Intruder Detection System Through Walking Pattern Analysis for Home Security

R.C.D. Peiris[1], R.Tharmikka[2], H.M.V.R. Herath[3], M.B. Dissanayake[4], U.S. Navaratne[5]

Department of Electrical and Electronic Engineering, University of Peradeniya,
Peradeniya, Sri Lanka.

[1]charitha.dhananjaya@eng.pdn.ac.lk, [2]tharmikka10@gmail.com, [3]vijitha@eng.pdn.ac.lk, [4]maheshid@ee.pdn.ac.lk, [5]sudheera@ee.pdn.ac.lk

***Abstract* – Modern home automation systems have home security enhancing features such as face detecting camera systems, and fire alarm systems. In this paper, a novel home security system is proposed which can detect intruders by analyzing walking patterns and update the owner immediately. The system architecture adopts the concept of Internet of Things (IoT), providing a network and user friendly system, which supports simple expansion through plug and play devices. The paper focuses mainly on the intruder detection system and design of the device is presented.**

***Keywords:* IoT, Intruder detection, Classification, Accuracy, Matching probability**

I. INTRODUCTION

Nowadays, people are highly concerned about the security of their homes and properties while they are away from home. The developments in IoT enables home automation using large number of IoT ready intelligent sensors [1]. In home automation CCTV camera systems, fire alarm systems, smart door locks, smart refrigerators, smart TV, and etc. are integrated into IoT base platform and the owner gains access to the system via telecommunication service providers [2]. There exists many mechanisms for intruder detection, for example a CCTV camera system can detect faces and by using intelligent processing determine whether the person is an authorized user or an intruder. In this work intruder detection is done using vibration sensors that capture walking patterns of people. Once the walking patterns of the home dwellers are registered in the system it can classify a person walking in the home as a dweller or an intruder. This information is then conveyed to the owner through his mobile phone as an alarm.

In following sections, the development of the sensor system, intruder detection algorithm and results are discussed. And finally further improvements to the system is also discussed.

II. DEVELOPMENT

When a person walks on a floor it generates vibrations on the surface of the floor. The characteristics of those vibrations on the surface are different from person to person as shown in Fig 1 and 3. In this system those characteristics are used to differentiate each person, hence the identification.

In our system, floor vibrations due to a person's walking is captures and converted to an electrical signal using a piezo plate sensor [3], which is connected to the floor. The piezo plate is covered with silicon adhesive and a weight is kept on it to avoid the surrounding noise from interfering with the actual signal. The sensor is connected to an offset amplifier to amplify the signal and set the fluctuation between 0-5V which is ideal for the analog to digital converter (ADC). To enable the Wi-Fi connectivity, Node-MCU is used and ADS1115 16 bit [4] is used as the ADC extension. For the signal classification and identification a Machine Learning (ML) algorithm is developed. A mobile application is developed to communicate and present the decision to the home user [5]. In order to train the ML algorithm hundred sets of signals are collected from each dweller. In order to test the system fifty sets of signals are used.

A single input signal captured from the piezo plate contains 6 steps in average and is collected from the sensor at a rate of 50 samples/s. The vibration signals from two people are plotted in Fig. 1 and 3. All sampled data is transferred to the main server by Node-MCU for processing. After receiving the data, server will process them for identification. The algorithm has three aspects, preprocessing, feature generation and intruder detection. In the preprocessing stage, digital signal is scaled to voltage levels and filtered in order to reduce the noise using FIR low pass filter.

Feature Generation stage generates the required characteristics from the processed signal. Then using the generated features the developed ML algorithm is trained and eventually tested with the test set. The intruder detecting algorithm is trained by the data collected from home dwellers (in actual scenario) and after the system satisfied the accuracy, it can detect any intruder walking through the surveillance area.

III. INTRUDER DETECTION ALGORITHM

As discussed in section II, the intruder detection algorithm has three parts; namely preprocessing, feature generation, and intruder detection, which operates independently in parallel. In preprocessing stage, the digital values are scaled to voltage values, using sensitivity of the ADC which is 1.25×10^{-4} mV/bit. After that the signal is filtered using FIR low pass filter (LPF) in order to reduce the noise as shown by the Fig. 2 and 4 and block diagram in Fig. 5. At the end of this stage the signal is fed to the feature generation stage. The feature generation stage calculates the energy, mean and standard deviation of the peaks, and time between peaks of the input signal as time domain features [3].

