Activity recognition using temporal evidence theory

Susan Mckeever[a*], Juan Ye[b] , Lorcan Coyle[c] , Chris Bleakley[a] and Simon Dobson[b]
[a]*Systems Research Group, School of Computer Science and Informatics, University College Dublin, Belfield, Dublin 4, Ireland*
[b]*School of Computer Science, University of St Andrews, Scotland*
[c] *Lero, the Irish Software Engineering Research Centre*

Abstract. The ability to identify the behavior of people in a home is at the core of Smart Home functionality. Such environments are equipped with sensors that unobtrusively capture information about the occupants. Reasoning mechanisms transform the technical, frequently noisy data of sensors into meaningful interpretations of occupant activities. Time is a natural human way to reason about activities. Peoples' activities in the home often have an identifiable routine; activities take place at distinct times throughout the day and last for predicable lengths of time. However, the inclusion of temporal information is still limited in the domain of activity recognition. Evidence theory is gaining increasing interest in the field of activity recognition, and is suited to the incorporation of time related domain knowledge into the reasoning process. In this paper, an evidential reasoning framework that incorporates temporal knowledge is presented. We evaluate the effectiveness of the framework using a third party published smart home dataset. An improvement in activity recognition of 70% is achieved when time patterns and activity durations are included in activity recognition. We also compare our approach with Naïve Bayes classifier and J48 Decision Tree, with temporal evidence theory achieving higher accuracies than both classifiers.

Keywords: context reasoning, activity recognition, evidence theory, dempster-shafer theory, temporal, smart home dataset, time

1. Introduction

The ability to recognize and monitor the behavior of occupants is a core premise of smart environments. Sensors embedded in these environments yield data about the occupants' behavior. To recognize activities, a reasoning process uses the sensor data to infer which activities are 'occurring' at a particular point in time. This involves matching sensor data, or a translated meaningful form of sensor data, against a pre-defined model of activities for the environment. Such models may be learned from training data via learning techniques [14], [27], hand crafted using rule-based or ontological approaches [15], or derived from a combination of both [33]. Once matched, an algorithm appropriate to the reasoning technique(s) selects the activities that are occurring.

Making sense of data is a complex task. Sensors are imprecise, the data is noisy, with missing values if sensor failures occur [5]. Learning approaches have been widely used for reasoning with activity information, because of their ability to automate the creation of the activity model from training data and to handle noisy sensor data. On the downside, training data can be difficult and costly to acquire [26]. Like learning techniques, evidence theory manages uncertain information. It also reduces the reliance on training data because it incorporates domain knowledge for evidential reasoning. It is widely used in the fields of medical diagnosis, risk management, robotics, image processing, speech recognition and engineering fault diagnosis [24]. It is recently gaining attention in the smart environment and general pervasive computing domain [9], [36].

---

* Corresponding author: Email: susan.mckeever@ucd.ie

At present, the use of temporal information in the reasoning process for activity recognition is still limited. Time is a natural human reasoning tool that provides knowledge about activities. For example, home-based activities often have a clear time pattern across separate days, such as 'breakfast' in the morning, 'sleeping' at night, and so forth. People's activities can have predictable time durations such as typical time taken for 'preparing a meal' or 'showering'. Activities may also have a sequential pattern, occurring in a particular order. Incorporating such temporal knowledge into activity recognition should allow activities to be more easily differentiated from each other, thus boosting recognition capabilities.

Evidence theory [25] provides a mathematical basis for determining belief in hypotheses (such as activities) by combining evidence from separate sources. Unlike machine learning techniques such as Bayesian schemes, it specifically quantifies and preserves uncertainty encountered in the inference process. Evidence theory provides a theoretically sound basis for incorporating domain knowledge, so it is suited to incorporating temporal knowledge into the activity recognition process.

This paper makes three contributions: (1) The extension of evidence theory to include temporal features. As part of this, a temporal version of Dempster's rule of combination is presented. The temporal version of the rule fuses evidence that is spread over time, as opposed to co-occurring. (2) An evidential reasoning framework that can be used to infer activities from sensor data is presented. In addition to the basic function of inferring activities from sensor evidence, the framework addresses a number of issues that can occur in evidential approaches, such as single sensor dominance. (3) We evaluate our framework using a widely-used third-party dataset: VanKasteren et al.'s home activity dataset [27], described in more detail in Section 4. The effectiveness of temporal extensions is evaluated. Results are also compared to Naïve Bayes classifier and J45 Decision Tree, and to published activity recognition results [31, 32] from other researchers using the same third-party dataset.

This paper is structured as follows: Section 2 explains the framework, covering each of the evidential operations that are used to recognize activities from sensor data. Section 3 covers the general use of the framework with an explanatory worked example. Section 4 contains the evaluation o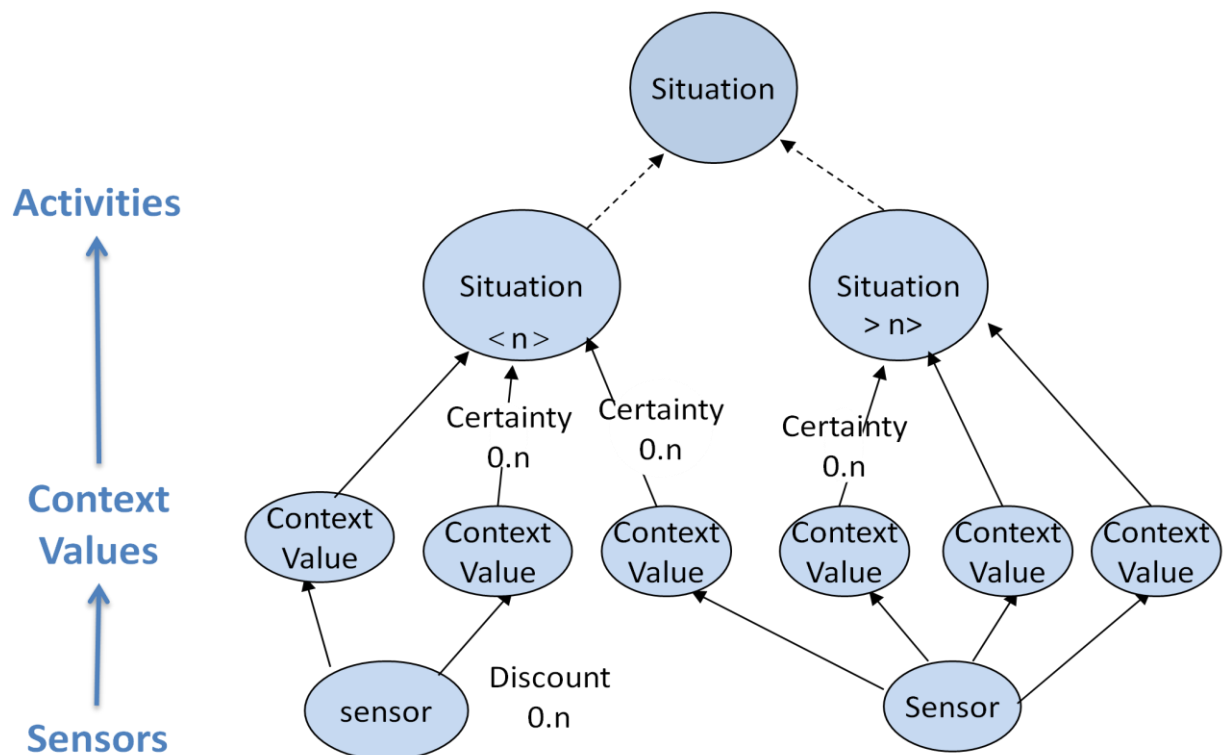f the framework, where temporal evidence theory is used to infer activities in a smart home dataset. Section 5 discusses related work in activity recognition. Summary and future work are presented in Section 6.

2. Evidence Theory for Activity Recognition

Evidence theory is a mathematical theory of evidence [25] which is used to combine separate pieces of information (evidence) to calculate the probability of an event. The basic premise of using evidence theory for activity recognition is as follows: Sensor readings are used as evidence of higher level states within an activity model. These states are fused to determine more complex and higher level states until the level of belief in the activities of interest is determined. For a specific domain such as a smart home, the structure of the activity model must be known in order to support the distribution and fusion of evidence. In section 2.1, situation directed acyclic graphs (DAGs) are explained as a tool for documenting activity models. This is followed in Section 2.2 by a description of the evidential operations that are used in activity reasoning.

