Human Activity Recognition in Smart Homes: Research Challenges Classification

Samaneh, Zolfaghari
Alzahra University
Tehran, Iran
e-mail: s.zolfaghari.ir@ieee.org

Mohammad Reza, Keyvanpour
Alzahra University
Tehran, Iran
e-mail: keyvanpour@alzahra.ac.ir

ABSTRACT
Nowadays, learning and understanding of human activities lie at the heart of many research fields. The decision on what a person does at a given time according to the type of activity, the environment, and the person, poses numerous challenges. Failure to identify these challenges leads to some barriers in method development. Due to the lack of a comprehensive platform for systematic identification aimed at removing development barriers, this paper offers a comprehensive classification for challenges in activity recognition systems. Also a comparison structure is provided to map the well-known machine learning techniques onto human activity recognition challenges. The proposed classification can provide different research topics and help to develop more accurate and efficient activity recognition methods.

Keywords
Human Activity Recognition, Challenges Classification, Machine Learning, Smart Homes.

1. INTRODUCTION
With the rapid growth of the elderly population in the world, rising health care costs have led to an increasing economic burden on governments and people [1]. Therefore, Ambient Assisted Living (AAL) systems are required to help the elderly live independently [2]. In this regard, individuals must be able to undertake daily activities. Thus, automatic recognition of activities is an important step to achieve this goal [4].

In this area there are many obstacles and challenges. The main objective and innovation of this review paper is threefold: first, considering a variety of existing challenges in activity recognition systems in smart homes, then a classification of challenging issues in such systems is proposed. This classification introduces a systematic structure for challenges facing the activity recognition in health smart homes; third, the applications of well-known machine learning techniques used to address some of these challenges in this field are then discussed.

The remainder of this paper is organized as follows: Section 2 reviews the related work. In Section 3 the architecture of activity recognition systems in smart home is described. The proposed classification of challenges involved in Human Activity Recognition (HAR) in smart homes is presented in Section 4. Section 5 provides a comparative analysis of algorithms. Conclusion is presented in Section 6.

2. RELATED WORK
Activity recognition in health care applications can be effective to patients' treatment [4]. Previous studies in HAR in smart homes just reported different challenges which are caused by their infrastructure or they are related to the nature of human activities and the fact of learning algorithms which depend on environments, residents and activities [4][5].

Some of them tried to classify different approaches to recognize human activities in smart homes. An example of this classification is proposed in [4] and shown in Fig.1.

Despite these wide researches, there has been no study to represent an all-around classification, which covers all the existing HAR challenges. Furthermore, we think the idea of a classification framework for challenges that can be applied to grasp the complexity of the topic, should be applied to ongoing and future work and support researchers getting new into the field to target their work, is original and needed. Therefore, in this study, a comprehensive classification of HAR challenges in smart homes is proposed, which tries to cover all the existing challenges in this field, and also some well-known machine learning solution, which tries to solve these challenges has been investigated and evaluated.

3. THE PROCESS OF HAR IN SMART HOMES
Smart homes sense the users' characteristics and their environment through sensors and store the acquired data streams in a database. This collected information plus domain knowledge and users profile are used in processing, management and activity inference [4][5][6]. For the formalization of the context, it is critical to achieve a clear understanding of the considered activities in the smart home and its special features for proper sensor selection and system design [7]. The activities performed by the user in such a smart environment are recognized by applying the activity recognition techniques on data acquisition then the appropriate services are selected and presented by user interfaces [5][7]. In the next section, HAR challenges in smart homes are discussed using our proposed classification.

4. THE PROPOSED CHALLENGES CLASSIFICATION OF HAR IN SMART HOMES
In general, there is no system which offers excellent accuracy in HAR. This is due to the fact that most systems have their own shortcomings, activities are complicated, and human behaviours are complex in nature. In this section, a framework for classification of systematic challenges involved in HAR in smart homes is proposed, as shown in Fig.2. It should be noted that the challenges presented in this paper are based on challenging issues directly or indirectly referred to in various studies.

