Elderly Action Prediction and Anomalous Activity Detection in Smart Homes through Profiling Residents' Behavior

Malihe Erfanmanesh [ID] [1], Hooman Tahayori [ID] [1, *] and Andrea Visconti [ID] [2]

[1]Department of Computer Science and Engineering and IT, Shiraz University, Shiraz, Iran
[2]Department of Computer Science, Universita Degli Studi di Milano, Milano, Italy

*Corresponding author: Department of Computer Science and Engineering and IT, Shiraz University, Shiraz, Iran. Email: tahayori@shirazu.ac.ir

Abstract

Background: Elderly healthcare is one of the important issues in an aging society. Smart homes in healthcare domain help the elderly to be continuously monitored, instead of being under supervision at expensive health centers, and hence enable them to live independently. This service requires detecting and monitoring of residents' normal activities of daily living in smart homes.

Objectives: By profiling residents' behavior and identifying changes in normal activities of the elderly over time, one can detect anomalous behavior and determine whether their health status declines. Hence, the possibility of preventive care for some elderly people or providing assistance to the elderly will be partly provided in case of occurrence of the anomalies.

Methods: In this paper, first a method was proposed for detection and prediction of elderly activities by extracting several features from available information. In the second step, statistical measures were applied on the features to profile the elderly's behavior. The AdaBoost learning algorithm was used for detecting the anomalies and modeling normal/abnormal behavior.

Results: For detection and prediction of the activity, the proposed method was tested using a dataset collected in the "eHealth Monitoring Open Data Project". The accuracy of 98.48% was obtained by considering features of start time, end time, duration, location, previous action, water, and electric device use. Anomalous behavior was detected in the same dataset with the average f-score of 90%.

Conclusions: Results of the present study revealed that, the proposed methods are effective for detecting abnormal actions of the residents in smart homes to a fairly good level. This enables the elderly to live independently while being under continuous monitoring and significantly reduces the elderly health care costs.

Keywords: Machine Learning, Action Recognition, Smart Home, Anomaly Detection, Behavior Profiling

1. Background

Aging naturally comes with various diseases and disabilities. Hence, with the increase in the elderly population, their long-term care imposes high costs on medical and healthcare systems around the world. It is estimated that by 2050 the total cost associated with the care of patients with Alzheimer reaches $1 trillion (1). Also, in the United States, heart diseases cost over $200 billion due to the required health care, medications, and lost productivity (2). Such high-risk patients, more or less, can lead their life independently, however, they require special attention, which is the most challenging task for the healthcare systems around the globe.

Due to the changes in the lifestyles of the elderly, they have to spend most of their time alone, which they often prefer to do. Therefore, helping them live independently is one of the main goals of smart homes in the domain of healthcare. It is expected that smart homes for patients, especially the elderly, bring some caring services like reminding them to take medicine, monitoring vital signs, predicting health issues, informing emergencies in case of an unexpected accident, etc. Increasing safety, promoting quality of life, and reducing the feeling of isolation are some aspects of a smart home, which may help the elderly (3).

Monitoring the health status of smart home residents is feasible through the use of various sensors. Such sensors range from camera, microphones, wearable sensors, temperature sensors, pressure sensors, and motion detectors. These sensors are categorized into obstructive-like wearable sensors, cameras and microphones- and none-obtrusive (4). Due to the privacy issues, costs of installation and maintenance, volume of generated data, and ease of processing the gathered data, the latter types of sensors are more preferred.

Monitoring the residents' health and mobility can be

accomplished by profiling the normal behavior of the residents and comparing them with their daily living activities. Normal behaviors and daily activities are detected through the data that is gathered by installed sensors in the smart home. If the person's health diminishes or any unexpected accident happens, it would be reflected in his/her daily activities. For example, the duration of the person's activity would last longer than usual, interruption between activities increases, and activities would be performed at a time that is not expected or even would be missed. Moreover, the presence of a person in a location that is normally not expected to be or vice versa might be another sign of anomaly. Detecting such cases helps to provide advice to the individual, informing those who are in relation with the person or even sending medical staff to attend to the patient (5).

2. Objectives

The contribution of this paper is related to the scope of remote health monitoring and assisted living through smart homes equipped with none-obtrusive sensors. In the paper, in the first stage, we propose an accurate and fast method to recognize the normal activities of elderly persons and predict their future occurrences. In the second stage, we use several statistical measures to determine the behavior profile of the elderly. Using the profile, we label activities as normal or abnormal. Finally, we detect anomalies and predict future occurrences of them with a high accuracy using AdaBoost classifier.

