Anomaly Detection of Elderly Patient Activities in Smart Homes using a Graph-Based Approach

Ramesh Paudel[1], William Eberle[1], and Lawrence B. Holder[2]
[1]Department of Computer Science, Tennessee Technological University, Cookeville, TN
[2]School or Electrical Engineering and Computer Science, Washington State University, Pullman, WA

Abstract - *Sensor-based smart home provide the ability to track resident activities without interfering in their daily routine. It is useful to detect and predict the behaviors of an elderly resident in order to improve the safety of the residents' home environment and provide aid for their caregiver. This paper presents a graph-based approach that successfully discover patterns and anomalies in resident activities. We analyze activity graphs constructed from smart home daily activities to detect normative patterns as well as temporal, spatial, and behavioral anomalies. We also present case studies for cognitively impaired participants and discuss how these anomalies can be linked to the decline in their cognitive abilities which will ultimately provide clinicians and care givers important knowledge regarding their patients.*

Keywords: Graph mining, Anomaly Detection, Cognitive Health Prediction, Smart Homes.

1 Introduction

According to the United States Census Bureau [1], 13% of the U.S. population in 2013 was of the age 65 and over and that figure is estimated to rise to 20 percent by the year 2030. Although people are living longer, that does not imply that they are at optimal health. Alzheimer's disease and other forms of dementia are prominent among the elderly population and are estimated to affect 5.2 million people above the age of 65 [2]. These forms of cognitive disabilities can limit their ability to perform day-to-day activities, requiring them to be dependent on caregivers. Smart homes can provide aid to elderly residents, especially those who are suffering from Mild Cognitive Impairment (MCI), and their caregivers [7]. Various intelligent systems such as smart homes have been used to provide aid in health-care monitoring [3,8]. These smart homes have sensors installed throughout the home to collect information about a resident's everyday activities without interfering with their routines. In addition to that, smart homes can also prompt users to perform particular activities, such as reminding a resident to take their medicine [8]. This allows the residents to be more independent as well as providing aid to their family and caregivers. In 2016, the total annual costs associated with the care of patients with Alzheimer's was estimated at $236 billion [2]. Using sensor data, the daily activities and behavioral patterns of residents can be monitored using various tools. Ideally, these behavioral anomalies can be used to better assess the individual's current state, rate of change, and the potential need for assistance, and ultimately a reduction in the costs of care and medical emergencies.

Typically, the data from smart homes are sensor logs generated while a participant performs daily activities. With this research, we will attempt to convert these daily activity logs into a *graph* by representing a sensor's spatial or temporal information as a node with edges. Then we can search for patterns in what are called activity graphs, and ultimately be able to analyze the patterns in a participant's daily activity so as to better understand their daily routine. According to Galvin and Sadowsky [6], some warning signs of Alzheimer's disease are memory loss, difficulty performing familiar tasks, misplacing things, and changes in behavior. For example, if a participant is "*preparing a cup of soup using a microwave*", a sensor can detect whether the participant forgot to pour water in a cup or forgot to boil it. (The sensor can capture if the participant turns on/off the tap or microwave.) Also a sensor can detect if the participant is in a particular room for an extended amount of time, which (if they are not sleeping) could indicate that the participant has fallen or other health risks have occurred. Thus, these traits (or patterns) can be used to define *anomalous behavior* among residents.

In order to further define what we are referring to as anomalous behavior, we will recognize this type of behavior in smart home data similar to what was defined in [10, 18]:

Temporal Anomaly: Abnormality in duration such as an inappropriately long period of time for performing a task.

Spatial Anomaly: Performing activities in the wrong places or wandering around. For example, if a person is performing an activity sweep the living room, but is in a doorway or some other room besides the living room while doing the activity.

Behavior Anomaly: Abnormality in a behavior pattern. Normal behavior for performing an activity is defined by a sequence of sub-activities and if the participant violates the expected sequence then it is an anomaly.

These anomalous behaviors represent possible scenarios where a participant could have a decline in their cognitive ability. It is our hypothesis that we can use a graph-based approach to detect anomalies in elderly patient activity in smart homes. We formally purpose two hypotheses.

Hypothesis I: A graph-based approach can successfully discover anomalies in elderly resident activity in smart homes.

Hypothesis II: Anomalies are potential indicators of a decline in the cognitive health of an elderly resident.

To validate our hypothesis, we will use the Kyoto dataset with 400 participants provided by Washington State University's CASAS program (http://casas.wsu.edu/). CASAS (Center for Advanced Studies in Adaptive Systems) smart homes have real-time data from sensors that record participants' everyday activities [3]. We will also use the GBAD tool [9] for our graph based anomaly detection.

The layout of this paper is described as follows: first, we present related work in the area of investigating elderly behavioral patterns in a smart home environment, followed by a description of the data and data preprocessing. This is then followed by a description of experiments and results. We then conclude with some discussion and our proposed future work.

