1st Conference on Healthcare Computing Systems and Technologies (CHEST 2019)

Feature-Based Elderly Behavior Detection and Prediction in Smart Homes

Malihe Erfanmanesh
Master Student
Department of Computer Science and Engineering and IT
Shiraz University, Shiraz, Iran
Email: m.erfanmanesh@shirazu.ac.ir

Hooman Tahayori
Assistant Professor
Department of Computer Science and Engineering and IT
Shiraz University, Shiraz, Iran
Email: tahayori@shirazu.ac.ir

Andrea Visconti
Assistant Professor
Department of Computer Science
Universita degli Studi di Milano
Milano, Italy, 20135
Email: andrea.visconti@unimi.it

Abstract— **One of the capabilities of smart home in the healthcare domain is helping elderly to live independently. This demands for detecting and monitoring their normal activities of daily living (ADLs). By considering changes in the occurrence of these normal activities we can decide about declining the health status of the elderly. Hence, the possibility of preventive care for some elderly people would be partly provided. In this paper, we propose a method for detection and prediction of elderly activities by extracting several features from available information. Obtained results reveal that the proposed features are more effective for detection and prediction of elderly behavior.**

Keywords- activity recognition; machine learning; classification; behavior prediction; smart home

I. INTRODUCTION

The population of elderly people is increasing and their long-term care is time-consuming and costly. Old people, often prefer to live independently in their home, so helping them to live on their own is one of the main goals of smart home in the domain of healthcare. In smart homes, recording residents' daily activities through sensors enables monitoring residents' activities and performance. Increasing safety for elderly, promoting quality of life and reducing the feeling of isolation are some aspects of a smart home which may help elderlies[1].

In smart home, monitoring the residents' health and mobility can be accomplished by finding the normal behavior of the residents and comparing them by their daily living activities. If the person's health diminishes, his/her activities would be affected e.g. the duration of the person's activity would last longer than usual, he/she would have difficulty completing his/her work or he/she would miss some activities. Moreover presence of a person in a location that normally is not expected to be or vice versa, might be a sign of anomaly. Detecting such cases help providing advice to the individual or sending someone to help them when needed[2].

The contribution of this paper is related to the scope of remote health monitoring and assisted living. In the paper, by considering specific features, we propose an accurate and fast method to recognize normal activities of an elderly person and predict their future occurrences.

The remainder of this paper is organized as follows. In section 2, related works are reviewed. In section 3, we introduce the dataset that we have used. In section 4, a theoretical description of our proposed behavior prediction model for elderly is presented. Section 5 provides experimental analysis in "eHealth Monitoring Open Data Project". Section 6 concludes the paper.

II. RELATED WORKS

Cook et al. in [3] have proposed a method to assess the quality of activities in smart environments. They modeled the normal activities using the Naïve Bayes and Markov models and by detecting the states that were ignored, they were able to detect anomalies.

Singla et al. in [4] have provided a method for detecting more than one person in an environment. They have exploited HMM model, which first considered both individuals. Then, by considering separate model for each individual, the accuracy of detecting the activities was improved.

Singla et al. in [5] have used CASAS dataset to recognize and track activities in complex situations e.g. when activities are interleaved or correlated. They utilized Naive Bayes, Hidden Markov Model (HMM) and extended versions of HMM techniques. Their proposed method, however could not detect missing steps.

Raeiszadeh et al. in [6] have proposed Uncertain Pattern-Discovery Method to recognize activities of daily livings in smart homes. In [6], 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 future occurrence of activities.

Rashidi et al. in [7] have proposed a method for health monitoring and assisting individuals which have difficulties living independently at home. They have used unsupervised method for discovery and recognition of activities. They could discover activities by Discontinuous Varied-Order Sequential Miner (DVSM) and k-mean clustering algorithms. Then, they could recognized activities by a boosted version of HMM. Their proposed method however, could not distinguish some similar activities and could not identify concurrent activities.

Amirjavid et al. in [8] have modeled activities as a chain of fuzzy events to predict the intension of smart home residents when they perform a few actions. The difficulty with their proposed method was that the beginning and ending points of the activities had to be clear.

Liouane et al. in [9] aimed to propose a prediction method called Recurrent Extreme Learning Machine (RELM), that provides the ability to learn human behavior and predict accurately and fast the future activities of the elderly inhabitant.

III. Data set

We use dataset of “eHealth Monitoring Open Data Project” which is Open data set for the monitoring and healthcare of dependent persons such as elderlies [10]. The dataset contains information of an elderly person for about a year.

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, we have just information about the start and ending time of the action in seconds (day and time of action are combined together), and action code. In this paper, we have used the coded version.

IV. Proposed Method

A. Definitions

Action vs. activity: An action is a more detailed concept than activity. Each activity can include several actions. For example, “make a tee”, “make coffee”, “make sandwich” and “wash dish” are separate actions but all of them relates to the activity of meal preparation. In this paper, most of our focus is on action concept rather that activity.

