

# Long term analysis of daily activities in a smart home

Labiba Gillani Fahad<sup>1</sup>, Arshad Ali<sup>2</sup> and Muttukrishnan Rajarajan<sup>1</sup>

1- School of Engineering and Mathematical Sciences,  
City University London, United Kingdom.

2- School of Electrical Engineering and Computer Sciences,  
National University of Science and Technology, Pakistan.

**Abstract.** In this paper, we propose an approach to monitor the change in the daily routine of a person living in a smart home using the long term analysis of the activities performed, where daily routine is the group of activities that can be performed in a single day and are repeated over a period of time. In the proposed approach, first the activity recognition is performed, in which the newly detected activity instances are labeled using the probabilistic neural network learning model. Next, the daily routine of the occupant is analyzed by exploiting the group of activities of a day performed over a period of time. We apply K-means clustering to separate the normal routine from unusual and suspected routines. The proposed approach is validated on a publicly available Kasteren dataset.

## 1 Introduction

A smart home is equipped with sensors and actuators in order to monitor the human activities. Activity monitoring has a number of prospective real world application in health-care and assistive living. Activity recognition is a challenging problem due to variations in performing an activity, unreliable nature of the sensor data, missing and faulty values and sparse feature set due to the spatially distributed nature of activities. The smart home projects such as AWARE Home at Georgia Tech [1] and CASAS Smart Home [2] have been developed to provide valuable functionalities in activity recognition and assistive living. Most of the classification approaches focus on capturing the sensor observations and recognizing the activities of daily life [3, 4, 5].

This paper describes an approach to monitor the activities of daily life of an individual living in a smart home and then detects the changes in the activity patterns in the daily routine of an occupant over time. The approach first recognizes the activities using probabilistic neural network (PNN) and then analyzes the daily routine by K-means clustering. The PNN classifier is selected for activity recognition, which is derived from Bayesian networks and has the ability to converge to an optimal classifier. PNN is less sensitive to noisy and sparse data [6]. In next step, we use the k-means clustering to cluster the days based on the frequency of performed activities. The cluster analysis finds deviations from the normal daily routine and detects the unusual trends.

The rest of the paper is organized as follows. Section II discusses the related work on activity recognition. In Section III, we discuss the proposed approach

based on PNN and K-means. In Section IV, the approach is validated using the publicly available Kasteren dataset [3]. Finally, Section V draws conclusions.

## 2 Related work

Over the past few years, a number of different machine learning approaches have been proposed to recognize the activities of daily life. The Decision Tree (DT) classifier has been used to recognize the activities such as walking and running [5]. The multilayer feed forward neural network is applied for the classification of static (sitting, standing) and dynamic activities (walking, running) [7]. The Support Vector Machine (SVM) widely being used in activity recognition is applied to recognize and discriminate the cough sounds [8]. Next, multiclass one-vs-one SVM is used to recognize the daily activities such as sleeping, eating by incorporating the prior knowledge and the performance is improved by utilizing the time information [4]. The performance of Hidden Markov (Model HMM) and Conditional Random Fields (CRF) is compared for activity recognition in [3]. In [9] the comparison of soft-margin multiclass SVM (C-SVM), CRF and linear discriminant analysis (LDA) is performed for the recognition of daily activities, where superior performance is shown by C-SVM in comparison to CRF and LDA, while CRF overfits the dominant class. A generalized activity recognition approach can be developed by using the common activities spanned over multiple environmental settings and different residents [10].

Moving a step further in activity recognition, monitoring the daily routine through long term analysis of recognized activities provides a useful insight of the smart home occupant. A behavior recognition approach is presented using Naive Bayes (NB) probabilistic model and k-means clustering [11]. The support vector data descriptors is exploited to analyze the elderly behavior [12]. A data mining approach based on growing self organizing maps is applied to monitor the daily activities [13]. An activity monitoring approach to access the blood glucose level of diabetic patients through daily activities is developed using Artificial Neural Network (ANN) [14].