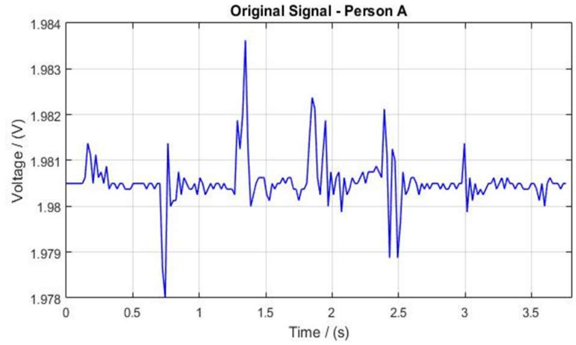

Fig. 1. Original Signal of Person A

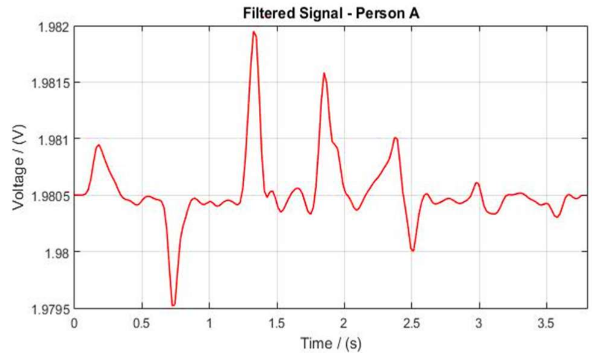

Fig. 2. Filtered Signal of Person A

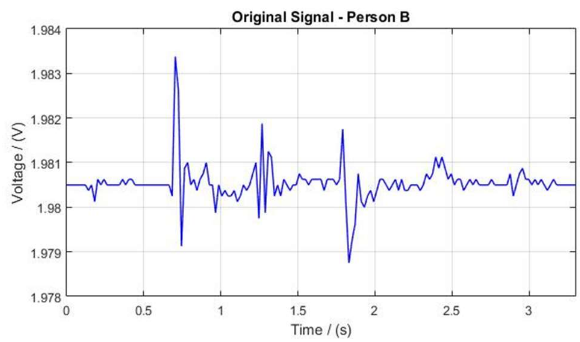

Fig. 3. Original Signal of Person B

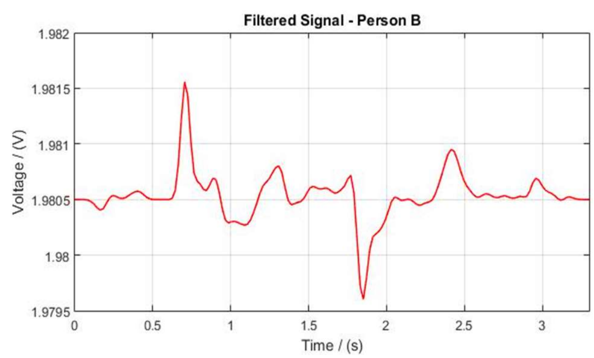

Fig. 4. Filtered Signal of Person B

In frequency domain, the spectral centroid, power spectral density, and amplitudes of peaks are generated [6] as shown by the block diagram in Fig. 6. Then the features are feed into a ML algorithm.

