Monitoring Health by Detecting Drifts and Outliers for a Smart Environment Inhabitant [1]

Gaurav Jain [a,2], Diane J. Cook [a] Vikramaditya Jakkula [a]
[a] *Department of Computer Science and Engineering*
University of Texas at Arlington
cook@cse.uta.edu

Abstract. To many people, home is a sanctuary. For those people who need special medical care, they may need to be pulled out of their home to meet their medical needs. As the population ages, the percentage of people in this group is increasing and the effects are expensive as well as unsatisfying. We hypothesize that many people with disabilities can lead independent lives in their own homes with the aid of at-home automated assistance and health monitoring. In order to accomplish this, robust methods must be developed to collect relevant data and process it to detect and/or predict threatening long-term trends or immediate crises.

The main objective of this work is to design techniques for using agent-based smart home technologies to provide this at-home health monitoring and assistance. Specifically, we address the following technological challenges: 1) identifying lifestyle trends, 2) detecting anomalies in current data, and 3) designing a reminder assistance system. We discuss one such smart environment implementation in the MavHome project and present results from testing these techniques in simulation and with a volunteer in an apartment setting.

Keywords. health monitoring, drift detection, outlier detection, reminder assistance

1. Introduction

Since the beginning, people have lived in places that provide shelter and basic comfort and support, but as society and technology advance there is a growing interest in improving the intelligence of the environments in which we live and work. The MavHome (**M**anaging an **a**daptive **v**ersatile **Home**) project is focused on providing such environments. Here we discuss methods by which we can adapt a smart home environment such as MavHome to perform health monitoring and assistance for persons with disabilities and for aging adults.

As Lanspery and Hyde [4] state, "For most of us, the word ‘home’ evokes powerful emotions [and is] a refuge". They note that older adults and people with disabilities want

[1] Proceedings of the International Conference On Smart homes and health Telematics. This work is supported by US National Science Foundation grant IIS–0121297.
[2] Correspondence to: Diane Cook, Department of Computer Science and Engineering, Box 19015, University of Texas at Arlington Tel.: +1 817 272 3606; Fax: +1 817 272 3784; E-mail: cook@cse.uta.edu

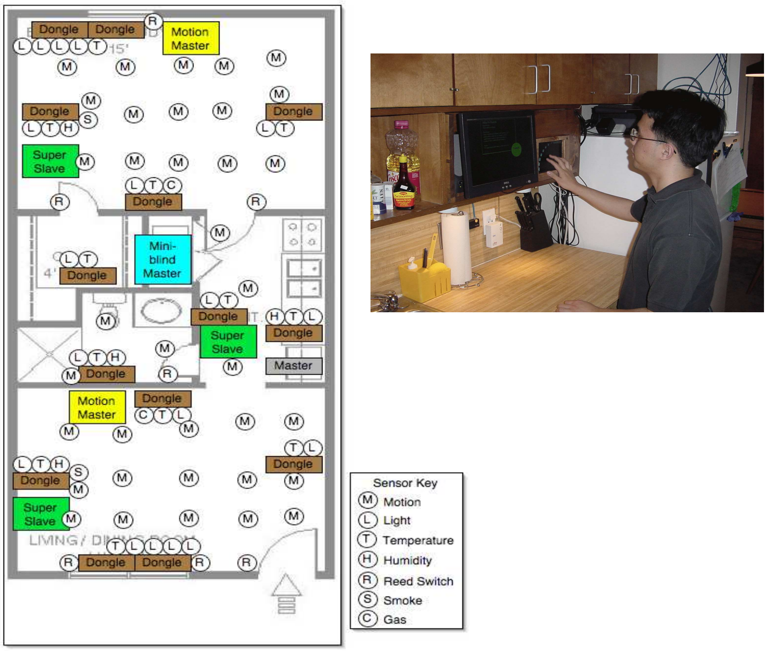

Figure 1. MavPad sensor layout and environment.

to remain in their homes even when their conditions worsen and the home cannot sustain their safety. In addition, the problems of aging and disability are converging. Improvements in medical care are resulting in increased survival into old age, thus problems of mobility, vision, hearing, and cognitive impairments will increase. An AARP report [1] strongly encourages increased funding for home modifications that can keep older adults with disabilities independent in their own homes. Our goal is to assist the elderly and individuals with disabilities by providing smart space capabilities that will monitor health trends and assist in the inhabitant's day to day activities in their own homes.