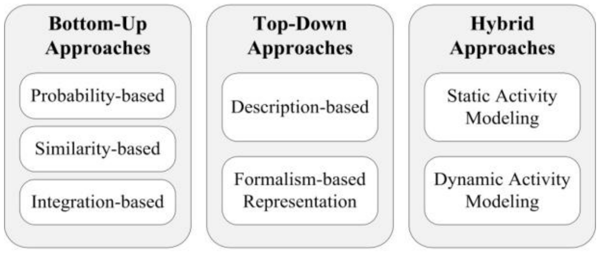

Fig. 1. Approaches Classification for HAR in Smart Homes [4]

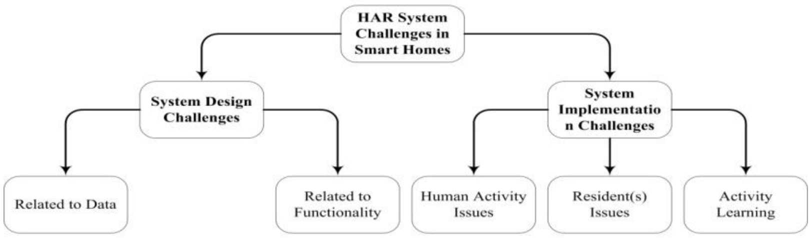

Fig. 1. Proposed HAR Challenges Classification

4.1. HAR System Design Challenges

This section discusses the challenges that need to be considered in design and development of smart home systems.

4.1.1. Data Challenges.

Most of HAR challenges can be considered in this category because it deals with sensory data. Sensory data are inherently noisy, inaccurate, have different sampling rates and complicated correlations [5][7]. To reduce these effects, there is a need to data cleaning techniques. These methods are filtering out artifacts and removing irrelevant information, in order to keep only the relevant information and smooth out raw data [7].

The probabilistic and statistic classifiers such as Naïve Bayesian (NB) [8], Support Vector Machine (SVM) [2], Conditional Random Field (CRF) [8], Hidden Markov Model (HMM) [1], and Dynamic Bayesian Network (DBN) [10] present a good framework for handling temporal and uncertain information which pose problems due to issues such as performing a specific activity in different ways by different people as well as uncertainty in activity duration.

Tracking a predefined list of activities requires a considerable amount of training data where collecting, labelling and annotating of such data in a smart environment is an extremely time-consuming and error-prone task [11]. There are always challenges between the annotation accuracy and the required annotation time. It is necessary to find ways to reduce the need of data labelling and provide acceptable accuracy. It seems that the use of semi-supervised [12] or unsupervised approaches [11] instead of supervised approaches is appropriate to recognize normal daily activities in a smart home setting.

However, some related work such as [22] showed that classification performance and accuracy can be improved by combining multiple classifiers together, instead of using one by ensemble methods [4].

A similar approach to semi-supervised learning is active learning. In this method, if a classifier is not very reliable on labelling, it needs human annotation, and as the human interaction is undesirable, so active learning is not appropriate in such application [12].

Skewed class distribution is another challenging issue [13]. Therefore, a small number of samples and imbalance in the resident activities in smart homes will lead to a gradual decline in the efficiency and accuracy of learning methods. Most machine learning algorithms are characterized by low performance in such problems [7]. There are several approaches to deal with this problem: sampling, instance reweighting, applying cost-sensitive learning or developing a specialized algorithm [13]. SVM classifier has shown its ability in the presence of imbalanced datasets because it only considers support vectors [14]. This classifier is computationally efficient and can achieve good performance in the presence of higher inter-class and lower intra-class changes [14][15].

To handle the missing value , the linear and the nearest neighbour and cubic interpolation can be used [4]. Class overlapping is another sensory data challenges in smart environment. This problem may lead to ambiguity [4][7].

4.1.2. System Functionality Issues

The AAL systems for older people should be adapted to their context and their needs [16]. Furthermore, most proposed models are based on specific daily activities that do not change over time and do not consider the fact that the patterns which define a daily activity can change due to the dynamic nature of human activity. This has led to adaptability and inconsistency reduction in real-world situations [17].

The proposed method in [21] can be adapted to environment, human habits and temporal information using Fuzzy logic and Learning Automaton. However, recognition accuracy is still reduced due to complex conditions such as interleaved activities and potential relationships between them. Using Transfer learning [18] and Evolving classifiers [17] is another solution to scalability, adaptability, and reusability and generality problems.

Another challenge in this field is the prediction reliability [6][17]. Additional pre-processing steps such as rankings features to select the most discriminative features, data balancing and defining the degree of confidence can ensure the reliability of HAR [14][15].

The most important functionalities in order of their importance are represented in Fig. 3.

After all it must be considered that older adults are not the only ones who use AAL systems but also other stakeholders should be taken into account [16].