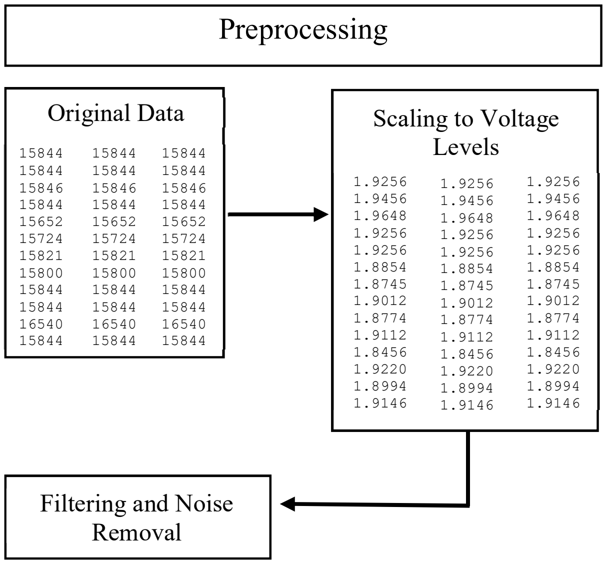

Fig. 5. Block Diagram of Preprocessing Stage

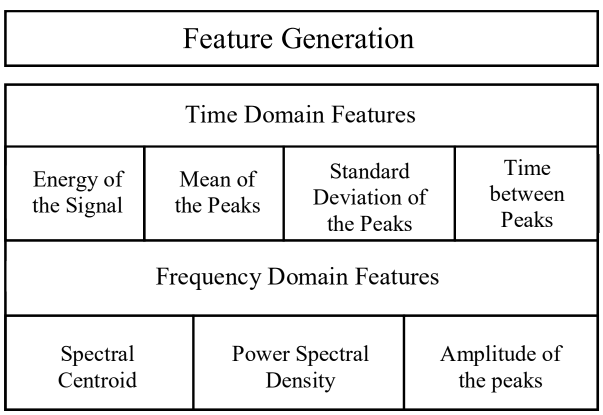

Fig. 6. Block Diagram of Feature Generation Stage

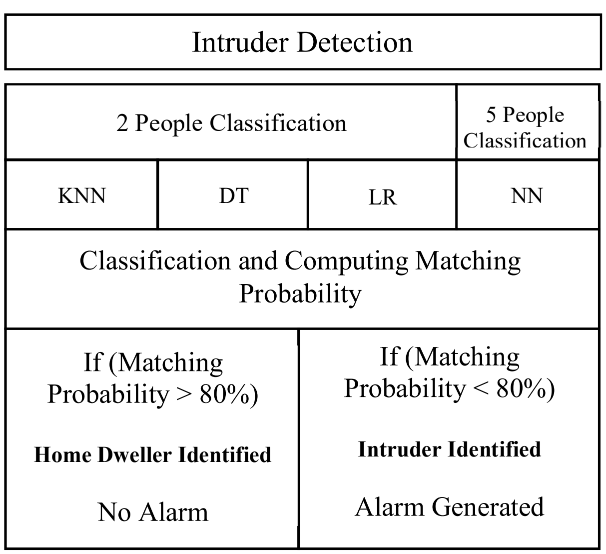

Fig. 7. Block Diagram of Intruder Detection Stage

Basically there are four algorithms implemented in order to get the maximum accuracy and performance at the intruder detection stage. Firstly, the algorithms are trained with two individual dwellers' data. From those data three types of classification algorithms are developed namely; K-Nearest Neighbors (KNN) [7] classification, Decision Tree (DT) [8] classification and Logistic Regression (LR) [9] classification as shown by the block diagram in Fig. 7. As a performance parameter the accuracy of the algorithm is calculated using test data set. Accuracy higher than 95% is considered as an optimized identification algorithm in this work. The accuracy of the algorithms obtained by the successive attempts taken by the algorithm with testing set. When the number of people increases more than two, the accuracy of all KNN, DT and LR methods decrease. To increase the accuracy and the capability of identification process for more than two individuals, a new algorithm was developed using a simple Neural Network (NN) with Multi-Layer Perceptron (MLP) classifier [10], which is a supervised learning algorithm and a basic level of classifiers in NN. In this approach 5 individual persons' data are collected as the data set, and to improve the performance of the algorithm each feature is scaled using standard scaling process. Each feature is set so that they have zero mean and unit variance. This new approach increases the performance of the algorithm than the previous attempt.

After the algorithm optimized for the highest accuracy that can achieve, the matching probability of the incoming signal to the each trained person is calculated. If the incoming signal is from a trained person, the matching probability is always higher than 80% to the respective person. If the signal is from an intruder the matching probability is lower than 80%. By setting the threshold level of the matching probability as 80%, it can be determine whether the person is an intruder or a home dweller. The performance of these algorithms and challenges are discussed in the following sections.