## 3 Proposed activity monitoring approach

The proposed approach identifies the change in the daily routine of a person by long term monitoring of the every day performed activities. Useful features are extracted from the pre-segmented activities. We use the learning algorithm PNN [15, 6] for activity recognition. The recognized activities are then utilized to group the days in which a person follows a different routine from normal. We use K-means clustering algorithm for grouping [16]. The block diagram of the proposed approach is shown in Fig. 1. Let  $A$  be the training set of  $K$  activities  $\{A_k\}_{k=1}^K$  and  $I_k = I_{1k}, \dots, I_{jk}, \dots, I_{Jk}$  be the set of  $J$  activity instances of  $A_k$ . Each  $I_{jk}$  is observed through  $R$  binary sensors. A feature set  $F_{jk}$  of  $R$  features is extracted from  $I_{jk}$  such that:

$$F_{jk} = \{f_{jk}^r\}_{r=1}^R. \quad (1)$$

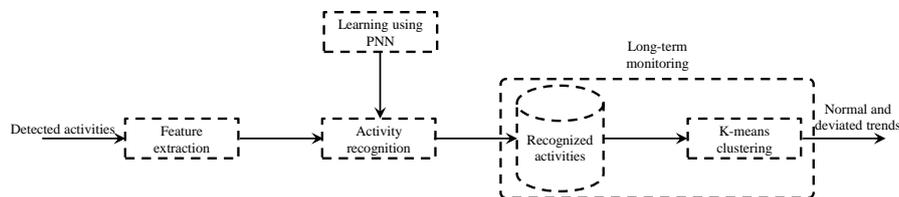


Fig. 1: Block diagram of the proposed approach.

Each feature represents the frequency of activated sensors within  $I_{jk}$ . PNN classifier is trained for activity recognition using the extracted features from multiple activity instances of each activity. In PNN, neurons are in four layers: input, pattern, summation and decision layer. The *input* layer of PNN distributes the feature set of the training sample  $F_{jk}$  to the pattern layer neurons. In the *pattern* layer, neurons  $x_{j'k'}$  is equal to the training samples and each pattern neuron belongs to a class  $A_k$ . Each pattern neuron receives the feature vector  $F_{jk}$  and computes the output  $\phi_{j'k'}(F_{jk})$  by

$$\phi_{j'k'}(F_{jk}) = \frac{1}{(2\pi)^{R/2}\sigma^R} \exp \left[ -\frac{(F_{jk} - x_{j'k'})^T (F_{jk} - x_{j'k'})}{2\sigma^2} \right], \quad (2)$$

where  $R$  denotes the dimension of the pattern vector  $F_{jk}$ ,  $\sigma$  is the smoothing parameter and  $x_{j'k'}$  is the neuron vector. The *summation* layer sums the output from the pattern layer corresponding to a particular class from which the training samples are selected. The summation layer neurons compute the maximum likelihood of pattern being classified into by summarizing and averaging the output of all neurons that belong to the same class.

$$p_k(F_{jk}) = \frac{1}{(2\pi)^{R/2}\sigma^R} \frac{1}{J} \sum_{j=1}^J \exp \left[ -\frac{(F_{jk} - x_{j'k'})^T (F_{jk} - x_{j'k'})}{2\sigma^2} \right], \quad (3)$$

where  $J$  denotes the total number of activity instances in  $A_k$ . In *decision* layer, from the output of summation layer while assuming an equal a priori probability of each activity class  $A_k$  and an equal losses associated with every incorrect decision for each activity class  $A_k$ , the input  $F_{jk}$  is assigned to the class in accordance with the Bayes's decision rule as

$$C^{F_{jk}} = \arg \max_k \{p_k(F_{jk})\}, k = 1, 2, \dots, K, \quad (4)$$

where  $C^{F_{jk}}$  denotes the estimated class of  $F_{jk}$  and  $K$  is the total number of activity classes in the training samples. The trained PNN is then used to recognize the new activity instances.

Next, the recognized activities of each day are used to define the daily routine of the person. Let a routine  $\Gamma_d$  is defined by the frequency of the activities performed in a day,  $\{\Gamma_d\}_{d=1}^D$ , where  $D$  is the total number of days.

$$\Gamma_d = \{\eta_{A_k}^d\}_{k=1}^K, \quad (5)$$

Table 1: The accuracy breakdown of activities using proposed approach

Activities	Leaving	Toileting	Showering	Sleeping	Breakfast	Dinner	Drink
Leaving	91.18	0	0	0	0	0	0
Toileting	5.8	99.12	0	2.08	0	0	10.00
Showering	0	0	100.00	2.08	0	0	0
Sleeping	0	0	0	93.75	0	0	0
Breakfast	0	0.88	0	0	75.00	25.00	0
Dinner	2.94	0	0	0	15.00	58.33	0
Drink	0	0	0	2.08	10.00	16.67	90.00