4.2. HAR System Implementation Challenges

This section discusses the challenges that need to be considered in HAR system implementation challenges in smart homes.

4.2.1. Human Activity Issues in AR.

In general, the human behaviours can be classified using different perspectives and granularity levels [1][3]. The number, order, and duration of different steps of an activity vary significantly; even if it is done by the same user [4]. In [23] to overcome these challenges, multi-tape fuzzy finite state automata are proposed. They claimed their proposed method can handle a number of simultaneous inputs with provisions for handling variations in number, order and duration of these inputs.

HMM has high performance for recognition of short-term activities and has the potential to keep temporal information about smooth outliers [1]. This classifier is good for identifying temporal patterns, recognizing interleaved activities and predicting activity labels when a slow transition from one activity to another is accomplished [11].

However, the HMM learning model, unlike CRF, might not

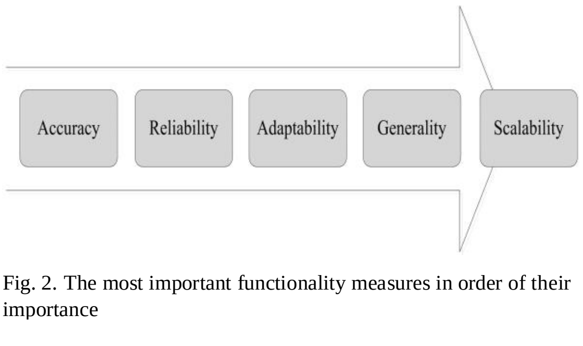

Fig. 2. The most important functionality measures in order of their importance

be able to capture long term dependency on the observed sensors due to its strong independence assumption [14].
Although, there is an over-fitting possibility in CRF learning model [20].
Generally, approaches which can model temporal relationships are graphical models. They can model complex activities [4][5]. Gu et al. in [19] also used frequent sensor mining to sequential, interleaved and concurrent activity recognition and design a scalable, noise-resistant to recognize both single and multiuser activities.

4.2.2. Resident(s) Issues.

It is clear that the ADL complexity increases with the number of residents in smart environments [3]. In the presence of several individuals many daily activities can be performed in parallel or in groups.
Currently, few studies have focused on the presence of multiple residents in the house such as [14] and [19] but this area of research is growing aimed to find proper models for capturing the interactions among individuals using accurate sensors and network design, processing multiple data streams of multiple users etc., which implies a scalability problem [19]. Machine learning techniques such as HMM, CRF, Coupled HMM (CHMM), Factorial Conditional Random Field (FCRF) etc., are activity learning algorithms which are used to resolve such challenges [3]. A great review article on this subject is [3].
Using some filters such as Bayes and particle filters for data pre-processing have the ability to noise filtering but they may struggle to track multiple occupants because other occupants do not behave like noise processes [7].
Thus, we need some manner to determine which occupant generated any observation which may become severe in this domain. In general, research in this area is slow and faces many problems. However, there are still numerous challenges when only a single resident lives in the house.

4.2.3. Related to Activity Learning.

Challenges based on learning are those which relate to applying learning algorithms and challenges which AmI faces including modelling and reasoning on the given volume of data, extracting and selecting features, choosing learning techniques and evaluating the measures for the suggested approach.
Datasets with different features lead to different accuracy rates in identifying activities. Choosing a dataset with inappropriate features increases the computational complexity as well as decreases the activity recognition accuracy [6].
On the other hand, uncertain information can handle by incorporating activity duration as a feature into the activity learning algorithm as it was done in [9]. Some well-known feature selection methods in this field include Information Gain (IG) based on entropy, Minimum Redundancy-Maximum Relevance (mRMR), Euclidean distance, correlation-based feature selection [6][9].
The AR approaches can be divided into Online and Offline approaches. Online approaches are able to extract patterns from continuous data streams and learning by incremental inferences which are less expensive than the construction of a new model [17].
Most approaches which have been presented in the field of online and real-time activity recognition are using fuzzy logic. In [21] Online behaviour recognition, is proposed based on automata learning and temporal windows. These methods also face challenges including the fact that it is impossible to discriminate between activities that are defined by the same sensor events, because fuzzy rules are not designed to deal with such hypotheses [17]. However, the Offline methods achieve acceptable accuracy, especially in

Table 1. Evaluation of well-known machine learning HAR methods based on classified challenges

