2018 6th Iranian Joint Congress on Fuzzy and Intelligent Systems (CFIS)

A Novel Method for Detecting and Predicting Resident's Behavior in Smart Home

1st Mahsa Raeiszadeh
Department of Computer Science and Engineering
Shiraz University
Shiraz, Iran
mahsa.raeiszadeh@shirazu.ac.ir

2nd Hooman Tahayori
Department of Computer Science and Engineering
Shiraz University
Shiraz, Iran
tahayori@shirazu.ac.ir

***Abstract*—By development of sensor technology, wireless communication, machine learning and artificial intelligence techniques, smart environments have improved into a significant research area. A kind of smart environment is smart home that provides an intelligent and integrated environment, equipped with sensors and actuators. Activity of Daily Living (ADL) in smart homes has helped in elderly care, health care, etc. Human Activity Recognition (HAR) is the task involved in reasoning within smart homes with the purpose of recognizing the ongoing activity of the resident and is inferred using machine learning and artificial intelligence methods out of sensors data. In this paper, we have proposed an activity mining and tracking method called Uncertain Pattern-Discovery Method that enables tracking regular activities and detect changes in an individuals behavioral pattern and lifestyle.**

***Index Terms*—Human Activity Recognition (HAR), smart home, Activity of Daily Living (ADL)**

I. INTRODUCTION

A smart home is a residential with miniaturized processors, mobile agents communicating with each other and multi-modal sensors that are embedded in various objects. By recording residents daily activities through sensor readings, an assistive system in a smart home can be devised which can take automated proper actions to assist residents. Modeling and recognizing human activities in smart home by using sensor data is a complex job. Individuals have various lifestyles, habits, or abilities and perform the same ADL in many different ways [1].

The main goal of this paper is to observe the behavior and activities of the residents and model them to discover interesting patterns. In this paper a supervised method for discovering human activities from sensor data is presented. Proposed method is consisted of two steps. First, detecting repeated patterns and generating the prediction model from supervised learning. Second, using the prediction model from the first step to predict and recognize the user activities.

978-1-5386-2836-2/18/$31.00 ©2018 IEEE

The remainder of this paper is organized as follows. In section II, related works are reviewed. In Section III, a theoretical description of the proposed method is presented. Section IV provides experimental analysis in the Massachusetts Institute of Technology (MIT) smart home testbed for ADLs. Last section concludes the paper.

II. RELATED WORKS

Despite the fact of complexity and diversity of ADLs, different researches exist that have proposed different methods for HAR [2]. Researchers, through HAR have explored a number of approaches in the IoT based smart home environment to enhance security, safety and comfort of the residents [3].

Reference [4] has proposed a sequential activity prediction using k-Nearest Neighborhood (KNN). Nearest neighbor has been used to detect every day activity such as walking, watching TV, etc. Support Vector Machines (SVM) are also used for detecting smart home residents behaviors [5]. Similarly, back propagation neural networks are used for the same purpose and it is shown that they are effective for mapping pre segmented sensor streams to activity labels.

Other researchers have employed decision trees to learn logical description of the activities. This approach often generates understandable rules and if weighs of the attributes are determined rationally, results may be more appealing than other methods [6].

Some HAR system uses as a probabilistic model in the case of generative approach to model the recognition system [7]. Naive Bayes classifier (NBC), Hidden Markov Models (HMM) and Dynamic Bayesian Networks (DBM) are the most popularly used generative modeling techniques. Naive Bayes classifier outputs the greatest probability activity given the observation using a probability function that models the dependencies between the observation and activity labels. Modeling temporal information also is done through HMM and DBN.

III. Proposed Method

Based on the fact that different people do jobs with similar objective in different manners, the capability to detect user behaviors is critical. Therefore, supervised learning with an obvious training phase offers a promising approach for HAR problem. Even though activities are performed in different order, we should be able to detect both continuous and discontinuous events. We take a supervised approach to activity tracking in this paper. We discover activities that naturally occur frequently in an individuals home. We will then build models to recognize these activities as they occur. Our proposed method is consisted of three phases that are schematically depicted in Fig. 1.

