Development of an Electrocardiogram Based Biometric Identification System: A Case Study in the University

Nur Izzati Mohammed Nadzri[1], Khairul Azami Sidek[1], Dedy H.B. Wicaksono[2]
[1]Department of Electrical and Computer Engineering,
International Islamic University Malaysia,
P.O. Box 10, Jalan Gombak, 50728 Kuala Lumpur.
[2]Faculty of Bioscience and Medical Engineering (FBME),
Universiti Teknologi Malaysia,
Building V01, Block A, 5th Floor, Room 05-14-01,
81310 UTM, Skudai, Johor,Malaysia.
izzatinadzri91@gmail.com

***Abstract*—This paper focuses on the electrocardiogram (ECG) based biometric identification system in the university scenario as an alternative to the traditional methods being used nowadays. There are a lot of researches and studies about ECG based biometric system where some of them showed positive result. However, ECG based biometric system in the university scenario is under-researched. Therefore, this issue will be the main focus of our study. A total of five subjects were used for experimentation purposes. A bandpass filter is used to remove unwanted portion of the signal. Unique features are extracted from these filtered ECG signals. Later, Multilayer Perceptron and Naïve Bayes are used to classify the subjects using the discriminant features. Based on the experimentation results, classification accuracies of 90% and 80 % were achieved which suggest the capability of our proposed system to identify individuals. The result provides an alternative mechanism to detect a person besides using the traditional methods.**

***Index Terms*—Electrocardiogram; Multilayer Perceptron; Naïve Bayes**

I. INTRODUCTION

Nowadays, identity theft and fraud are becoming a serious issue to most countries. According to UK's Fraud Prevention Service by National Fraud Database, identity fraud has the highest percentage with 49.8% as compared to the others from the overall case [1] which can be conclude as in Figure 1. While in America, Bob Sullivan the columnist of NBC News reported in his column that 12.6 million of American citizen became a victim of identity theft in 2012.

In another report by the Justice Department's Bureau of Justice Statistics (BJS), identity theft causes financial losses of $24.7 million from 16.6 million people that experienced identity theft in 2012 [2]. Therefore, in order to prevent this problem, a system which recognizes a person based on their behavioural and physical characteristics is needed. The system that fulfills these criteria is called as biometric identification system.

Biometric is used to solve the problem faced by traditional methods which are categorized as token based and password based. A token based method uses a device or material used to identify someone, for example an ID card and smart card. This technique is not secure as the card may be lost, stolen, or duplicated. On the other hand, a password based approach basically requires a person to memorize certain combination of alphanumeric representation. However, passwords can be forgotten, shared, observed and broken. Thus, biometric is an alternative solution to overcome this aforementioned issue. Biometric uses characteristic which are available from our body to identify a person.

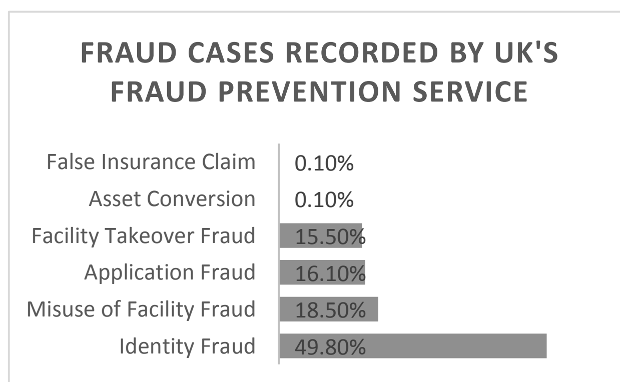

Figure 1: Type of fraud cases recorded by the UK's Fraud Prevention Service in 2012 [1]

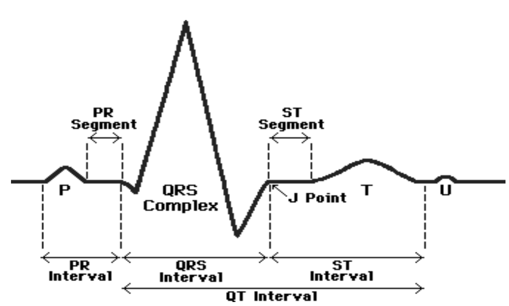

Figure 2: The electrical signal from the heart [3]

Biometric factors which are universal, unique, permanence, collectability and acceptability will help in averting fraud and theft. There are many types of biometric modalities such as fingerprint, deoxyribonucleic acid (DNA), face, retina, iris, ear and keystroke. Access system with fingerprint, face, and iris biometric authentications were widespread but are prone to falsification. Recently, ECG which is another feasible form of identification for access control is getting the attention from researchers. ECG as in Figure 2 is the voltage variation graph, which detects the electrical changes on the skin. It is produced when the heart muscle depolarizes during each heartbeat. ECG is common in medical diagnosis, however, researchers found that ECG can also be used as a person identification mechanism.