Challenges Classification	**Methods Example**	**HAR System Design Challenges**									**HAR System Implementation Challenges**					
		Sensor Data Challenges					**System Functionality Issues**				**Human Activities Issues in AR**					**Resident Issues**
		Noise Handling	Uncertainty	Need Large volume of data	Labeled Data	Imbalancing	Accuracy	Adaptability	Scalability	Generality	Short Term Activities	Long Term Activities	Similar Activities	Concurrent Activities	Overlapped Activities	Multi-Resident
Offline Activity Learning Methods	NB [8]	*	*	*	*	*L*	*		*	*	*		*L*			*L*
	HMM [1]	*	*	*			*	*	*	*	*	*				
	CHMM[25]	*					*	*	*	*	*	*				*
	DBN [10]		*	*			*	*	*	*	*					*
	PCA-SVM [2]			*L*	*	*	*			*	*		*L*			*
	CSVM [20]			*	*	*	*		*L*		*		*L*			
	CRF[8]	*	*		*		*	*		*	*	*	*L*	*	*	*
	FCRF [25]	*	*		*		*	*		*	*	*		*	*	*
	DT [24]				*		*				*					*
	PNN [15]			*		*L*	*				*		*L*			*
	KNN [15]			*	*	*L*	*				*		*L*			*
	ET-KNN[15]			*	*	*L*	*				*		*L*			*
	ARSH-SV[14]	*	*	*L*	*	*	*				*		*			*
Online Activity Learning Methods	Evolving Classifiers [17]		*		*L*		*	*	*	*	*		*L*			
	FALA + Fuzzy [21]		*		*L*		*	*	*	*	*		*L*			

the presence of multi resident and recognition of complex activities in smart homes.

5. EVALUATION OF THE PROPOSED CLASSIFICATION FOR CHALLENGES IN HAR

In this section, we analyze the efficiency of different machine learning techniques in addressing some of the HAR challenges in a smart home. Many of these algorithms assert that their recommended methods are applicable in different kinds of situations for sensor-based activity recognition. But their results list one or a limited number of challenges presented earlier, as presented in Table 1. As can be seen, there are still many open issues that need to be addressed.

Compared to Offline activity recognition, little strong researches have been conducted on Online activity recognition (the mentioned researches in Table 1 are just some of strong ones.).

It should be noted, that the cells with "*" represented solvable challenges by the specified machine learning approach and the cell filled with "L" had a limited ability to handle that kind of challenges. However, despite the significant steps in this field, these solutions often suffer from the issues of scalability, security and privacy [16].

Large volumes of data collected by AmI systems can be useful in various fields and at the same time may cause many security issues. Particularly for health care systems, security and privacy were already very complex issues, and in addition to the large number of sensors and devices will lead to more challenges in this area [6].

6. CONCLUSION

In this paper, a comprehensive platform for systematic identification aimed at removing the development barriers, a classification for challenges in activity recognition systems which tries to cover all the existing challenges in this field, are presented. Also a comparison structure is provided to map some well-known machine learning techniques onto HAR challenges. These approaches and the classified challenges in this area were evaluated. The results of this study show that unlike many of these algorithms, their recommended methods are applicable to solve different challenges, but as represented in Table 1, their results list one or a limited number of challenges. Therefore, the proposed classification can generate different research topics and help to develop more accurate and efficient activity recognition methods as well as they can play an effective role in further development of this area.

REFERENCES

[1] H. Alemdar, C. Tunca, C. Ersoy, "Daily life behavior monitoring for health assessment using machine learning: bridging the gap between domains," Personal and Ubiquitous Comput., vol. 19, pp. 1–13, 2014.

[2] A. Fleury, M. Vacher, N. Noury, N., "Svm-based multimodal classification of activities of daily living in health smart homes: sensors, algorithms, and first experimental results," IEEE Trans. Inf. Technol. in Biomed. pp. 274-283, 2010.

[3] A. Benmansour, A. Bouchachia, M. Feham, "Multioccupant Activity Recognition in Pervasive Smart Home Environments," ACM Comput. Surveys (CSUR), pp. 34, 2015.

[4] S. Zolfaghari, M.R. Keyvanpour, "SARF: Smart activity recognition framework in Ambient Assisted Living," In: Federated Conf. on Comput. Science and Info. Sys. (FedCSIS). pp. 1435–1443, Gdansk, 2016.