2 Related Work

Smart homes collect information and monitor the health of residents by using sensors embedded in various locations. These data from smart homes can be analyzed to understand the behavior of residents which will help us to improve the living of people with medical issues like cognitive disabilities. There are various research works regarding the analysis of behavior and health monitoring from a smart home.

Supervised machine learning approaches for predicting cognitive health on an elderly patient has been used by [11, 14, 15]. Lotfi et al. [11] investigated elderly residents living independently in real home environments who were diagnosed with dementia. They applied recurrent neural networks to predict sensor activity in order to inform the caregiver of any anomalous behavior that can be expected in the future. This worked better for residents with more routine activities such as senior citizens rather than younger residents. Dawadi et al. [14] proposed a Clinical Assessment using an Activity Behavior (CAAB) approach to model a smart home resident's daily behavior and predict clinical assessment scores in hopes to help clinicians make decisions regarding diagnosing patients. However, in this case, most of the 18 residents analyzed were cognitively healthy. Cook et al. [15] used machine learning techniques on data from older adults using smart home and wearable sensors while they performed complex activities of daily living and concluded that it was possible to automatically recognize a difference in behavior between healthy, older adults versus adults with Parkinson's disease. However, it is not easy to always get a labelled dataset and semi-supervised or unsupervised approaches are required to detect anomalies on such data.

Zhu et al.'s [10] semi-supervised learning approach using maximum-likelihood estimation and Laplace smoothing for anomaly detection demonstrates promising results. They used a mock environment of wearable sensors and based their anomalies on location, time, duration, type of activity, and the transition of activities. Unsupervised machine learning approaches are also used for anomaly detection in smart home sensor data. Jakkula and Cook [12] used a one-class support vector machine (SVM) to detect anomalous behavior in smart home data. Novák et al. [13] used self-organizing maps for detecting anomalies based on the duration of an activity like an unusually long inactivity or changes in daily activities. They also used a first order Markov model to detect 75% of the artificially injected anomalies.

Graph-based approaches have been successfully implemented in a smart home environment [16, 17]. Akter and Holder [17] represented a smart home as a graph where motion sensors are considered as a vertex and movements as edges to perform activity recognition. The graph-based features are then extracted and used as input for a Support Vector Machine. This method when compared with three other approaches, Naive Bayes, Hidden Markov Model (HMM) and Conditional Random Fields (CRF) outperformed each one of them for activity recognition. However, they do not consider the temporal information. Long and Holder [16] use three different graph-based approaches for an activity prediction by representing time-based sensor data as a graph. Although none of the graph-based approaches outperformed the non-graph SVM, they provided an uncorrelated error which demonstrates that the graph-based approach is capable of correctly classifying graphs which cannot otherwise be classified correctly. Hence, the ensemble built using all of these approaches gains 6.5% over the best classifier alone.

Although there have been several works using various machine learning techniques for predicting the cognitive health of a patient in smart home, little has been done using a graph-based approach. To the best of our knowledge, there is not any published research that deals with graph-based anomaly detection on a resident's activity that would lead to the indication of a cognitive health decline. In this paper, we address the problem of anomaly detection in elderly resident's activity in smart homes that has potential for identification of cognitive health decline using an unsupervised graph-based approach. It is not always possible to get labelled data, and predicting the cognitive health using unlabeled data can be very valuable to any smart home health monitoring system.

3 Dataset

The dataset used for analysis is the Kyoto dataset with 400 participants provided by Washington State University's CASAS program. CASAS aims to improve the comfort, safety and/or productivity of the residents with the use of smart home technology [3]. They use real-time data from sensors to analyze and monitor residents' health and behavior.

3.1 Sensor and Floorplan Information

The CASAS website provides a raw sensor log dataset (snapshot shown in Fig. 2) for each participant containing time (HH:MM:SS), sensor identification (e.g., "*M017*" represents a motion sensor and the number "017" indicates a particular area in the home), sensor value (e.g., sensors like motion, item and door that have binary states of ON/OFF, PRESENT/ABSENT or OPENED/CLOSED, etc.), and an activity number to show the activity is being executed (e.g., "*2-start, 2.1*" represents participant starting activity 2 and performing step 2.1). The details about activities are discussed in next section. The layout of the sensor setup and the apartment floorplan are shown in Fig. 1. The sensors include wide-area infrared motion sensors, temperature sensors, item sensor for selected items in the kitchen, burner sensor, hot and cold water sensor and magnetic door sensors. These sensors are set up throughout the house and record the activity of a resident, such as the location (room) they are in, the item they are using (like oven, refrigerator, water tap, etc.), activity they