2.1. Situation directed acyclic graphs

The situation DAG documents inference knowledge: the evidence sources used, how their evidence is fused, and the hierarchy of activities in the environment. Looking at Fig 1, sensors are the root nodes at the base of the diagram. At the next level up, sensor information is abstracted to one or more context values. Context values are human understandable descriptions of sensor states that are useful in the reasoning process. For example, a binary fridge door sensor may generate two context values of 'fridge used' or 'fridge not used'. Moving up the hierarchy, activities are inferred from one or more context values. Higher level activities may also be inferred from lower level activities. Uncertainty of inference rules is captured numerically as a number between 0 and 1 against the inference path. For example, if the freezer is used 7 out of 10 times in dinner preparation, the inference path from the freezer context value to the 'preparing dinner 'activity will be annotated with 0.7.
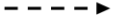
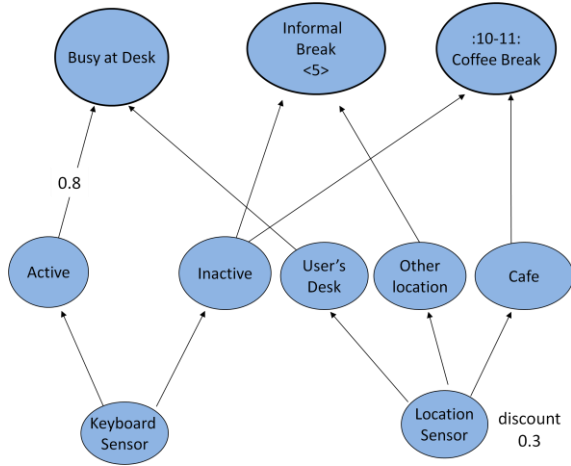Fig 1 Situation directed acyclic graph

Fig 2 Sample Situation DAG

If an activity is determined from a choice of lower level states, the "is a type of" notation is used. For example, a 'leave home' activity might be detected from either of two observed states: 'Front Door used' OR 'No sensors in use'. A sample situation DAG for two sensors and three office situations is shown in Fig 2. The location sensor is discounted by 30%. The keyboard sensor being active is 'usually' indicative that the user is busy at their desk, with 80% certainty.

2.1.1. Temporal features on situation DAGS

Evidence that accumulates over time is represented by a time period enclosed in '< >' brackets within the time-distributed situation node. This number indicates the typical duration of the activity. Where the actual sequence of evidence occurrence is also relevant, the duration is enclosed by '> >' brackets. The time at which an activity occurs, termed absolute time [32] is documented above the activity title between the ': :' symbols. This can be a semantic description such as 'morning' or a numeric specification such as '10-11' (occurs between the hours of 10 and 11 each day). Looking at Fig 2, the 'informal break' situation has a typical duration of 5 minutes and the 'coffee break' situation occurs between 10 and 11 in the morning.

2.2. Evidential Concepts in the framework

Once the situation DAG is defined, evidential operations are used to propagate and fuse evidence from sensor readings through to activity level. A variety of evidential operations from evidence theory are involved in this process. The core concepts of evidence theory are the frame of discernment, mass functions and Dempster's combination rule. These are briefly described. The new temporal extensions to evidence theory presented in this paper are then explained. Additional operations such as evidence propagation that are used in the evidence framework and that are taken from the existing body of research on evidence theory are also included.

2.3. Core Concepts

2.3.1. Frames of Discernment

An evidence source (e.g. sensor) assigns belief across a possible set of choices or hypotheses (e.g. context values). This combined set of n hypotheses $\{\{h_1, h_2 ... h_n\}$ is called the Frame of Discernment, θ. This frame θ has a power set, 2^θ, allowing evidence to be applied to single hypotheses and sets.

2.3.2. Mass functions

Mass functions are used to assign belief from a sensor across its context values (the frame of discernment for the sensor). Each belief assignment is a number between 0 and 1, and total belief assigned across the Frame must sum to 1. Formally, mass functions for evidence sources must satisfy the following conditions

$$m(\varphi) = 0 : \varphi \text{ is the empty set} \qquad (1)$$

$$\sum_{A \subseteq \Theta} m(A) = 1 : A \text{ is a subset of } \theta \qquad (2)$$

An evidence source can quantify its ignorance or uncertainty by assigning belief to the full set of hypotheses. For example, a door sensor detects whether the door is open or closed. The frame of discernment for the door sensor is $\{open, closed, \theta\}$ where θ represents uncertainty, (open or closed). If the sensor has a known accuracy of 80%, and is firing as open, the mass function will assign 0.8 mass to 'door open' and 0.2 mass to θ: $\{0.8, 0, 0.2\}$.

2.3.3. Dempster's rule of combination.

Where multiple evidence sources assign belief across the same frame of discernment, their evidence is fused in order to get a collective picture

of the evidence. For example, if five kitchen-based sensors are used to detect the 'preparing breakfast' activity, their evidence will be combined to determine the belief that the 'preparing breakfast' activity is occurring. Dempster's rule of combination is the defacto fusion rule in evidence theory. It fuses the evidence in agreement, and normalizes out evidence that is in conflict. m_1 and m_2 represent mass functions from two separate independent evidence sources. The fusion of m_1 and m_2 is calculated as:

$$m_{12}(A) = \frac{\sum_{\forall X,Y:X\cap Y=A} m_1(X).m_2(Y)}{1 - \sum_{\forall X,Y:X\cap Y=\varphi} m_1(X).m_2(Y)}$$

Where $m_{12}(A)$ is the fused belief for a hypothesis A.   (3)

2.4. New temporal evidence theory extensions

In the evidence framework, our aim is to include time in the reasoning process. The hypothesis is that inclusion of temporal features in the evidence framework will improve the accuracy of activity recognition. Two temporal features are incorporated into the framework to enhance reasoning: (1) The fusion of time-distributed evidence for activities that have a time duration. This is evidence that is not necessarily happening at the same time, such as the step-by-step triggering of various kitchen sensors when preparing a meal. (2) Using the absolute time at which an activity usually takes place.
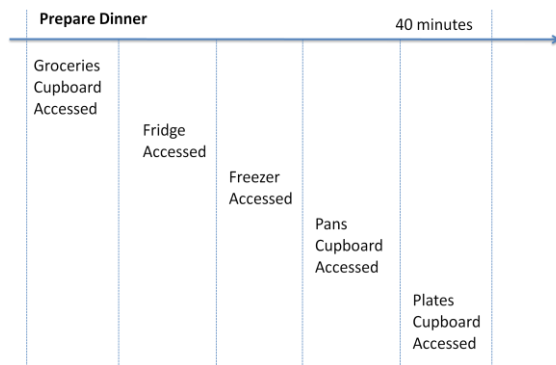

Fig 3 Transitory evidence for an enduring activity 'preparing dinner'

2.4.1. Time-Distributed evidence

Existing approaches in evidence theory for smart homes assume that all evidence is co-occurring. For example, the sensors used to infer kitchen activities in the framework of [9] are fused as if they are all triggered at the same time. In reality, evidence may be spread out over time, co-occurring or not, and in with no particular sequence as shown in Fig 3 and Fig 4. Looking at Fig 3, a 'preparing dinner' activity may typically endure for about 40 minutes, with indicative evidence of 'grocery cupboard used', 'fridge used' and so on. None of this evidence is necessarily occurring at the same time. The events may occur in any sequence, with no particular order expected. Events may co-occur and/or occur separately, with gaps between events, such as the example shown in Fig 4. The user opens the plate cupboard and fridge in the same sampling period, *then* uses the pans cupboard and freezer, then retrieves groceries. Such evidence for a higher level state that does not endure for the full time duration of the state is termed *transitory* evidence.
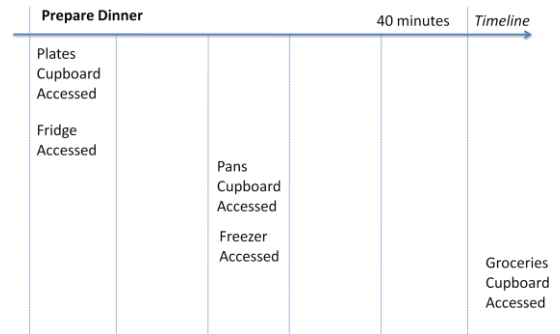

