Using Temporal Relations in Smart Environment Data for Activity Prediction

Vikramaditya R. Jakkula vjakkula@eecs.wsu.edu
Washington State University, EECS, Spokane Way, Pullman, Washington – 99164, U.S.A.
Diane J. Cook cook@eecs.wsu.edu
Washington State University, EECS, Spokane Way, Pullman, Washington – 99164, U.S.A.

Abstract

In this paper, we aim to experiment and develop a general framework representing the temporal relations of actions and events in smart home datasets, that support a wide range of reasoning tasks, with primary focus on prediction. In earlier work, we performed prediction based solely on the sequence of observed activities. In this work, we supplement evidence for a particular action using the temporal relation information. We compare the predictive accuracy with and without temporal information, and illustrate the benefit of temporal relationships for prediction of smart home events and activities.

1. Introduction

All smart homes provide rich datasets for analyzing, but they also introduce unique challenging features. A smart home dataset includes a timestamp, indicating when a particular activity occurred or a sensor was triggered. Such datasets incorporate the concept of time, which describe an event with respect to its start and end times. The temporal nature of this data when analyzed and used provides us with a much better observation of trends and helps us learn patterns of resident behavior over time. While smart home actions and events are instantaneous, most of them occur over an interval of time. As a result, it is more effective to describe activities using time intervals rather than time points. To accomplish this, we investigate methods of using Allen's temporal logic to analyze smart home data and perform related tasks such as prediction and anomaly detection.

From Allen's original thirteen temporal relations (Allen et. al., 1994) we represent and identify the nine relations shown in Table 1. These subsets of relationships relate a particular event with the next observed event, and thus are useful for event prediction. To analyze smart home data, we first find the temporal relations among the data and then apply associate rule mining to identify frequent event sequences. Based on the relationships that are found we build a probability model using the temporal relations to calculate the probability of the event most likely to occur. Let us focus now on how to calculate the probability that event Z will occur (in this case, the start of the event interval). Evidence for this probability is based on the occurrence of other events that have a temporal relationship with Z, and is accumulated over all such related events. First consider the probability of Z occurring given that the start of the temporal interval for event Y has been detected. The formula to calculate the probability of event Z based on the occurrence of event Y and its temporal relationship with Z is given by Equation (1). Note that the equation is based on the observed frequency of the observed temporal relationships between Y and Z as well as the number of occurrences of Y in the collected event history.

Evidence of P(Z|Y) =|After (Y,Z) + During(Y,Z) + OverlappedBy(Y,Z)+ MetBy(Y,Z) + Starts(Y,Z) + StartedBy(Y,Z)+Finishes(Y,Z)+FinishedBy(Y,Z)+ Equals(Y,Z)|/|Y| (1)

The need for a robust model is essential for prediction for any intelligent smart home to function in a dynamic world. For an agent to perform prediction, it should be capable of applying the limited experience of environmental event history to a rapidly changing environment, where event occurrences are related by temporal relations.

In earlier work, we performed this prediction based solely on the sequence of observed activities. In this work, we supplement evidence for a particular action using the temporal relation information. We compare the predictive accuracy with and without temporal information, and illustrate the benefit of temporal relationships for prediction of smart home events. Based on results generated from synthetic and real smart home data, we conclude that temporal logic provides substantial benefits for smart home tasks. Identification of temporal relations provides key insights to smart home activities and aids with prediction and anomaly detection in a smart home or other smart environment.

Appearing in *Proceedings of the 24dth International Conference on Machine Learning*, Corvallis, OR, 2007.

Table 1. Temporal relations for prediction.

TEMPORAL RELATION	VISUAL DIAGRAM	CONSTRAINT	USABLE
Y **After** X	X Y	Start(X)<Start(Y); End(X)<Start(Y)	✓
X **During** Y	X Y	Start(X)>Start(Y); End(X)<End(Y)	✓
Y**Overlapped-by** X	X Y	Start(X)<Start(Y); Start(Y)<End(X); End(X)<End(Y)	✓
Y **Met-by** X	X Y	Start(Y) = End(X)	✓
X**Finishes**Y	X Y	Start(X)≠start(Y); End(X) = End (Y)	✓
Y**Finishedby**X	X Y	Start(X)≠start(Y); End(X) = End (Y)	✓
X **Starts** Y	X Y	Start(X)=Start(Y); End(X)≠End(Y)	✓
Y**Startedby** X	X Y	Start(X)=Start(Y); End(X)≠End(Y)	✓
X **Equals** Y	X Y	Start(X)=Start(Y); End(X)=End(Y)	✓

2. Related Work

Morchen argued that Allen's temporal patterns are not robust and small differences in boundaries lead to different patterns for similar situations (Morchen, 2006). Morchen presented a Time Series Knowledge Representation, which expresses the temporal concepts of coincidence and partial order. Although this method appears feasible and computationally sound, it does not suit our smart home application due to the granularity of the time intervals in smart homes datasets. His approach does not involve ways to eliminate noise and the datasets are so huge that computational efficiency would not be the only factor to be considered.

