Mehran University Research Journal of Engineering & Technology
Vol. 37, No. 3, 467-482, July 2018
p-ISSN: 0254-7821, e-ISSN: 2413-7219
DOI: 10.22581/muet1982.1803.03

Adaptive HAR System to Improve Recognition Accuracy

MUHAMMAD ARSHAD AWAN*

RECEIVED ON 23.02.2017 ACCEPTED ON 29.05.2017

ABSTRACT

HAR (Human Activity Recognition) system becomes complex, inefficient and less accurate as we keep on adding new activities into the system; because it follows a specific procedure for activity recognition, from raw data collection to classification. In this study, we discuss an adaptive system to improve recognition accuracy. We developed a mathematical model to categorize the activities based on their data pattern. It observed that as we group the activities; although a separate classification model is required for each group, but it increases the recognition accuracy and efficiency of the system. The experiments on the data of eleven activities gathered from 10 volunteers proved the usability, scalability and effectiveness of our proposed methodology. The recognition accuracy of eleven activities was increased in total about 9-37% and reached up to 90% in different cases, using different number of groups and classification algorithms.

Key Words: **Activity Recognition, Adaptive Human Activity Recognition, Context-Awareness, Ubiquitous Computing.**

1. INTRODUCTION

HAR has become a prominent research field because of its inclusion in daily life. This field is growing due to progress in sensors and wireless network technologies. Creation of Smartphones with diverse sensors produced a boom to this research area. In the past, a variety of sensors for recognizing human activities were used on the body, causing an inconvenience to the user [1]. But now-a-days the Smartphone is being utilized in everyday life. From children to old age people are using the Smartphone comfortably and the use of Smartphone for HAR instead of several body-worn sensors has become common.

But the use of Smartphone in HAR is at naïve state. This field is a combination of signal processing and classification. When a user performs some physical activity, the horizontal and vertical readings from the accelerometer are being recorded for further processing. Although a number of HAR systems have been developed to RHA (Recognize Human Activities), but new activities cannot be added to a system to provide the future needs without making major changes in the overall infrastructure of the system [2]. This study is a follow up of our ongoing research on HAR.

Awan et. al. [2] have developed a method to include new activities into the system that makes it flexible for the user, but the system becomes complex, less accurate and inefficient as the activities are kept on increasing. The reason of this limitation is due to the fact that a specific

Authors E-Mail: (m99arshad@aiou.edu.pk)
* Department of Computer Science, Allama Iqbal Open University, Islamabad.

procedure is followed in an HAR system, from data collection/feature extraction to classification.

Usually, a single model is used for classification for a defined set of activities. In order to recognize the new set of activities, there is a need to upgrade the existing model or develop a new model for the complete set of activities, which makes the classification model a complex, less accurate and inefficient model because the classification model depends on the working mechanism of the machine learning algorithm [2]. Some machine algorithms work well on a limited number of classes, but as we keep on increasing the number of activities, the number of classes to be classified by the machine learning algorithm will also be increased and making the system complex as already described.

A similar kind of study to compare statistical features that are typically extracted from the sensors' raw data, and on classification algorithms, which are used to construct classification models, was done in [3], to examine the data behavior. Based on the study [3], it can be concluded that certain statistical features give an insight regarding the data pattern of an activity. The data pattern analysis is an approach that can be used to divide the activities into various groups. Then a different classification model can be built for each group, which reduce the complexity and increase the efficiency.

In this study, we present a mathematical model that make the groups of activities based on the standard deviation values extracted from the activities, using a defined threshold value. Classification models are developed according to the group of activities. We tried to prove the working of our proposed model by collecting the data of eleven activities, including; descending escalator laying, ascending escalator, descending elevator, ascending elevator, descending stairs, ascending stairs, jogging, walking, standing, and sitting from 10 volunteers.

The recognition accuracy was increased in total about 9-37% in different cases. The detail is given below in the proceeding paragraphs. We showed the usability, scalability and effectiveness of our approach.

SVM (Support Vector Machine), J48 (Decision Tree), KNN (K-Nearest Neighbor), MLP (Multilayer Perceptron), LR (Logistic Regression), NB (Native Bayes), and BN (Bayesian Network), are used classification, which are recommended classification algorithms for recognizing activities using Smartphone accelerometer.

Rest of the paper is organized as follows: Some related work is presented in Section 2, Section 3 describes the adaptive HAR system, along-with brief description of each component. In Section 4, we discussed our proposed classification approach as opposed to conventional approach. The methodology, including mathematical model, algorithm and flowchart is presented in Section 5. Section 6 is experimentation section, which tells the detail about experiments and their results and Section 7 concludes our paper.

2. RELATED WORK

In this section, some related work on HAR in general and particularly using Smartphone accelerometer described. There is no standard or benchmark to analyze the HAR system using Smartphone accelerometer. Various systems have been developed, based on a selective set of data features and classification algorithms [4-29], but are not flexible to adapt changes easily within the system, e.g. adding new activities and making new models to improve recognition accuracy and efficiency of the system. This study is a follow up of our ongoing research on HAR using Smartphone accelerometer, and we are providing a new approach to improve the efficiency and recognition accuracy of an HAR system, by introducing the concept of activity categorization based on data pattern, and then making classification models based on the categorization, as opposed to already developed HAR systems.

The HAR is a mixture of signal processing and classification. The data from Smartphone accelerometer processed through various steps and classification models are generated using machine learning algorithms to infer the user current activity. According to Guan et. al. [5] HAR system can be divided into two categories, including; video and physical sensor based activity recognition, while physical sensor based recognition can also be divided into wearable and object usage based activity recognition.Activity recognition using Smartphone accelerometer is considered in the wearable sensors approach, and is the focus of our study.

There is a continuous debate in the research community regarding less vs. more number of sensors on the body for better recognition, while considering the convenience of the user. Bao and Intille [7] used 5 accelerometers on the different parts of the body of the user and got an accuracy of 84% with an assumption that high accuracy can only be achieved with more sensors on the body of the user, while Kwapisz et. al. [8] conducted a study using Smartphone accelerometer and achieved 90% accuracy by the data of 29 users with an assumption that large data can improve accuracy. These and other related works only considered a limited number of activities, less than 10, and showed a reasonable accuracy with different classification algorithms. However, there is no optimal or standard approach to get the same amount of accuracy or efficiency when the number of activities is increasing with the requirement.

Awan et. al. [2] have presented a method to include various activities into the systemthat causes the system to cater new needs of the user, but makes the system complex, less accurate and inefficient. We provided an approach in this study to cater this limitation.

Various surveys [8-10], including the survey of Lara and Labrador [11] on HAR using wearable sensors have highlighted the several issues, including the inflexibility of current HAR systems. And there is no optimal solution still existed for an adaptive HAR system. Several studies [12-16] have attempted to tackle various issues in this field, for example, position- or orientation-independent activity recognition.

