exhaustively all the possible relations among sensors and activities. Even more, in this particular case it seems sound to assume that a sensor evidence about an activity being performed is independent of the other sensors information given the activity. Consequently, our model gives very acceptable results relying mostly in rules modelling independently the influence of a sensor value to recognise an activity. We show below some of the rules composing the MLN model having the highest weights after the process of weight learning :

4. `liris.cnrs.fr/advene/`

Table 2. Overall Accuracy, precision and recall

Method	SVM		Naive Bayes		MLN	
Accuracy	59.64		66.1		85.25	
	Precision	Recall	Precision	Recall	Precision	Recall
Eating	64.8	71.0	75.1	79.8	90.4	91.9
Tiding Up	40.0	39.0	58.3	56.4	75.1	86.9
Hygiene	55.8	57.4	67.4	61.7	82.6	80.9
Communicating	83.7	71.9	40.9	47.4	100.0	82.5
Dressing Up	32.3	41.1	13.3	11.8	85.3	56.9
Sleeping	57.6	60.1	60.1	74.3	84.7	82.2
Resting	81.5	73.5	82.4	70.8	90.2	92.2
Unknown	10,2	8.2	19.7	17.9	63.6	25.0

$2.17 : PercentageLocationBedroom(win, High) => class(win, Sleeping)$
$2.14 : PercentageAgitationElectricity(win, Medium) => class(win, TidyingUp)$
$1.94 : SpeechStudy(win, High) => class(win, Communication)$
$1.85 : SoundBedroomMic1(win, Medium) => class(win, Dressing)$

This subset of rules characterises the implicit knowledge learnt by the model. Many of the rules having a high weight are concerned with location of the inhabitant, what is easily explained by the fact that most activities are performed in specific rooms. Indeed, several previous studies highlighted the importance of location for activity recognition. However, other information appropriately complement location, as in the above example a medium electricity consumption helps to detect the vacuum cleaner usage and consequently the tidying up activity.

Table 3 shows the confusion matrix for the recognition with MLN taking the unknown class into account. The matrix exhibits high values in the confusion of eating and tidying up since they are often performed in the same location (tidying up is also preformed in the bedroom) and complementary information are similar in some occasions : water consumption, sounds, opening of placard doors. Moreover, the confusion of these two activities represents the main difference in accuracy results among SVM and MLN. In general, confusion is higher when activities share a common location as : Tidying up/sleeping and Tidying up/resting. From these results we can conclude that MLN models better the complementary information to discriminate activities performed under similar settings. Actually, the set of learnt rules in our MLN model characterises any detail of a sensor value that can be useful to recognise an activity. Additionally, note that a weighted set of logic rules is a good model to represent knowledge. Being highly readable and clear, its analysis can help to understand core differences between sensor information related to activities in pervasive environments. It also can be noticed that the use of microphones, which is an essential feature of the Sweet-Home project, proved to be very important for activity recognition. Indeed, detection of speech is fundamental to recognise communication activity. This explains the high accuracy of this activity shown in the confusion matrix. Sound plays also an important role; for instance, the amount of sound in the bedroom can resolve the ambiguity between dressing up and sleeping, which are both performed in the same location.

An important finding drawn from these results is that SVM and Naive bayes models are more easily affected by the unbalanced number of instances. Thus, MLN performs better on classification of classes having few instances such as dressing up and hygiene. In real daily routines, some activities have a short duration and are carried out just a few times, then the capacity of MLN to deal with unbalanced training data seems appropriate to be applied in smart homes.

Table 3. Confusion Matrix for MLN results with Unknown class)

True class/Prediction	1	2	3	4	5	6	7	8
Eating(1)	546	47	0	0	0	1	0	0
Tiding Up(2)	23	373	3	0	1	18	4	7
Hygiene(3)	14	6	114	0	2	0	5	0
Communicating(4)	0	2	0	47	0	0	8	0
Dressing Up(5)	2	8	6	0	29	1	5	0
Sleeping(6)	17	9	11	0	0	221	11	0
Resting(7)	0	22	0	0	1	9	458	5
Unknown(8)	2	30	4	0	1	11	17	21

6 Discussion and perspective

The findings of our study allow us to assert that MLN based models have features that make them more adapted than traditional classifiers for activity recognition. The main advantages of the presented MLN model are : the completeness to represent every detail of sensor values and activity relations ; the readability of the model to understand which sensors explain better the occurrence of an activity ; the possibility of including additional knowledge to the model through logic rules, as it was done with the temporal relations among instances ; and the ability to learn from unbalanced data sets.

Another great advantage of MLN rely on its formal logic nature. In Ambient Intelligent systems, the domain knowledge is often represented as logic formulae (e.g., Description Logic in the case of the OWL format). When possible, this permits translation from one representation to another to performs, for instance, consistency checking of MLN models or addition of relational knowledge as a priory knowledge in the MLN structure learning. In this perspective, the use of a formal domain knowledge description and logic-based recognition method could leads to a higher re-usability of the model learnt in one home to another home. Given the difficulty and cost of acquiring training data in the smart home domain this way seems promising to alleviate the need of large volumes of training data of purely statistical methods.

On the uncertainty modelling, the MLN being a probabilistic model, the outputs of the MLN inference can be exploited and fused with other uncertain evidence in the context-aware decision as it is the case in the intelligent controller described in Section 3. This ability might be highly interesting in other challenging application such as recognition of interleaving activities. In addition, we believe that some characteristics of MLN for activity recognition go beyond the mere classification. Thus, the analysis of set of logic rules composing the MLN and their weights can give us rich information to be exploited during the design of the Sweet-Home context aware system.

