Enhancing Smart Home Algorithms Using Temporal Relations

Vikramaditya R. JAKKULA[1] and Diane J COOK

School of Electrical Engineering and Computer Science

Abstract. Smart homes offer a potential benefit for individuals who want to lead independent lives at home and for loved ones who want to be assured of their safety. We have designed algorithms to detect anomalies and predict events based on sensor data collected in a smart environment. In this paper we explain how representing and reasoning about temporal relations improves the performance of these algorithms and thus enhances the ability of smart homes to monitor the well being of their residents.

Keywords. Smart homes, anomaly detection, event prediction, temporal reasoning

Introduction

Temporal rule mining and pattern discovery applied to time series data has attracted considerable interest over the last few years. In this paper we consider the problem of learning temporal relations between time intervals in smart home data, which includes physical activities (such as taking pills while at home) and instrumental activities (such as turning on lamps and electronic devices). Our long-term goal is to keep individuals functioning independently at home longer using smart home technologies. The objective of this work is to enhance smart home anomaly detection and prediction algorithms using temporal relations extracted from raw sensor data in a smart home environment.

We propose one such framework to derive temporal rules from a time series representation of observed inhabitant activities in a smart home, and validate the algorithm using both synthetic datasets and real data collected from the MavHome smart environment. This framework is based on Allen's temporal logic [1]. Allen suggested that it is more common to describe scenarios by time intervals rather than by time points, and listed thirteen relations formulating a temporal logic (before, after, meets, meet-by, overlaps, overlapped-by, starts, started-by, finishes, finished-by, during, contains, and equals). These temporal relations play a major role in identifying temporal activities which occur in a smart home. Consider, as an example, a case where an elderly person takes pills after eating food. We notice that these two activities, taking pills and eating, share the temporal relation "after" between them. When this relationship is violated, the relationship type is updated to "meets" and an anomaly in activity is noted. The objective of this research is to identify temporal relations among

[1] Corresponding Author: Washington State University, Box 642752, Pullman, WA 99164-2752, USA; E-mail: vjakkula,cook@eecs.wsu.edu.

daily activities in a smart home and to enhance prediction and decision making with these discovered relations.

Temporal interval discovery based on Allen's interval relations has several disadvantages when used for knowledge discovery and pattern recognition. One of the major disadvantages is its ambiguous nature. As seen in Figure 1, by applying the notion of temporal relations we can identify these relations as A (turn on range top) "before" B (turn on oven) and B "before" C (turn on toaster). Finding the best representation for the identified temporal interval is a current challenge. We can see that A "before" B "before" C is a possible relationship label. However, an alternative representation consistent with the events is A "before" B; B "finishes-by" C. The second interpretation actually changes our perspective of the scenario. In this case when we use the relation B "before" C we know that the event B just occurs before C. In contrast, when we interpret the relationship as B finished-by C, an anomaly can be flagged in cases where B and C do not finish at the same time. If we were to use the earlier relation of B "before" C, such anomalies would not be captured. Thus the relation of B "finished-by" C is a better fit for the relationship illustrated in Figure 1 between events A, B, and C.

Morchen argued that Allen's temporal patterns are not robust and small differences in boundaries lead to different patterns for similar situations [2]. As a possible solution, Morchen presented a Time Series Knowledge Representation, which expresses the temporal concepts of coincidence and partial order. Although this method appears feasible, it does not suit our smart home application due to the granularity of the time intervals in smart homes datasets. His approach does handle noise elimination, which is a problem with the large datasets generated by smart home sensors. Björn, et al. [3] also reason that space and time play essential roles in everyday lives. They offer qualitative approaches for spatiotemporal reasoning in smart homes which are not yet presented in an implementation.

1. Data Collection

We treat a smart environment as an intelligent agent [4], which perceives the state of the residents and their physical surroundings using sensors and acts upon the environment using device controllers. This approach is implemented in our MavHome smart home project. We have collected two months of data on volunteer resident

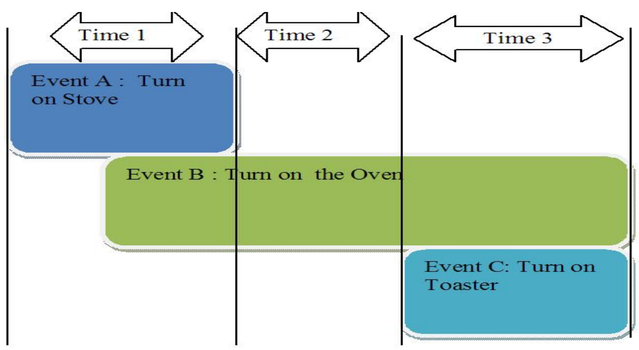


Figure 1. Temporal intervals are labeled as A "before" B "before" C or A "before" B "finishes-by" C.