Fig 4 Transitory evidence with some evidence co-occurring

Activities with duration that are inferred from transitory evidence are documented on the situation DAG, denoted using the '<>' identifier under the activity name. During the inference process, the occurrence of any evidence for that activity will trigger the start of that activity duration. Looking at the 'preparing dinner' example in Fig 5 (based on the Fig 3 example), if any of the groceries cupboard, fridge, freezer, pans cupboard or plates cupboard sensors are fired, the reasoning system will 'start' the dinner activity. The lifetime of the triggered sensor evidence for that activity will be extended to last for the activity duration stored for that activity. As inference continues over time, the lifetime of any further evidence for the activity will be extended for the duration that is left of the activity (activity

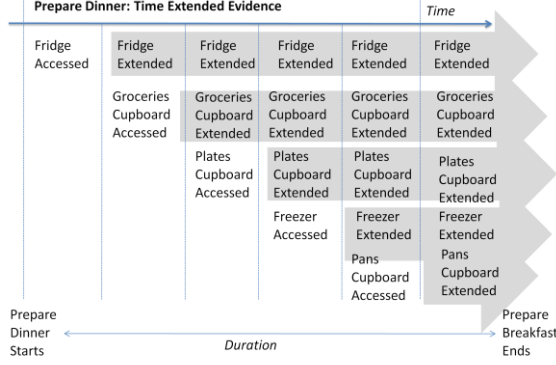duration less elapsed time). Once the full duration of the activity is reached, the evidence will expire.

Fig 5 Time extension of transitory evidence: 'preparing dinner'

By extending the lifetime of the evidence, at any point in time, the evidence sources can be fused as if they are co-occurring.

Sensors that provide transitory evidence for more than one activity will trigger more than one activity to start. For example, 'preparing breakfast' and 'preparing drink' are also inferred from the fridge sensor. If this fires, the duration will kick off for 'preparing diner', 'preparing breakfast' and 'preparing drink'. The 'fridge used' context value lifetime for each of the three activities will be separately extended for the lifetime of each of the three durations. That is, it will expire after 3 minutes as evidence of 'preparing drink', after 15 minutes for 'preparing breakfast' and after 40 minutes for 'preparing dinner'.

If multiple simultaneous sensor events happen at the same time, where the events are evidential of different activities, the evidence is allocated to the relevant activity as per the situation DAG. For example, if a toaster sensor activates in the kitchen in the same sampling period as a sensor in the bathroom, evidence will be allocated to the 'preparing breakfast' and 'showering' activities respectively. The interpretation of these activities as co-occurring or not will be environment specific. If, for example, there are multiple inhabitants of the house, both 'breakfast' and 'showering may be recognized as co-occurring as it is possible that two activities happening at the same time. In this case, a belief threshold may be used to filter situations, with situations that have belief levels exceeding the threshold as 'occurring'. In an environment where activities can only occur one at a time, as in the case of the smart home dataset used in our evaluation, the activity with the greatest evidence (highest belief) is deemed occurring.

To use time extension of transitory evidence in the evidence framework, definition of mass and the fusion rule for masses from multiple sources require this time extension. Formally, a frame of discernment, Θ, contains one or more hypotheses, h, of time duration $t_{dur,}$. Belief from evidence sources that provide transitory evidence are assigned a lifetime of the duration of the enduring hypothesis. If the hypothesis has already been detected by earlier evidence, the lifetime of the mass is the remainder $t_{rem.}$ of the duration, where remainder is calculated as hypothesis duration less elapsed time, $t_{dur\text{-}}$ $t_{elapsed.}$. When mass is assigned to hypothesis, h, of time duration, t_{dur} at time t, the mass assigned to h at time t, $m_t(h)$ will continue to exist for the remaining time t_{rem} of the hypothesis duration. This 'extended' mass, $m_{t+t_{rem}}(h)$ for hypothesis h that exists during the remaining duration of h is represented as:

$$m_{t+t_{rem}}(h) = m_t(h) \qquad (4)$$

Where

$$t_{rem} = t_{dur} - t_{elapsed}$$

To fuse extended mass, the combination rule is used. To fuse evidence for two extended masses for enduring hypothesis, $h_{t_{dur}}$ during their lifetime $t_{rem,}$, fuse the evidence at each point in time, t, as if they are co-occurring. Dempster's combination rule for two transitory extended evidence sources for a hypothesis, $h_{t_{dur}}$ is:

$$m_{12t_{rem}}\left(h_{t_{dur}}\right) =$$

$$\frac{\sum_{\forall X,Y:X\cap Y=h_{t_{dur}}} m_{1t_{rem}}(X).m_{2t_{rem}}(Y)}{1 - \sum_{\forall X,Y:X\cap Y=\varphi} m_1 t_{rem}(X).m_2 t_{rem}(Y)}$$

Where

$$t_{rem} = t_{dur} - t_{elapsed} \qquad (5)$$

2.4.2. Using absolute time

Activities in the home often have an identifiable absolute time, such as taking breakfast in the morning, sleeping at night time. Evidential reasoning can easily incorporate domain knowledge, so is suited to the inclusion of absolute time as part of the inference process. This can be done by treating 'time' as a virtual evidence source with its own mass function. A virtual time will be included on the situation DAG and inferences rules used to connect the time context values to activities. This will be useful if there is some uncertainty involved such as 'breakfast usually takes place in the morning". If no uncertainty is included, absolute time can be used directly to filter the set of possible activities that can be occurring for a particular point in time t. For example, if "preparing breakfast" 'always' takes place in the morning, the activity will only be considered as possible to occur outside of the times defined as within 'morning'.

2.5. Additional Evidence Concepts for activity recognition

For the evidence framework, the following additional evidence operations are used to support activity recognition: evidence propagation, Murphy's alternative rule of combination, alternative evidence combination and sensor discounting.

2.5.1. Evidence propagation

Evidence propagation is used to transfer evidence from context values through to higher level activity beliefs. Compatibility relations [16] define maps between frames of discernment, by defining which hypotheses in the frames are true simultaneously. Evidence propagation, as used by [9], is then used to transfer evidence along compatible paths defined using compatibility relations. For example, in the smart home dataset used for our evaluation, a bathroom door sensor has a frame of discernment {opened, closed, θ}. The opening of the bathroom door indicates the 'showering' activity which is part of a frame of discernment {showering, ¬showering, θ}. Bathroom door 'opened' is compatible with 'showering' (i.e. they are both true simultaneously) and so on for the remaining elements in both frames. The mass of belief for bathroom door 'opened' is propagated as belief to the 'showering' activity.

2.5.2. Murphy's Alternative Combination Rule

Using Dempster's rule of combination, a single contradictory sensor can overrule other agreeing sensors [20]. If the conflicting sensor assigns all of its belief to a contradictory hypothesis, the evidence from the others sensors is lost. Binary sensors are particularly affected by this because such sensors tend to assign all belief to a single hypothesis (i.e. 0 or 1). To overcome this, Murphy proposed an alternative rule of combination [20]. Evidence is averaged prior to combining it using Dempster's rule of combination. This eliminates the dominance of a single sensor. Use of Murphy's combination rule will also eliminate Zadeh's paradox [35]. This is a well documented problem with Dempster's rule of combination whereby a minority opinion can be selected from conflicting evidence sources.

2.5.3. Alternative evidence combination

For scenarios where evidence sources are combined in an 'OR' scenario, the highest belief from the evidence sources will be selected. For example, a 'leave home' activity may be detected as 'front door used' OR 'all sensors inactive'. The belief of 'leave home' will be the maximum belief assigned to either 'front door opened' or 'all sensors inactive'. Formal representation of this maximization approach is described in [8].

2.5.4. Sensor discounting

Evidence theory uses a discount factor to weight evidence sources [25]. Discounting is useful when quality information about a sensor is available. A sensor discount is applied as a weight between 0 and 1. For example, a door sensor that is 80% reliable will have a discount of 0.8 applied to its evidence. When a sensor is discounted, the uncertainty of its evidence increases. The formal representation of sensor discounting is explained in [25]. The combination of static and dynamic quality information via sensor discounts is explained in more detail in [18]

3. Applying the Evidential Framework to Activity Recognition

To apply the evidential framework to real-life smart environments, we need to capture the activity model in a Situation DAG. Given a set of sensor readings,

an activity will be inferred in the following steps: (1) calculating sensor mass functions; (2) propagating evidence to activities; (3) fusing multiple pieces of evidence; and (4) determining the occurring activities according to their belief scores.