3. Methods

3.1. Related Works

In this section, we review researches in two categories, those works that are focused on modeling and predicting behaviors of residents in a smart home, and the papers on anomalous behavior detection in smart homes.

As considerable works in the prestigious CASAS project on smart home, Singla and Cook (6), tried to recognize and track activities in complex situations e.g. when activities are interleaved or correlated. They utilized Naive Bayes, HMM, and extended versions of HMM techniques. Later Singla et al. (7), provided a method for detecting more than one person in an environment. They exploited the Hidden Markov Model (HMM), which first considered both individuals. Then, by considering a separate model for each individual, the accuracy of detecting the activities was improved. In the same project of CASAS, Cook and Schmitter-Edgecombe (8), proposed a method to assess the quality of

activities in smart environments. They modeled normal activities using the Naïve Bayes and Markov models and by detecting the states that were ignored, they were able to detect anomalies.

Chen et al. (9), designed a data mining framework for activity recognition in smart homes. They extracted useful features from raw sensor data, which was collected from the CASAS smart home environment and then applied feature selection methods to select the most important and optimal features. They utilized Bayes Belief Networks, Artificial Neural Network, Sequential Minimal Optimization, and LogitBoost machine learning techniques; however, their obtained accuracy was not better than 90%. Nazerfard and Cook (10), used Bayesian networks for daily activity prediction. They proposed an approach to predict next activity and its starting time based on modeling the relative start time using continuous normal distribution and outlier detection. In their work, however, they could not reach the prediction accuracy of more than 74%. Dawadi et al. (11), developed a framework to perform an automated cognitive assessment of an individual by analyzing the individual's performance on activities of daily living in a smart home. They extracted features to indicate how well participants performed the activities and then used Naïve Bayes, J48, SMO, Neural Network machine learning approach for assessment. The highest accuracy that they reached was 88.63%.

Rashidi et al. (12), proposed a method for health monitoring and assisting individuals having difficulties living independently at home. They used the unsupervised method for discovery and recognition of activities. They discover activities by Discontinuous Varied-Order Sequential Miner (DVSM) and k-mean clustering algorithms. Then, they recognize activities by a boosted version of HMM. Their proposed method, however, could not distinguish some similar activities and could not identify concurrent activities.

Lotfi et al. (13), proposed a method to support the independent living of the elderly with dementia by detection and prediction of abnormal behavior. They used a recurrent neural network to predict future activities. Finally, they informed caregivers about any anomalies that may occur in the future.

Raeiszadeh et al. (14, 15), have proposed a pattern discovery method to recognize activities of daily livings in smart homes. After converting sensor data into event sequences, frequent and repeated sequential patterns are mined by using PrefixSpan and LCS algorithms and then an activity recognition model is created to predict the future occurrence of activities. They have applied their work on CASAS and MIT datasets and have achieved accuracies between 80% and 95%.

Amirjavid et al. ([16](#)), modeled activities as a chain of fuzzy events to predict the intention of smart home residents when they perform a few actions. The difficulty with this method is that the beginning and ending points of the activities had to be clear.

Liouane et al.'s studies ([17](#), [18](#)) can be counted as notable works on eHealth dataset. Liouane et al. ([17](#)), proposed a prediction method called Recurrent Extreme Learning Machine (RELM), that provides the ability to learn human behavior and accurately predicting the future activities of the elderly inhabitant. They applied their method on eHealth dataset. Liouane et al. ([18](#)), had introduced an algorithm for detecting abnormal activities through calculating the normal time length for performing each activity.

In this paper we propose an approach that incorporates easy to interpret features that can be simply extracted from cheap and easy-to-install sensors with supervised learning methods, and provide higher accuracy than existing methods, both for action and anomaly detection.

3.2. Data Set

We use the dataset of "eHealth Monitoring Open Data Project", which is an open data set for monitoring and healthcare of dependent persons such as elderlies ([19](#)). The dataset contains scenarios that describe daily living activities of an elderly person for about one-year with and without changes in dependency level.

This dataset exists in two versions of human-readable and coded. In the human-readable version, the starting and ending dates and time of the action (in the format of hour: minute: second) and the action name are given. In the coded version, information regarding the starting and ending time of the action in seconds (day and time of action are combined together), and action code is available. In this paper, we use the coded version.