In universities, traditional access systems are used. For example, when the students in the faculty come to class, they need to sign the attendance sheet as a proof that they were present to the class. However, other people can sign for their friend that are absent on that day. Therefore, as compared to traditional method, biometric is an alternative solution for this concern. ECG signals is a better substitute for a biometric system as it provides the proof of the subject being present at the time of recognition.

II. RELATED WORKS

In the past, several studies on the development of an ECG based biometric identification system have been conducted. Kyoso et al. proposed an ECG identification system based on the comparison of the subject's ECG with previously documented ECG feature parameters from nine normal subjects. P wave, PQ interval, QRS interval and QT interval were used in the study [4]. The research was divided into parts which are measurement of an ECG signal and AD conversion block. Next, unique features are extracted to accomplish the identification process and lastly discriminant analysis is performed by selecting the minimal Mahalanobis generalized distance to calculate the optimum pattern. Result of the experiment shows the highest accuracy value is 99.6% from QRS interval and PQ interval while the lowest accuracy is 4.1% from the P wave.

Low et al. introduced the PC based ECG monitoring system that acquire ECG signal and transfer the signal to remote terminal through a telemedicine network [5]. In this study, three electrodes were used to get the acquisition data. The first two electrodes were placed at the right and left wrist while the last electrode was placed on the ankle of the leg as the ground. A bandpass filter and a high Q notch filter were used to remove the noise. In this study, a 16 bit DAQ card is used to get the output from a filter to the PC. A graphical user interface (GUI) was developed using Labview to get the resultant waveform. However, the study does not mention about the total number of subjects that were used to test the proposed system and the method of classification. Result of the experiment shows that only the peak R-wave is clearly shown.

Azim et al. presented the development of an annotated ECG database. The objective of this study was to evaluate the performance of ECG processing and data compression algorithm. In order to acquire an annotated database, there were five processing steps that need to be followed. First and foremost, digital filters are used to eliminate the noise. Second step is to detection the QRS complexes. The next step was to extract the features to discriminate ECG samples. Finally, classification of the ECG signals into different categories. For this step, neural network classifier was used for classification. And finally is the annotated process. ECG wave of fifty five subjects were tested and obtained. Result of this study shows the difference between unfiltered and filtered signals using FMH filter which can be observed [6].

Belgacem, et al. presented an ECG based identification mechanism using wavelets and random forest. The purpose of this study was to introduce a novel personal identification approach using ECG signals. A total of 80 individuals for ECG database were acquired from the Physionet database which are (MIT-BIH, ST-T, NSR and PTB). The cardiac signals were used to identify from ECG database. Discrete Wavelet Transform was used for feature extraction which has proven to give a good result for extracting the signals. Finally, Random Forest was used for ECG signal verification. The outcome of this study states that the average training and the verification rate was 100% for healthy subjects whereas the false acceptance and false rejection rates were 0.60% and 0.58% respectively [7].

In a different paper, Belgacem, et al. elaborated the identification system based on ECG signal using Labview which depends on the P, QRS, and T waves and implemented Fast Fourier Transform (FFT) to normalize autocorrelation coefficient. Measurement of the study was performed on the palm for forty normal healthy people for identification and the outcome from this study, ECG signal wave from the palm had more noise as compared to torso but the waveform morphologies were the same as Lead I. However, the signal-to-noise ratio of the palm ECG signal was lower than of the chest ECG signals which result in an improvement of the system performance [8].

These researches were conducted with different approaches and methods using ECG for person identification. To conclude, some of these approaches were able to obtain high accuracy for recognition and the other does not focus on the result for an ECG biometric identification. Therefore, in this study, an ECG based biometric identification system is proposed in an environment that has been less discussed which is the university scenario. The capability of the implementation of such system in identifying student will assist the teaching instructor for attendance purposes.

III. METHODOLOGY

For this study, there are four stages involved to develop an ECG based biometric identification system. The four different stages are data acquisition, preprocessing, feature extraction and classification. Each of these stages will be explained in detail in the next sub-sections.

A. Data Acquisition

In the first stage, the raw ECG data were collected from five individuals consisting of four female and one male student from IIUM and University Technology Malaysia (UTM) in a resting condition. The ECG readings were collected from KL-75001 Electrocardiogram ECG Module that operates with KL-72001 Biomedical Measurement System Module, digital storage oscilloscope and the ECG electrodes.

The electrodes were used as a medium for signal acquisition and the ECG samples were collected when the clamp electrodes are connected to the wrist of the subject in a resting condition. It is connected to the ECG Module which displays the signal waveform on the oscilloscope window. The acquired ECG data are saved in an Excel document spreadsheet. The next sub-chapter will discuss further about the devices used for this study.