2. Overview of the MavHome Smart Home

We define an intelligent environment as one that is *able to acquire and apply knowledge about its inhabitants and their surroundings in order to adapt to the inhabitants and meet the goals of comfort and efficiency*. The MavHome (**M**anaging an **a**daptive **v**ersatile **Home**) project is focused on providing such an environment [7,8]. We view our environment as an intelligent agent, which perceives the state of the environment using sensors and acts upon the environment using device controllers.

The MavHome project is unique in that we learn a decision policy to control an environment in a way that optimizes a variety of possible criteria, including minimizing manual interactions, improving operating efficiency, and ensuring inhabitant health and safety. We also ensure that our software need not be redesigned as new devices are registered, new spaces are tested, or new inhabitants move into the environment.

All of the MavHome components are implemented and are being tested in two physical environments, the MavLab workplace environment and an on-campus apartment, the

MavPad (shown in Figure 1). Powerline control automates all lights and appliances, as well as HVAC, fans, and miniblinds. Perception of light, humidity, temperature, smoke, gas, motion, and switch settings is performed through a sensor network. For the health monitoring study described below, we also captured systolic, diastolic, and heart rate data using a wrist wearable device.

Communication between high-level components is performed using the Common Object Request Broker Architecture (CORBA), and each component registers its presence using zero configuration (ZeroConf) technologies. Implemented services include a PostgreSQL database that stores sensor readings, prediction components, data mining components, and logical proxy aggregators. Resource utilization services monitor current utility consumption rates and provide usage estimates and consumption queries.

3. Core Technologies

To automate our smart environment, we collect observations of manual inhabitant activities and interactions with the environment. We then mine sequential patterns from this data using a sequence mining algorithm. Using this information, we create a hierarchical Markov model, then use this model to learn an action policy for the environment.

3.1. Mining Sequential Patterns Using ED

A smart home inhabitant typically interacts with various devices as part of his routine activities. These interactions may be considered as a sequence of events. We characterize each inhabitant-home event as a triple consisting of the sensor or device manipulated, the resulting change that occurred in that device, and the time of interaction. We move a window in a single pass through the history of events or inhabitant actions, looking for episodes (sequences) within the window that merit attention.

When evaluating candidate episodes, the Episode Discovery (ED) algorithm [2] looks for patterns that minimize the description length of the input stream using the Minimum Description Length (MDL) principle. The MDL principle targets patterns that can be used to minimize the description length of a database by replacing each instance of the pattern with a pointer to the pattern definition. With this formula, it is easily seen that finding the model that yields the minimum description length of the data is equivalent to finding the patterns that provide the largest compression value. Our MDL-based evaluation measure thus identifies patterns that balance frequency, length, and periodicity.

In this way, ED identifies patterns of events that can be used to better understand the nature of inhabitant activity. Once the data is compressed using discovered results, ED can be run again to find an abstraction hierarchy of event patterns.

3.2. Decision Making Using ProPHeT

To automate an environment, we apply reinforcement learning to the problem which is modeled as a as a Partially Observable Markov Decision Process (POMDP). Recently, there have been many published hierarchical extensions that allow for the partitioning of large domains into a tree of manageable POMDPs [6]. Research has shown that strategies for new tasks can be learned faster if policies for subtasks are already available. Current approaches generally require *a priori* construction of the hierarchical model. In contrast,

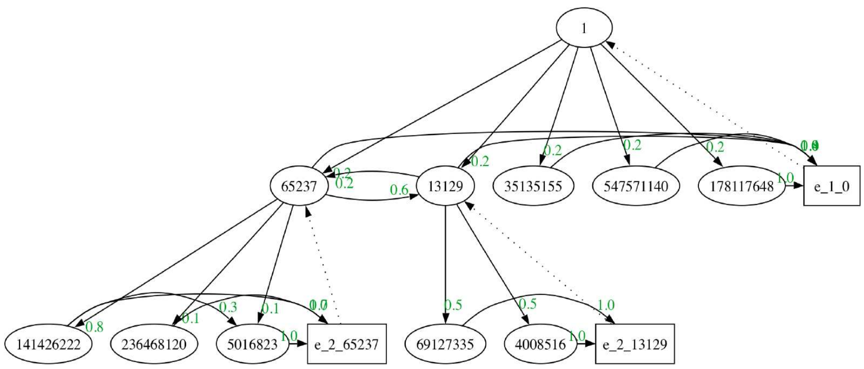

Figure 2. ProPHeT-generated hierarchical POMDP (only the higher levels are shown). Eight abstract tasks are found in the first iteration of ED and two meta-tasks (nodes 65237 and 13129) are found in the second iteration.

our decision learner, ProPHeT, actually automates model creation by using the ED-mined sequences to represent the nodes in the higher levels of the model hierarchy.