An excerpt of the records of the dataset in both versions is shown in Table 1. Some of the actions that are mentioned in the dataset with their corresponding code are shown in Table 2.

3.3. Proposed Method

In the first phase, we focused on the detection and prediction of actions of the elderly and in the second phase, through labeling normal and abnormal actions, we detect and predict anomalies by the use of a supervised learning algorithm.

3.3.1. Action Detection and Prediction Phase
3.3.1.1. Definitions

Action vs. activity: An action is a more detailed concept than activity. Each activity can include several actions.

For example, "make some tea", "make some coffee", "make a sandwich", and "wash the dishes" are separate actions, however, all of them can be related to the activity of meal preparation.

In this paper, most of our focus is on actions rather than activities.

Features: It should be noted that when an action happens, there are a set of environmental conditions and characteristics associated with it, which we refer to as features. Time and location are some well-known features. Some features are related to a specific action and some of them are common between different actions. For example, using an electrical device like a TV remote control is just related to the "watching TV" action but location as a feature can be seen in several actions like "making coffee", "making tea", "washing the dishes" and "making a sandwich". We will show that selecting more detailed and dedicated features increases the chance of correct recognition of actions.

3.3.1.2. Supervised Learning

In this paper, we explore a supervised learning approach to detect and predict smart home residents' actions. For this purpose, we use Random Forest Classifier (RFC), Gaussian Naive Bayes (GNB), Decision Tree (DT), and KNN Classifier models, since they are fast and easy to train.

A series of features are explicitly extractable from the dataset, while some other features can be extracted from the dataset implicitly.

Initially, we used the time duration feature from the dataset for action detection and prediction. We also considered two features of start time and end time for each action.

Since most actions usually occur in specific locations, we added the location feature for each action. We divided the home environment into the hypothetical regions including kitchen, bathroom, bedroom, restroom, office, living room, and hallway, and we added the corresponding location to each action.

We also considered the previous action as another feature. The previous action represents the action that precedes the current action. Finally, two other features were considered that were related to the use of water and electrical equipment. Regarding water usage, three codes were devised, one relating to the actions that necessarily require water usage, another relating to actions that do not require water, and the third code refering to the actions that we were not sure if they needed water or not. Subsequently, each code was assigned to the corresponding action. In terms of power consumption, two codes were considered, one for actions that use an electrical device and the other for actions that do not; then, the related code was assigned to the corresponding action.

Table 1. Sample Records of E-Health Coded and Human Readable Version

Human Readable Version			Coded Version		
Activity	**Start Time**	**End Time**	**Activity Code**	**Start Time**	**End Time**
Washing (take shower)	01 day - 08:03:32	01 day - 08:22:40	3	29012	30160
Hair dray	01 day - 08:23:46	01 day - 08:26:53	5	30226	30413
Change clothes	01 day - 08:28:50	01 day - 08:38:39	2	30530	31119
Toileting	01 day - 08:40:37	01 day - 08:50:24	7	31237	31824
Washing hand/face	01 day - 08:52:12	01 day - 08:55:38	4	31932	32138

Table 2. Some of Dataset Actions' with Related Code

Action Code	Action Label
1	Eating
2	Wear/take off shoes
3	Take shower
4	Washing hand/face
5	Hair dry
6	Make up
12	Wash dish
13	Make coffee
14	Make tea
15	Make sandwich
16	Make hot food

3.3.2. Anomaly Detection Phase

At this stage of our work, for detecting and predicting anomalous actions, we used part of a eHealth dataset, which was related to elderly's information by changing the level of dependency during one year.

Features taken from the previous phase that are used in this stage include start time, end time, and the duration of the action. In addition to the mentioned features, the start day of the action, the interval between the end of the previous action to the starting time of the current action, and the current performed action are also considered. Based on these features user profiles are made. An action with respect to the profile of the person is considered as an anomaly if:

- Its duration lasts more or less than normal.

- An unauthorized delay occurs between consecutive actions.

- The action starts or ends at times that are not expected.

- An invalid action occurs before the current action.

Initially, we used several statistics to separate normal actions from abnormal and label them accordingly. In the following, we will discuss each of the used methods.