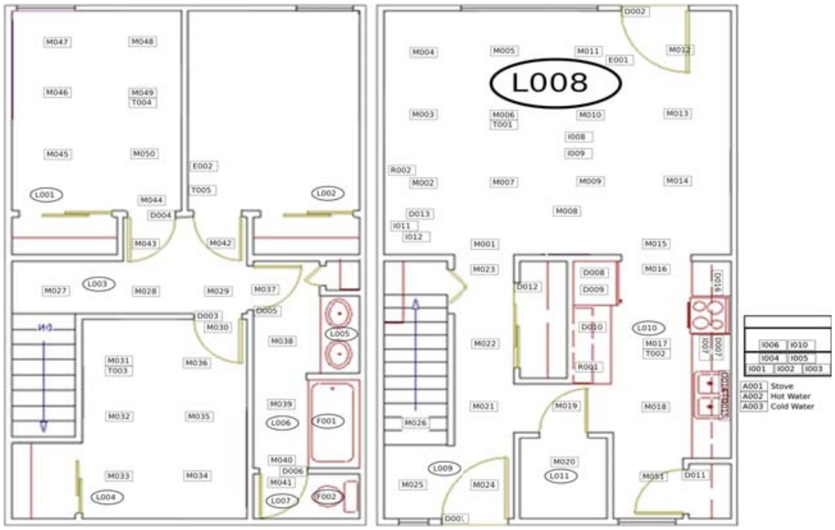

Fig. 1: Smart Home Floorplan with Sensor Layout

```
13:54:04.271642 M009 ON
13:54:04.644604 M052 OFF
13:54:06.304539 M008 ON 6-start,6.1
13:54:06.323086 M006 OFF
13:54:06.337327 M009 OFF
13:54:08.88587 M007 OFF
13:54:09.261463 M009 ON 6.5
13:54:10.08014 M052 ON
13:54:10.450021 M008 OFF
13:54:13.128099 M052 OFF
13:54:13.522625 M009 OFF
13:54:14.133687 M014 ON 6.5
13:54:15.105225 M015 ON 6.5
13:54:18.861099 M015 OFF
13:54:23.205409 A001 2.43832
13:54:25.671188 A001 2.4841
13:54:29.449646 M014 OFF
13:54:36.051103 M052 ON
13:54:38.272318 M052 OFF
13:54:40.595442 M014 ON 6.5
13:54:43.60144 M052 ON
13:54:44.640538 P001 2051
13:54:45.960414 M052 OFF
13:54:47.972658 A001 2.45207
13:54:50.407441 A001 2.48223
13:54:50.556613 M014 OFF
13:54:50.642797 P001 1697
13:54:53.330694 M014 ON 6.5
13:55:00.279838 A001 2.52059
13:55:02.740132 A001 2.49254
13:55:04.856631 M014 OFF
```

Fig. 2: Sample Sensor Data

are doing etc. Based on these sensors we can recognize the location of a participant inside a house. For example, “*M001*”, “*M002*”, “*T001*”, “*D013*”, “*I011*”, etc., are sensors installed in the living room (Labelled as “*L008*” in Fig. 1). The movement inside the house can then be mapped into a graph that can then be used for identifying anomalous activity of the participant.

3.2 Activity Information

All 400 participants were asked to perform a list of tasks. For our experiments, we chose to include the data related to eight activities called Instrumental Activities of Daily Living (IADLs) [5] as shown in Table 1. The successful completion of IADLs is required by a person to live independently because the completion of IADLs requires a high level of cognitive and functional ability [4]. Based on a resident’s ability to perform these activities, clinicians can characterize their daily behavior and find out whether they have cognitive or physical difficulties. Researchers believe that a decline in an ability to perform IADLs is often related to the decline of cognitive ability [6,7]. Sensors can keep track of the progress made while performing these activities and we can study the changes in functional and cognitive ability of a person using sensor data [4]. To complete each of these activities, the participants need to perform various corresponding steps (we will call these steps sub-activities going forward). For example, to complete activity 6, the participant has to complete 6 steps: *receive phone* (6.1), *answer questions* (6.2), *sit down during conversation* (6.3), *stand in one place* (6.4), *walk around* (6.5), *and hung up the phone* (6.6). This process of performing specific activities using their corresponding steps can be mapped using a graph because activities and steps can be considered entities and represented as nodes and the sequence and their relationship to the participant can be represented as edges. Since the behavior will vary among each activity, like the behavior for filling the medicine dispenser is different from the behavior for making a cup of soup, we will analyze the data by making separate graphs for each activity in order to discover patterns and trends based on the progress made while performing that specific activity.

TABLE I. ACTIVITY USED FOR EXPERIMENT

1.	Sweep the kitchen and dust the living room.
2.	Obtain a set of medicines and fill medicine dispenser.
3.	Write a birthday card and enclose a check.
4.	Find the appropriate DVD and watch the news clip.
5.	Obtain a watering can and water living room’s plant.
6.	Answer the phone and respond to questions.
7.	Prepare a cup of soup using the microwave
8.	Pick a complete outfit for an interview.