[5] S. Zolfaghari, R. Zall, M.R. Keyvanpour, "SOnAr: Smart Ontology Activity recognition framework to fulfill Semantic Web in smart homes," In: Second Int. Conf. on Web Research (ICWR), pp. 139–144, Tehran, 2016.

[6] D.J. Cook, N.C Krishnan, "Activity learning: discovering, recognizing, and predicting human behavior from sensor data," John Wiley & Sons, 2015.

[7] Q. Ni, A.B. Garca Hernando, I. Pau de la Cruz, "The elderlys independent living in smart homes: A characterization of activities and sensing infrastructure survey to facilitate services development," Sensors, vol. 15, pp. 11312-11362, 2015.

[8] T.V. Kasteren, G. Englebienne, B. Krose, "Human activity recognition from wireless sensor network data: Benchmark and software," In: Proc. Of Activity recognition in Perv. Intell. environments. (2011).

[9] C. Chen, B. Das, D.J. Cook, "A data mining framework for activity recognition in smart environments," In: Proc. Of IEEE Int. Conf. on Intell. Environments, Kuala Lumpur, 2010.

[10] F.J. Ordonez, G. Englebienne, P.D. Toledo, T.V. Kasteren, A. Sanchis, B. Krose, "In-home activity recognition: Bayesian inference for hidden Markov models," IEEE Perv. Comput., pp. 67–75, 2014.

[11] P. Rashidi, D.J. Cook, L.B. Holder, M. Schmitter-Edgecombe, "Discovering Activities to Recognize and Track in a Smart Environment," IEEE Trans. Knowledge and Data Eng. 23, pp. 527–539, 2011.

[12] M. Stikic, K.V. Laerhoven, B. Schiele, "Exploring Semi-Supervised and Active Learning for Activity Recognition," pp. 81–88, 2008.

[13] K.D. Feuz, D.J. Cook, "Modeling Skewed Class Distributions by Reshaping the Concept Space," 2017.

[14] L.G. Fahad, A. Khan, M.Rajarajan,"Activity recognition in smart homes with self-verification of assignments,"Neurocomputing,vol.149, pp.1286-1298, 2015.

[15] L.G. Fahad, A. Ali, M. Rajarajan, "Learning models for activity recognition in smart homes," In: Proc. Of Int. Conf. on Inf. Sci. Appl.,pp. 819-826, Pattaya, 2015.

[16] P. Rashidi, A. Mihailidis, "A survey on ambient-assisted living tools for older adults," IEEE J of Biomed. and Health Informatics, vol. 17, pp. 579–590, 2013.

[17] J. Ordez, J. Iglesias, P.D. Toledo, "Online activity recognition using evolving classifiers," Expert Sys. with App., vol. 40, pp. 1248-1255, 2013.

[18] D.J. Cook, K. Feuz, N.C. Krishnan, "Transfer learning for activity recognition: A survey," Knowledge and Info. Sys., vol. 36, pp. 537-556, 2013.

[19] T. Gu, L. Wang, H. Chen, X. Tao, J. Lu, "Recognizing multiuser activities using wireless body sensor networks," IEEE Trans. on Mobile Comput., vol. 10, pp. 1618-1631 2011.

[20] M. Abidine, B. Fergani, "Evaluating c-svm, crf and lda classification for daily activity recognition," In: Proc. of the IEEE Int. Conf. on Multimedia Comput. Syst., Tangiers, 2012.

[21] M. Ros, M.P. Cullar, M. Delgado, A. Vila, "Online recognition of human activities and adaptation to habit changes by means of learning automata and fuzzy temporal windows," Info. Sciences, vol. 220, pp. 86–101, 2013.

[22] L.G. Fahad, M. Rajarajan, "Integration of discriminative and generative models for activity recognition in smart homes," Applied Soft Comput., 2015.

[23] H.K. Ali, D.G. Amalarethinam, "Activity Recognition with Multi-tapeFuzzy Finite Automata," Inter. J. of Modern Education and Computer Science, vol. 5, 2013.

[24] L. Bao, S. Intille, "Activity recognition from user-annotated acceleration data," In: Proc. Perv., pp. 1-17. (2004).

[25] L. Wang, T. Gu, X. Tao, H. Chen, J. Lu, "Multi-user activity recognition in a smart home," In: Perv. Intell. Environments Activity Recognition, pp. 59-81, Atlantis Press, 2011.