3.3.2.1. Min-Max Range Based on First Season

In this method, because the dependency of the elderly is still constant during the first season (according to the information provided by the dataset owners), we chose the first season as the basis to determine the normal range for the features of each action. More precisely, we assumed all actions that are performed in the first season i.e., data gathered in the first three months, are normal actions. Using first season records for each action, we calculate the normal range for all the features. Given the feature f of the jth sample of the ith action be represented as $Action_j^i(f)$, then

$$Action^i(start\,time)$$
$$= \left[\min_j \left\{ Action_j^i(start\,time) \right\}, \max_j \left\{ Action_j^i(start\,time) \right\} \right] \tag{1}$$

$$Action^i(end\,time)$$
$$= \left[\min_j \left\{ Action_j^i(end\,time) \right\}, \max_j \left\{ Action_j^i(end\,time) \right\} \right] \tag{2}$$

$$Action^i(duration)$$
$$= \left[\min_j \left\{ Action_j^i(duration) \right\}, \max_j \left\{ Action_j^i(duration) \right\} \right] \tag{3}$$

$$Action^i(time\,from\,previous\,action)$$
$$= \left[\begin{matrix} \min_j \left\{ Action_j^i(time\,from\,previous\,action) \right\}, \\ \max_j \left\{ Action_j^i(time\,from\,previous\,action) \right\} \end{matrix} \right] \tag{4}$$

Moreover, we prepare the list of possible actions that are performed before each action. For determining the list, if an action is performed only one or two times before the specific action during the season, it will not be added to the list. After determining the valid range for all features of each of the actions, we reviewed all samples of other seasons. If the action had at least one feature out of the valid range of the underlying feature, we labeled that as an anomaly otherwise, we considered the action as a normal one.

3.3.2.2. Mean +/- 3stdev Range Based on First Season

In this method, similar to the previous one, we considered the first season as the basis for extracting normal intervals for the features of the actions. By considering all samples of the first season, we extracted the time valid intervals for all the features of each action as,

$$Action^i \left(start\ time\right) = \underset{j}{mean} \left\{ Action^i_j \left(start\ time\right) \right\}$$
$$\pm\ \underset{j}{stdev} \left\{ Action^i_j \left(start\ time\right) \right\} \tag{5}$$

$$Action^i \left(end\ time\right) = \underset{j}{mean} \left\{ Action^i_j \left(end\ time\right) \right\}$$
$$\pm\ \underset{j}{stdev} \left\{ Action^i_j \left(end\ time\right) \right\} \tag{6}$$

$$Action^i \left(duration\right) = \underset{j}{mean} \left\{ Action^i_j \left(duration\right) \right\}$$
$$\pm\ \underset{j}{stdev} \left\{ Action^i_j \left(duration\right) \right\} \tag{7}$$

$$Action^i \left(time\ from\ previous\ action\right)$$
$$= \underset{j}{mean} \left\{ Action^i_j \left(time\ from\ previous\ action\right) \right\} \tag{8}$$
$$\pm\ \underset{j}{stdev} \left\{ Action^i_j \left(time\ from\ previous\ action\right) \right\}$$

Then, by examining and comparing the samples of other seasons, if all the values of the features for each action are within the valid range, we give the normal label to the action, and otherwise, labeling the action as abnormal.

3.3.2.3. Mean +/- 3 stdev Range Based on Total Year

In this method, instead of placing just one season as a basis, the whole year is considered for determining the normal range of features. In this method, for labeling each sample, possible ranges of features for each action is calculated based on all other records within one year. If all the values of the sample features were within the valid range, then we give the action a normal label, and otherwise, we label it as abnormal.

3.3.2.4. Inter-Quartile Range Based on First Season

In this method, the first season is considered for calculating the normal range of features. After calculating the possible range of $(Q_1 - 1.5 \text{*IQR}, Q_3 + 1.5 \text{*IQR})$ for all features of each action, records of the remaining three seasons are labeled according to the specified intervals. Valid ranges are calculated with respect to the inter-quartile range of the values for each feature, i.e.

$$Action^i \left(start\ \ time\right) \tag{9}$$
$$= \begin{bmatrix} Q1 \left(Action^i \left(start\ time\right)\right) - 1.5 \times IQR \left(Action^i \left(start\ time\right)\right), \\ Q3 \left(Action^i \left(start\ time\right)\right) + 1.5 \times IQR \left(Action^i \left(start\ time\right)\right) \end{bmatrix}$$

$$Action^i \left(end\ \ time\right) \tag{10}$$
$$= \begin{bmatrix} Q1 \left(Action^i \left(end\ time\right)\right) - 1.5 \times IQR \left(Action^i \left(end\ time\right)\right), \\ Q3 \left(Action^i \left(end\ time\right)\right) + 1.5 \times IQR \left(Action^i \left(end\ time\right)\right) \end{bmatrix}$$