3.3 Data Preprocessing

The sensor log dataset for each participant are raw text files as shown in Fig. 2. We built a python-based parser tool to convert text log files into graphs. The basic layout of the graph used for our experiments is shown in Fig. 3. For our analysis, we considered *participant*, *room*, *activity*, and *sub-activity* as nodes. The location of a participant can be considered as an attribute of a participant node and was represented by the edge “*is_at*” between *participant* and *room*. In addition, a participant starts or continues doing an activity, hence, the relationship between these two entities are represented as either a “*start*” edge or a “*continue*” edge, where the participant can move from one room of the apartment to another while doing a specific activity. While a single room can have multiple sensors installed, we consider the movement of the participant only if he/she has changed rooms (i.e., out of the sight of all sensors in a room) rather than just a change in an individual sensor. In this case, movement is represented as “*move*” edge between two *participant* nodes (i.e., participants in two different rooms). While performing a sub-activity in a room several motion sensors can be activated. We counted the number of sensors activated and calculated the time duration the participant spent doing a sub-activity in that location and bucketized values into three bins labelled “*low*”, “*mid*” and “*high*”. We came up with these three bins by calculating the mean (μ) and standard deviation (σ) of count and duration for sensors for each sub-

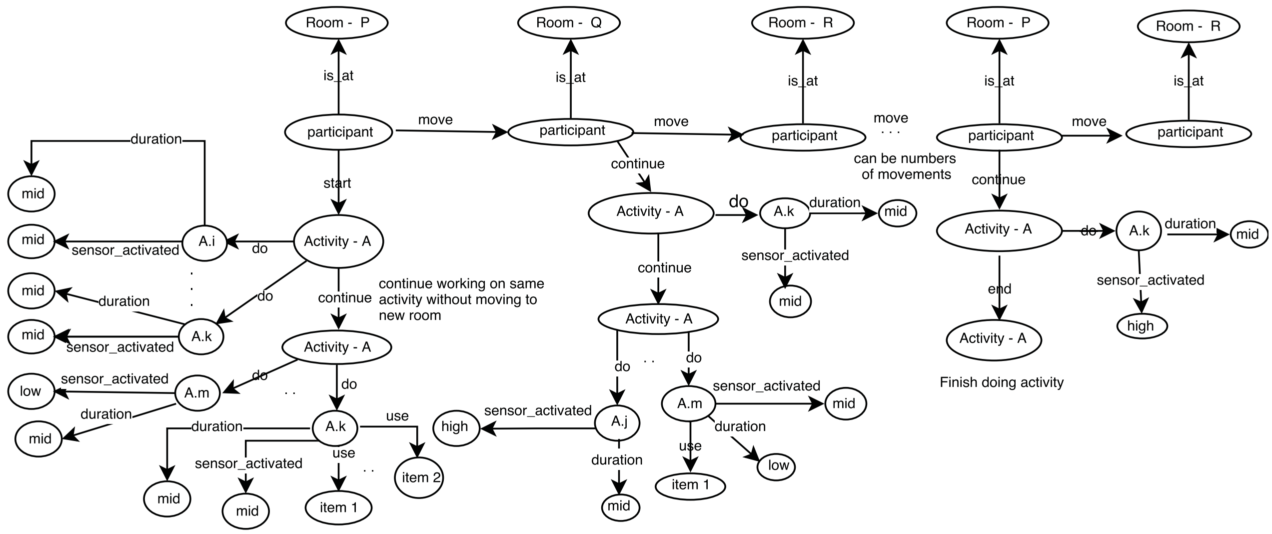

Fig. 3: General Graph Layout

activity across all 400 participants and using the following formula: "*low*", $if\ val < (\mu\text{-}\sigma)$, "*mid*" $if\ (\mu\text{-}\sigma)\ val < (\mu + \sigma)$, and "*high*", $if\ val > (\mu + \sigma)$. Both "*sensor_activated*" and "*duration*" were added as attributes to the *sub-activity* node by creating an edge from the *sub-activity* node and its value (with label "*low*", "*mid*", or "*high*") node. Furthermore, a participant might use various items while doing a sub-activity. For example, to perform sub-activity 2.3 "*fill dispenser with medication"*, the participant uses the medicine dispenser represented by sensor id, "*I006*" or "*I010*". To represent this, we added an attribute to the node "*sub-activity*" by making an edge called "*use*" between item "*I006*" or "*I010*" and the "*sub-activity*".

Our proposed graph layout is based on a sequence of activities and their duration for a corresponding participant. This proposed graph topology can then be used to discover interesting patterns, such as when a resident forgets the required sequence of activities, or the duration is unusually long. Our hypothesis is that by discovering normal and anomalous patterns in these activity graphs, medical personnel will be better able to determine whether a resident is in the stages of developing cognitive disabilities.