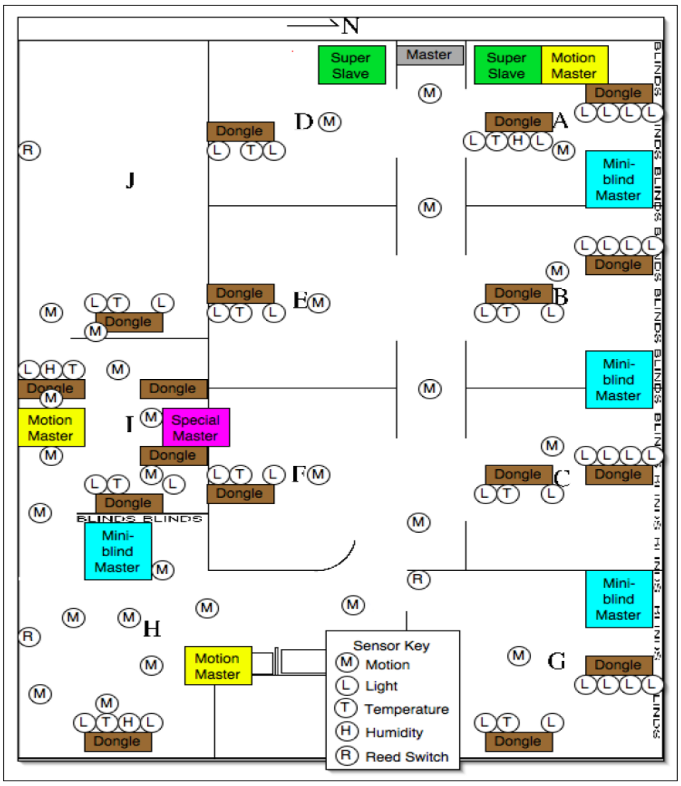


Figure 2. MavHome environment sensor layout.

activities in the MavLab (shown in Figure 2), resulting 4000 sensed events and representing one of the first projects to offer long-term inhabitant and modeling algorithms. The MavHome data collection system consists of an array of motion sensors which collect information using X10 devices and our in-house sensor network. The lab consists of a presentation area, kitchen, student desks, and a faculty room. There are over 100 sensors deployed in the MavLab that include light, temperature, humidity, and reed switches. In addition, we created a simulator which generates event data corresponding to an activity pattern spanning several rooms and interacting with eight devices.

2. Experimental Evaluation

Modeling temporal events in smart homes is an important problem and has great advantage to people with disabilities and the elderly. We see that temporal constraints can model normal activities; if a temporal constraint is not satisfied then a potential "abnormal" or "critical" situation may occur. The goal of this experiment is to identify temporal relations in smart home datasets and later use them for prediction. There are two major problems associated with using Allen's temporal relations. The first problem is the failure of Allen's approach to identify a single most descriptive relation between a pair of events. The second challenge is how to process event relationships in smart home data, which by its nature has a minute time granularity. In our implementation we try to resolve these problems and provide an alternate solution as how the temporal relations can be identified and associated on smart home datasets.

The best way to eliminate ambiguity in identifying the temporal relations is to identify and define the boundary conditions for the thirteen defined intervals before we use it in our algorithm. We illustrate these boundary conditions in Figure 3, using

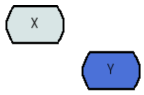
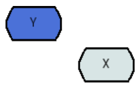
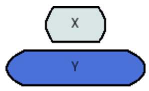
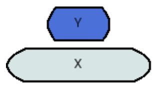
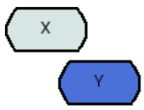
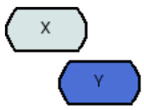
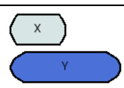
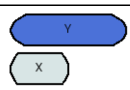
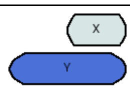
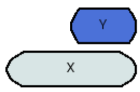
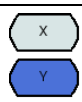
Temporal Relations	Pictorial Representation	Interval Constraints
X Before Y		Start(X)<Start(Y); End(X)<Start(Y)
X After Y		Start(X)>Start(Y); End(Y)<Start(X)
X During Y		Start(X)>Start(Y); End(X)<End(Y)
X Contains Y		Start(X)<Start(Y); End(X)>End(Y)
X Overlaps Y		Start(X)<Start(Y); Start(Y)<End(X); End(X)<End(Y)
X Overlapped-By Y		Start(Y)<Start(X); Start(X)<End(Y); End(Y)<End(X)
X Meets Y		Start(Y) = End(X)
X Met-by Y		Start(X)= End (Y)
X Starts Y		Start(X)=Start(Y); End(X)≠End(Y)
X started-by Y		Start(Y)=Start(X); End(X)≠End(Y)
X Finishes Y		Start(X)≠start(Y); End(X) = End (Y)
X Finished-by Y		Start(X)≠Start(Y); End(X)=End(Y)
X Equals Y		Start(X)=Start(Y); End(X)=End(Y)

Figure 3. Boundary conditions for Allen's temporal intervals.

events X and Y as example events. These conditions are used in our algorithm for identifying temporal intervals. The first step of the algorithm applies the Apriori algorithm [5] as shown in the pseudo code below to identify the most frequent event pairs whose relationships fit one of the thirteen categories. This step addresses the puzzle of handling minute time granularity.