Another study [17] was presented that achieve device orientation independence. The study was conducted by placing the Smartphone at waist-mounted position and data was collected by keeping the phone in different orientations. Awan et. al. [18] presented a solution to achieve the subject-independence.

Normally, a single model is used for classification for a defined set of activities. In order to recognize the new set of activities, there is a need to upgrade the existing model or develop a new model for the complete set of activities, which makes the classification model a complex, less accurate and inefficient model because the classification model depends on the working mechanism of the machine learning algorithm. Some machine algorithms work well on a limited number of classes, but as we keep on increasing the number of activities, the number of classes to be classified by the machine learning algorithm will also be increased and making the system complex as already described. We utilized the technique of data behavior or pattern analysis to categorize the activities into different groups, and then making a different classification model for each group, which not only reduce the overall complexity of the classification model, but also increase the efficiency and recognition accuracy of the system.

3. ADAPTIVE HAR SYSTEM

Adaptive we mean a system capable of making changes in the system in order to handle the new situation and especially on feedback. Feedback mechanism can be within the system, third party application, and/or a user of the system. Adaptive HAR system should behave to make the system flexible depending upon feedback. Conceptual system presenting an adaptive HAR system can be shown in Fig. 1.

Sensor acquisition layer is the bottom most layer, which can be used to get the raw data. Let S be a set of sensors, then mathematically: $S = \{S_1, S_2, S_3, \ldots, S_n\}$, where $S_i = (x_i, y_i, z_i)_{i=1 \text{ to } n}$, and (x_i, y_i, z_i) are three axis of the accelerometer.

The next is preprocessing layer to remove noise from data. Normally, a low-pass filter is used to make raw data noise free. Let the preprocessing gives noise free data by using low-pass filter on set S, then $P = \{P_1, P_2, P_3, \ldots, P_n\}$, where $P_j = (x_j, y_j, z_j)_{j=1 \text{ to } n}$.

Segmentation is used to divide data in chunks. Various methods are used for segmentation and the most commonly used method is sliding window. The smaller and larger window sizes, both have pros and cons. Smaller size is recommended to be faster, needs less resources and energy. Larger size gives better accuracy. There are two techniques, i.e. sequential and overlapped sliding window method. The overlapped technique is believed to be good as compared to sequential approach in some cases.

The sequential approach considering the mobile platform can be better, as it requires fewer resources. The data segments from preprocessed data can be written as: $D = \{D_1, D_2, D_3, \ldots, D_m\}$.

The statistical features are extracted in the next step. The statistical data features provides the data behavior and that can be utilized to improve recognition accuracy. Feature extraction process can be demonstrated as: $F = \{F_1, F_2, F_3, \ldots, F_n\}$, where $F_k = f(D_k)$, and f is a statistical function.

Lastly, the process is to make classification models to recognize new instances. New instances are classified by matching the defined/known patterns. Let C represents set of classification models, we can write it as: $C = \{C_1, C_2, C_3, \ldots, C_n\}$, where $C_L = \{a\,(F_L)\}_{L=1 \text{ to } n}$, and a is a machine learning algorithm.

Fig. 2 can be seen for an overall execution flow of an Adaptive HAR system.

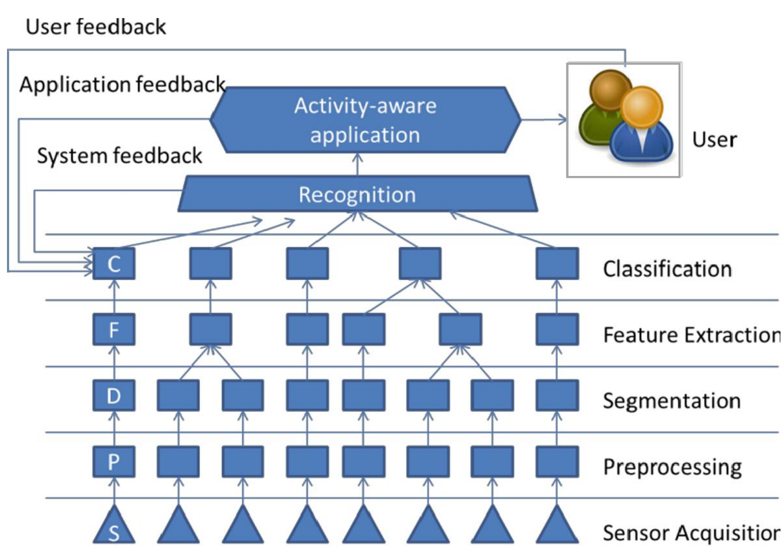

FIG. 1. AN ADAPTIVE HAR SYSTEM

4. CONVENTIONAL VS. PROPOSED CLASSIFICATION APPROACH

A process that makes classification models by applying machine learning algorithms and by training those models with training data set of activities is known as classification. Then activities are recognized based on the defined patterns. In conventional approach, single model is built based on defined machine learning algorithm for all set of activities, means making a single group for all activities. The approach is good for a smallnumber of activities, but as we keep on increasing the number of activities, the model becomes complex and performance reduces.

Therefore, a new classification approach is proposed that works well from medium to a large number of classes for classification. The mechanism behind this approach is to predict the data patterns of various activities and sort them into certain groups. A different classification model is built for each group of activities. The results showed the improvement in recognition accuracy.

A comparativeanalysis between conventional and our proposed approach can be depicted in Fig. 3. After preprocessing, we analyze the data behavior/patterns and make certain groups of activities based upon data pattern, as shown in Fig. 3.