The lowest-level nodes in our model represent a single event observed by ED. Next, ED is run multiple iterations on this data until no more patterns can be identified, and the corresponding abstract patterns comprise the higher-level nodes in the Markov model. The higher-level *task* nodes point to the first event node for each permutation of the sequence that is found in the environment history. Vertical transition values are labeled with the fraction of occurrences for the corresponding pattern permutation, and horizontal transitions are seeded using the relative frequency of transitions from one event to the next in the observed history. As a result, the n-tier hierarchical model is thus learned from collected data. Given the current event state and recent history, ED supplies membership probabilities of the state in each of the identified patterns. Using this information, ProPHeT maintains a belief state and selects the highest-utility action.

To learn an automation strategy, the agent explores the effects of its decisions over time and uses this experience within a temporal-difference reinforcement learning framework to form control policies which optimize the expected future reward. The current version of MavHome receives negative reinforcement (observes a negative reward) when the inhabitant immediately reverses an automation decision (e.g., turns the light back off) or an automation decision contradicts safety and comfort constraints.

3.3. Initial Case Study

As an illustration of the above techniques, we have evaluated a week in an inhabitant's life with the goal of reducing the manual interactions in the MavLab. The data was restricted to motion and lighting interactions which generate 1400 events per day.

Figure 2 shows the four-tier HPOMDP that is automatically constructed from the ED patterns. As a point of comparison, we automated the environment using a hierarchical Markov model with no abstract nodes. This single-level model reduced interactions by 38.3%, and the ProPHeT-generated model reduced interactions by 76%.

4. Using a Smart Home to Assist Elderly and People with Disabilities

The data mining, prediction, and multiagent technologies available in MavHome can be employed to provide health care assistance in living environments. Specifically, models can be constructed of inhabitant activities and used to learn activity trends, detect anomalies, and provide automation assistance for inhabitants with special needs.

Our smart environment can identify patterns indicating or predicting a change in health status and can provide inhabitants with needed automation assistance. Collected data includes movement patterns of the individual, periodic vital signs (blood pressure, pulse, body temperature), water and device usage, use of food items in the kitchen, exercise regimen, medicine intake (prescribed and actual), and sleep patterns.

4.1. Capability 1: Identify lifestyle trends.

Many smart space algorithms can provide particular benefit to individuals with health needs who are living independently. One such benefit is to process the captured data in order to identify lifestyle *trends* that may highlight a growing need for the individual.

As a motivating example, consider a scenario involving an elderly man recuperating at home alone after hospitalization. The patient's son lives several hundred miles away but wants to be informed of his father's state of health. If the patient is a smart space inhabitant, he can be regularly monitored for changes in health measurements including heart rate, blood pressure and body temperature. However, this data may not provide a complete picture of his health status. As such, the data needs to be integrated with information on changes in other parameters such as the room temperature and humidity and the individual's movement around the house, eating patterns, medicine intake, and adherence to his daily routine. The smart environment algorithms learn the inhabitant behaviors and start reporting timely information about changes in his health. A few weeks later the son notices in a system report that his father has a sudden decrease in his movements around the house. He calls his father and finds out that in fact his father has not been feeling well the last few days.

A variety of approaches have been investigated in recent years to automate caregiver services. Many of the efforts offer supporting technologies for specialized devices or for specialized tasks such as detecting falls. Little work has combined remote monitoring capabilities with prediction for the purpose of health monitoring, although that has been progress toward using behavior patterns to provide emergency notifiers or inhabitant reminders, particularly useful for the elderly and patients suffering from dementia [3,5].

Collected data can be analyzed not only to provide automation but also to assess activity and health trends. In particular, MavHome algorithms currently classify slow changes in collected data as one of a number of types of pattern *drifts*: cyclic, increasing, decreasing, chaotic, and stable. The size of sample windows is chosen in such a way that it is approximately four times the length of the longest detectable cycle.

Tests for various classes of drifts are performed using temporal autocorrelation plots, which measure the correlation between time-shifted values in a time series. The test for a stable pattern is performed first. This describes data which is nearly constant (within a tolerance threshold) for the entire window of data. A cyclic trend, which is checked next, shows high upward peaks in the autocorrelation graph because correlation between cylic values is high. In Figure 3 (left), frequencies of an action are shown and the cor-

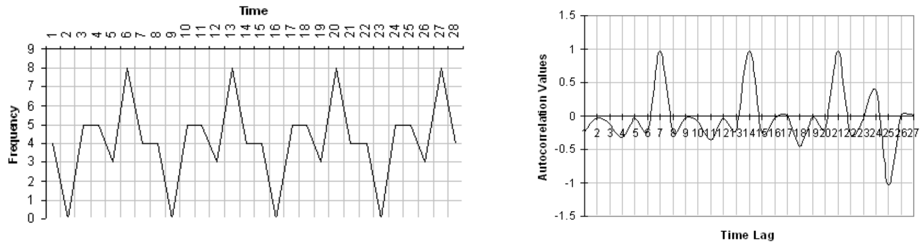

Figure 3. An example of cyclic data (left) and the corresponding autocorrelation plot (right).

responding autocorrelation plot (Figure 3 (right) shows upward-facing peaks at intervals of seven. This indicates that the length of the cycle is seven.