$$Action^i \left(duration\right) \tag{11}$$
$$= \begin{bmatrix} Q1 \left(Action^i \left(duration\right)\right) - 1.5 \times IQR \left(Action^i \left(duration\right)\right), \\ Q3 \left(Action^i \left(duration\right)\right) + 1.5 \times IQR \left(Action^i \left(duration\right)\right) \end{bmatrix}$$

$$Action^i \left(time\ from\ previous\ action\right)$$
$$= \begin{bmatrix} Q1 \left(Action^i \left(time\ from\ previous\ action\right)\right) \\ -1.5 \times IQR \left(Action^i \left(time\ from\ previous\ action\right)\right), \\ Q3 \left(Action^i \left(time\ from\ previous\ action\right)\right) \\ +1.5 \times IQR \left(Action^i \left(time\ from\ previous\ action\right)\right) \end{bmatrix} \tag{12}$$

3.3.2.5. Inter-Quartile Range Based on Total Year

In this method, for labeling each record, we calculate the range of $(Q_1 - 1.5 \text{*IQR}, Q_3 + 1.5 \text{*IQR})$ for all the features considering the records and label the current record accordingly.

4. Results

4.1. Action Detection and Prediction

We have implemented our proposed method with Python in Jupyter environment. To test the first stage of our proposed method, the discussed features have been added to the dataset in a stepwise manner. For evaluating the accuracy of our model, we considered 70% of 30 days (21 days) of the dataset as training sample and 30% of the remaining days (9 days) as test sample.

Table 3 shows that by considering only the duration feature, the accuracy of action prediction with Random Forest Classifier model was 43.33%, however after adding all features, the accuracy has reached 98.48% (Figure 1). By considering all the features, after calculating the accuracy of detection and prediction of activities every 12 months of the year by the use of Random Forest classifier, we obtained a mean accuracy of 97.96% and a standard deviation of 1.21.

We also compared our proposed method with Liouane et al. (17). Similar to them we have used the first three weeks of the eHealth dataset for the training phase and nine days for the test purposes.

Table 4 shows the results of the comparison by using three metrics of RMSE (Equation 14), cosine similarity (Equation 15) and percentage error (Equation 16) for each test day. Similar to Liouane et al. study (17), for calculating RMSE, we first normalized the values using (Equation 13).

$$\bar{x} = \frac{x - min\left(x\right)}{max\left(x\right) - min\left(x\right)} \tag{13}$$

Table 3. Achieved Accuracy with the Used Features and Classification Models of RF, GNB, DT and KNN

Feature	Accuracy			
	RF	**GNB**	**DT**	**KNN**
Time Duration	43.33%	47.27%	43.03%	47.88%
Time Duration, Start Time, End Time	62.42%	45.45%	60.0%	50.91%
Time Duration, Start Time, End Time, Location	85.45%	74.85%	85.45%	50.91%
Time Duration, Start Time, End Time, Location, Previous Action	88.18%	74.24%	86.06%	50.91%
Time Duration, Start Time, End Time, Location, Previous Action, Water Use and Electrical Device Use	98.48%	91.21%	97.27%	50.91%
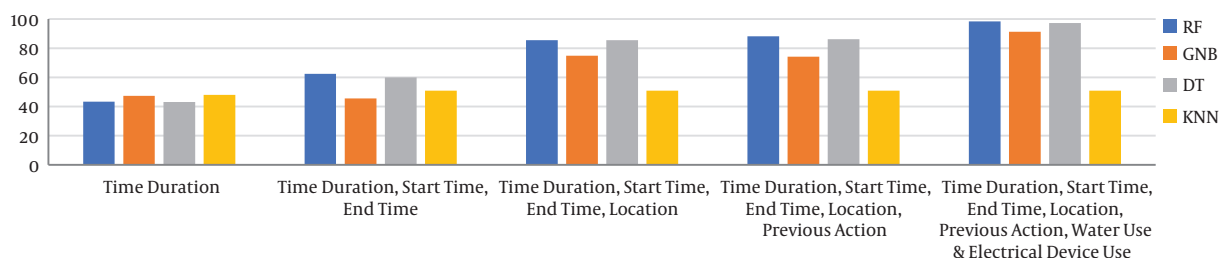
Figure 1. Achieved accuracy for each group of features

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^{N} (y_i - \hat{y}_1)^2} \qquad (14)$$

$$Cosine\ similarity = \frac{y.\hat{y}}{||y|| \cdot ||\hat{y}||} \qquad (15)$$

$$Percentage\ error = \frac{|y - \hat{y}|}{|y|} \qquad (16)$$

In Equations 14 - 16 y is the vector of the main action labels of a day, and $\hat{y}$ denotes the vector of the predicted labels of the given day.