IV. RESULTS

The four algorithms are developed in such a way that they are computationally efficient and achieve highest possible accuracy in real time executions. The performance of the algorithms can be increased by improving the quality of training data set. In this work the algorithms are set to their highest possible accuracy since it need to ensure the security of the home.

According to the table III, results from the KNN, LR and DT methods are similar, when the system is limited to identify two individuals. In KNN method the accuracy of the data set is set to an optimum level by setting the number of neighbors considered for the classification. Table I shows the variation of accuracy of the KNN method with the number of neighbors considered in the application.

In LR method the features are regularized to improve the accuracy of the algorithm. Regularization avoids the overfitting [11] of the data set to the training model. With the regularization of the features the accuracy of the algorithm increases. Table II shows the variation of accuracy of the LR method with respect to the strength of regularization. In DT method also the overfitting affects negatively to the results. Since the importance of the features to the DT algorithm is different, the overfitting is reduced by controlling the maximum depth of the DT algorithm.

The results obtained from NN method justifies the capability of the system to identify multiple number of people through this system. The standard scaling increases the calculation performances and further improvement is achieved by tuning each feature in the preprocessing stage. Table III gives a comparison of the performance of each algorithm with respect to number of people to be identified.

A signal obtained from person F, who is an intruder to the system is also feed into the algorithm. Results obtained are shown in the Table IV. Based on this results, in each occasion the algorithm could identify the person correctly. Since the F's signal's matching probability is not exceeding the 80% threshold, F could be identify as an intruder.

TABLE I
VARIATION OF ACCURACY OF KNN ALGORITHAM WITH NUMBER OF NEIGHBORS

	Number of Neighbors				
	4	5	6	7	8
Accuracy (%)	94	95.6	97.8	95.1	94.6

TABLE II
VARIATION OF ACCURACY OF LR ALGORITHM WITH STRENGHT OF REGULARIZATION

	Strength of the Regularization				
	0.1	1	10	50	100
Accuracy (%)	93	94.1	95.6	96.7	97.9

TABLE III
COMPARISON PERFORMANCE OF EACH ALGORITHM WITH RESPECT TO NUMBER OF PEOPLE TO INDENTIFY

Algorithm	Accuracy (%) w.r.t Number of People				
	1	2	3	4	5
KNN	95.5	95.2	94.8	94.1	93.8
LR	95.6	95.1	94.9	94.2	93.9
DT	95.8	95.6	95.2	94.8	94.1
NN	96.7	96.4	96.5	96.4	96.3

TABLE IV
VARIATION OF MATCHING PROBABILITY WITH RESPECT TO THE INPUT SIGNAL WHEN NN ALGORITHM USED

Trained Person	Variation of Matching Probabilities (%) w.r.t Input Signal					
	A	B	C	D	E	F
A	86.6	0.9	0.4	0.2	1.3	14.8
B	2.8	88.4	0.3	0.8	1.6	9.6
C	6.6	9.7	97.8	11.9	10.6	46.8
D	2.2	0.2	0.9	86.1	1.1	22.6
E	1.8	0.8	0.6	1.0	85.4	6.2

V. CHALLENGES AND IMPROVEMENTS

Although each algorithm work fair enough to identify a person, there are challenges to the performance of the whole system. The main challenge of this system is the dependency of the accuracy and the matching probability upon the mix of data set. That is the accuracy of identifying person A, from the algorithm which is trained from person A and B differs from the algorithm trained from person A and C. This situation occurs with the similarities of the characteristics of the walking patterns of the people. In the above mentioned case person A and C have average body structure and weight, but the person B is much weighs more than A and C, hence the accuracy and matching probability are differed.

In this work the vibration signals are taken from the wooden surface when a person is walking with his regular walking pattern. In the experiment the surface texture and material contribute to the characteristics of the signal. In a concrete floor the attenuation of the signal and the noise from other structural vibrations are also higher than the wooden floor. Since the surface factor is not considered, if the device is placed in a different surface the entire system need to be trained again using data collected from the new test environment. In essence, the system need to be trained for each surface environment, where it is being used.