For increasing or decreasing trends, a high degree of autocorrelation is seen between adjacent and near-adjacent observations. For this type of drift, the autocorrelation plot will show a high correlation at lag one and will steadily decrease as the lag increases. The direction of the change can be determined by calculating the sum of the deviation in the adjacent data points. Any pattern in the sample window which is not classified as another type of drift is classified as choatic. This type of drift may be caused by a large number of irregular changes, by a change in the type of drift, or by noise in the data.

Pattern drifts are reported by MavHome if their urgency is high. Urgency is calculated as a combination of the confidence in the drift and the criticality of the analyzed data (drifts involving blood pressure are more critical than those based on changes in television-watching schedules). Confidence is a numeric value between 0 and 1, and reflects the strength of the detected drift. The confidence value of a trend varies according to the type of trend. For cyclic patterns, confidence is calculated as the average height of the first two peaks of the cycle instances, minus a constant multiplied by the average variation between cycles peaks.

We analyzed seven weeks of MavPad inhabitant data for drifts and made the following observations. For most of the collected activity data, patterns were classified as stable or chaotic. Increasing and decreasing trends in motion data was detected at points due to the increased (or decreased) amount of time the inhabitant was spending at home. Lights were on an increased amount of time curing the study, possibly due to longer night hours as winter approached. Cyclic drifts were the rarest. Although two three-day cycles were detected, they only involved the use of lights and both were assigned a low criticality. In the case of health data, a decreasing trend was found throughout much of the collected time window. The inhabitant in this scenario is young and fairly healthy. We would expect different results when monitoring an elderly individual at home.

4.2. Capability 2: Detect anomalies in current data.

MavHome employs two techniques to detect outliers or anomalies in activity and health data. We define an outlier as an extremely high or low value when compared to the rest of the data in the sample window. We use a z-score, or standard score, to detect such outliers. This check is performed before looking for possible drifts.

In the case of the MavPad inhabitant, outliers were detected on day 31 for three different actions. As the graph in Figure 4 shows, the inhabitant's systolic value is zero in

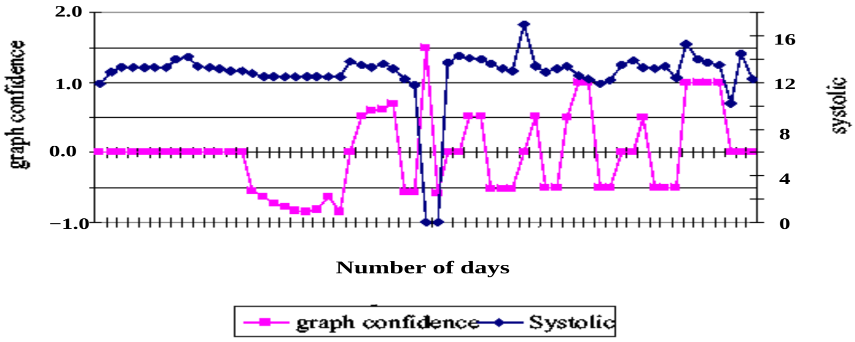

Figure 4. MavPad data graph confidence with systolic values vs. number of days.

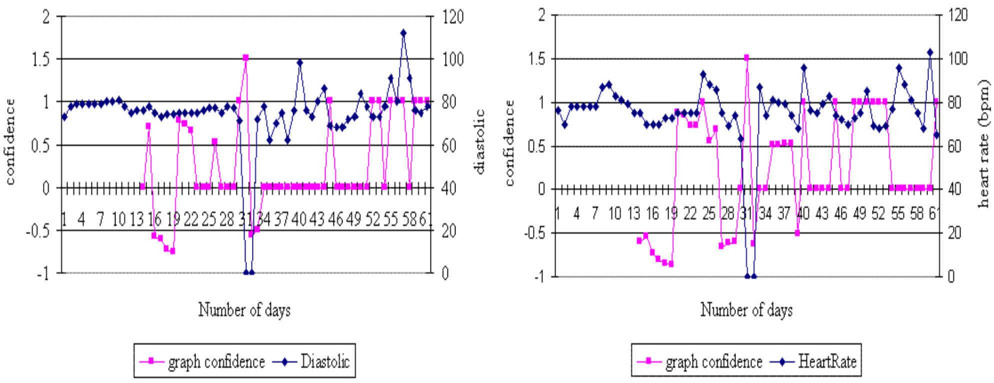

Figure 5. MavPad data graph confidence with diastolic (left) and heart rate (right) values vs. number of days.