Björn (Björn. et al., 2006) also reasons that space and time play essential roles in everyday lives. They discuss several AI techniques for dealing with temporal and spatial knowledge in smart homes, mainly focusing on qualitative approaches to spatiotemporal reasoning.

3. Temporal Relations

Activities in a smart home include physical activities as well as instrumental activities. These may include walking, sitting on a couch, turning on a lamp, using the coffeemaker, and so forth. We see that these activities are not instantaneous, but have distinct start and end times. We also see that there are well-defined relationships between time intervals for different activities. These temporal relations can be represented using Allen's temporal relations and can be used for knowledge and pattern discovery in day-to-day activities. These discoveries can be used for developing systems which can act as reminder assistants and also help detect anomalies and aid us in taking preventive measures (Jakkula & Cook, 2007).

Let us consider a scenario which involves a cooker, oven and a lamp being used in a smart home. We see that the inhabitant turns on the cooker and after some period of time turns on the oven. As time progresses, oven is turned off and the individual continues using the cooker. Later on, the cooker is turned off and the individual turns on the lamp to illuminate the room. We see that this scenario involved three activities each defined by interaction with a single device, namely a cooker, an oven and a lamp. Now we apply Allen's logic to establish the temporal relations among the activities which occurred. The scenario is illustrated in figure 1. These activities can be represented as cooker "contains" oven and "meets" lamp. We can also represent these relationships as cooker "meets" lamp and oven "before" lamp.

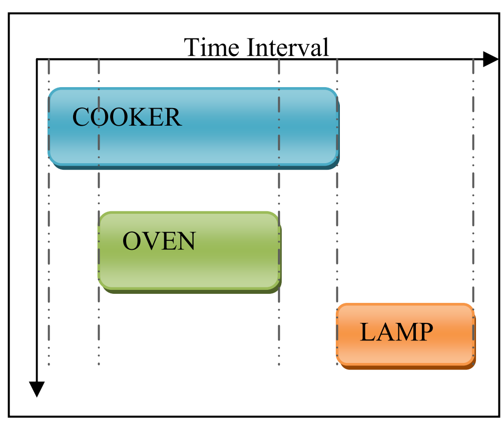

Figure 1: Illustration of temporal relations using activity intervals of occurrence.

4. Experimentation Results

The data collection system consists of an array of motion sensors, which collect information using X10 devices and the in-house sensor network. Our dataset is collected for an inhabitant working in the MavLab (see Figure 2) (Youngblood. et. al., 2005) and consists of two months of data. The lab consists of a presentation area, kitchen, student desks, and faculty room. There are over 100 sensors deployed in the MavLab that include light, temperature, humidity, and reed switches. In addition, we created a synthetic data generator to validate our approach. We developed a model of a user's pattern which consists of a number of different activities involving several rooms and eight devices. For this paper we generated a data set containing about 4,000 actions representing two months of activities.

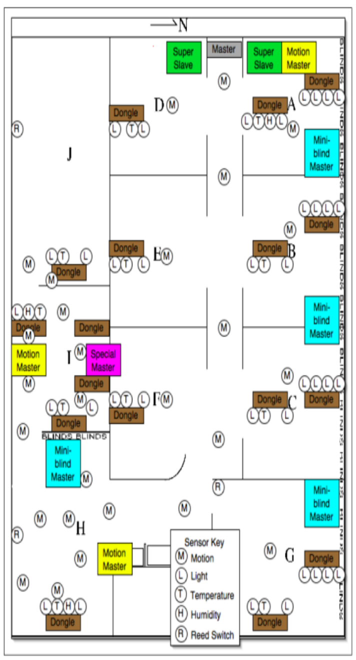

Figure 2: Illustration of Temporal Intervals

The parameters of the datasets (real and synthetic) used are given in table 2 below.

Table 2. Parameter settings of the dataset used for finding rules.

Data Set	No of days	No of Events	Intervals Found
Real	60	17	1640
Synthetic	60	08	1738

The first step of the experimentation is to process the raw data and find the intervals in the data which is done by a simple tool which takes the timestamp of the event that occurred and based on the state (ON/OFF) forms the intervals and later this data is passed through a temporal analyzer tool which identifies the temporal intervals based on the constraints formulated. The pseudo code for the temporal analyzer tool is described in algorithm 1 given below.

Algorithm 1 Temporal Interval Analyzer

Input: data timestamp, event name and state
 repeat
  while [Event && Event + 1 found]
   Find paired "ON" or "OFF" event in data to determine temporal range.
   Read next event and find temporal range.
   Take both events and look up kind of relation from possible relation types based on the constraint (see Table 1).
   Write out relation type and related data.
   Increment Event Pointer
Loop until End of Input.