The identification process can expand analyzing signals from two sensors for a surveillance area as an improvement. From this approach common mode noise can be cancelled out. As an added advantage this method will generate more data for the feature generation. The home based centralized processing unit can expand to work with several sensor modules placed on different places in home and different surfaces.

VI. CONCLUSIONS

In this paper an intruder detection system through walking pattern analysis, which developed as an IoT support system for home automation, is presented. In the proposed system, several ML algorithms were tested for the identification which ensures efficient computation and high accuracy. Due to dependencies and practical limitations, the proper identification achieved by NN algorithm higher than other algorithm tested. The similarities of walking patterns and the physical characteristics of people highly influence the end result of the identification. Since the accuracy of the algorithm achieved more than 90% the method of intruder detection through walking pattern analysis enables a new IoT home security aspect in addition to the current methods.

ACKNOWLEDGEMENTS

This work is supported by Sri Lanka Telecom PLC.

REFERENCES

[1] Mohamed Moubarak, *Internet of Things for Home Automation*, Bachelor Thesis, 2016. [Online]. Available: https://www.researchgate.net/publication/314508660_Internet_of_Things_for_Home_Automation

[2] Surinder Kaur, Rashimi Singh, Neha Khairwal and Pratyk Jain, "Home Automation and Security System" *Advanced Computational Intelligence: An International Journal (ACII)*, Vol.3, No.3, July 2016; [Online]. Available: http://aircconline.com/acii/V3N3/3316acii03.pdf . [Accessed: 10 Mar. 2018]

[3] Shijia Pan, Ningning Wang, Yuqiu Qian, Irem Velibeyoglu, Hae Young Noh and Pei Zhang, *Indoor Person Identification through Footstep Induced Structural Vibration*, in Proceedings of the 16th International Workshop on Mobile Computing Systems and Applications, 2015, Santa Fe, New Maxico,USA.

[4] Texas Instruments, "ADS111x Ultra-Small, Low-Power, I2C-Compatible,860-SPS,16-BitADCs with Internal Reference, Oscillator and Programmable Comparator," ADS1115 datasheet, May 2009 [Revised January 2018]. [Online]. Available: http://www.ti.com/lit/ds/symlink/ads1115.pdf. [Accessed: 11 Jan. 2018]

[5] Ovidu Vermesan and Peter Friess, *Internet of Things-From Research and Innovation to Market Development*, 2014.

[6] Unjung Nam, "Special Area Exam Part II," April 28, 2001. [Online]. Available: https://ccrma.stanford.edu/~unjungs/AIR/areaExam.pdf [Accessed: 21 May 2018].

[7] Documentation of scikit-learn 0.19.2, "sklearn.neighbors.KNeighborsRegressor". [Online]. Available: http://scikit-learn.org/stable/modules/generated/sklearn.neighbors.KNeighborsRegressor.html. [Accessed: 11 June 2018].

[8] Documentation of scikit-learn 0.19.2, "sklearn.tree.DecisionTreeClassifier". [Online]. Available: http://scikit-learn.org/stable/modules/generated/sklearn.tree.DecisionTreeClassifier.html#sklearn.tree.DecisionTreeClassifier. [Accessed: 11 June 2018].

[9] Documentation of scikit-learn 0.19.2, "sklearn.linear_model.LogisticRegression". [Online]. Available: http://scikit-learn.org/stable/modules/generated/sklearn.linear_model.LogisticRegression.html. [Accessed: 11 June 2018].

[10] Documentation of scikit-learn 0.19.2, "sklearn.neural_network.MLPClassifier". [Online]. Available: http://scikit-learn.org/stable/modules/generated/sklearn.neural_network.MLPClassifier.html. [Accessed: 11 June 2018].

[11] Jason Brownlee, "Overfitting and Under fitting with Machine Learning Algorithms," March 21,2016. [Online]. Available: https://machinelearningmastery.com/overfitting-and-underfitting-with-machine-learning-algorithms/. [Accessed: 11 June 2018].