this day and the corresponding graph correlation is 1.5, which is identified as an outlier. We also see that the systolic values slowly decrease between days 10 and 23, which was identified as a decreasing drift of eleven days in length. The detected outlier is most likely due to an error in measurement, as the inhabitant was healthy on that day.

Between days 25 and 28, an increasing drift is reported despite the slight decrease in systolic values on days 25 and 26. This is because in the larger window of a few weeks, these values are still higher than earlier. This observations indicates that our algorithm can account for small amounts of noise in the system. We also see that data is classified as chaotic when there is a transition from increasing to decreasing trends. Sudden changes not marked as outliers (e.g., systolic values on days 23, 33, 40, 49, and 59) are also classified as chaotic because for this short time the distribution is too skewed to able to detect a drift. Similar observations are made for diastolic and heart rate data, shown in Figure 5. As a result, a chaotic drift following another type of drift may indicate a change in the distribution. For health monitoring, this could be a situation that bears closer investigation.

As with detected drifts, anomalies of a high criticality are identified for reporting. When a critical anomaly occurs, the home will first try to contact the inhabitant (through the interactive display for a lesser critical anomaly, or through the sound system for a more critical anomaly). If the inhabitant does not respond and the criticality of the anomaly is high, the caregiver will be made aware of the situation.

4.3. Capability 3: Design reminder assistance system.

Reminders can be triggered by two situations. First, if the inhabitant queries the home for his next routine activity, the activity with the highest probability will be given based on the ALZ prediction. Second, if a critical anomaly is detected, the environment will initiate contact with the inhabitant and remind him of the next typical activity. Such a reminder service will be particularly beneficial for individuals suffering from dementia.

As described in the initial MavHome design, automation assistance is always available for inhabitants, which is beneficial if some activities are difficult to perform. A useful feature of the architecture is that safety constraints are embedded in the MavHome software. If the inhabitant or the environment is about to conflict with these constraints, a preventative action is taken and the inhabitant notified. This can prevent accidents such as forgetting to turn off the water in the bathtub or leaving the house with doors unlocked.

5. Conclusion

We have demonstrated that the MavHome software architecture can successfully monitored and provided automation assistance for volunteers living in the MavPad site. However, there is much work to be done to enhance and test the benefits of the smart space algorithms for assisting the elderly and people with disabilities. We are currently collecting health-specific data in the MavHome sites and will be testing in the living environments of recruited residents at the C.C. Young Retirement Community in Dallas, Texas.

References

[1] AARP. These four walls... Americans 45+ talk about home and community, 2003.

[2] E. O. Heierman and D. J. Cook. Improving home automation by discovering regularly occurring device usage patterns. In *Proceedings of the International Conference on Data Mining*, 2003.

[3] H. Kautz, L. Arnstein, G. Borriello, O. Etzioni, and D. Fox. An overview of the assisted cognition project. In *Proceedings of the AAAI workshop on automation as caregiver*, 2002.

[4] S. Lanspery, J. J. C. Jr, J. R. Miller, and J. Hyde. Introduction: Staying put. In S. Lanspery and J. Hyde, editors, *Staying Put: Adapting the Places Instead of the People*, pages 1–22. Baywood Publishing Company, 1997.

[5] M. E. Pollack, L. Brown, D. Colbry, C. E. McCarthy, C. Orosz, B. Peintner, S. Ramakrishnan, and I. Tsamardinos. Autoreminder: An intelligent cognitive orthotic system for people with memory impairment. *Robotics and Autonomous Systems*, 44:273–282, 2003.

[6] G. Theocharous, K. Rohanimanesh, and S. Mahadevan. Learning Hierarchical Partially Observable Markov Decision Processes for Robot Navigation, 2001. IEEE Conference on Robotics and Automation.

[7] G. M. Youngblood, E. O. Heierman, L. B. Holder, and D. J. Cook. Automation intelligence for the smart environment. In *Proceedings of the International Joint Conference on Artificial Intelligence*, 2005.

[8] G. M. Youngblood, L. B. Holder, and D. J. Cook. Managing adaptive versatile environments. In *Proceedings of the International Conference on Pervasive Computing*, 2005.