Thus we see that the temporal relations dataset is now formed and we proceed to the next step of experimentation where we identify the association rules which can be used for prediction. Weka implementation of Apriori-type algorithm is used, which iteratively reduces the minimum support until it finds the required number of rules within the given minimum confidence. The table 3 given below gives us the different parameters set and the number of rules generated with a given specified minimum confidence for real dataset and table 4 for synthetic data.

Table 3. Parameter settings and rules generated using Apriori-type algorithm in Weka for real dataset.

Run #	Minimum support	minimum confidence	No of Best Rules Found
1	0.00	0.5	100
2	0.01	0.5	006
3	0.02	0.5	002
4	0.05	0.5	001

Table 4. Parameter settings and rules generated using Apriori-type algorithm in Weka for synthetic dataset.

Run #	Minimum support	minimum confidence	No of Best Rules Found
1	0.00	0.5	100
2	0.01	0.5	006
3	0.02	0.5	002
4	0.05	0.5	001

In the tables above, the confidence level above 0.5 and support above 0.05 could not be used, as they could not result in any best rules, due to the smaller datasets being used. As we see that the datasets are small, run 2 is chosen for this experimentation, with minimum confidence of 0.5 and minimum support of 0.01 and their resulting best runs. The final step involves using these rules with the existing sequential predictor (Gopalratnam & Cook, 2005) and compare the performance without the rules. This was tested on a single day of smart environment data. The enhanced prediction is given by the Algorithm 2 below.

Algorithm 2 Temporal Rules Enhanced prediction.

Input: Output of ActiveLezi Predictor a, Best Rules r, Temporal Dataset
repeat
If a != *null*
repeat
Find first event in the relation rule and set it to r1
If (r1 ==a) **Then**
If (Relation != "After") **Then**
Calculate Evidence (use equation 1) & if high evidence is noted then
Make event related to r1 in the best rule as next predictor output ;
else
*Get next predicted event and look for there temporal relation in the temporal relations database based on the frequency.
If again the relation is after Then goto * Until no more after relations found then calculate evidence if high then predict;
Else
Calculate evidence and if high then predict this event based on the relation;
Continue.
End if.
Until end of rules.
End if.
Loop until End of Input.

Table 5. Comparing ActiveLezi based prediction with and without temporal rules found with 0.5 confidence.

DATA SET	PERCENTAGE ACCURACY	PERCENTAGE ERROR
WITHOUT RULES		
REAL	55	45
SYNTHETIC	64	36
WITH RULES		
REAL	56	44
SYNTHETIC	69	31

The table 5 above discusses the observed accuracy of the prediction performance on real and synthetic datasets which are from an smart environment (We observed percentages were rounded-off). We see that there was 1.86% improved and 7.81% improved real and synthetic dataset predictions respectively. One of the main reason for higher error rate is the use of a smaller dataset and also the temporal relations are based on activities which occur and the activities or events in the test dataset need not be of the same pattern of some earlier day. As we have larger datasets we see that the performance of the temporal relations enhanced prediction would also improve drastically over time.

5. Conclusion

Temporal rule based knowledge discovery is a new area in smart home research. We notice that the use of temporal relations provides us a unique new approach for prediction. The current approach is currently experimented on small datasets, but we will next validate the performance of our algorithm on larger datasets. Some future directions of this work also includes the expansion of the temporal relations by including more temporal relations, such as until, since, next, and so forth, to create a richer collection of useful temporal relations.

Acknowledgments

This work is supported by NSF grant IIS-0121297.

References

James F. Allen, & George Ferguson (1994). *Actions and Events in Interval Temporal Logic*. Technical Report *521*.

K. Gopalratnam & D. J. Cook (2004). *Active LeZi: An Incremental Parsing Algorithm for Sequential Prediction*. International Journal of Artificial Intelligence Tools. 14(1-2):917-930.

Mörchen, F. (2006) *A better tool than Allen's relations for expressing temporal knowledge in interval data*. In Proceedings the Twelveth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. Philadelphia, PA, USA.

Björn Gottfried, Hans W. Guesgen, & Sebastian Hübner. (2006) *Spatiotemporal Reasoning for Smart Homes*. Designing Smart Homes. Lecture Notes in Computer Science, Springer Berlin / Heidelberg, Pg 16-34, Volume 4008/2006.

Vikramaditya Jakkula, & Diane J. Cook. (2007). *Learning temporal relations in smart home data*. Proceedings of the second International Conference on Technology and Aging. Canada.(To appear).

G. Michael Youngblood, Lawrence B. Holder, & Diane J. Cook (2005). *Managing Adaptive Versatile Environments*. Proceedings of the IEEE International Conference on Pervasive Computing and Communications.