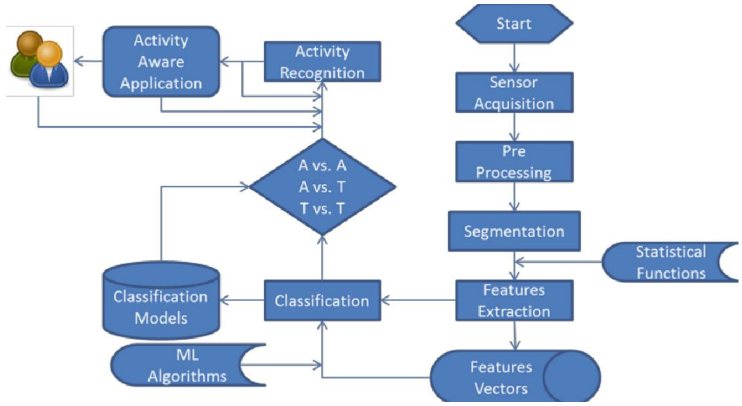

FIG. 2. OVERALL EXECUTION FLOW OF AN ADPATIVE HAR

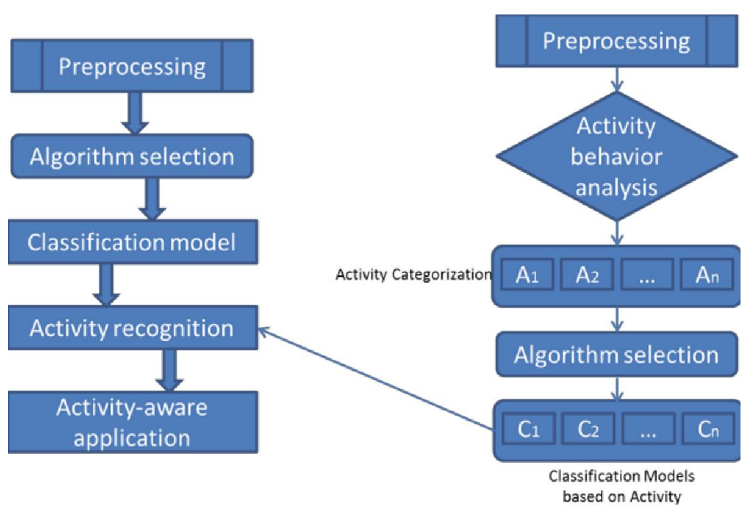

FIG. 3. CONVENTIONAL VS. PROPOSED CLASSIFICATION APPROACH

Fig. 4 shows an overview of data behavior. Some activities show similar patterns, even with different values. The x-axis represents the time series of data, while y-axis representing the data value of x, y, and z-axis for different activities. These 3D (Three Dimensional) values are the accelerometer data of a Smartphone collected for different activities.

After experimentations on data pattern/behavior, we came to conclusion that SD (Standard Deviation) can be utilized to group various activities. Since the standard deviation, by definition is a measure of the dispersion of a set of data from its mean, so different activities that even have different mean values, can have same or nearly equal standard deviation values. The activities; laying, sitting, and standing have different values but nearly equal SD value. The SD feature is used to build a method of grouping activities and make classification models to improve the recognition accuracy of the system.

5. METHODOLOGY

A mathematical model is used to elaborate the proposed approach. Step by step detail is given beneath:

- Let us define a set of activities represented by A. So $A = \{A_1, A_2, A_3, \ldots, A_n\}$.
- In our case, set of activities are: {Laying, Sitting, Standing, Walking, Jogging, AscStairs, DescStairs, AscElevator, DescElevator, AscEscalator, DescEscalator}.
- A double value of x-, y-, and z-axis of accelerometer is obtained.
- Let us assume the data set by D, then $D = \{D_1, D_2, D_3, \ldots, D_n\}$, and $D_i = (x_i, y_i, z_i)$, where $i = 1$ to n.
- All steps before classification is considered as preprocessing.
- Classification algorithm selection is the next step tomake a classification model "C" for activity recognition.

Proposition: The SD feature can be used to predict the data pattern of different activities even with different data values to make categorization of activities.

Definition: The SD is a measure of the dispersion of a set of data from its mean.

Proof: we will prove our statement by logical deduction.

- As by definition of SD, the data pattern of activities can be determined, and thus can be used for activities categorization.
- Let we have a set of activities as stated above, $A = \{A_1, A_2, A_3, \ldots, A_n\}$.
- And SD(A) denotes the set based on SD values of the activities, then $(A) = \{SD(A_1), SD(A_2), SD(A_3), \ldots, SD(A_n)\}$.
- Let SLA represents a sorted list of activities, then $Ord\{SD(A)\} \rightarrow SD(SLA)$, and $SD(SLA) = \{SD(A_i): 1 <= i <= n\}$, for all $SD(A_i) <= SD(A_{i+1})$.

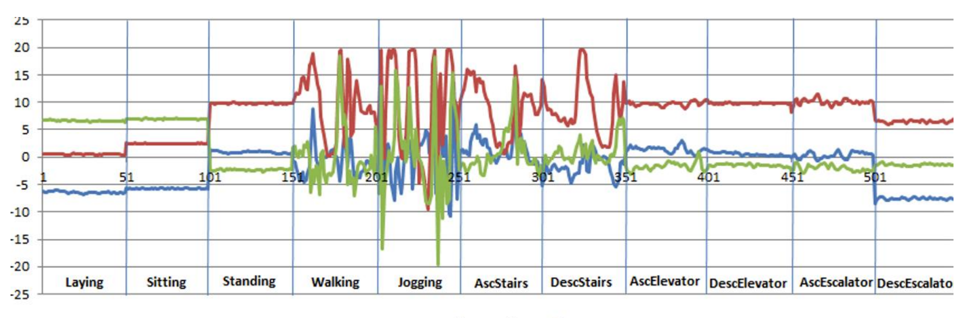

FIG. 4. DATA BEHAVIOR/PATTERN OF VARIOUS ACTIVITIES

- Similarly, Ord(A) → (SLA), and SLA = {A_i: 1 <= i <= n | A is an activity}, for all SD(A_i) <= SD(A_{i+1}).
- Now we can group the activities according to certain threshold value.
- Let G(A) denotes the group of activities with at least two activities, then G(A) = {G(A_i): 1 <= i <= n | G(A) is a group of activities}, for all G(A_{i+1}) – G(A_i) <= θ (Theta), where θ is a threshold value.

Now let us consider an example to elaborate our methodology and mathematical proof. This example is based on real data, which was collected from 10 volunteers, for a set of eleven activities defined before.

The data was collected from the accelerometer of Smartphone by keeping it in the left trousers pocket, facing upward direction by using an android application.

Example: Consider the statistical data feature values of eleven activities, as given in Table 1.

- It can be seen from Table 1, that laying and standing activities have same SD value, but totally different values of mean, min and max and similarly, in the case of ascending and descending stairs.
- Let A = {Laying, Sitting, Standing, Walking, Jogging, AscStairs, DescStairs, AscElevator, DescElevator, AscEscalator, DescEscalator}
- SD(A) = {0.07, 0.06, 0.07, 3.08, 6.10, 2.23, 2.52, 0.39, 0.24, 0.17, 0.17}
- SD(SLA) = {0.06, 0.07, 0.07, 0.17, 0.17, 0.24, 0.39, 2.23, 2.52, 3.08, 6.10}
- SLA = {Sitting, Laying, Standing, AscEscalator, DescEscalator, DescElevator, AscElevator, AscStairs, DescStairs, Walking, Jogging}
- Now let us consider the threshold value, θ = 0.16, then
- G(A_1) = {Laying, Sitting, Standing, AscElevator, DescElevator, AscEscalator, DescEscalator}
- G(A_2) = {AscStairs, DescStairs}, and
- G(A_3) = {Walking, Jogging}

The algorithm that we used for implementation.