4 Graph Based Approach

In order to lay the foundation for this effort, we hypothesize that a real-world, meaningful definition of a graph-based anomaly is an unexpected deviation to a normative pattern, which we define as follows:

Definition 1. *A labeled graph G = (V, E, F), where V is the set of vertices (or nodes), E is the set of edges (or links) between the vertices, and the function F assigns a label to each of the elements in V and E.*

Definition 2. *A subgraph SA is anomalous in graph G if (0 < d(SA, S) < TD) and (P(SA|S) < TP), where P(SA|S) is the probability of an anomalous subgraph SA given the normative pattern S in G. TD bounds the maximum distance (d) an anomaly SA can be from the normative pattern S, and TP bounds the maximum probability of SA.*

Definition 3. *The score of an anomalous subgraph SA based on the normative subgraph S in graph G is d(SA, S) * P(SA|S), where the smaller the score, the more anomalous the subgraph.*

The advantage of graph-based anomaly detection is that the relationships between entities can be analyzed for structural oddities in what could be a rich set of information, as opposed to just the entities' attributes. However, graph-based approaches have been prohibitive due to computational constraints, because graph-based approaches typically perform subgraph isomorphism, a known NP-complete problem. Yet, in order to use graph-based anomaly detection techniques in a real-world environment, we need to take advantage of the structural/relational aspects found in dynamic, streaming data.

In order to test our approach, we will use the publicly available GBAD test suite, as defined by [9]. Using a greedy beam search and a minimum description length (MDL) heuristic, GBAD first discovers the "best" subgraph, or normative pattern, in an input graph. The MDL approach is used to determine the best subgraph(s) as the one that minimizes the following:

$$M(S, G) = DL(G|S) + DL(S), \qquad (1)$$

where G is the entire graph, S is the subgraph, DL(G|S) is the description length of G after compressing it using S, and DL(S) is the description length of the subgraph. The complexity of finding the normative subgraph is constrained to be polynomial by employing a bounded search when comparing two graphs. Previous results have shown that a quadratic bound is sufficient to accurately compare graphs in a variety of domains [9]. For more details regarding the GBAD algorithms, the reader can refer to [9]. In summary, the key to the GBAD approach is that anomalies are discovered based upon small deviations from the norm – not outliers, which are based upon significant statistical deviations from the norm.

5 Experimentation

Out of the 400 participants in Kyoto dataset, we selected all healthy participants (239 in total) as well as randomly selected 3 participants who have been previously diagnosed with cognitive impairment. We made this decision based upon the fact that while this particular dataset has an even distribution of healthy patients and non-healthy patients, in the real-world, many more patients will be healthy, as the anomalies would be the non-healthy ones. (And our choice of 3 was arbitrary.) We created separate graphs for each of the eight IADL activities for all 242 participants which we will call *activity graph*. With 21,293 vertices and 21,050 edges, *activity graph 1* (i.e., *sweep kitchen and dust living room*) is the biggest graph while *activity graph 3 (i.e., write birthday card)* is the smallest (only 4367 vertices and 4125 edges).

After running GBAD on all 8 *activity graphs*, we discovered various patterns and anomalies. Fig. 4 (b) shows the example of a temporal anomaly (a node with a white background) obtained while running GBAD on the activity 5 (*water plants*) graph. The normative pattern (as seen in Fig. 4 (a)) indicates that the time duration required for filling the water can is "*mid*", but the anomalous instant shows that this particular participant took a longer time to fill a water can, showing a duration of "*high*" as an anomaly. Fig. 5 (b) shows an example of a spatial anomaly (marked as a white) obtained while running GBAD on activity 8 (*pick outfit for interview*) graph. The normative pattern (as shown in Fig. 5 (a)) indicates that the participant is in a closet to pick out an outfit while the anomalous instance indicates that the participant is trying to pick out an outfit from a wrong location, i.e., the living room. Hence, this is a spatial anomaly. Similarly, Fig. 6 (b) shows the example of a behavior anomaly (marked as a white node) on the activity 6 (*answer the phone and respond to questions*) graph. The normative pattern (as shown in Fig. 6 (a)) shows the snapshot of the activity 6 graph where the participant is in the living room while he/she starts answering the phone; they first receive the call and then continue talking (indicating other steps are followed after this). But the anomalous instance indicates that the participant ends the call right after receiving it. This tells us that remaining steps like *answer questions*, *sit down*, *stand in one place*, *walk around*, and *hang up the phone* are not performed – a deviation from normal behavior. Hence, this is a behavior anomaly.

These three examples demonstrate our initial hypothesis: a *graph based approach can successfully discover anomalies in elderly resident activity in smart homes*. Fig. 4, 5, and 6 only show individual examples of anomalous and normative patterns. It should be noted that while we have represented the graph using this particular topology, the choice of topology was somewhat arbitrary. There are many other ways this data could be represented as a graph, and additional experiments that were performed showed similar results, but were too numerous to represent within the constraints of this paper. To empirically test our second hypothesis, we will test whether three participants, P1, P2, and P3, with cognitive impairments, have instances of the three anomaly types in their *activity graphs*. Table 2 shows the performance of GBAD for participants P1, P2, and P3, and will be discussed in detailed in the following sections.