To document the situation DAG, knowledge is needed about which sensors are used and how sensors map to activities via inference rules. This knowledge can be obtained from domain knowledge of experts and users. Training data, if available, can also be used to supplement the knowledge. Sensors and interpretation of sensor readings is the domain of experts. User interviews or observation may be used to glean information about how activities are conducted, time patterns of activities and typical durations. Uncertainty in inference rules can be defined when users identify uncertainty such as "I sometimes use frozen food for making dinner". This can be quantified informally, or limited amounts of training data if available can be used to quantify the uncertainty of the inference rule. For example, in the evaluation of the framework, a third of the dataset is used to generate mass functions, and two thirds held back for training.

3.1. Activity Recognition Worked Example

Using the evidential operations described, a simple worked example is provided from the smart home dataset used in our evaluation. For each activity, a frame of discernment `{activity,¬activity,θ}` is defined. Table 1 shows two timeslices from the dataset, during which the occupant is preparing a drink. The fridge and dup sensors are used to detect the 'preparing drink' activity. The fridge sensor has a frame of discernment `{FridgeUsed,¬FridgeUsed,θ}` and the cup sensor `{CupUsed,¬CupUsed,θ}`. The occupant always uses the fridge and 'usually' uses a cup, with 80% frequency of using the cup for a drink. Typical duration of the 'preparing drink' activity is three minutes (obtained from user interviews, observation or training data), with both fridge and cup as transitory evidence sources. The inference steps for each timeslice are as follows:

Table 1 Sample timeslice evidence for 'preparing drink' activity

Timeslice	Sensor events	Preparing Drink Evidence
9:49	Fridge, Cup	Fridge, Cup (0.8)
9:50	Fridge	Fridge, Cup (0.8)

At a time of **9:49**, the fridge and cup sensors fire. Both of these events are indicative of the 'preparing drink' activity, which is not currently in progress. The elapsed time of drink is set to 1 minute (length of timeslice).

Step 1: *Use sensor mass functions to obtain context value beliefs*. Both the fridge and cup sensors fired:

```
{FridgeUsed=1, ¬FridgeUsed=0}
{CupUsed=1, ¬CupUsed=0,θ=0}
```

Step 2*: Transfer belief from context values to activities*. The fridge and cup sensor evidence is propagated to the 'preparing drink' frame using compatibility relations and evidence propagation:

```
{FridgeUsed=1,¬FridgeUsed=0}
→{PrepDrink=1,¬PrepDrink=0}
```

A cup is used with certainty of 0.8 when preparing a drink, with the remainder classified as uncertainty.

```
{CupUsed = 1,¬CupUsed=0,θ=0}
→{PrepDrink=0.8,¬PrepDrink=0,θ=.2 }
```

Step 3: *Combine evidence using Murphy's combination rule* to obtain belief for the 'preparing drink' frame. As Murphy's version of the combination rule is being used, the evidence is averaged prior to combining:

```
{PrepDrink=0.9,¬PrepDrink=0,θ=0.1}
```

Then, the averaged evidence is fused using Dempster's rule of combination, to obtain belief for the 'preparing drink' frame of discernment at time 9:49 as:

```
{PrepDrink=0.99,¬PrepDrink=0,θ=0.01}
```

At the next timeslice **9:50**, the fridge sensor fires again.

```
{FridgeUsed=1,¬FridgeUsed=0}
→{PrepDrink=1,¬PrepDrink=0}
```

The cup sensor does not fire, but the cup context values from the previous timeslice are extended as they are within the 3 minute duration of the 'preparing drink' activity. The lifetime, t_{rem}, of the cup context values is calculated as the 'preparing drink' time duration (3 minutes) less the elapsed time of 'preparing drink' (1 minute), as per equation (4):

```
{CupUsed=1,¬CupUsed=0,θ=0}
→{PrepDrink=0.8,¬PrepDrink=0,θ=.2}
```

Using the extended evidence of the cup and the fridge sensor, the evidence is fused using the temporal version of Dempster's combination rule in equation (5). Evidence is averaged prior to fusion as per Murphys' variation on the combination rule, resulting in belief at time 9:50 for 'preparing drink' as

```
{PrepDrink=0.99,¬PrepDrink=0,θ=0.01}
```

This inference process is also conducted for all other activities in the smart space. At time t, the activity with the highest belief is selected (assuming that only one activity can be happening at one time). If more than one activity can be occurring at the same time, a belief threshold approach can be used to establish which activities are occurring.

4. Evaluation

Evidence theory with temporal extensions for activity recognition is evaluated with the use of a third party smart home dataset, captured in a real-life home environment. The main purpose of our evaluation is check whether accuracy of activity recognition is improved using temporal features of evidence theory, when compared to not using temporal features. Another aim is to compare inference results using evidence based inference to those using established learning techniques. To meet these aims, three experiments are run. In the first experiment, activity recognition accuracy using evidence theory with absolute time, versus not using time is conducted. Results will show an improvement in accuracy with the use of absolute time. Secondly, activity recognition accuracy using

time-extended evidence (and absolute time) versus absolute time only is compared. Results will demonstrate that the time-extended evidence approach recognizes activities that are derived from transitory evidence more accurately than without use of time extension. Finally, evidential reasoning using both time-extended evidence and absolute time will be compared to two classic machine learning techniques, Naïve Bayes and J48 Decision Tree. Absolute time is added to the training and test sets for both learning techniques to allow a more direct comparison. Results will show that the temporal evidential framework outperforms these two techniques when limited training data is used.

4.1. Dataset

In order to evaluate our temporal extensions, we required a smart home dataset that contains situations with discernible time durations over a time period. Our requirement was to use a real-life smart home dataset rather than one captured in a laboratory environment. We also wanted to use a dataset that has been used by other researchers to test activity recognition techniques, so that we can compare our evidential approach with existing published results. Availability of published smart home datasets is still a challenge in the pervasive computing field, particularly where published results are desirable, using transparent, repeatable methodologies [34]. The Placelab dataset [10] has been used extensively by researchers for testing recognition techniques. However, researchers typically use subsets of the dataset, making it difficult to compare results when a full cycle, such as a month, is under examination as in our temporal evidence theory evaluation.

VanKasteren's dataset [27] is a public third party dataset that originates from the intelligent autonomous systems group in the University of Amsterdam. It has been widely used by other researchers for smart home experimental evaluations [31], [13], [27], [37], [28]. The data is recorded in the home of a 26 year old man over 28 days in his apartment. Annotation was done by the occupant via voice recognition from a headset. Over the 28 days, 2120 activities were annotated, resulting in 245 activity instances. Seven different activities were recorded: 'sleeping', 'leave home' 'toileting', 'showering', 'sleeping', 'preparing breakfast', 'preparing dinner' and 'preparing a drink'. Only one activity is defined as occurring at any point in time.

14 state change digital sensors were installed in doors, kitchen cupboards and kitchen appliances. Each sensor transmits binary values only. A '0' indicates the sensor is not in use, a '1' indicates that the sensor is firing, such as a cupboard sensor indicating that the cupboard is open.

Clearly Van Kasteren's data set provides only a small and limited view onto the activities occurring in the home, and a larger sample would be desirable. It does, however, provide a common and widely-used reference for comparing different approaches to activity recognition. We note in passing that there are very few data sets available for such comparative study: a point to which we return in section 6.

4.2. Set up

Inference knowledge is used to establish the situation DAG. In a real-life environment, the relationship between sensors and activities can involve user interviews. Questions such as "what do you do when preparing breakfast" will establish which sensors are being triggered for each activity. As we are using a generated dataset, we use a limited amount of training data, combined with common sense domain knowledge to establish our situation DAG. A common practice in machine learning is to use two thirds for training with a third

for testing. These proportions are reversed to illustrate the limited dependence on training data. Using a third of the dataset, the sensors that are triggered for each activity are identified.

In addition, common sense domain knowledge of home activities enables the following assumptions: activities in the kitchen (breakfast, dinner, drink) only involve sensors in the kitchen; No occupant activated sensors will be firing when 'leave home' and 'sleeping' are happening'; door sensors are of interest when their state is changing, but a door left open (with an ongoing value of '1') is not useful for inference. A situation DAG is established for each activity. The situation DAGs for 'preparing breakfast' and 'preparing a drink' are shown in Fig 6. Inference rule uncertainty is annotated on the DAG, but actual values will depend upon which portion of the dataset is used for training so will be assigned during experiment runs. No sensor discounting is used because there are no known quality issues with the sensors for the dataset.