As can be seen in Table 4, our proposed method shows considerable improvement over the method of Liouane et al. (17). It is notable that in Liouane et al.'s study (17), comparison of the predicted actions with real values was done using Equations 14 - 16 which is based on comparing the codes assigned to the actions, i.e. each action should be encoded and then these numerical codes can be used in the equations. We argue that such a comparison highly depends on the codes assigned to the actions. More precisely, if a predicted action and the real one were respectively coded as 10 and 11, then the error in the metrics e.g. RMSE or percentage error would be different than when the actions were coded as 10 and 19. Therefore, we propose using Equation 17 for calculating RMSE.

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^{N} (y_i - \hat{y}_1)^2} \qquad (17)$$

$$where \quad y_i - \hat{y}_i = \begin{cases} 0; y_i = \hat{y}_i \\ 1; otherwise \end{cases} \qquad (18)$$

In Equation 17, y_i is the real action label, and $\hat{y}_i$ denotes the predicted action label. With our proposed method, using Equation 17, the accumulated RMSE for the 9 test days sums up to 0.1584, which is calculated independent of the codes assigned to the actions.

4.2. Anomaly Detection Phase

After labeling the activities as normal or abnormal, for detecting abnormal actions and predicting future occurrences of them, we used AdaBoost supervised learning algorithm, which works well in detecting anomalies due to its sensitivity to noisy data and outliers. While forming the training and testing data we ensured that 70% of the anomalies be in the training dataset and 30% of them in the testing dataset. The number of abnormal class samples was much less than the number of normal class instances, hence it is called the minority class and normal class is called the majority class. Due to the large differences in the number of normal and abnormal class samples, we encountered with an imbalanced dataset. Before applying the AdaBoost learning algorithm on the dataset, we used sampling techniques to make dataset more balanced. We down-sampled major class for this purpose, as can be seen in Table 5, we calculated precision, recall, and f_1 score for each season for all labeling proposed methods. As can be

Table 4. Comparison of Our Method with Liouane et al. (17) for 9 Test Days

	Cosine Similarity		Percentage Error		RMSE	
	Liouane et al. (17)	Our Method	Liouane et al. (17)	Our Method	Liouane et al. (17)	Our Method
Day1	0.9982	1.0	9.7757	0.0	$2.2776 * 10^{-4}$	0.0
Day2	0.9981	1.0	9.0219	0.0	$2.2818 * 10^{-4}$	0.0
Day3	0.9978	1.0	8.8479	0.0	$2.2630 * 10^{-4}$	0.0
Day4	0.9981	1.0	9.3887	0.0	$2.2911 * 10^{-4}$	0.0
Day5	0.9984	0.9953	10.0653	1.0982	$2.3236 * 10^{-4}$	0.0418
Day6	0.9983	1.0	10.0648	0.0	$2.3230 * 10^{-4}$	0.0
Day7	0.9979	1.0	8.1065	0.0	$2.2146 * 10^{-4}$	0.0
Day8	0.9982	1.0	8.9060	0.0	$2.2860 * 10^{-4}$	0.0
Day9	0.9993	1.0	12.0979	0.0	$2.3457 * 10^{-4}$	0.0

seen, from the whole year based methods, which we preferred them to be season based, the Mean +/- 3stdev has given better results. By the use of inter-quartile method, we reached the precision and recall value of 86.33% and 51% respectively, while by the use of mean +/- 3stdev method, we reached the value of 100% and 83%. The average value of metrics for each used labeling method can be seen in Figure 2.

5. Discussion

In order to detect and predict resident's actions in a smart home, we chose a stepwise feature selection as is shown in Table 3. The features of start time, end time, duration, location, previous action, water, and electric device use, although are simply available in smart homes, proved to be reasonable for action detection and prediction. We applied our proposed method on eHealth Monitoring dataset and reached the accuracy of 98.48%. Comparing our work with Liouane et al. work (17) showed that our proposed method reached higher accuracy with lower percentage error in predicting actions. In the next step, for the anomaly detection, we first applied statistical measures to label actions as normal/abnormal and construct the user behavior profile. Afterwards by downsampling the major class, we succeed to make a more balanced dataset. Finally, we detected anomaly by the use of the AdaBoost supervised learning algorithm with the average precision value of 100% and average recall value of 83% for anomaly class by the use of mean +/- 3stdev based on the whole year labeling technique. By considering the fact that individual behavioral habits may vary in different seasons, the methods that made one season a basis for profiling the resident's behavior did not take into account individual behavioral changes between different seasons; however, we argue that considering the whole year for deriving the valid ranges embeds individual behavioral changes among seasons.