TABLE 1. STATISTICAL DATA FEATURES OF 11 ACTIVITIES

Activity	Maximum	Minimum	Mean	Standard Deviation
DescEscalator	-7.07	-8.43	-7.69	0.17
AscEscalator	1.30	-0.61	0.29	0.17
DescElevator	1.49	-1.02	0.79	0.24
AscElevator	2.93	-1.67	1.20	0.39
DescStairs	16.41	-12.39	0.24	2.52
AscStairs	14.02	-7.13	0.87	2.23
Jogging	19.46	-14.65	1.19	6.10
Walking	14.50	-7.35	-0.14	3.08
Standing	1.42	0.50	0.95	0.07
Sitting	-5.54	-6.03	-5.77	0.06
Laying	-6.07	-6.99	-6.51	0.07

ALGORITHM:

Data: list of all activities along-with their SD values

Result: group of activities based on certain threshold (θ) value

1. get the SD values for each activity
2. create a list to store SD values
3. call a method to sort the SD values
4. call a method to sort activities according to SD values
5. set a threshold (θ) value for activities grouping
6. create a group and put first activity into it
7. **do while** list is not empty (for each activity)
8. **if** next activity lies within threshold (θ) value
9. put into the same activity group
10. **else**
11. create a new group
12. add activity into this group
13. **end if**
14. **end do**
15. call a method to send the activity groups for classification
16. **select** classification algorithm
17. make a classification model for each activity group
18. **end select**
19. recognize activities

A flowchart for the proposed methodology is shown in Fig. 5.

6. EXPERIMENTATION

In order to prove the usability, scalability and effectiveness of our proposed approach, a number of experiments were performed on a set of 11 activities using Smartphone.

Proposition: The accuracy of RHA can be improved by grouping activities based on SD value and making different classification model for each group.

The assumption was proved through experiments. Below is the description along-with some snapshots and related detail. WEKA [30] tool is used for experiments.

Data of 11 activities was gathered from 10 participants. Seven most commonly used classification algorithm's were used for classification.

Table 2 shows the confusion matrix for 11 activities by applying J48 classification algorithm. 11 activities are shown using "a to k" symbols. Overall accuracy of about 81% is achieved in case of J48 algorithm. A summary using all classification algorithms is presented later.

A comparison summary by using one three classification models is shown in Table 3, after grouping activities.

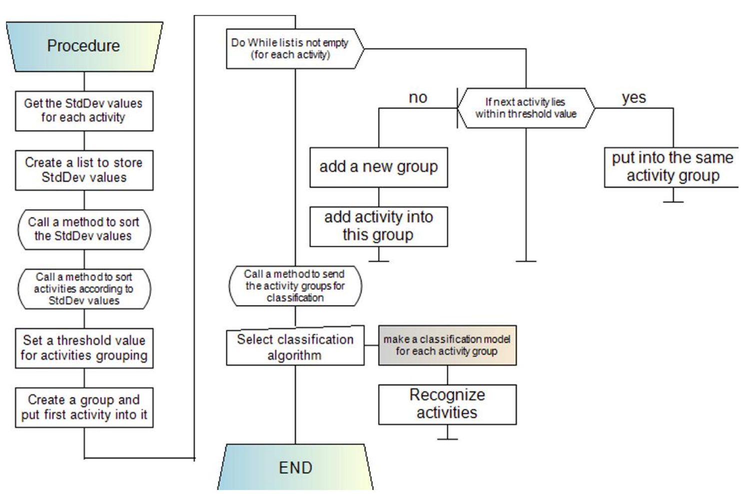

FIG. 5. PROPOSED METHODOLOGY, A FLOW-CHART

Seven most commonly used classification algorithms were used for classification. Column-1 of Table 3 shows the name of classification algorithms, and the average percentage of accuracy for 11 activities, which are mentioned with symbols from "a to k", is presented in the last column. The symbol 1 and 3 with the names of the algorithms is used to represent1 and3 groups respectively, e.g. NB1 and NB3, means Naïve Bayes algorithm for 1 and 3 groups respectively.

Graphs in Figs. 6-7 show a comparison of recognition accuracy for 1 vs. 3 groups of activities.

TABLE 2. 11 ACTIVITIES' CONFUSION MATRIX USING J48 ALGORITHM

Activity	a	b	c	d	e	f	g	h	I	j	k
k	0	0	0	0	0	0	0	0	0	0	870
j	0	0	6	1	0	24	2	15	24	778	0
i	0	0	255	0	0	17	4	113	736	25	0
h	0	0	138	0	0	62	5	1337	143	15	0
g	0	1	1	430	257	550	1438	8	9	6	0
f	0	0	3	329	98	2143	441	32	26	28	0
e	4	1	0	203	1783	145	310	1	3	0	0
d	0	0	0	1777	140	362	417	4	0	0	0
c	0	0	3407	0	0	2	0	15	67	9	0
b	0	4698	0	0	2	0	0	0	0	0	0
a	3748	0	0	0	2	0	0	0	0	0	0

TABLE 3. RECOGNITION ACCURACY COMPARISON, 1 VS. 3 GROUPS OF ACTIVITIES

Algorithm	A	b	c	d	e	f	g	h	i	j	k	(%)
SVM1	100	100	99.4	70.9	76.0	73.4	57.9	72.2	55.3	94.2	100	81.75
SVM3	100	100	100	70.3	76.2	76.7	57.5	74.5	87.0	96.2	100	85.31
J48-1	99.9	100	97.3	65.8	72.8	69.1	53.3	78.6	64.0	91.5	100	81.12
J48-3	100	100	100	66.7	72.7	72.3	52.9	83.1	86.2	93.6	99.9	84.31
KNN1	100	100	89.7	64.1	72.4	60.6	47.5	79.5	60.9	89.5	100	78.56
KNN3	100	100	100	64.5	71.8	63.5	48.3	83.6	76.5	92.8	100	81.91
MLP1	100	100	98.3	54.6	52.1	42.7	41.6	70.4	36.0	91.6	99.9	71.56
MLP3	100	100	100	25.7	72.9	69.7	20.3	73.4	86.1	95.3	100	76.67
LR1	100	100	100	28.5	27.1	21.4	5.60	0.00	0.00	0.00	100	43.87
LR3	100	100	100	42.2	34.1	67.1	3.30	75.6	74.9	94.2	100	71.95
BN1	100	100	97.5	48.2	68.4	65.8	37.1	69.8	57.9	91.6	99.9	76.02
BN3	100	100	100	47.8	68.9	71.1	35.2	72.2	82.1	94.2	100	79.23
NB1	99.9	100	98.7	32.9	63.4	77.8	8.50	71.5	47.2	94.9	99.2	72.18
NB3	100	100	100	33.3	63.4	62.0	9.10	73.4	80.1	95.1	100	76.04

In order to prove the robustness of our approach we tested the data set of eleven activities for different threshold values, and found to be correct in each case. Table 4, and Figs. 8-9 show the comparison of recognition accuracy for one vs. 2 and 3 groups of activities.