In order to test our method and show its privileges, we have relied on the Massachusetts Institute of Technology (MIT) smart home data set. Using this data set has imposed some specific requirements that we will discuss them in the following. Table I shows a part of the data set as an example.

A. Phase 1- Preprocessing

In the preprocessing phase sensor data will be converted into the event sequences. To this aim sensor-id, sensor activation and deactivation time and the number of activation or deactivation of the sensor are considered. For example, consider Preparing Lunch activity that is shown in Table I.In the preprocessing phase, as is shown in Fig. 2, raw sensor data , without date and time are converted to the xxxyzzz format.

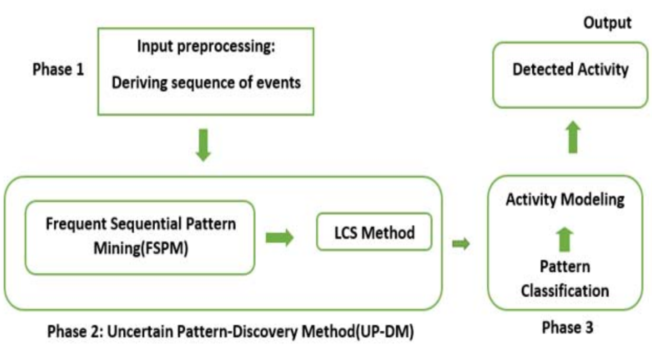

Fig. 1. Architecture of our proposed method for discovering and tracking activities .

TABLE I
SAMPLE OF DATA SET

Preparing lunch	*Day1*	*11:21:17*	*11:38:22*	
140	137	131	53	131
Door	Freezer	Toaster	Cabinet	Toaster
11:23:04	11:23:55	11:24:08	11:34:59	11:35:12
11:23:07	11:24:03	11:24:14	11:35:01	11:35:22
Dressing	*Day2*	*8:25:31*	*8:26:46*	
57	75	139		
Dress cabinet	Drawer	Jewelry box		
8:24:49	8:26:02	8:26:14		
8:25:50	8:26:11	8:26:21		

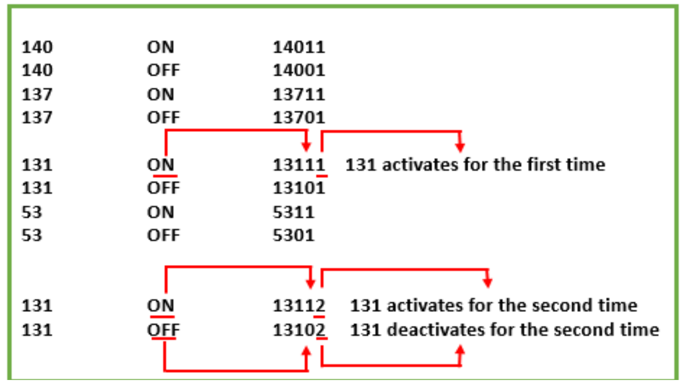

Fig. 2. Preprocessing the sensor data .

xxx represents, sensor-id, y represents sensor state which can be 1(=activate) or 0 (=deactivate) and zzz represents the number of times the sensor is activated or deactivated.

B. Phase 2- UP-DM

Phase 2 is consisted of two sub-phases, namely Frequent Sequential Pattern Mining Algorithm FSPMA and Longest Common Subsequence(LCS).

Frequent patterns are subsequences that appear in a data set with frequency greater than or equal to some specific threshold [8]. In FSPMA first we will identify the frequent and repeated patterns of event sequences. Then we can create a model to recognize the activity. PrefixSpan algorithm as a sequential frequent pattern mining, discovers sequential patterns in sequence databases [9]. We use PrefixSpan to find sequence patterns from discontinuous instances that might also display varied-order events. To reduce the number of sequential patterns from the PrefixSpan output, we apply Longest Common Subsequences (LCS) method to all sequential patterns [10].

C. Phase 3- Predictive Model

After finding frequent sequential patterns of activities in UP-DM phase, the patterns are classified and a model is built which will recognize future occurrence of the activity. In our approach, we make use of RandomForest model to recognize activities. It is trained to recognize the patterns that correspond to the classification representatives found by pattern classification models.