Apriori Algorithm Pseudo-code:

2.1. C_k*: Candidate itemset of size k*
2.2. L_k *: frequent itemset of size k*
2.3. L_1 *= {frequent items};*
2.4. ***for*** *(*k *= 1;* L_k *!=∅;* k++*)* ***do begin***
2.5. C_{k+1} *= candidates generated from* L_k*;*
2.6. ***for each*** *day* t *in datasets do*
2.7. *increment the count of all candidates in* C_{k+1} *that are contained in* t
2.8. L_{k+1} *= candidates in* C_{k+1} *with min_support*
2.9. ***return*** $\cup_k L_k$*;*

The next step of the algorithm applies a weighting to the mined temporally-related pairs. This weighting helps us handle ambiguous situations where a temporal relation can be labeled as two or more types. To resolve the issue, we select the temporal label that occurs the most often in the observed data between the targeted pair of events.

$$\text{Weight} = \frac{\text{No of Occurrence of specific Relation}}{\text{Total Identified Relations}} \qquad (1)$$

When validating the algorithm, we apply it to real and synthetic data, whose features are specified in Table 1. Note that the algorithm is constrained in this case to only identify event pairs that co-occur in the same day. Table 2 describes the results of varying the minimum support to find the frequent temporal relations for the real and synthetic datasets.

Table 1. Parameter settings for experimentation.

	Parameter Setting			
Datasets	**#Days**	**#Events**	**#Intervals**	**Database size**
Synthetic	60	8	1729	106KB
Real	60	17	1623	104KB

Table 2. Minimum support vs. Number of frequent temporal relations identified.

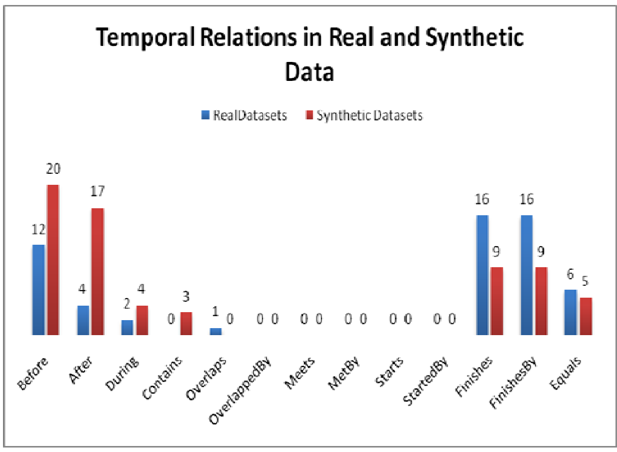
	No of Frequent Itemsets	
Minimum Support	**Real Dataset**	**Synthetic Dataset**
8	3	2
4	5	3
2	10	4

We then further analyzed the frequent itemsets identified with minimum support of value 2 to identify key temporal relations by using the weight based temporal analyzer algorithm Figure 4 illustrates the identified temporal relations in these datasets. Table 3 displays samples of identified frequent relation pairs.

3. Enhancements to Anomaly Detection and Prediction Algorithms

Not only does temporal relationship discovery provide insights on patterns of resident behavior, but it is also beneficial for constructing other smart homes algorithms such as those for anomaly detection and event prediction. The ability to detect anomalies in a smart home is critical for ensuring the well-being of the home's residents. Consider our scenario in which a son remotely monitors his father's health with smart home assistance. A few weeks later the son may notice in a report that there his father has an anomalous lack of movements around certain parts of the house. The smart home can alert the son and the physician when such an anomaly is detected. The caregiver will then investigate, finding that in fact the father has not been feeling well the last few days and has been mostly staying in the bedroom or bathroom.