Similarly, Table 5, and Figs. 10-11 show the comparison of recognition accuracy for one vs. 2, 3 and 4 groups of activities.

A comparison was also done on the time to build the classification models. This comparison was made in order to check the overhead of making different classification model for each group of activities. Results proved that the overall time taken to build different classification model for each group with a small number of activities in every group is smaller than building the one classification model for a large number of activities. Table 6 can be seen for a comparison summary.

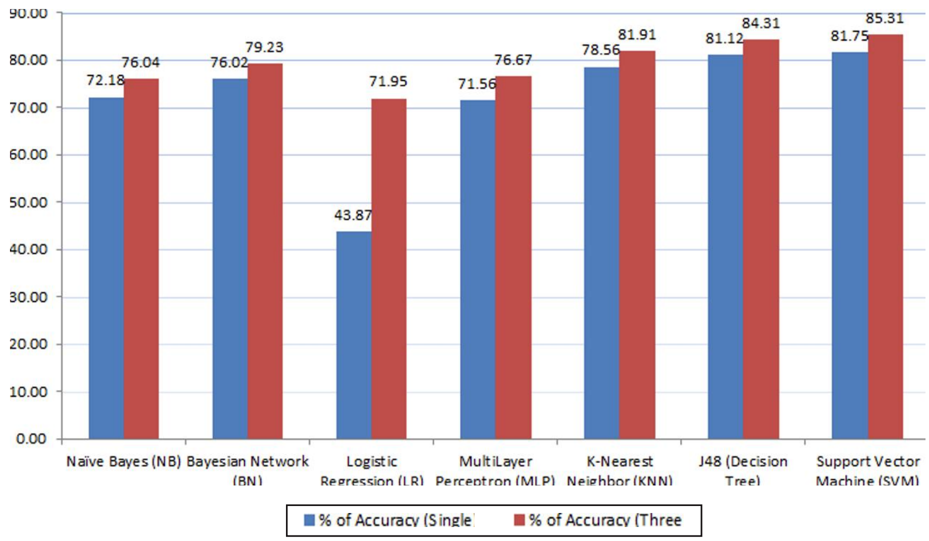

FIG. 6. RECOGNITION ACCURACY COMPARISON, 1 VS. 3 GROUPS OF ACTIVITIES

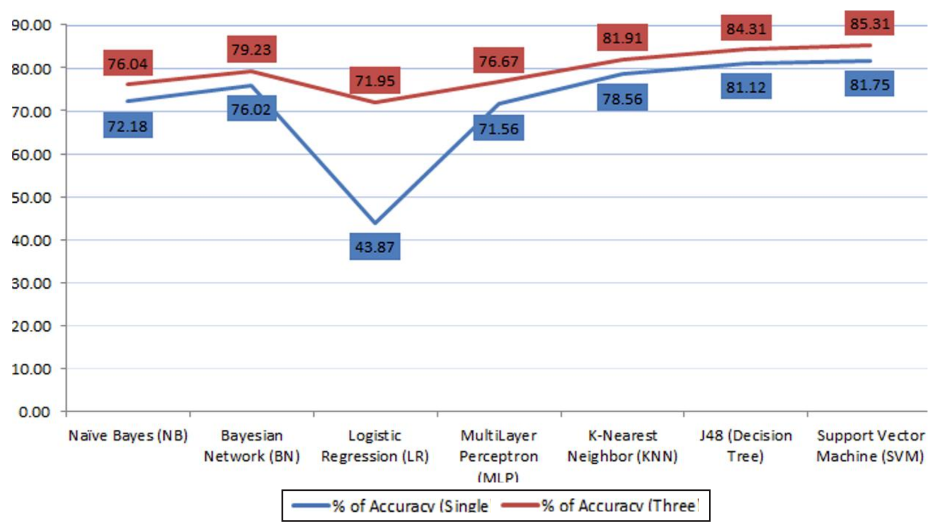

FIG. 7. RECOGNITION ACCURACY COMPARISON, 1 VS. 3 GROUPS OF ACTIVITIES

TABLE 4. RECOGNITION ACCURACY COMPARISON, 1 VS. 2 AND 3 GROUPS OF ACTIVITIES

Algorithm	k	j	i	h	g	F	e	d	c	b	a	(%)
SVM1	100	94.2	55.3	72.2	57.9	73.4	76.0	70.9	99.4	100	100	81.75
SVM2	100	96.0	54.9	74.0	57.5	76.7	76.2	70.3	99.4	100	100	82.27
SVM3	100	96.2	87.0	74.5	57.5	76.7	76.2	70.3	100	100	100	85.31
J48-1	100	91.5	64.0	78.6	53.3	69.1	72.8	65.8	97.3	100	99.9	81.12
J48-2	99.8	94.5	62.7	82.3	52.9	72.3	72.7	66.7	97.1	100	100	81.91
J48-3	99.9	93.6	86.2	83.1	52.9	72.3	72.7	66.7	100	100	100	84.31
KNN1	100	89.5	60.9	79.5	47.5	60.6	72.4	64.1	89.7	100	100	78.56
KNN2	100	91.8	62.4	82.2	48.3	63.5	71.8	64.5	89.4	100	100	79.45
KNN3	100	92.8	76.5	83.6	48.3	63.5	71.8	64.5	100	100	100	81.91
MLP1	99.9	91.6	36.0	70.4	41.6	42.7	52.1	54.6	98.3	100	100	71.56
MLP2	99.9	92.4	60.5	75.9	20.3	69.7	72.9	25.7	95.7	100	100	73.91
MLP3	100	95.3	86.1	73.4	20.3	69.7	72.9	25.7	100	100	100	76.67
LR1	100	0.00	0.00	0.00	5.60	21.4	27.1	28.5	100	100	100	43.87
LR2	100	92.2	41.6	68.8	3.30	67.1	34.1	42.2	99.1	100	100	68.04
LR3	100	94.2	74.9	75.6	3.30	67.1	34.1	42.2	100	100	100	71.95
BN1	99.9	91.6	57.9	69.8	37.1	65.8	68.4	48.2	97.5	100	100	76.02
BN2	100	94.1	56.4	71.9	35.2	71.1	68.9	47.8	97.5	100	100	76.63
BN3	100	94.2	82.1	72.2	35.2	71.1	68.9	47.8	100	100	100	79.23
NB1	99.2	94.9	47.2	71.5	8.50	77.8	63.4	32.9	98.7	100	99.9	72.18
NB2	100	95.1	47.1	73.4	9.10	82	63.4	33.3	98.5	100	100	72.90
NB3	100	95.1	80.1	73.4	9.10	62.0	63.4	33.3	100	100	100	76.04