5.1. Conclusions

In aging societies, assistive tools and mechanisms are needed for caring those elderlies or patients who can lead their life on their own. Such facilities help reduce considerable costs that are imposed on healthcare systems. Equipping homes with none-obtrusive sensors to monitor residents' behavior is a preferred choice since they continuously monitor residents without any intervention. However, profiling residents' behavior is a challenging task that we tried to address in this paper. We showed that identification of appropriate features and learning algorithm, have a profound effect on the accuracy of smart home residents' action and anomalous behavior detection. By selecting proper features, we showed that reaching high accuracy in detection and prediction of the residents' actions is feasible.

Footnotes

Authors' Contribution: Study concept and design: Malihe Erfanmanesh and Hooman Tahayori; analysis and interpretation of data: Malihe Erfanmanesh; drafting of the manuscript: Malihe Erfanmanesh and Hooman Tahayori; critical revision of the manuscript for important intellectual content: Andrea Visconti and Hooman Tahayori; study supervision: Hooman Tahayori.

Conflict of Interests: The authors declare that there was no conflict of interest.

Ethical Considerations: This paper is an extension to "Feature-Based Elderly Behavior Detection and Prediction in Smart Homes" published in the 1st Conference on

Table 5. Result of Anomaly Detection Based on Methods 1-5

Used Methods for Labeling	Used Metric	Season 2	Season 3	Season 4	Average
Method 1: Min-max range based of first season					
	Precision	97	90	99	95.33
	Recall	84	74	99	85.66
	F$_1$ score	90	81	99	90
Method 2: Inter-quartile based on first season					
	Precision	89	96	99	94.66
	Recall	78	80	95	84.33
	F$_1$ score	83	87	97	89
Method 3: Inter-quartile based on whole year					
	Precision	62	100	97	86.33
	Recall	45	14	94	51
	F$_1$ score	53	25	95	57.66
Method 4: Mean +/- 3stdev based on first season					
	Precision	92	96	99	95.66
	Recall	75	80	95	83.33
	F$_1$ score	83	87	97	89
Method 5: Mean +/- 3stdev based on whole year					
	Precision	100	100	100	100
	Recall	88	67	94	83
	F$_1$ score	93	80	97	90
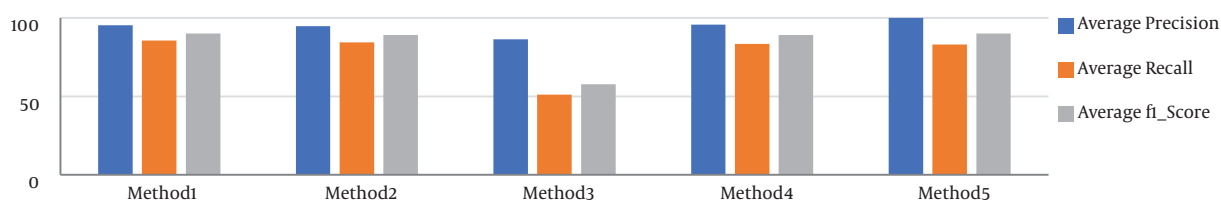
Figure 2. Average precision, recall and f$_1$ score for methods 1-5

Healthcare Computing Systems and Technologies (CHEST 2019).