where  $\eta_{A_k}$  is the number of instances of  $A_k$  recognized in a day  $d$ .  $\Gamma_d$  is further grouped into three clusters  $S_z$ , where  $z = 1, 2, 3$  using unsupervised learning K-means clustering so as to minimize the within-cluster sum of squares

$$\arg \min_S \sum_{z=1}^3 \sum_{\Gamma_d \in S_z} \|\Gamma_d - \mu_z\|^2, \quad (6)$$

where  $\mu_z$  is the mean of points in  $S_z$ . Next, we find the inter-difference between three clusters. The two cluster with the maximum difference between each other are labeled as normal and deviated trends in daily routine. Normal is the one with more  $\Gamma_d$  and  $A_k$  in a day while the other cluster represents the deviation from normal. The third cluster in the middle is the suspected cluster which contains undecided routines. In suspected cluster most of the activities are performed but with less frequency.

## 4 Experimentation

The performance of the proposed approach is evaluated in-terms of recognition accuracy of the activities and the detection of unusual trends in the daily routine of the smart home occupant on the Kasteren dataset [3]. The dataset is collected in a three bedroom apartment for a period of 28 days using 14 state change sensors. The Dataset comprises of leaving, toileting, showering, sleeping, breakfast, dinner and drink activities. The information used for the formation of feature set includes sensor location and frequency of sensor (de)activation. We apply leave one out cross validation, where one day data is used for testing while remaining days data is used for training.

The average recognition accuracy achieved by the PNN is 88.62%. Table. 1 shows the accuracy breakdown of the activities on the Kasteren dataset. The PNN recognized the Leaving, Toileting, Showering and Sleeping activities with an accuracy of 91.18%, 99.12%, 100% and 93.75% respectively. The activities of Breakfast, Dinner and Drink are recognized with an accuracy of 75.00%, 58.33% and 90.00% respectively. The activity of Breakfast sends its 15% errors to Dinner and 10% errors to the Drink activity. The Dinner activity sends its 25% errors to Breakfast and 16.67% to Drink activity. As the Breakfast, Dinner and Drink activities are performed in the same location and share similar objects therefore, PNN confuses in the recognition of these activities.

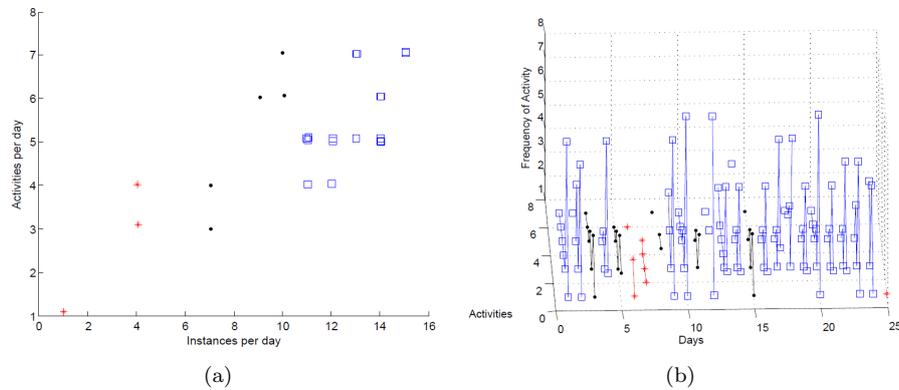


Fig. 2: (a) Clusters presenting normal (box), suspected (dot) and unusual (asterisk) trends in activities performed per day. (b) The frequency of performed activities per day.

In the next step, the output of the PNN is utilized for long term monitoring of activities performed by the smart home occupant. In Fig. 2-(a), the clusters on the extremes of top and bottom represents respectively, the normal and the unusual trends of smart home occupant, whereas the middle cluster presents the suspected routine based on the less frequency of the executed activities. Out of the total 25 days, normal routine is followed in 17 days. The five days presents the suspicious trend, in which activities are performed, but less frequently and three days presents the unusual trends. Fig. 2-(b) shows the frequency of performed activities over the days. For example, it can be observed, that on the sixth day, the occupant only performed the toileting, showering, sleeping and breakfast activities and stayed mostly at home, indicating an unusual trend because of less frequency of activities performed. The above results shows that long term analysis of the activities can provide a detailed insight into the daily routine of smart home occupant and any change in the trends can be observed.