To use the time series data, it is first divided into timeslices of equal duration. This timeslice duration is long enough to be discriminative and short enough to provide high accuracy labeling results [27]. Timeslices where no activity is annotated are excluded. A total of 25,680 annotated timeslices of data are generated, where each slice captures the sensor values and annotated activity that occurred during that minute.
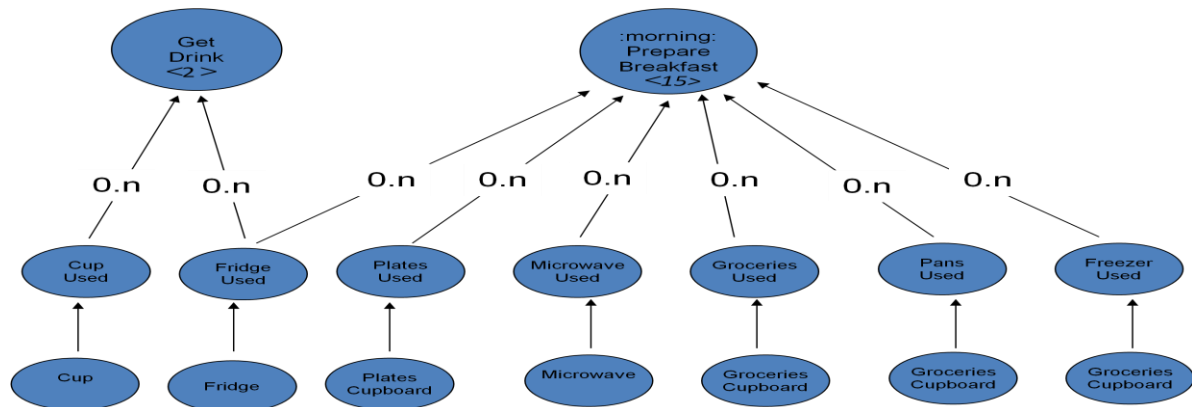

Fig 6 Situation DAG for 'preparing drink' and 'preparing breakfast'

Once the situation DAG has been established, the inference process analyses sensor readings for each timeslice as follows:

At time t:

- Sensor mass functions define belief in context values based on available sensor readings for time t.
- Evidence for any activities with time duration and transitory evidence is extended by the remaining lifetime of the activity.
- Evidence from context values are propagated to higher level activity states
- Evidence is fused where multiple context values or activities are used to detect higher level states.
- The activity with the highest belief is deemed to be occurring, assuming only one activity is happening at any one time.
- The durations of all activities 'in progress' is reduced by the timeslice length so that time-extended lifetimes are updated.

This process continues for the next time: t + *timeslice*, to produce continual activity recognition spread over time.

4.2.1. Methodology

The timesliced dataset is divided into thirds. Using cross validation, each third is used for generating mass functions and inference rule uncertainty, with the remaining two thirds of the data held back for testing as explained in Section 4.1. Table 2 shows the inference rule uncertainties generated for the 'preparing breakfast' activity for one of the dataset thirds. Looking at the table, the pan cupboard sensor triggering is 0.3 indicative of the 'preparing breakfast' and 0.7 of uncertainty.

Table 2 Sample inference rule certainties for 'preparing breakfast'

Context Value	Inference Rule certainty
Microwave	01
Cups	0.1
Fridge	1.0
Plates	1.0
Pans	0.3
Freezer	0.4
Groceries	0.6

In the dataset, only one activity is occurring at any point in time. Therefore, the activity with the highest belief is deemed to be occurring (subject to absolute time filtering). If two or more activities have equal belief, the activity with the least uncertainty is selected.

Table 3 Absolute Times for Dataset Activities

Activity	Absolute time
Breakfast	Morning
Dinner	Evening
Showering	Morning
Leave home	Daytime
Sleeping	Nighttime

For experiments where we compare with other learning techniques, we divide the data in two ways (1) Cross validation, holding back one third of the data for testing, two thirds for training. This is to illustrate the use of 'limited' training data for evidence theory (2) The commonly used 'leave one day out' technique for time series data [29], where one day is used for testing, and the remaining 27 days for training.

Three measures are used to identify the performance of activity recognition (1) Precision is the ratio of the times that an activity is correctly inferred $N_{infCorr}$ to the times that it is inferred N_{inf} (2) Recall is the ratio of the times that a situation is correctly inferred $N_{infCorr}$ to the times that is actually occurs in the dataset N_{act}:

$$Precision = N_{infCorr}/N_{inf}$$
$$Recall = N_{infCorr}/N_{act}$$

(3) F-measure is the weighted mean of precision and recall and is used to summarize inference accuracy.

4.3. Experiment 1 – Absolute Time of Day

In this experiment, the impact of using absolute time in the inference process is examined. The absolute times for activities are shown in Table 3. 'Preparing drink' occurs at various times during the day and night so no particular time pattern is evident. Fig 7 shows the inference results comparing evidence theory used without absolute time, and with absolute time. The use of absolute time improves the inference accuracy for all activities

that have an absolute time, as listed in Table 3. 'Preparing drink', for which absolute time is not used is slightly lower. 'Leave home' and 'sleeping' activities are derived from the same evidence (no sensors active), so cannot be distinguished unless time is used (i.e. nighttime for 'sleeping', daytime for 'leave house'. Therefore, when absolute time is not is used in inference, both 'leave house' and 'sleeping' have equal belief and certainty and are indistinguishable. 'Leave house' is selected by default and 'sleeping' activity is never recognized. When absolute time is included, 'sleeping' activity can be inferred.

4.4. Experiment 2 – Time extension of Evidence

In this experiment, the impact of time-extended evidence for the duration of the higher level activity is examined. Durations are used for 'breakfast', 'dinner', 'drink', 'showering' and 'toileting' as each of their context events can be spread over time. No sensor is usually fired during 'leave home' and 'sleeping' activities so no time extension of evidence is used for these activities. Activity durations are calculated as the average of the
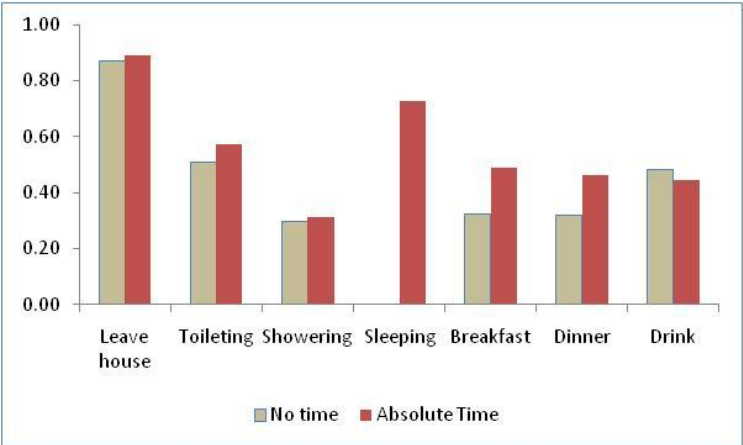

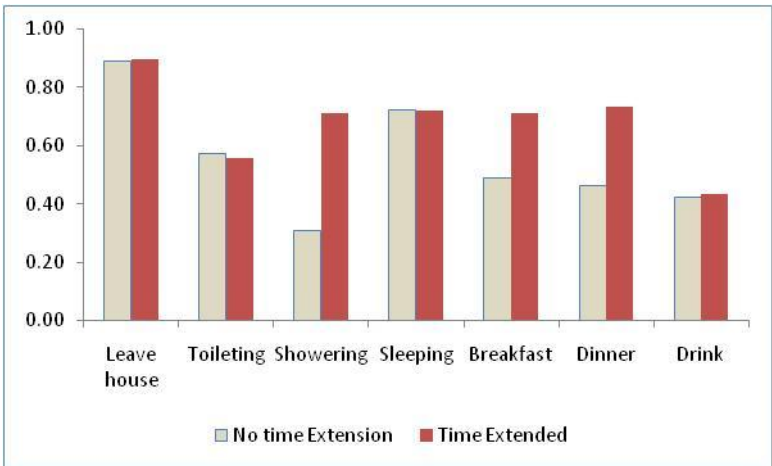Fig 7 F-measure Using 'No Time' versus 'Absolute Time'

Fig 8 Comparison of F-measure using time-extended evidence versus no time extension

activity duration from the training data sample. Alternatively, a user interview might include questions such as "how long does it typically take you to prepare breakfast?"