A sequence of sensor activation and deactivation depicts each individuals activity. Each sequence of events that is labeled as an individuals activity is considered as a sample record and the sequential pattern that is resulted from UP-DM is considered as a feature in our database. Then we have considered activities as a class label. For relating the sample of records with features the Levenshtein (edit) distance [11] - sim(A,B)- was used to define a similarity measure between the two patterns.

IV. Experimental Result

We have examined our proposed method on the data that is collected in the MIT smart home testbed. Fig. 3 describes the layout of the two smart homes. Both subjects lived alone in one-bedroom apartments. 77 sensors were installed in the first subject's apartment and 84 in the the second subject's

Fig. 3. (a) Sensors distribution for subject one. (b) Sensor distribution for subject two [12] .

apartment. We will test our proposed method in two different realistic cases to detect ADLs and evaluate the performance of the proposed approach.

A. Subject 1- ADL discovery

The first subject was a professional 30-year-old woman who spent her free time at home. Data were recorded for each of activities performed by resident. To test our method, dataset is being fed to the preprocessing phase to be converted to event sequences. Then we have applied the UP-DM algorithm to dataset. In the UP-DM phase, we discovered frequent sequential patterns in the sensor event data by FSPMA. We experimentally set the minimum support thresholds -minSup- of frequent event, to 0.8 and the threshold of max pattern length for each activity to the minimum length of sensor sequences for intended activity. When we analyzed the sensor events, the algorithm in total discovered 186 patterns with the lengths varying from 1 to 6 events. Then LCS method removed patterns that are subset of another pattern. Finally, UP-DM's output was 52 patterns with the lengths varying from 1 to 6 events. Using the above data as a training set, in the phase 3 we have exploited the RandomForest classification model, to recognize future activities of the users. We obtained the accuracy level of 97.45% in this model for the first subjects activities.

B. Subject 2- ADL discovery

The second subject was a 80-year-old woman who spent most of her time at home. Similar to the first experiment,we applied FSPMA and discovered 101 sequential frequent patterns for all activities, with the lengths varying from 1 to 10 events. The parameter, minSup and max pattern length for each activity were defined as in the previous experiment. After FSPMA, LCS method reduced 101 patterns to 45 patterns, with the same lengths of events in FSPMA. In the next step we built our model based on the discovered activities and applied the algorithm to the remaining data to identify when the discovered activities occurred. Fig. 4, demonstrates that, the accuracy of predicting the activity Bathing, Grooming, Preparing a beverage, and Preparing lunch got a maximum value of 100% in RandomForest model and Cleaning activity, unlike to previous experiment got the minimum accuracy prediction. The accuracy result for this subject reached to 91.37%. Moreover, Fig. 4 shows the accuracy of each activity by prediction model.

We have also compared our method with other works who have used the same data set as what we have used for the same purpose. The results are reported as Table II. In [12], a supervised method for recognizing activities based on MIT real smart home testbed is introduced. This approach has evaluated a probabilistic technique like Naive Bayes classifier for HAR. The aim of this approach was providing a tool to automatically monitor and assist elderly individuals. However, its accuracy to detect both subject 1 and subject 2 ADL in contrast our method is lower.