Our approach to anomaly detection makes use of the discovered temporal relationship information. From the frequency of the nine relationship events which affect anomaly detection (before, contains, overlaps, meets, starts, started-by, finishes, finished-by, and equals), we calculate the probability that an event *X* will or will not occur given the observed occurrence of other events that are temporally related to event *X*. Given the distribution of probabilities over the set of possible events, we calculate the outlier threshold as the mean probability +/- 3 standard deviations. If the event has a high probability of occurring given the observed occurrence of other events, yet does not occur, an anomaly is reported. A similar approach can be taken for detecting the absence of an event as an anomaly. We applied this approach to detecting anomalies in the real and synthetic datasets. The anomaly values, defined as the inverse of the corresponding event probabilities, are shown for these datasets in Figure 5. Our algorithm did not detect any anomalies in the real data. This is not surprising, given that the resident was a healthy adult and did not report any unusual activities during



Real data:
Lamp J16 BEFORE Radio J11
Lamp I14 AFTER Lamp C9
Lamp I4 EQUALS Lamp I4

Synthetic data:
Cooker BEFORE Oven
Fan AFTER Cooker
Lamp BEFORE Cooker

Figure 4. Types of identified temporal relations (left) and samples of identified frequent relations (right).

that time. In the synthetic dataset we embedded three anomalous events in order to validate our algorithm. Our algorithm caught all three anomalies, which in a health care situation would be reported to the caregiver.

In a similar fashion, we can use temporal relationship information to predict the occurrence of events in a smart home. This is useful as a basis for automating interactions with the home that may be difficult for individuals with physical limitations. In addition, predicting commonly-occurring events forms the basis of providing reminders to individuals who need the cues to complete key Activities of Daily Living.

To predict events, we compute the probability of an event *X* based on observed occurrences of other events which typically precede or co-occur with *X*. As a result, we are interested in the frequencies of the nine temporal relationships: after, during, overlapped-by, met-by, starts, started-by, finishes, finished-by, and equals. The combination of frequencies of these relationships forms the basis of calculating the probability for upcoming events. As events are observed, the likelihood of future events is constantly updated. When the probability of a predicted event exceeds a calculated threshold, the event can be "primed" by the smart home by preparing to automate the event or reminding the resident to perform the event for himself.

4. Conclusions

Temporal analysis in smart homes would be a boon for elderly residents, as this would enable them to lead a more independent life. The discovery of temporal patterns would help them associate their daily activity and if a violation is found, such as, forgetting to take pills, a quick reminder can be sent and also would help them identify the next set of activities which might occur when we have a prediction component added to this temporal analyzer. We are continuing to evaluate the performance of our anomaly detection and event prediction algorithms on real data collected in smart living and smart workplace environments, and will in the future consider other supporting algorithms that can be enhanced through the addition of temporal reasoning.

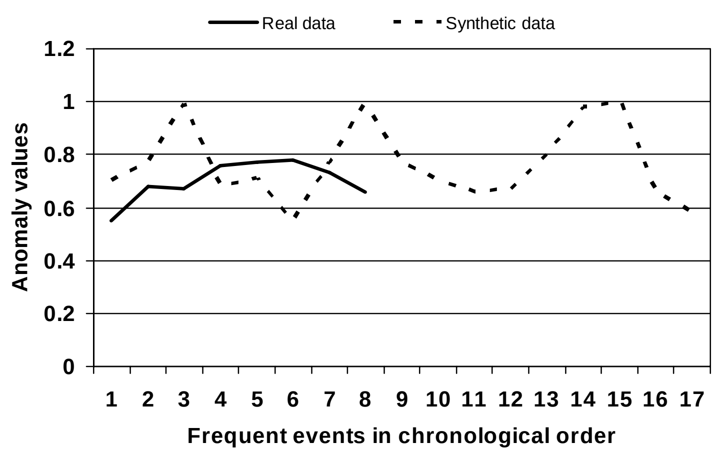


Figure 5. Anomaly values for events in the real and synthetic datasets.

Acknowledgements

This work is supported by NSF grant IIS-0121297.

References

[1] J.F. Allen and G. Ferguson, Actions and Events in Interval Temporal Logic, *Journal of Logic and Computation*, 4 (1994), 531-579.

[2] F. Mörchen, Algorithms for time series knowledge mining, *Proceedings of the Twelveth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, Philadelphia, PA (2006), 668-673.

[3] B. Gottfried, H.W. Guesgen, and S. Hübner, Spatiotemporal Reasoning for Smart Homes, *Designing Smart Homes,* Lecture Notes in Computer Science, Springer Berlin / Heidelberg, 4008 (2006), 16-34.

[4] G.M. Youngblood, L.B. Holder, and D.J. Cook. Managing Adaptive Versatile Environments, *Proceedings of the IEEE International Conference on Pervasive Computing and Communications*. (2005), 351-360.

[5] R. Agrawal and R. Srikant, Fast Algorithms for Mining Association Rules, *Proceedings of the International Conference on Very Large Data Bases*, Morgan Kaufmann (1994), 487--499.