Feature: When an action happens, there are a set of environmental conditions and characteristics associated with it, that we refer to them 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 multiple actions, for example, using an electrical device like TV remote control is just related to “watching TV” action but location feature can be the same in several actions like “making coffee”, “making tea”, “washing dishes” and “making a sandwich”. We will show that, selecting more detailed and dedicated features, increases the chance of correct recognition of actions.

B. Supervised learning

In this paper, we explore a supervised learning approach to detect and predict smart home residents’ actions. For this purpose, we have used Random Forest Classifier (RF), 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. In our proposed method we add features in a step-by-step manner and tune their influence over the accuracy of the action detection. Finally, we reach a set of features with which we can recognize actions with a high accuracy.

Initially, we used time duration feature from dataset for action detection and prediction. It seems reasonable that a model with one feature do not work well. We added two more features of start time and end time in seconds for each action. As expected, the accuracy of the model showed an increase.

In the next step, 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, rest room, office, living room and hallway. Then we added the corresponding location to each action. We expect significant improvement of the accuracy by considering features of start time, end time, duration and location of each action.

We also considered previous actions. The previous action represents the action that occurred before the current action. Finally two other features were considered that were related to the use of water and electrical equipment. Regarding the water use, three codes were devised, one relating to the actions that would necessarily require water use, another relating to actions that do not require water, and the third code refer 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 did not use an electric device and then the related code was allocated to the corresponding action.

V. Experimental Result

We have implemented our proposed method with Python in Jupyter environment. To test our method, the discussed features have been added to the dataset in a stepwise manner. For evaluating accuracy of our model, we considered 70% of 30 days (21 days) of the dataset as training sample and 30% of remaining days (9 days) as test sample. The accuracy was measured as the number of recognized actions that were correctly labeled.

Figure 1, corresponding to Table 1 shows that by considering only duration feature, the accuracy of action prediction with Random Forest Classifier model was 43.33%, however after adding all features, the accuracy reached 98.48%.

We also compared our proposed method with Liouane et al.'s [9] that has used the same data set for the same purpose. In [9], authors have proposed an algorithm based on neural network for elderly activity prediction using time series. Similar to their work, we have used the first three weeks of the dataset for the training phase and nine days for the test purposes.

Table 2 shows the results of the comparison by using three metrics of RMSE (2), cosine similarity (3) and percentage error (4) for each test day. Similar to [9], for calculating RMSE, we first normalized the values using (1).

$$\bar{x} = \frac{x - \min(x)}{\max(x) - \min(x)} \quad (1)$$

TABLE 1: STEPWISE FEATURE ADDITION AND RELATED ACHIEVED ACCURACY WITH THE USED 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 & Electrical Device Use	98.48%	91.21%	97.27%	50.91%

$$\text{RMSE} = \sqrt{\frac{1}{N} \cdot \sum_{i=1}^{N} (y_i - \hat{y_i})^2} \quad (2)$$

$$\text{Cosine similarity} = \frac{y.\hat{y}}{||y||.||\hat{y}||} \quad (3)$$

$$\text{Percentage error} = \frac{|y - \hat{y}|}{|y|} \quad (4)$$

In equations (2), (3) and (4), 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 2, our proposed method shows a considerable improvement over [9].

It is notable that in [9], after predicting activities with an extended version of neural network, comparison of the predicted activities with real values was done using (2) - (4) which are based on comparing the codes assigned to the activities, i.e. each activity 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 activities. More precisely, if a predicted activity and the real one were respectively codded as 10 and 11, then the error, in the metrics e.g. RMSE or percentage error would be different than when the activities were coded as 10 and 19. So we propose using (5) for calculating RMSE.

$$\text{RMSE} = \sqrt{\frac{1}{N} \cdot \sum_{i=1}^{N} (y_i - \hat{y_i})^2} \quad (5)$$

$$\text{Where } y_i - \hat{y_i} = \begin{cases} 0 & y_i = \hat{y_i} \\ 1 & \text{other wise} \end{cases}$$

Figure 1: Accuracy of using RF, GNB, DT and KNN classifiers on different features