5.1 Case Study – P1

When inspecting the results of GBAD on all eight *activity graphs*, P1 was flagged with all three types of anomalies. As shown on Table 2, GBAD reported a temporal anomaly on *activity graph* 1, 2 and 7, a spatial anomaly on *activity graph* 8, and a behavior anomaly on *activity graph* 6. The instances of temporal anomalies showed that P1 was taking a longer time to perform various sub-activities. Inspecting each *activity graph* of P1, we discover the following temporal anomalies:

1. On *activity graph* 1, duration and sensor activated was "*high*" every time when P1 was performing sub-activity *dust dining room* or *dust dining room.*
2. On *activity graph* 2 P1 had two instances of temporal anomaly. P1 took a longer time performing sub-activities *fill medicine dispenser* while took less time performing *read instruction* (i.e., duration was "*low*").
3. On *activity graph* 7 (i.e., *make a cup of soup*), duration was "*high*" for *retrieve materials*. Inspecting this graph we saw that P1 could only complete 2 sub-activities out of 9 and never finishes making a cup of soup.

Further inspecting the spatial anomaly on *activity graph* 8, we found that P1 was choosing the outfit for an interview in

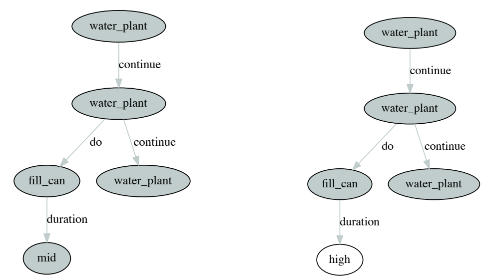

Fig. 4. a) Normative Pattern b) Temporal Anomaly

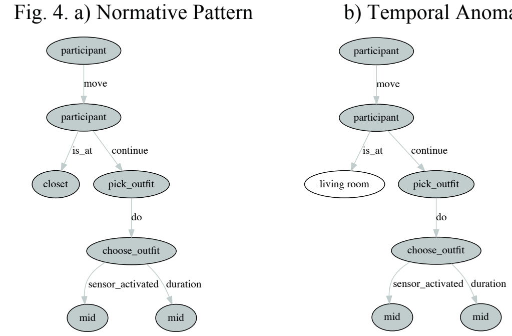

Fig. 5. a) Normative Pattern b) Spatial Anomaly

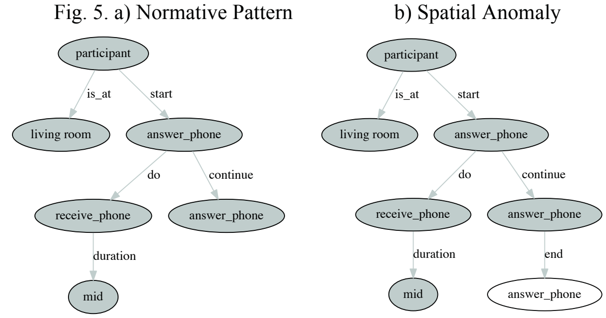

Fig. 6. a) Normative Pattern b) Behavior Anomaly

TABLE II. GBAD RESULTS FOR PARTICIPANTS P1, P2, AND P3

Activity Graphs	Anomaly		
	Temporal	Spatial	Behavior
1. Sweep kitchen & dust living room	P1,P2,P3		
2. Fill medicine dispenser	P1, P2		
3. Write a birthday card			
4. Watch news clip	P2		
5. Water plants in living room	P2, P3		P2, P3
6. Answer phone	P3		P1
7. Prepare a cup of soup	P1, P2		
8. Choose outfit for an interview	P3	P1,P2,P3	

the living room, instead of in the closet, as indicated by the normative pattern for this specific sub-activity (graph shown in Fig. 5 (b)). Similarly, the behavior anomaly on *activity graph* 6 shows that P1 ended the phone call just after receiving it. This tells us that sub-activities like *answer questions*, *sit down*, *stand in one place*, *walk around*, and *hang up the phone* were not performed (shown in Fig. 6 (b)).

5.2 Case Study – P2

For participant P2, GBAD, after running on all eight *activity graph*s, was able to successfully detect all three types of anomaly. As shown in Table 2, P2 had a temporal anomaly on *activity graph* 1, 2, 4, 5 and 7, and a spatial anomaly on *activity graph* 8, and a behavior anomaly on *activity graph* 5. Further inspecting temporal anomaly results, we discover that:

1. P2 took a longer time (i.e., duration was "*high*") and wandered a lot (i.e., sensor activated was "*high*") performing sub-activities *sweep kitchen* as well as *dust dining room* on *activity graph* 1.
2. On *activity graph* 2 (i.e., fill medicine dispenser), P2 duration was "*high*" for sub-activity *read instruction*.
3. P2 took a longer time to read instructions for watching a DVD in *activity graph* 4, and could only perform 2 sub-activities out of 8.
4. On *activity graph* 5, P2 took a longer time performing the sub-activity *water table plant*.
5. Similarly, on *activity graph* 7, the duration for sub-activity *boil water* was "*high*" for P2.

The spatial anomaly on *activity graph* 8 resembles what was observed with P1 i.e. choosing an outfit by staying in the living room, instead from the closet (graph shown in Fig. 5 (b)). The behavior anomaly for P2 was found while running *activity graph* 5 on GBAD, and with further inspection of the anomalous instance, we found that P2 was using the stove (or maybe leaving stove open) while doing sub-activity *water table plant* which seems unusual based on the normal behavior where no stove is used for watering the table plant. This tells us that P2 is deviating from normal patterns for watering table plants (i.e., activity 5).