Fig 8 compares the inference results of using extended evidence against not. Absolute time is included in both. The result was that recognition accuracy improved for four out of the five enduring activities. Time extension is not used for 'leave house' and 'sleeping' activities, and as expected, their inference accuracy is almost identical. For the remaining five time-extended activities, the biggest improvements is shown in 'showering', 'preparing breakfast' and 'preparing dinner'. These activities are longer in duration than the 'preparing drink' and 'toileting' activities, so their evidence is sparser throughout the duration. Therefore, they benefit more from the extension of their transitory evidence. The 'toileting' activity recognition actually decreases very slightly with the use of time-extended evidence. This is because the sensors used in 'toileting' overlap with those for 'showering' and the two activities were often performed sequentially.

Table 4 Comparison of average F-measure for evidence theory with no time, absolute time and extended time

	No Time	Absolute time	Time Extension (and Absolute)
F-measure	0.40	0.56	0.68

The impact of time on evidential reasoning is summarized in Table 4. This shows average F-measure for all activities when no time is used in reasoning, when absolute time is used, and when both time extension *and* absolute time are used. F-measure improves by 70% with the use of both time reasoning techniques.

4.5. Experiment 3 – Comparison with other inference techniques

In this experiment, temporal evidence theory (using absolute time and time-extended evidence) is compared to two machine learning techniques: Naïve Bayes and J48 Decision Tree). Absolute time is incorporated as an attribute into the datasets for Naïve Bayes and Decision Tree to make the comparison as equal as possible. The comparisons are done in two ways (1) using limited training data

(one third) with the remainder held back for testing (2) Using a 'leave one day out' cross validation approach as described in the methodology. As shown in Table 5 and Fig 9, with the use of one third training, time-extended evidence theory outperforms both Naïve Bayes and J48 Decision Tree. The gap is greatest for activities with longer duration: 'preparing dinner', 'preparing breakfast' and 'showering'. Minimal difference is shown for 'preparing drink' which is just 3 minutes long, so benefits less from time extension of evidence than the longer activities.

Table 5 Average F-Measure for time-extended evidence, Naïve Bayes and J48 decision tree using one third training data

	Average F-measure
Time -extended Evidence	0.68
Naïve Bayes	0.49
J48 Decision Tree	0.34

For the second approach, using 'leave one day out', the results as shown in Fig 10 from the three techniques are much closer than when one third training data is used. Time-extended evidence outperforms or matches the two learning approaches, with greater performance shown on two of the enduring activities, 'showering' and 'preparing dinner'. The average F-Measure distributions differ to those from the one third training data results because there are days on which some activities do not occur, recording a zero F-Measure for the activity for that day. This effect applies equally to all three techniques so does not affect the relative performance of the techniques.

4.6. Discussion

This section shares the experience of using our evidence theory with temporal features, and discusses its strength and limitation.

4.6.1. Impact of absolute time

Greater time patterns will yield greater activity recognition. Our first experiment shows that the use of absolute time in our evidence theory inference improves the accuracy of activity inference. In the smart home dataset, five of the seven activities had an identifiable time pattern. Inference accuracy improved for all five activities when absolute time was used, with an improvement of average F-

measure of 40% overall. The usefulness of absolute time depends on how much activities follow an identifiable time pattern. Activities in VanKasteren's dataset occur at regular times throughout the day so using absolute time is beneficial. Greater time patterns will yield greater activity recognition. Time patterns will be applicable in home environments where people have an identifiable pattern of when they take their meals, shower, and so forth.

4.6.2. Impact of time-extended evidence

Longer duration activities have more sparsely spread out evidence, so they will benefit from extension of evidence to cover 'gaps' in evidence during the activity. Our second experiment tested the impact on activity recognition accuracy when time extension of transitory evidence was used. Average F-measure improved by 28% when extended time evidence was used in addition to absolute time, when compared to using absolute time only. Recognition accuracy improved for four out of the five enduring activities, with the greatest improvement seen for the longer duration activities.

4.6.3. Temporal evidence theory versus other inference techniques

With the incorporation of temporal knowledge, evidence theory outperforms the classic machine learning techniques when they are purely training-based. In our third experiment, the temporal evidence approach was compared with two classic machine learning techniques, Naïve Bayes and J48 decision tree. The experiments were run using limited training data (one third, cross validated) and then using a 'leave one day out' cross validation approach. Absolute time was included in the data for both Naïve Bayes and J48 Decision Tree. Looking at Fig 9, our results showed that temporal evidence theory clearly performed better than the other two techniques when limited training data was used. This improvement was much less marked when using the 'leave one day out' approach as shown in Fig 10, although evidence theory is still the best performing of the three techniques.

Since evidence theory is suited to the incorporation of domain knowledge, this result is encouraging. Evidence theory will be useful when training data is not easily available and where domain knowledge can be gleaned from expert knowledge and user knowledge. These sources can be used to obtain inference knowledge in a piecemeal approach, with users providing information on absolute times, activity descriptions and durations, and experts providing knowledge of sensor mass functions and sensor quality.

4.6.4. Comparison with published results

Temporal evidence theory inference results were also compared to those published by VanKasteren et al. in [27]. They use Hidden Markov Models, to recognize occurring activities. The evaluation method is the 'leave one day out' technique. They use a class accuracy measure calculated as average percentage of correctly recognized timeslices per activity. Using VanKasteren et al.'s class accuracy measure calculation and 'leave one day out' evaluation technique, time-extended evidence achieves an average class accuracy of 69% against VanKasteren's HMM class accuracy of 49.2%. This comparison is made using the raw sensor representation published with the VanKasteren dataset. VanKasteren et al.'s work also uses three other more informative sensor representations that encode temporal information. The highest accuracy achieved is a class accuracy of 79.4% using a 'changepoint plus last' sensor representation as described in [27]. However, since raw sensor representations are published in the dataset, this evaluation compares directly with results from raw sensor representation only.

Ye [31] uses situation lattices to infer activities in the VanKasteren dataset. Ye's results yield a class accuracy of 88.3% using raw sensor representations and the 'leave one day out' cross validation technique. This is higher than the results from the temporal evidence framework (69%) and VanKasteren et al.'s HMM results (49.2%). Ye's lattice method includes absolute time in the inference method, and combines both training and domain knowledge. However, timeslices in which no sensor changes take place are excluded. These timeslices are hard to infer because of the lack of sensor information so the dataset is likely to yield improved results to some degree.
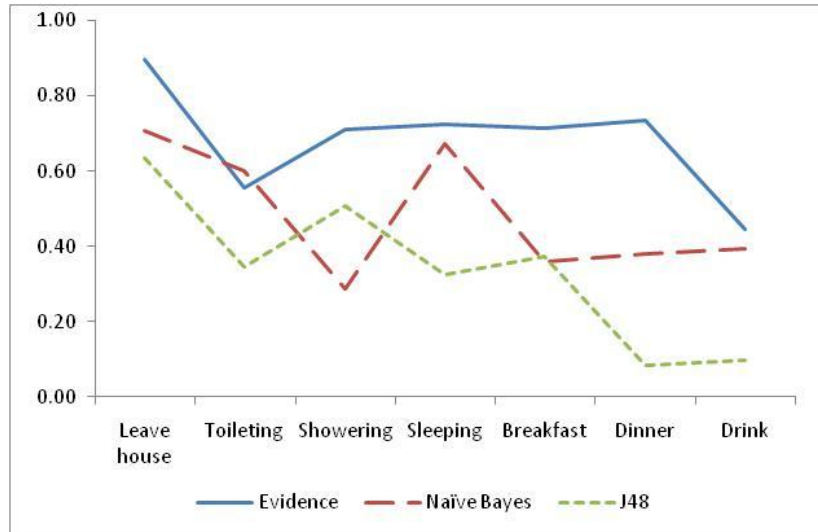
Fig 9 F-measure by activity for time-extended evidence, Naïve Bayes and J48 Decision Tree using one third training
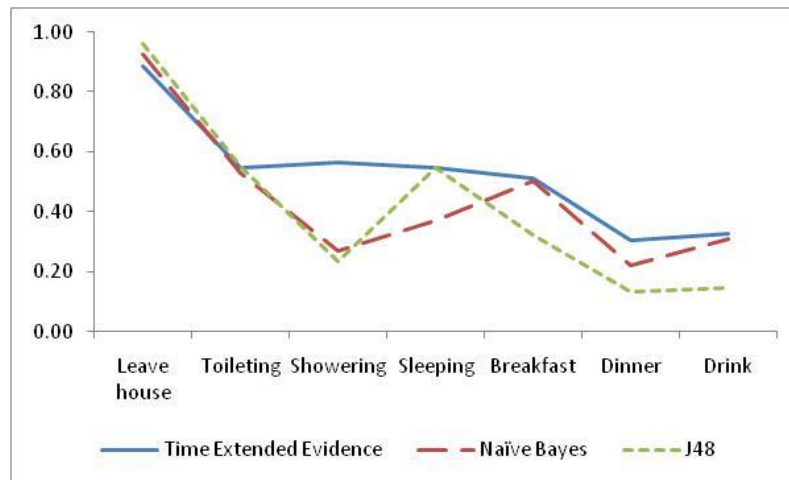