Références

1. Tapia, E.M., Intille, S.S., Larson, K. : Activity recognition in the home using simple and ubiquitous sensors. Pervasive Computing **2** (2004) 158–175
2. Storf, H., Becker, M., Riedl, M. : Rule-based activity recognition framework : Challenges, technique and learning. In : Pervasive Computing Technologies for Healthcare, London, UK (2009) 1–7
3. van Kasteren, T., Krose, B. : Bayesian activity recognition in residence for elders. In : 3rd IET International Conference on Intelligent Environments, Ulm, Germany (2007) 209–212
4. Coutaz, J., Crowley, J.L., Dobson, S., Garlan, D. : Context is key. Communications of the ACM **48**(3) (2005) 49–53
5. Aggarwal, J., Ryoo, M. : Human activity analysis : A review. ACM Comput. Surv. **43** (April 2011) 1–43
6. Portet, F., Vacher, M., Golanski, C., Roux, C., Meillon, B. : Design and evaluation of a smart home voice interface for the elderly — acceptability and objection aspects. Personal and Ubiquitous Computing (to appear)
7. Helaoui, R., Niepert, M., Stuckenschmidt, H. : A statistical-relational activity recognition framework for ambient assisted living systems. In : International Symposium on Ambient Intelligence ISAmI, Guimarães, Portugal (2010)
8. Zappi, P., Lombriser, C., Stiefmeier, T., Farella, E., Roggen, D., Benini, L., Tröster, G. : Activity recognition from on-body sensors : Accuracy-power trade-off by dynamic sensor selection. In Verdone, R., ed. : European conference on Wireless Sensor Networks. Volume 4913 of Lecture Notes in Computer Science., Springer (2008) 17–33
9. Fleury, A., Vacher, M., Noury, N. : SVM-based multi-modal classification of activities of daily living in health smart homes : Sensors, algorithms and first experimental results. IEEE Transactions on Information Technology in Biomedicine **14**(2) (march 2010) 274 –283
10. Duong, T., Phung, D., Bui, H., Venkatesh, S. : Efficient duration and hierarchical modeling for human activity recognition. Artificial Intelligence **173**(7-8) (May 2009) 830–856
11. Naeem, U., Bigham, J. : Activity recognition using a hierarchical framework. Pervasive Computing Technologies for Healthcare (2008) 24 – 27
12. Chen, L., Nugent, C.D. : Ontology-based activity recognition in intelligent pervasive environments. IJWIS **5**(4) (2009) 410–430

13. Augusto, J.C., Nugent, C.D. : The use of temporal reasoning and management of complex events in smart homes. European Conference on Artificial Intelligence (2004) 778–782
14. Getoor, L., Taskar, B. : Introduction to Statistical Relational Learning. The MIT Press (November 2007)
15. Natarajan, S., Bui, H.H., Tadepalli, P., Kersting, K., Wong, W. : Logical hierarchical hidden markov models for modeling user activities. In : Proceedings of the 18th international conference on Inductive Logic Programming, Prague, Czech Republic, Springer-Verlag (2008) 192–209
16. Liao, L. : Location-based Activity Recognition. PhD thesis, University of Washington (2006)
17. Tran, S.D., Davis, L.S. : Event modeling and recognition using markov logic networks. In : Proceedings of the 10th European Conference on Computer Vision : Part II, Marseille, France, Springer-Verlag (2008) 610–623
18. Edwards, W., Grinter, R. : At home with ubiquitous computing : Seven challenges. In Abowd, G., Brumitt, B., Shafer, S., eds. : Ubicomp 2001 : Ubiquitous Computing. Volume 2201 of Lecture Notes in Computer Science., Springer Berlin / Heidelberg (2001) 256–272
19. Klein, M., Schmidt, A., Lauer, R. : Ontology-centred design of an ambient middleware for assisted living : The case of soprano. In Kirste, T., König-Ries, B., Salomon, R., eds. : Towards Ambient Intelligence : Methods for Cooperating Ensembles in Ubiquitous Environments (AIM-CU), 30th Annual German Conference on Artificial Intelligence (KI 2007), Osnabrück, September 10, 2007. (2007)
20. Katz, S. : Assessing self-maintenance : activities of daily living, mobility, and instrumental activities of daily living. Journal of the American Geriatrics Society **31**(12) (1983) 721–727
21. Vacher, M., Fleury, A., Portet, F., Serignat, J.F., Noury, N. In : Complete Sound and Speech Recognition System for Health Smart Homes : Application to the Recognition of Activities of Daily Living. Intech Book (2010) 645 – 673
22. Chahuara, P., Portet, F., Vacher, M. : Fusion of audio and temporal multimodal data by spreading activation for dweller localisation in a smart home. In : STAMI, Space, Time and Ambient Intelligence, Barcelona, Spain (Jul. 16-22 2011) 17–22
23. Richardson, M., Domingos, P. : Markov logic networks. Machine Learning **62**(1-2) (2006) 107–136
24. Lowd, D., Domingos, P. : Efficient weight learning for markov logic networks. In Kok, J.N., Koronacki, J., de Mántaras, R.L., Matwin, S., Mladenic, D., Skowron, A., eds. : PKDD. Volume 4702 of Lecture Notes in Computer Science., Springer (2007) 200–211
25. Cortes, C., Vapnik, V. : Support-vector networks. Machine Learning **20**(3) (1995) 273–297
26. Cristianini, N., Shawe-Taylor, J. : An Introduction to Support Vector Machines and Other Kernel-based Learning Methods. 1 edn. Cambridge University Press (2000)
27. Vacher, M., Portet, F., Fleury, A., Noury, N. : Development of audio sensing technology for ambient assisted living : Applications and challenges. International Journal of E-Health and Medical Communications **2**(1) (January-March 2011) 35–54