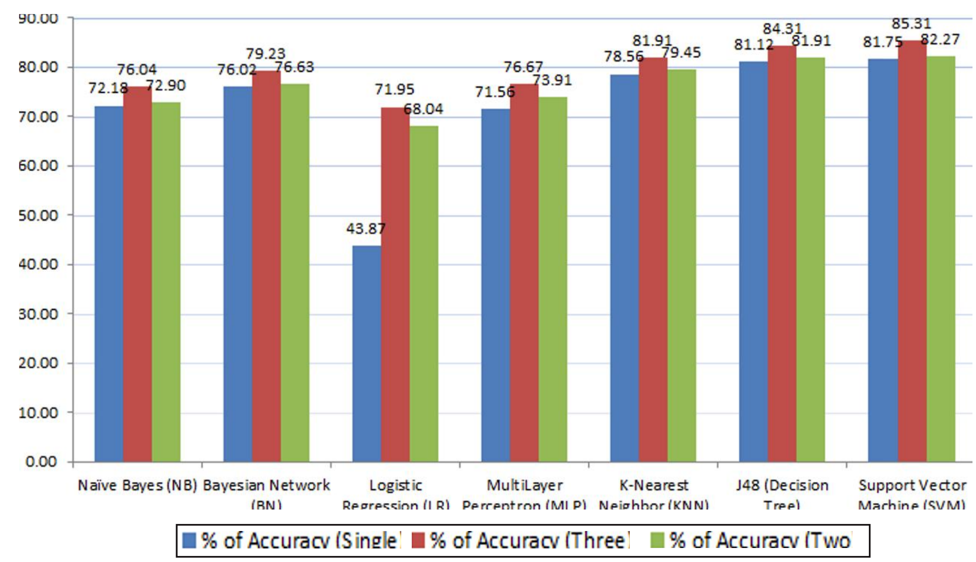

FIG. 8. RECOGNITION ACCURACY COMPARISON, 1 VS. 2 AND 3 GROUPS OF ACTIVITIES

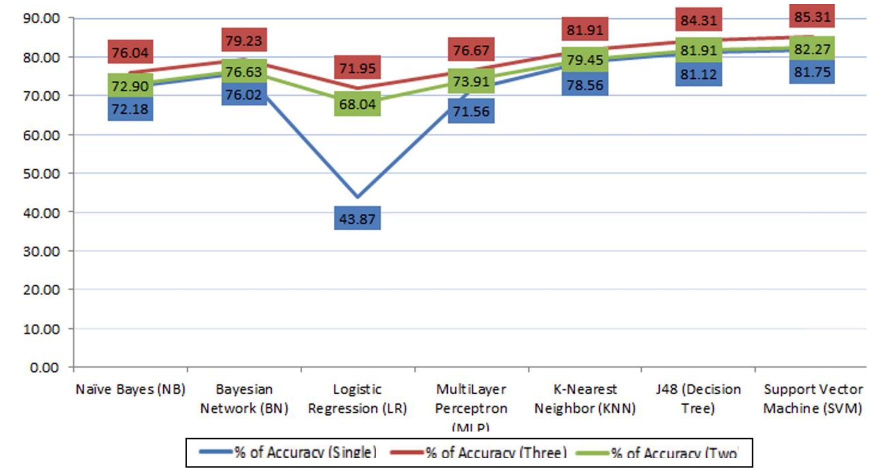

FIG. 9. RECOGNITION ACCURACY COMPARISON, 1 VS. 2 AND 3 GROUPS OF ACTIVITIES

TABLE 5. RECOGNITION ACCURACY COMPARISON, 1 VS. 2, 3 AND 4 GROUPS OF ACTIVITIES

Algorithm	k	j	i	h	g	F	e	d	c	b	a	(%)
SVM1	100	94.2	55.3	72.2	57.9	73.4	76.0	70.9	99.4	100	100	81.75
SVM2	100	96.0	54.9	74.0	57.5	76.7	76.2	70.3	99.4	100	100	82.27
SVM3	100	96.2	87.0	74.5	57.5	76.7	76.2	70.3	100	100	100	85.31
SVM4	100	96.2	87.0	74.5	75.5	82.6	86.2	93.1	100	100	100	90.46
J48-1	100	91.5	64.0	78.6	53.3	69.1	72.8	65.8	97.3	100	99.9	81.12
J48-2	99.8	94.5	62.7	82.3	52.9	72.3	72.7	66.7	97.1	100	100	81.91
J48-3	99.9	93.6	86.2	83.1	52.9	72.3	72.7	66.7	100	100	100	84.31
J48-4	99.9	93.6	86.2	83.1	75.0	79.8	82.6	93.7	100	100	100	90.35
KNN1	100	89.5	60.9	79.5	47.5	60.6	72.4	64.1	89.7	100	100	78.56
KNN2	100	91.8	62.4	82.2	48.3	63.5	71.8	64.5	89.4	100	100	79.45
KNN3	100	92.8	76.5	83.6	48.3	63.5	71.8	64.5	100	100	100	81.91
KNN4	100	92.8	76.5	83.6	68.4	75.0	86.6	88.7	100	100	100	88.33
MLP1	99.9	91.6	36.0	70.4	41.6	42.7	52.1	54.6	98.3	100	100	71.56
MLP2	99.9	92.4	60.5	75.9	20.3	69.7	72.9	25.7	95.7	100	100	73.91
MLP3	100	95.3	86.1	73.4	20.3	69.7	72.9	25.7	100	100	100	76.67
MLP4	100	95.3	86.1	73.4	66.0	68.3	56.4	92.7	100	100	100	85.29
LR1	100	0.00	0.00	0.00	5.60	21.4	27.1	28.5	100	100	100	43.87
LR2	100	92.2	41.6	68.8	3.30	67.1	34.1	42.2	99.1	100	100	68.04
LR3	100	94.2	74.9	75.6	3.30	67.1	34.1	42.2	100	100	100	71.95
LR4	100	94.2	74.9	75.6	34.7	78.3	49.9	72.6	100	100	100	80.02
BN1	99.9	91.6	57.9	69.8	37.1	65.8	68.4	48.2	97.5	100	100	76.02
BN2	100	94.1	56.4	71.9	35.2	71.1	68.9	47.8	97.5	100	100	76.63
BN3	100	94.2	82.1	72.2	35.2	71.1	68.9	47.8	100	100	100	79.23
BN4	100	94.2	82.1	72.2	59.3	80.3	75.7	84.2	100	100	100	86.18
NB1	99.2	94.9	47.2	71.5	8.50	77.8	63.4	32.9	98.7	100	99.9	72.18
NB2	100	95.1	47.1	73.4	9.10	82	63.4	33.3	98.5	100	100	72.90
NB3	100	95.1	80.1	73.4	9.10	62.0	63.4	33.3	100	100	100	76.04
NB4	100	95.1	80.1	73.4	33.3	85.1	64.3	88.1	100	100	100	83.58