Fig 10 F-measure by activity for time-extended Evidence, Naïve Bayes and J48 using Leave One Day Out

4.6.5. Summary,

The temporal aspect of evidence theory is useful for data where there is a discernable time pattern of activities (absolute time) or where transitory evidence is used to determine enduring activities (time-extended evidence). Evidence theory, in general, is good for scenarios where training data is at a premium, and where domain knowledge is available from experts and users. It is less suitable for scenarios where mapping of sensors to activities cannot be hand crafted or easily observed.

5. Related work

Related work in the field of activity recognition is covered in two parts (1) the application of temporal reasoning to activity recognition and (2) general activity recognition techniques, including evidential approaches.

5.1. Temporal reasoning for activity recognition

In recent years, the use of temporal knowledge has been employed in both learning and rules-based approaches to enhance activity recognition.

Looking firstly at learning approaches, Hidden Markov Model (HMM) is a statistical learning technique that has been widely applied to activity recognition [27], [19], [4]. HMMs take account of sequences of states. The system is assumed to be a Markov chain that is a sequence of events. The probability of each event is dependent on the event immediately preceding it. Modayil et al. [10] use an interleaved HMM to better predict transition probabilities by recording the last object observed in each activity. This approach achieves very low error rates, though it requires an approximation for the inference

process. Clarkson et al. [4] used HMMs for context recognition methods for wearable computers with the means of a wearable camera, and environmental audio signal processing. For a simple set of situations, they achieved recognition rates between 85 and 99%. They conclude that their results are not exposed to any drift from the trained models and that the contexts used are simple. As discussed in this work, VanKasteren et al. [27] use HMM for activity recognition of smart home activities. Their recognition accuracies ranged from a class accuracy of 49.2% using raw sensor representations to 79.4% using a sensor representation that contains more temporal information than the raw sensor state. HMMs are usable where training data is available to build a statistical model of the activity model for the environment, and where state sequences have a discernible pattern.

HMMs consider short term sequences only, based on the previous state. Choujaa and Dulay [3] observe that long term sequences (such as activities from an earlier part of the day) are also useful, and employ both short term and long term sequences in their activity inference approach, using a probabilistic framework obtained from training data. Their approach also caters for gaps in the data. They evaluate on a mobile phone dataset. With eight weeks of training, user activities can be inferred with over 70% accuracy when every other hour is missing in the day.

Jakkula and Cook [11] apply temporal knowledge about activities in order to detect anomalies in real time in a smart home, as a precursor to monitoring resident safety. They use training data to discover frequent sequences of sensor patterns, and temporal relations between sequences. Their approach supported the detection of anomalies occurring over a day, using 59 training days from their MavHome smart home environment.

Palmes et al. [21] use an object data mining approach to activity discovery that does not assume any particular sequence of activities. They note that activities may have a distinct series of steps but with no particular sequence. They note that in such cases, relying on sequence of events for activity recognition may significantly limit the accuracy and applicability of models that rely particularly on object sequence.

Ye et al. [32] use a situation lattice as a classifier method for activity data. The lattice can utilize both training data to establish the lattice and domain knowledge to tune the lattice. They use both absolute time and activity sequences in inference. Preliminary experiments show that more accurate classifiers are produced when absolute and relative time is used.

In additional to using temporal knowledge with learning approaches as described, temporal operators have been incorporated into rule-based approaches, such as the work of Augusto et al. [2]. In their reasoning approach, they use time dependent rules that consider the sequence and co-occurrence of events. Jakkula and Cook [12] use Allen's temporal logic relations [1] as the basis for defining temporal rules across activities. They then compare the predictive accuracy of activities with and without the temporal rules, noting an improvement when temporal rules are applied.

Time has been used directly or indirectly to treat the certainty of sensor readings. Sensor readings are usually time-stamped so time can be applied as part of a decay function, as done by [23] and [17]. For an evidence based model, use of decay for sensor readings can be done via the sensor mass functions as described in [17].

Interestingly, Partridge et al. [22] study the applicability of time-use study data for ubiquitous activity-inference systems. The time-use study covers all the human activities performed by the participants over a certain period, which could be a day or weeks. Partridge et al. analyse how well the time-use study predicts activities using time, location, demographics, and previous activity. They argue that the study data are useful in the sense that they enable cheap and comprehensive classifiers. One of their results is that, when combined with absolute time, the accuracy of activity prediction is increased up to 70%.

5.2. General approaches to activity recognition

Bayesian classifiers recognize higher level context states, based on the probabilities of lower level causal contexts in the network, and there are various examples in the literature of their use for inference [23] [14]. Ranganathan et al. [23] used a Bayesian network to determine the activity of a room, based on detecting contexts such as lighting level and presence of people. They achieve almost 84% true positives although they point out that their set up follows easily learnable and distinct patterns. They do not explain the 16% false readings. Korpipaa et al. [14] developed a multi-layer context-processing framework for mobile
devices which uses a Bayesian classifier for activity identification. Their results indicate that situations were extracted with 96% true positives in restricted scenarios of 9 situations.

However, in real-world situations where they encountered context transitions, situation transitions and undefined phenomena, the recognition accuracy fell to 87% true positives. Bayesian networks are useful for capturing discrete higher level contexts. The disadvantage of this approach is that they do not explicitly support knowledge about state (i.e. situation) transition. Also, they require training data to deduce prior and conditional probabilities, so they are not suitable in scenarios where training data is too difficult or expensive to obtain.

Fuzzy decision trees are used by Guan et al. [7] to deduce contexts from uncertain sensor data. Decision trees require advance knowledge of rules, similar to evidential networks. In contrast to learning techniques such as Bayesian and HMMs, they can reveal intelligible decision paths to the user if required.

Evidence theory has been applied to context or activity recognition, but no temporal knowledge is included in current approaches. Hong et al. [9] define an evidence based activity model, and apply a set of evidential operations to derive activity belief from sensor mass functions. Their work does not include temporal factors, assuming evidence of activities to be co-occurring. Wu [30] used Dempster-Shafer theory for sensor fusion of context. This work included a dynamic discount factor for sensors that changes over time. However, the weighting is reliant on ground truth availability shortly after fusion takes place which is not a workable assumption for activity recognition. Zhang et al. [36] use evidence theory for reasoning about activities. Alternative fusion rules are tested, and conflict resolution strategy for Zadeh's paradox is proposed. Similar to [9] and the work in this paper, an evidence model (CRET) that propagates evidence from sensor level to activity level is described. Temporal knowledge is not included in the CRET model.

6. Conclusion and Future Work

This paper presents an evidential framework with extension of temporal knowledge for reasoning about activities. The framework achieved 70% improvement of recognition accuracy with temporal information in the evaluation on a real-life smart home dataset, which outperformed classical machine learning techniques.

As future work, further temporal information on activity transitions will be incorporated into the framework. This will aim to provide a similar capability to that of Hidden Markov Models in allowing activity sequence patterns to improve recognition. As part of this, we would like to use our approach on a dataset captured over a longer period, with longer term temporal patterns.

A second aim is to investigate the intelligibility of using the evidential framework. One of the challenges in pervasive computing is the user's need to understand the decision making process of the system. Intelligibility is a crucial usability requirement in smart environments [6]. The reasoning process using evidence theory is quite transparent, and indeed, is illustrated via the situation DAG. Therefore, it should be possible to generate explanations for reasoning.