References

1. Alzheimer's Association. 2016 Alzheimer's disease facts and figures. *Alzheimers Dement.* 2016;**12**(4):459–509. [PubMed: 27570871].
2. Nichols M, Townsend N, Scarborough P, Rayner M. Cardiovascular disease in Europe 2014: Epidemiological update. *Eur Heart J.* 2014;**35**(42):2950–9. doi: 10.1093/eurheartj/ehu299. [PubMed: 25139896].
3. Demir E, Köseoğlu E, Sokullu R, Şeker B. Smart home assistant for ambient assisted living of elderly people with dementia. *Proc Comput Sci.* 2017;**113**:609–14. doi: 10.1016/j.procs.2017.08.302.
4. Bakar UABUA, Ghayvat H, Hasanm SF, Mukhopadhyay SC. Activity and anomaly detection in smart home: A survey. *Next Gener Sens Sys.* 2016;**16**:191–220. doi: 10.1007/978-3-319-21671-3_9.
5. Lotfi A, Langensiepen C, Moreno PA, Gómez EJ, Chernbumroong S. An ambient assisted living technology platform for informal carers of the elderly. *eHealth.* 2017;**3**(9):438–42. doi: 10.1007/978-3-319-49655-9_-53.
6. Singla G, Cook D. Interleaved activity recognition for smart home residents. *Intelligent Environments 2009 - Proceedings of the 5th International Conference on Intelligent Environments.* Barcelona, Spain. 2009. p. 145–52.
7. Singla G, Cook DJ, Schmitter-Edgecombe M. Recognizing independent and joint activities among multiple residents in smart environments. *J Ambient Intell Humaniz Comput.* 2010;**1**(1):57–63. doi:

10.1007/s12652-009-0007-1. [PubMed: 20975986]. [PubMed Central: PMC2958106].

8. Cook DJ, Schmitter-Edgecombe M. Assessing the quality of activities in a smart environment. *Methods Inf Med*. 2009;**48**(5):480–5. doi: 10.3414/ME0592. [PubMed: 19448886]. [PubMed Central: PMC2759863].

9. Chen C, Das B, Cook DJ. A data mining framework for activity recognition in smart environments. *2010 Sixth International Conference on Intelligent Environments*. Kuala Lumpur, Malaysia. IEEE; 2010. p. 80–3.

10. Nazerfard E, Cook DJ. Using Bayesian networks for daily activity prediction. *Proceedings of the 13th AAAI Conference on Plan, Activity, and Intent Recognition*. 2908246: AAAI Press; 2013. p. 32–8.

11. Dawadi P, Cook D, Parsey C, Schmitter-Edgecombe M, Schneider M. An approach to cognitive assessment in smart home. *Proceedings of the 2011 Workshop on Data Mining for Medicine and Healthcare*. San Diego, California, USA. 2011. 56 p.

12. Rashidi P, Cook DJ, Holder LB, Schmitter-Edgecombe M. Discovering activities to recognize and track in a smart environment. *IEEE Trans Knowl Data Eng*. 2011;**23**(4):527–39. doi: 10.1109/TKDE.2010.148. [PubMed: 21617742]. [PubMed Central: PMC3100559].

13. Lotfi A, Langensiepen C, Mahmoud SM, Akhlaghinia MJ. Smart homes for the elderly dementia sufferers: Identification and prediction of abnormal behaviour. *J Ambient Intell Humaniz Comput*. 2011;**3**(3):205–18. doi: 10.1007/s12652-010-0043-x.

14. Raeiszadeh M, Tahayori H. A novel method for detecting and predicting resident's behavior in smart home. *2018 6th Iranian Joint Congress on Fuzzy and Intelligent Systems (CFIS)*. 28 Feb.-2 March 2018. IEEE; 2018. p. 71–4.

15. Raeiszadeh M, Tahayori H, Visconti A. Discovering varying patterns of Normal and interleaved ADLs in smart homes. *Appl Intell*. 2019:1–14. doi: 10.1007/s10489-019-01493-6.

16. Amirjavid F, Bouzouane A, Bouchard B. Activity modeling under uncertainty by trace of objects in smart homes. *J Ambient Intell Humaniz Comput*. 2012;**5**(1):159–67. doi: 10.1007/s12652-012-0156-5.

17. Liouane Z, Lemlouma T, Roose P, Weis F, Messaoud H. An improved extreme learning machine model for the prediction of human scenarios in smart homes. *Appl Intell*. 2017;**48**(8):2017–30. doi: 10.1007/s10489-017-1062-5.

18. Liouane Z, Lemlouma T, Roose P, Weis F, Messaoud H. A genetic neural network approach for unusual behavior prediction in smart home.In: Madureira A, Abraham A, Gamboa D, Novais P, editors. *International Conference on Intelligent Systems Design and Applications*. Intelligent Systems Design and Applications; 2017. p. 738–48.

19. Mshali HH, Lemlouma T, Magoni D. Context-aware adaptive framework for e-health monitoring. *2015 IEEE International Conference on Data Science and Data Intensive Systems*. 11-13 Dec. 2015. IEEE; 2015. p. 276–83.