5.3 Case Study – P3

For participant P3 sensor records for activity 2, 3, 4 and 7 were missing in the dataset so the experiment was done only on the remaining four *activity graphs* (i.e., 1, 5 6, and 8). As shown in Table 2, GBAD was able to successfully detect all three types of anomaly on all four remaining *activity graphs*. P3 had a temporal anomaly on *activity graphs* 1, 5, 6 and 8, a spatial anomaly on *activity graph* 8 and a behavior anomaly on *activity graph* 5. By studying individual *activity graphs* for the temporal anomalies we found that:

1. Like P1 and P2, on *activity graph* 1, P3 was taking a longer time to sweep kitchen and dust the dining room.
2. On *activity graph* 5, the duration for three sub-activities *fill can, water table plant* and *water window plant* was "*high*" which in a normal scenario was "*mid*".
3. On *activity graph 6*, P3 walks around for a longer time, i.e., duration was "*high*" for sub-activity *walk around*.
4. On *activity graph* 8, duration for sub-activity *choose outfit* and *goto closet* was "*high"* instead of "*mid*".

Like P1 and P2, the spatial anomaly on *activity graph* 8 for P3 resembles the anomalous instance shown on Fig. 5 (b). P3 is also choosing the outfit for an interview by staying in the living room, instead of choosing it from closet. Finally, GBAD showed a behavior anomaly for P3 on *activity graph* 5. Inspecting the anomalous instance, we found that the normal behavior (as the normative pattern suggested) was to empty the extra water and return the watering can to the supply closet after watering the plants. But P3 did not return the watering can to the supply closet after emptying the extra water. This tells us that P3 might have forgotten this step, showing a deviation from the normal routine. The normal sequence for watering plant is as follows: *retrieves watering can from supply closet* (5.1), *fills watering can* (5.2), *water window plants* (5.3), *waters coffee table plants* (5.4), *empties extra water* (5.5), and *returns can to supply closet* (5.6).

6 Discussion

After running GBAD on all eight IADL *activity graphs*, various activities for participant P1, P2, and P3 were flagged as anomalous. GBAD could not find any anomalous activities in *activity graph 3* (i.e., write a birthday card) but all other seven *activity graphs* showed anomalies related to P1, P2, and P3. The most common anomaly demonstrated by the three impaired participants was a temporal anomaly (i.e., having duration "*high*" or "*low*" for performing a sub-activity) which was observed on six activity graphs out of eight. This indicates that participants are struggling to complete the tasks in a normal amount of time. All three participants showed this anomalous behavior while performing at least 3 tasks out of 8. Inspection individual graphs, we saw that P1 was unable to complete activity 7 (make a cup of soup) and P2 was unable to complete activity 4 (watch news DVD). All three participants were also flagged for spatial anomalies in *activity graph 8* for choosing an outfit from the wrong place (i.e., in the living room instead of the closet). P2 (using stove while watering plant) and P3 (missing to return water can to supply) were flagged for a behavior anomaly in *activity graph 5*. Also P1 was flagged for a behavior anomaly in *activity graph* 6 for ending a phone call without answering questions.

All of these anomalous behaviors were discovered during our experiments suggesting that the participants are struggling to successfully complete Instrumental Activity of Daily Living (IADL). They are either taking a longer time to complete the task, missing steps while doing the task, or in some cases not even able to finish the task. Since researchers believe that the decline in an ability to perform IADLs is often related to the decline of cognitive ability [6,7], the results shown by our experiments suggest that these anomalies (temporal, spatial and behavior) in a participant's activity graph are indeed indicators of a decline in cognitive ability. The fact that P1, P2 and P3 are known to have mild cognitive impairments supports the argument further, and our second hypothesis "*Anomalies are potential indicators of a decline in the cognitive health of an elderly resident*". It should be noted that healthy participants were also present in the anomaly list returned by GBAD. But, further inspecting their *activity graphs,* there was no evidence these participants had cognitive decline because they either showed single instances as an anomaly or showed only one type of anomaly.

7 Conclusions & Future Work

In recent years, sensor-based smart home environments have been successfully used with the intention of improving the independent living of elderly residents. Smart homes aim to not interfere with the normal activities of the residents and hope to reduce the cost of health care associated with caring for the resident. Since elderly residents are more susceptible to having cognitive health issues, understanding their everyday behavior using some kind of automatic tool on sensor data can provide important knowledge on the status of their health. In this work, we demonstrated that one aide in achieving this can be realized through using a graph-based approach on data collected from residents' activities. We have represented activity data from smart home sensors as a graph and used an unsupervised graph approach to find temporal, spatial and behavior anomalies in elderly resident's daily activities. We also theorized that these anomalous behaviors represent possible scenarios where a participant could have a decline in cognitive ability.