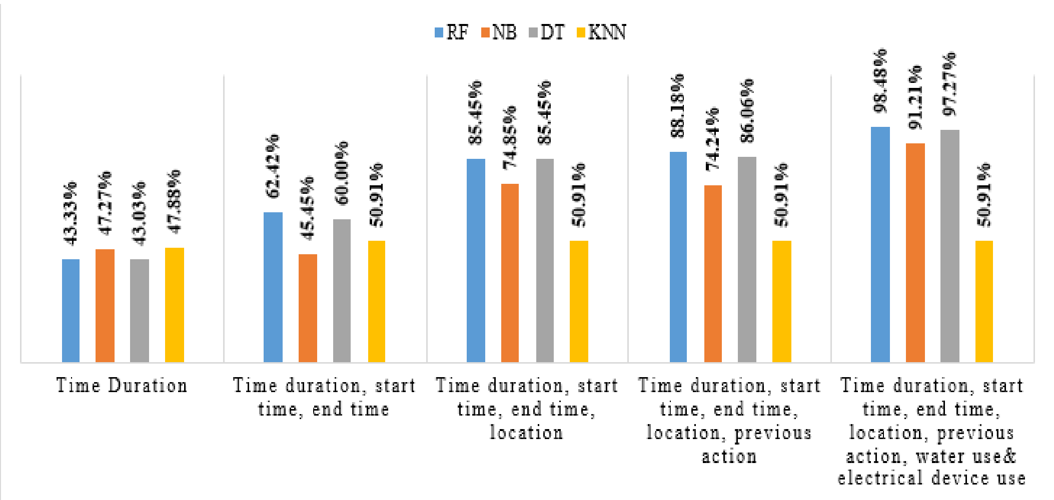

TABLE 2: COMPARISON OF OUR METHOD WITH LIOUANE ET AL. [9] FOR 9 TEST DAYS

	Cosine similarity		**Percentage error**		**RMSE**	
	Liouane et al.[9]	***Our method***	***Liouane et al.[9]***	***Our method***	***Liouane et al.[9]***	***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	٢,٢٦٣٠ * 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

In equation (5), y_i is the real action label, and $\hat{y}_i$ denotes the predicted action label. With our proposed method, using (5) the accumulated RMSE for the 9 test days would sum up to 0.1584 which is calculated independent of the codes assigned to the activities.

VI. CONCLUSION

In this paper, in a stepwise manner we identified effective features to detect and predict actions in smart environments. We have applied our proposed method to eHealth Monitoring dataset. We reached the accuracy of 99.70% by considering features of start time, end time, duration, location, previous action, water and electric device use. We have compared our work with [9] which has used the same data set and the results show that our proposed method reached higher similarity of prediction and real action labels with lower percentage error in predicting actions.

REFERENCES

[1] E. Demir, E. Köseoğlu, R. Sokullu, and B. Şeker, “Smart Home Assistant for Ambient Assisted Living of Elderly People with Dementia,” *Procedia Comput. Sci.*, vol. 113, pp. 609–614, 2017.

[2] A. Lotfi, C. Langensiepen, P. A. Moreno, E. J. Gómez, and S. Chernbumroong, “An Ambient Assisted Living Technology Platform for Informal Carers of the Elderly,” vol. 3, no. 9, pp. 438–442, 2017.

[3] D. J. Cook and D. J. Cook, “Assessing the Quality of Activities in a Smart Environment Corresponding,” vol. 48, no. 5, pp. 480–485, 2009.

[4] G. Singla, D. J. Cook, and M. Schmitter-Edgecombe, “Recognizing independent and joint activities among multiple residents in smart environments,” *J. Ambient Intell. Humaniz. Comput.*, vol. 1, no. 1, pp. 57–63, 2010.

[5] G. Singla and D. J. Cook, “Interleaved Activity Recognition for Smart Home residents,” *Ambient Intell. Smart Environ. Intell. Environ. 2009*, pp. 145–152, 2009.

[6] M. Raeiszadeh and H. Tahayori, “A novel method for detecting and predicting resident’s behavior in smart home,” *2018 6th Iran. Jt. Congr. Fuzzy Intell. Syst. CFIS 2018*, vol. 2018–Janua, pp. 71–74, 2018.

[7] P. Rashidi, D. J. Cook, L. B. Holder, and M. Schmitter-Edgecombe, “Discovering activities to recognize and track in a smart environment,” *IEEE Trans. Knowl. Data Eng.*, vol. 23, no. 4, pp. 527–539, 2011.

[8] F. Amirjavid, A. Bouzouane, and B. Bouchard, “Activity modeling under uncertainty by trace of objects in smart homes,” *J. Ambient Intell. Humaniz. Comput.*, vol. 5, no. 1, pp. 159–167, 2014.

[9] Z. Liouane, T. Lemlouma, P. Roose, F. Weis, and H. Messaoud, “An improved extreme learning machine model for the prediction of human scenarios in smart homes,” *Appl. Intell.*, pp. 1–14, 2017.

[10] H. H. Mshali, T. Lemlouma, and D. Magoni, “Context-Aware Adaptive Framework for e-Health Monitoring,” *Proc. - 2015 IEEE Int. Conf. Data Sci. Data Intensive Syst. 8th IEEE Int. Conf. Cyber, Phys. Soc. Comput. 11th IEEE Int. Conf. Green Comput. Commun. 8th IEEE Inte*, pp. 276–283, 2015.