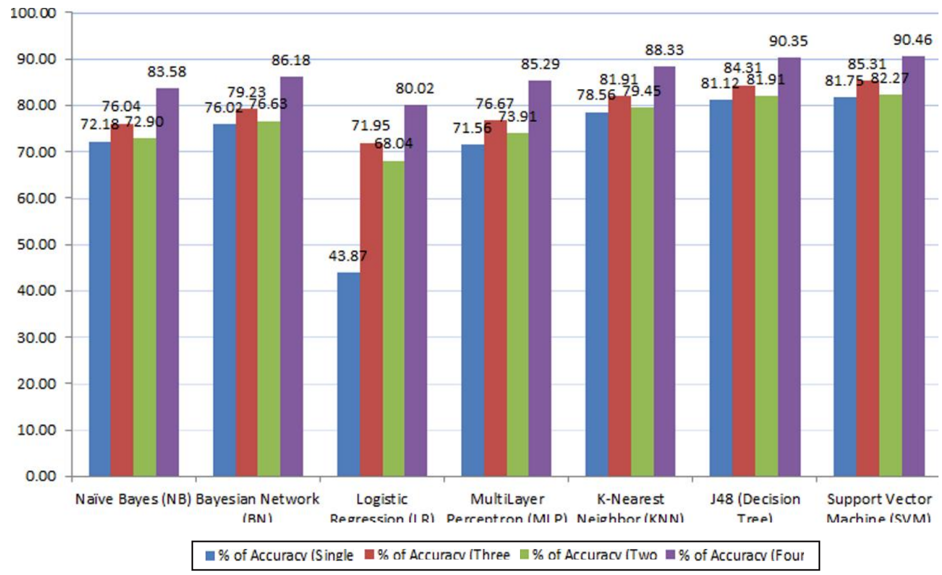

FIG. 10. RECOGNITION ACCURACY COMPARISON, 1 VS. 2, 3 AND 4 GROUPS OF ACTIVITIES

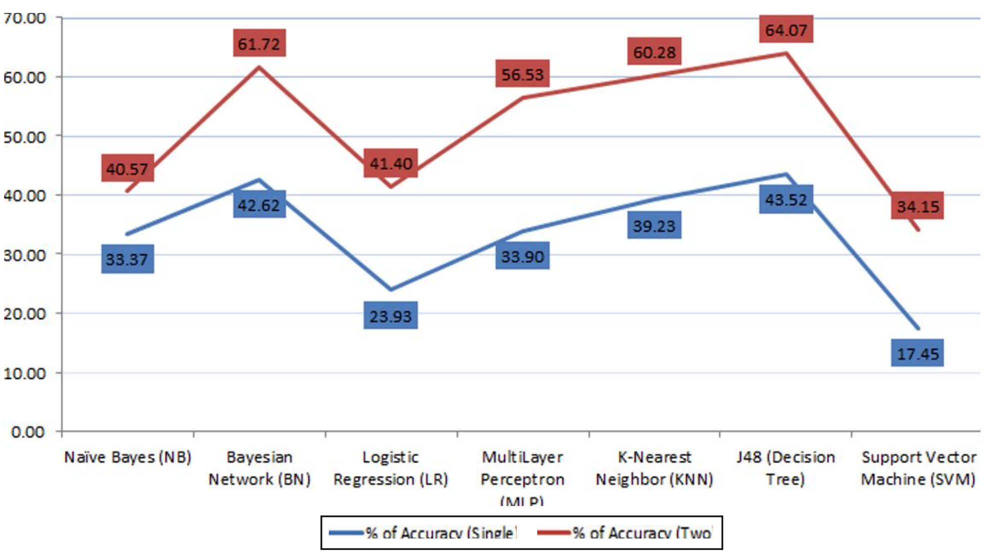

FIG. 11. RECOGNITION ACCURACY COMPARISON, 1 VS. 2, 3 AND 4 GROUPS OF ACTIVITIES

TABLE 6. CLASSIFICATION MODELS, 1 VS. 2, 3 AND 4 GROUPS (TIME TAKEN TO BUILD MODEL)

Time (sec)	SVM	J48	KNN	MLP	LR	BN	NB
Single	25.64	2.88	0.02	80.49	10.05	0.54	0.15
Two	17.17	1.34	0.01	35.02	4.06	0.34	0.13
Three	15.73	0.89	0.01	23.68	1.47	0.32	0.14
Four	8.74	0.46	0.02	17.76	1.28	0.39	0.12

7. CONCLUSION

The study demonstrates an adaptive HAR system to improve accuracy in recognition process. A mathematical model was developed that can be used to group activities into distinct categories depending upon their behavior/ pattern. The patterns of activities were analyzed through SD feature. During experimentation, it was detected that making groups of activities on smaller threshold values, not only increase the recognition accuracy of overall system, but also of individual activities, although there may be more groups and a different classification model is required for each group. Set of eleven activities were under observation to prove the usability, scalability and effectiveness of our proposed methodology. The recognition accuracy of eleven activities was increased in total about 9-37% and reached up to 90% in different cases, using different number of groups and classification algorithms.

ACKNOWLEDGEMENT

This research was supported by the Allama Iqbal Open University, Islamabad, Pakistan and Yonsei University, Seoul, South Korea.

REFERENCES

[1] Thuy, N., and Grenville, A., "A Survey of Techniques for Internet Traffic Classification using Machine Learning", IEEE Communications Surveys and Tutorials, Volume 10, No. 4, pp. 56-76, November, 2008.

[2] Awan, M.A., Guangbin, Z., and Kim, S.D., "A Dynamic Approach to Recognize Activities in WSN",International Journal of Distributed Sensor Networks, [online] Volume 2013, 2013.

[3] Awan, M.A., Guangbin, Z., Kim, C.G., and Kim, S.D., "Human Activity Recognition in WSN: AComparative Study",International Journal of Networked and Distributed Computing, Volume 2, No. 4, pp. 221-230, October, 2014.

[4] Trussell, H., Nilsson, A., Patel, P., and Wang, Y., "Estimation and Detection of Network Traffic", Proceedings of of 11th Digital Signal Processing Workshop, pp. 246-248, January 12-16, 2004.

[5] McGregor, A., Perry Lorier, M.H., and Brunskill, J., "Flow Clustering using Machine Learning Techniques", Proceedings of 5th International Workshop on Passive and Active Network Measurement,pp.205-214, June 2-7, 2004.