## 5 Conclusions

In this paper, we presented an approach for long term monitoring of activities in smart home to support independent living of the elderly. The proposed approach is based on feature extraction, learning on these features to recognize activities using probabilistic neural network and then recognized activities are used to find out the daily routine based on K-means clustering. The proposed approach is able to differentiate between the normal daily routine and the deviations from the normal. The future research direction of this work would be in investigating and modeling an online activity monitoring system, which may also utilize the time information.

**Acknowledgment** The author, Labiba Gillani Fahad, acknowledges the support

provided by National University of Science and Technology, Pakistan, for this research work.

## References

- [1] Gregory D. Abowd, Christopher G. Atkeson, Aaron E Bobick, Irfan A. Essa, and Blair Macintyre Elizabeth. Living laboratories: the future computing environments group at the georgia institute of technology. In *Proc. of Int. Conf. on Human Factors in Computing Systems*, pages 215–216, Hague, The Netherlands, April 2000.
- [2] Chao Chen, Barnan Das, and Diane J. Cook. A data mining framework for activity recognition in smart environments. In *Proc. of Int. Conference on Intelligent Environments*, pages 80–83, Kuala Lumpur, Malaysia, 2010.
- [3] Tim Van Kasteren, Athanasios Noulas, Gwenn Englebienne, and Ben Krose. Accurate activity recognition in a home setting. In *Proc. of Int. conf. on Ubiquitous computing*, pages 1–9, Seoul, Korea, September 2008.
- [4] Anthony Fleury, Norbert Noury, and Michel Vacher. Introducing knowledge in the process of supervised classification of activities of daily living in health smart homes. In *Proc. of IEEE Int. Conf. on e-Health Networking Applications and Services*, pages 322–329, Lyon, France, July 2010.
- [5] Yan Tang, Shuangquan Wang, Yiqiang Chen, and Zhenyu Chen. Ppcare: A personal and pervasive health care system for the elderly. In *Proc. of IEEE Int. Conf. on Ubiquitous Intelligence Computing on Autonomic Trusted Computing*, pages 935–939, Fukuoka, Japan, September 2012.
- [6] Donald F. Specht. Probabilistic neural networks. *Neural Networks*, 3(1):109–118, 1990.
- [7] Jeen-Shing Wang Jhun-Ying Yang and Yen-Ping Chen. Using acceleration measurements for activity recognition: An effective learning algorithm for constructing neural classifiers. *Pattern Recognition Letters*, 29(16):2213–2220, 2008.
- [8] Rafik Goubran Payam Moradshahi, Hanieh Chatzarrin. Improving the performance of cough sound discriminator in reverberant environments using microphone array. In *Proc. of IEEE Int. Conf. on Instrumentation and Measurement Technology*, pages 20–23, Graz, Austria, May 2012.
- [9] M. hamed Bilal Abidine and Belkacem Fergani. Evaluating c-svm, crf and lda classification for daily activity recognition. In *Proc. of IEEE Int. Conf. on Multimedia Computing and Systems*, pages 272 –277, Tangier-Morocco, May 2012.
- [10] Diane J. Cook. Learning setting-generalized activity models for smart spaces. *IEEE Intelligent Systems*, 27(1):32 –38, 2012.
- [11] M.T.Quazi N.K.Suryadevara and S.C.Mukhopadhyay. Intelligent sensing systems for measuring wellness indices of the daily activities for the elderly. In *Proc. of IEEE Int. Conf. on Intelligent Environments*, pages 347–350, Guanajuato, Mxico, June 2012.
- [12] Jae Hyuk Shin, Boreom Lee, and Kwang Suk Park. Detection of abnormal living patterns for elderly living alone using support vector data description. *IEEE Transactions on Information Technology in Biomedicine*, 15:438–448, 2011.
- [13] Huiru Zheng, Haiying Wang, and Norman Black. Human activity detection in smart home environment with self-adaptive neural networks. In *Proc. of IEEE Intl Conf. on Networking, Sensing and Control*, pages 1505–1510, Hainan, China, April 2008.
- [14] Qi Xie Samir Chatterjee and Kaushik Dutta. A predictive modeling engine using neural networks: Diabetes management from sensor and activity data. In *Proc. of IEEE Int Conf. on e-Health Networking, Applications and Services*, pages 230–237, October 2012.
- [15] K. Z. Mao, K.-C. Tan, and W. Ser. Probabilistic neural-network structure determination for pattern classification. *IEEE Transactions on Neural Networks*, 11(4):1009–1016, 2000.
- [16] J. A. Hartigan and M. A. Wong. A k-means clustering algorithm. *Applied Statistics*, 28(1):100–108, 1979.