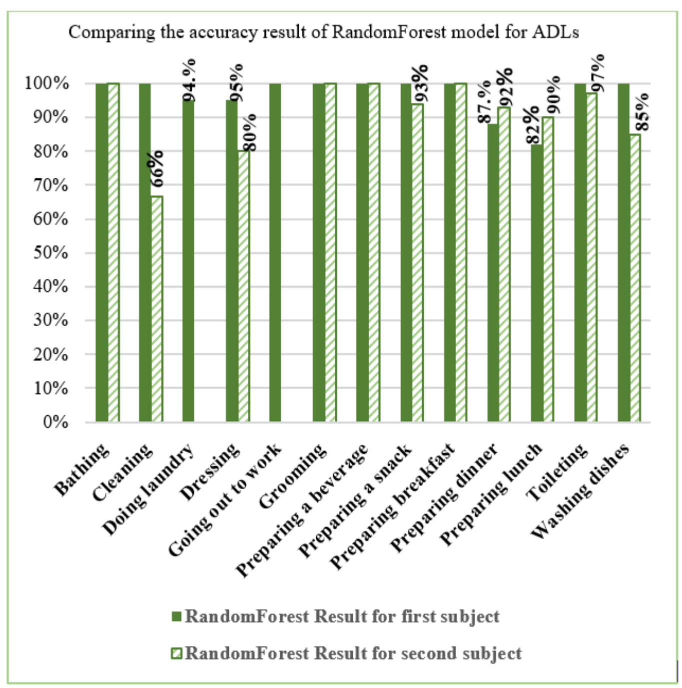

Fig. 4. Comparison accuracy of RandomForest technique for first and second ADLs .

TABLE II
COMPARISON OF OUR METHOD WITH [12]

	Dataset	Approach	Result
Proposed Method	MIT	Supervised Method (UP-DM+RandomForest)	Subject 1: 97.45% Subject 2: 91.37%
Tapia [12]	MIT	Supervised Method (Naive Bayes Classifier)	Subject 1: 60.6% Subject 2: 41.09%

CONCLUSION

In this paper, we introduced a method to recognize ADLs in smart environments. Our proposed method consists of three phases namely preprocessing, UP-DM and predictive model. We have applied our proposed method on subject 1 and subject 2 of MIT smart home data set. The tests shows that the proposed method is effective. We reached the accuracy of 97.45% in first subjects ADLs and 91.37% in second subjects ADLs. We have compared our work with other work which has used the same data set of MIT for HAR and the results shows our proposed method reach higher accuracy.

REFERENCES

[1] P. Mahya, H. Tahayor,i and A. Sadeghian, "An Online Demand Response EMS with Anomaly Usage Detection," 5th IEEE Int. Conference on Smart Energy Grid Engineering, pp. 271-275, 2017.

[2] E. Kim, S. Helal, and D. Cook, "Human Activity Recognition and Pattern Discovery, Pervasive Comput. IEEE, vol. 9, no. 1, pp. 48-53, 2010.

[3] L. Chen, C. D. Nugent, and H. Wang, "A Knowledge-Driven Approach to Activity Recognition in Smart Homes, IEEE Trans. Knowl. Data Eng., vol. 24, no. 6, pp. 961-974, 2012.

[4] C. Lombriser, N. B. Bharatula, D. Roggen, and G. Trster, "On-body activity recognition in a dynamic sensor network, Proc. ICST 2nd Int. Conf. Body Area Networks, pp. 1-6, 2007.

[5] O. Brdiczka, J. Maisonnasse, and P. Reignier, "Automatic detection of interaction groups, 7th Int. Conf. Multimodal interfaces - ICMI 05, pp. 32-36, 2005.

[6] U. Maurer, A. Smailagic, D. P. Siewiorek, and M. Deisher, "Activity recognition and monitoring using multiple sensors on different body positions, Int. Work. Wearable Implant. Body Sens. Networks, pp. 4-7, 2006.

[7] L. Chen, J. Hoey, C. D. Nugent, D. J. Cook, Z. Yu, and S. Member, "Sensor-Based Activity Recognition, 790 IEEE Trans. Syst. MAN, Cybern., vol. 42, no. 6, pp. 790-808, 2012.

[8] Z. Zhao, D. Yan, and W. Ng, "Mining probabilistically frequent sequential patterns in large uncertain databases, IEEE Trans. Knowl. Data Eng., vol. 26, no. 5, pp. 1171-1184, 2014.

[9] J. Pei et al., "PrefixSpan: mining sequential patterns efficiently by prefix-projected pattern growth, Int. Conf. Data Eng., pp. 215-224, 2001.

[10] L. Bergroth, "A Survey of Longest Common Subsequence Algorithms", 7th Int. Symp. String Process. Inf. Retrieval., pp. 39-8, 2000.

[11] 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.

[12] E. M. Tapia, S. S. Intille, and K. Larson, "Activity recognisation in Home Using Simple state changing sensors, Pervasive Comput., vol. 3001, pp. 158-175, 2004.