[6] Zander, S., Nguyen, T., and Armitage, G., "Self-Learning IP Traffic Classification Based on Statistical Flow Characteristics", Proceedings of 6th International Workshop on Passive and Active Measurement, pp. 325-328, March 23-27, 2005.

[7] Zander, S., Nguyen, T., and Armitage, G., "Automated Traffic Classification and Application Identification using Machine Learning", Proceedings of IEEE Conference on Local Computer Networks, pp. 250-257, February 11-12, 2005.

[8] Anguita, D., Ghio, A., Oneto, L., Parra, X., and Reyes-Ortiz, J.L., "Human Activity Recognition on Smartphones using a Multiclass Hardware-Friendly Support Vector Machine", International Workshop of Ambient Assisted Living, Vitoria-Gasteiz, Spain, December, 2012.

[9] WEKA (Waikato Environment for Knowledge Analysis), [online], http://www.cs.waikato.ac.nz/ml/weka/ (Accessed: 28thOctober 2013).

[10] van Kasteren, T.L.M., Englebienne, G., and Kröse, B.J.A., "Human Activity Recognition from Wireless Sensor Network Data: Benchmark and Software", Activity Recognition in Pervasive Intelligent Environments, Volume 4, pp. 165-186, 2011.

[11] Wang, S., Yang, J., Chen, N., Chen, X., and Zhang, Q., "Human Activity Recognition with User-Free Accelerometers in the Sensor Networks", Neural Networks and Brain, Volume 2, pp. 1212-1217, 2005.

[12] Bao, L., and Intille, S.S., "Activity Recognition from User-Annotated Acceleration Data", Lecture Notes in Computer Science, Pervasive Computing, Volume 3001, pp. 1-17, 2004.

[13] Kwapisz, J.R., Weiss, G.M., and Moore, S.A., "Activity Recognition Using Cell Phone Accelerometers", ACM SIGKDD Explorations Newsletter, Volume 12, No. 2, pp. 74-82, 2010.

[14] Oh, K., Park, H.S., and Cho, S.B., "A Mobile Context Sharing System Using Activity and Emotion Recognition with Bayesian Networks", Proceedings of 7thInternational Conference on Autonomic & Trusted Computing, pp. 244-249, 2010.

[15] Grosse-Puppendahl, T., Berlin, E., and Borazio, M., "Enhancing Accelerometer-Based Activity Recognition with Capacitive Proximity Sensing",Lecture Notes in Computer Science, Ambient Intelligence, Volume 7683, pp. 17-32, 2012.

[16] Liang, Y., Zhou, X., Yu, Z., Guo, B., and Yang, Y., "Energy Efficient Activity Recognition Based on Low Resolution Accelerometer in Smart Phones",Lecture Notes in Computer Science, Advances in Grid and Pervasive Computing, Volume 7296, pp. 122-136, 2012.

[17] Nam, Y., and Park, J.W., "Child Activity Recognition Based on Cooperative Fusion Model of a Tri-Axial Accelerometer and a Barometric Pressure Sensor", IEEE Journal of Biomedical and Health Informatics, Volume 17, No. 2, pp. 420-426, 2013.

[18] Bourke, A.K., van de Ven, P., Gamble, M., O'Connor, R., Murphy, K., Bogan, E., McQuade, E., Finucane, P., O'Laighin, G., and Nelson, J., "Evaluation of Waiste-Mounted Tri-Axial Accelerometer Based Fall-Detection Algorithms during Scripted and Continuous Unscripted Activities", Journal of Biomechanics, Volume 43, pp. 3051-3057, 2010.

[19] Guan, D., Ma, T., Yuan, W., Lee, Y.K., and Jehad Sarkar, A.M., "Review of Sensor-Based Activity Recognition Systems", IETE Technical Review, Volume 28, No. 5, pp. 418-433, 2011.

[20] Baldauf, M., Dustdar, S., and Rosenberg, F., "A Survey on Context-Aware Systems", International Journal of Ad Hoc and Ubiquitous Computing, Volume 2, No. 4, pp. 263-277, 2007.

[21] Alemdar, H., and Ersoy, C., "Wireless Sensor Networks for Healthcare: ASurvey", Computer Networks, Volume 54, No. 15, pp. 2688-2710, 2010.

[22] Scanaill, C.F., Carew, S., Barralon, P., Noury, N., Lyons, D., and Lyons, G.M., "A Review of Approaches to Mobility Telemonitoring of the Elderly in their Living Environment", Annals of Biomedical Engineering, Volume 34, No. 4, pp. 547-563, 2006.

[23] Lara, O.D., and Labrador, M.A., "A Survey on Human Activity Recognition using Wearable Sensors", IEEE Communications Surveys & Tutorials, Volume 15, No. 3, pp. 1192-1209, 2013.

[24] Friedman, A., Chehade, N.H., Chien, C., and Pottie, G., "Estimation of Accelerometer Orientation for Activity Recognition", Proceedings of 34th Annual International Conference of the IFFF EMBS, pp. 2076-2079, San Diego, California, USA, 2012.

[25] Song, W., Ade, C., Broxterman, R., Barstow, T., Nelson, T., and Warren, S., "Activity Recognition in Planetary Navigation Field Tests using Classification Algorithms Applied to Accelerometer Data", Proceedings of 34th Annual International Conference of the IFFF EMBS, pp. 1586-1589, San Diego, California, USA, 2012.

[26] Kaghyan, S., and Sarukanyan, H., "Activity Recognition using K-Nearest Neighbor Algorithm on Smartphone with Tri-Axial Accelerometer", International Journal of Information Models and Analyses, Volume 1, pp. 146-156, 2012.

[27] Huikari, V., Koskimaki, H., Siirtola, P., and Roning, J., "User-Independent Activity Recognition for Industrial Assembly line–Feature vs. Instance Selection", Proceedings of 5th International Conference on Pervasive Computing and Applications, pp. 307-312, Maribor, 2010.

[28] Khan, A.M., Lee, Y.K., Lee, S., and Kim, T.S., "Accelerometer's Position Independent Physical Activity Recognition System for Long-Term Activity Monitoring in the Elderly", Medical & Biological Engineering & Computing, Volume 48, No. 12, pp. 1271-1279, 2010.

[29] Thiemjarus, S., "A Device-Orientation Independent Method for Activity Recognition", Proceedings of International Conference on Body Sensor Networks, pp. 19-23, Singapore, 2010.

[30] Awan, M.A., Guangbin, Z., Kim, H.C., and Kim, S.D., "Subject-Independent Human Activity Recognition using Smartphone Accelerometer with Cloud Support", International Journal of Ad Hoc and Ubiquitous Computing, Volume 20, No. 3, pp. 172-185, 2015.