A final aim is to investigate the use of transferrable activity models from one environment to another. With machine learning approaches, training data must be collected for any change in environment. With the evidential framework, the situation DAG from one environment may be used as the basis for another similar environment. Adjustments to the situation DAGs for known changes in sensors, or activity definitions can be applied and the framework re-used

Studies of the kind reported here rely on the public availability of high-quality annotated data sets from real-world smart environments, something that is notably lacking in the field. Our own experience has been that collecting such data sets is enormously time-consuming and expensive, requiring access to a highly instrumented, populated facility - and even then often yields only low-quality data. The collection and publication of data sets is something that needs to be prioritized within the pervasive research community in order to support standardized evaluation of techniques for data interpretation.

7. References

[1] J. F. Allen and G. Ferguson. Actions and events in interval temporal logic. *Journal of Logic and Computation*, 4:531–579, 1994.

[2] J. C. Augusto, J. Liu, P. McCullagh, H. Wang, and J.-B. Yang. Management of uncertainty and spatio-temporal aspects for monitoring and diagnosis in a smart home. *International Journal of Computational Intelligence Systems*, 1(4):361–378, 2008.

[3] D. Choujaa and N. Dulay. Activity inference through sequence alignment. In *LoCA '09: Proceedings of the 4th International Symposium on Location and Context Awareness*, pages 19–36, Berlin, Heidelberg, 2009. Springer-Verlag.

[4] B. Clarkson, A. Pentland, and K. Mase. Recognizing user context via wearable sensors. *Wearable Computers, IEEE International Symposium*, page 69, 2000.

[5] D. J. Cook, J. C. Augusto, and V. R. Jakkula. Ambient intelligence: Technologies, applications, and opportunities. *Pervasive and Mobile Computing*, 5(4):277 – 298, 2009.

[6] A. Dey. Modeling and intelligibility in ambient environments. *Journal of Ambient Intelligence and Smart Environments*, 1(1):57–62, 2009.

[7] D. Guan, W. Yuan, A. Gavrilov, S. Lee, Y. Lee, and S. Han. Using fuzzy decision tree to handle uncertainty in context deduction. In *Second International Symposium on Ubiquitous Computing Systems (UCS'04)*, volume 4114/2006 of *Lecture Notes in Computer Science*, pages 63–72, 2006.

[8] X. Hong, C. Nugent, W. Liu, J. Ma, S. McClean, B. Scotney, and M. Mulvenna. Uncertain information management for adl monitoring in smart homes. In *Intelligent Patient Management*, volume 189/2009 of *Studies in Computational Intelligence*, pages 315–332, 2009.

[9] X. Hong, C. Nugent, M. Mulvenna, S. McClean, B. Scotney, and S. Devlin. Evidential fusion of sensor data for activity recognition in smart homes. *Pervasive and Mobile Computing*, 5(3):236–252, 2009.

[10] S. S. Intille, K. Larson, E. M. Tapia, J. Beaudin, P. Kaushik, J. Nawyn, and R. Rockinson. Using a live-in laboratory for ubiquitous computing research. In *Proc. of PERVASIVE06*, pages 349–365, 2006.

[11] V. R. Jakkula, A. S. Crandall, and D. J. Cook. Enhancing anomaly detection using temporal pattern discovery. *Advanced Intelligent Enviroments*, pages 175–194, 2009.

[12] V. R. Jakkula and D. J.Cook. Using temporal relations in smart environment data for activity prediction, 2007.

[13] A. Jehad Sarkar, Y.-K. Lee, and S. Lee. A smoothed naive bayes-based classifier for activity recognition. *IETE Technical Review*, 27(2):107–119, 2010.

[14] P. Korpipaa, J. Mantyjarvi, J. Kela, H. Keranen, and E. Malm. Managing context information in mobile devices. *Pervasive Computing, IEEE*, 2(3):42–51, 2003.

[15] S. W. Loke. Representing and reasoning with situations for context-aware pervasive computing: a logic programming perspective. *Knowledge Engineering Review*, 19(3):213–233, 2004.

[16] J. D. Lowrance, T. D. Garvey, and T. M. Strat. A framework for evidential-reasoning systems. *Readings in uncertain reasoning*, pages 611–618, 1990.

[17] S. McKeever, J. Ye, L. Coyle, and S. Dobson. A context quality model to support transparent reasoning with uncertain context. In *Quality of Context*, volume 5786 of *Lecture Notes in Computer Science*. Springer Verlag, 2009.

[18] S. McKeever, J. Ye, L. Coyle, and S. Dobson. Using dempster shafer theory of evidence for situation inference. In *Smart Sensing and Context*, volume 5741/2004 of *Lecture Notes in Computer Science*, pages 149–162. Springer Verlag, 2009.

[19] J. Modayil, T. Bai, and H. Kautz. Improving the recognition of interleaved activities. In *UbiComp '08: Proceedings of the 10th international conference on Ubiquitous computing*, pages 40–43, New York, NY, USA, 2008. ACM.

[20] C. K. Murphy. Combining belief functions when evidence conflicts. *Decision Support Systems*, 29(1):1–9, 2000.

[21] P. Palmes, H. K. Pung, T. Gu, W. Xue, and S. Chen. Object relevance weight pattern mining for activity recognition and segmentation. *Pervasive Mob. Comput.*, 6(1):43–57, 2010.

[22] K. Partridge and P. Golle. On using existing time-use study data for ubiquitous computing applications. In *UbiComp '08: Proceedings of the 10th international conference on Ubiquitous computing*, pages 144–153, New York, NY, USA, 2008. ACM.

[23] A. Ranganathan, J. Al-Muhtadi, and R. Campbell. Reasoning about uncertain contexts in pervasive computing environments. *Pervasive Computing, IEEE*, 3(2):62–70, 2004.

[24] K. Sentz and S. Ferson. Combination of evidence in dempster shafer theory. Technical Report SAND2002-0835, Sandia National Laboratories, Albuquerque, New Mexico, April 2002.

[25] G. Shafer. *A mathematical theory of evidence*. Princeton Unverisity Press, 1976.

[26] E. Tapia, S. Intille, and K. Larson. Activity recognition in the home using simple and ubiquitous sensors. In *Pervasive Computing*, volume 3001 of *LNCS*, pages 158–175, 2004.

[27] E. VanKasteren, Noulas and Krose. Accurate activity recognition in a home setting. In *Proc' of 10th International Conference on Ubiquitous Computing*, South Korea, Sept 2008.

[28] L. Wang, T. Gu, X. Tao, and J. Lu. Sensor-based human activity recognition in a multi-user scenario. In *Ambient Intelligence*, volume 5859, chapter 10, pages 78–87. Springer Berlin Heidelberg, Berlin, Heidelberg, 2009.

[29] I. H. Witten and E. Frank. *Data Mining: Practical Machine Learning Tools and Techniques*. Morgan Kaufmann, 2nd edition, 2005.

[30] H. Wu, M. Siegel, and R. Stiefelhagen. Sensor fusion using dempster-shafer theory. In *Proc' of IEEE Instrumentation and Measurement Technology Conference*, 2002.

[31] J. Ye. *Exploiting Semantics with Situation Lattices in Pervasive Computing*. PhD dissertation, University College Dublin, 2008.

[32] J. Ye, A. Clear, L. Coyle, and S. Dobson. On using temporal features to create more accurate human-activity classifiers. In *Proceedings of the 20th Conference on Artificial Intelligence and Cognitive Science*, pages p.274–283, UCD, Dublin, 2009.

[33] J. Ye, L. Coyle, S. Dobson, and P. Nixon. Representing and manipulating situation hierarchies using situation lattices. *Revue d'Intelligence Artificielle*, 22(5), 2008.

[34] J. Ye, L. Coyle, S. McKeever, and S.Dobson. Dealing with activities with diffuse boundaries. In *Proceedings of Pervasive 2010 workshop on How to do good activity recognition research: Experimental methodologies, evaluation metrics, and reproducibility issues*, page to appear, Helsinki, Finland, May 2010.

[35] L. A. Zadeh. A simple view of the dempster-shafer theory of evidence and its implication for the rule of combination. *AI Mag.*, 7(2):85–90, 1986.

[36] D. Zhang, M. Guo, J. Zhou, D. Kang, and J. Cao. Context reasoning using extended evidence theory in pervasive computing environments. *Future Gener. Comput. Syst.*, 26(2):207–216, 2010.

[37] V. W. Zheng, D. H. Hu, and Q. Yang. Cross-domain activity recognition. In *Ubicomp '09: Proceedings of the 11th international conference on Ubiquitous computing*, pages 61–70, New York, NY, USA, 2009. ACM.