Smart home activity data can be generated in real time, i.e., in the form of a data stream. In the future, we would like to extend our experiments to a real time data stream. We plan to convert real time sensor logs into graph streams and look for anomalies in graph streams which could support a real-time health monitoring tool for residents and aide clinicians or nurses. For our experiment, we randomly chose three cognitively impaired individuals. In the future, we would like to vary the sample and run multiple experiments to see if similar anomalies can be detected. We would also like to investigate the robustness of the graph topology to see how much the change of graph topology would affect the outcome of our anomaly detection. In addition, we hope to involve a clinician as a domain expert so that we can validate our hypothesis that these anomalies are actual indicators of cognitive health decline (or MCI).

Acknowledgment

We would like to thank Diane Cook (Washington State University) and the CASAS project for providing the dataset and subsequent aide with understanding the data.

8 References

[1] Ortman, J. M., Velkoff, V. A., Hogan, H. *et al.*, *An aging nation: the older population in the United States*. United States Census Bureau, Economics and Statistics Administration, US Department of Commerce, 2014.

[2] Association, A. *et al.*, "2016 alzheimer's disease facts and figures," *Alzheimer's & Dementia*, vol. 12, no. 4, pp. 459–509, 2016.

[3] Cook, D. J., Crandall, A. S., Thomas, B. L., and Krishnan, N. C., "Casas: A smart home in a box," *Computer*, vol. 46, no. 7, pp. 62–69, 2013.

[4] Dawadi, P., Cook, D., Parsey, C., Schmitter-Edgecombe, M., and Schneider, M., "An approach to cognitive assessment in smart home," in *Proceedings of the 2011 workshop on Data mining for medicine and healthcare*. ACM, 2011, pp. 56–59.

[5] Lawton, M. and Brody, E., "Instrumental activities of daily living scale (iadl)," *Gerontologist*, vol. 9, pp. 179–186, 1969.

[6] Galvin, J. E. and Sadowsky, C.H., "Practical guidelines for the recognition and diagnosis of dementia," *The Journal of the American Board of Family Medicine*, vol. 25, no. 3, pp. 367–382, 2012.

[7] Martin, S., Kelly, G., Kernohan, W. G., McCreight, B., and Nugent, C., "Smart home technologies for health and social care support," *The Cochrane Library*, 2008.

[8] Wilson, C., Hargreaves, T., and Hauxwell-Baldwin, R., "Smart homes and their users: a systematic analysis and key challenges," *Personal and Ubiquitous Computing*, vol. 19, no. 2, pp. 463–476, 2015.

[9] Eberle, W. and Holder, L., "Anomaly detection in data represented as graphs," *Intelligent Data Analysis*, vol. 11, no. 6, pp. 663–689, 2007.

[10] Zhu, C., Sheng, W., and Liu, M., "Wearable sensor-based behavioral anomaly detection in smart assisted living systems," *IEEE Transactions on Automation Science and Engineering*, vol. 12, no. 4, pp. 1225–1234, 2015.

[11] Lotfi, A., Langensiepen, C., Mahmoud, S. M., and Akhlaghinia, M. J., "Smart homes for the elderly dementia sufferers: identification and prediction of abnormal behaviour," *Journal of ambient intelligence and humanized computing*, vol. 3, no. 3, pp. 205–218, 2012.

[12] Jakkula, V. R. and Cook, D. J., "Detecting anomalous sensor events in smart home data for enhancing the living experience." *Artificial intelligence and smarter living*, vol. 11, no. 201, p. 1, 2011.

[13] Nova ́k, M., Bin ̌as, M., and Jakab, F., "Unobtrusive anomaly detection in presence of elderly in a smart-home environment," in *ELEKTRO, 2012*. IEEE, 2012, pp. 341–344.

[14] Dawadi, P. N., Cook, D. J., and Schmitter-Edgecombe, M., "Automated cognitive health assessment from smart home-based behavior data," *IEEE journal of biomedical and health informatics*, vol. 20, no. 4, pp. 1188– 1194, 2016.

[15] Cook, D. J., Schmitter-Edgecombe, M., and Dawadi, P., "Analyzing activity behavior and movement in a naturalistic environment using smart home techniques," *IEEE journal of biomedical and health informatics*, vol. 19, no. 6, pp. 1882–1892, 2015.

[16] Long, S. S. and Holder, L.B., "Using graphs to improve activity prediction in smart environments based on motion sensor data," in *International Conference on Smart Homes and Health Telematics*. Springer, 2011, pp. 57–64.

[17] Akter, S. S. and Holder, L. B., "Activity recognition using graphical features," in *Machine Learning and Applications (ICMLA), 2014 13th International Conference on*. IEEE, 2014, pp. 165–170.

[18] Tran, A. C., Marsland, S., Dietrich, J., Guesgen, H. W., and Lyons, P., "Use cases for abnormal behaviour detection in smart homes," in *International Conference on Smart Homes and Health Telematics*. 2010, pp. 144–151.