1. *KL-75001 Electrocardiogram ECG Module and KL-72001 Biomedical Measurement System Module*

KL-75001 ECG Module and KL-72001 Biomedical Measurement System Module boards are as shown in Figure 3. This modules measure the heartbeat of a person by detecting the ECG signals. Both KL-75001 ECG Module and KL-72001 Biomedical Measurement System Module were connected to the oscilloscope and the electrodes to display the ECG wave shape of a subject.

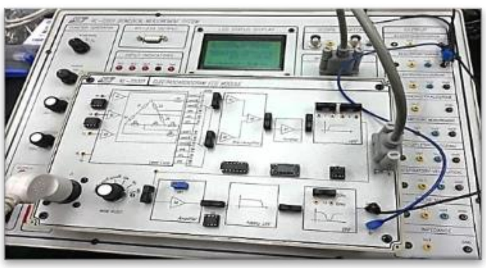

Figure 3: KL-75001 Electrocardiogram ECG Module and KL-72001 Biomedical Measurement System Module board

2. *Digital Storage Oscilloscope TBS 1062*

A digital storage oscilloscope used in this study is called the TBS 1062 Tektronix digital storage oscilloscope as in Figure 4. It is used as a platform to display the ECG signal of the subject. Besides displaying the ECG signal, it also provides USB connectivity which can transfer the ECG signal on a pen drive in an image format. Figure 5 shows the interface of the digital storage oscilloscope TBS 1062 with features of saving either as an image or storing it in the database.

Figure 4: Digital storage oscilloscope TBS 1062

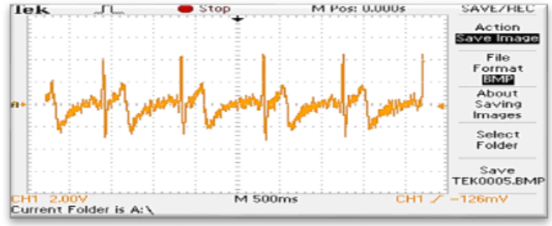

Figure 5: Interface of Digital storage oscilloscope TBS 1062

3. *Clamp Electrodes*

In order to capture ECG signal, clamp electrodes were used in this study. These electrodes consist of three main probes which can be placed at the ankle or wrist of the subject. The examples of the clamp electrodes are shown in Figure 6 whereas the placements of these electrodes on the subject's body are shown in Figure 7.

Figure 6: Clamp electrode

Figure 7: Example for the placement of clamp electrode

B. *Pre-processing*

In the second stage, data preprocessing ensures that there is no external noise, outliers or artifacts in the data. Initially, raw ECG signals which contain noise are collected. The KL-75001 ECG Module applies bandpass filter for preprocessing as shown in Figure 8.

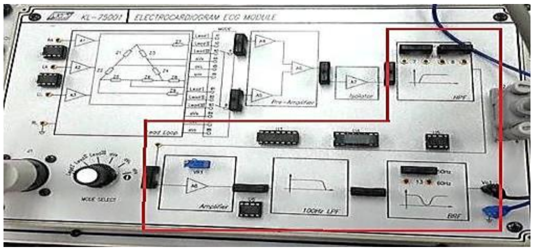

Figure 8: Filtering segment on the KL-75001 ECG Module

Bandpass filter is a filter that allows certain range of frequencies to pass through it and rejects frequencies that is outside the desired range [9]. Figure 9 shows the function block of the ECG module. In the first stage, the clamp electrode will convert the ECG signals into electrical voltage. Then, the amplifier will strengthen the ECG signal. Next, opto-coupler will isolate the in-amp and output. And finally, the signal is filtered using a bandpass filter with a frequency of 0.1 Hz until 100Hz by cascading a low pass and a high pass filter.

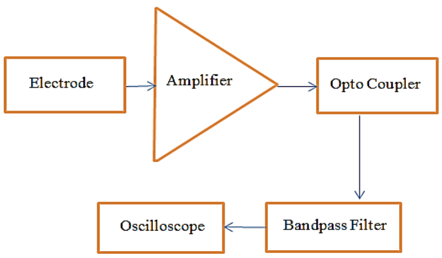

Figure 9: Function block of the KL-75001 ECG Module

C. *Feature Extraction*

The next phase is feature extraction where QRS complex are extracted as unique features. This can be done by analyzing the ECG data using various signal processing techniques. From the previous studies, examples of algorithms used for this stage are Mallat’s algorithm, morphological characteristic, statistical based algorithm, Morlet Meyer, and Daubeachies.

In this study, Pan and Tompkins algorithm will be used to extracts the QRS complexes by analyzing the width, slope and amplitude of the ECG signal. Figure 10 shows an example of a QRS complex which will be extracted and applied in the next stage.

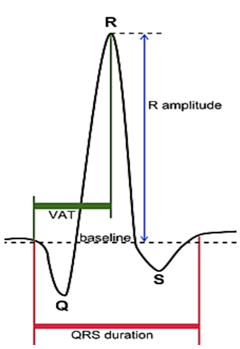

Figure 10: QRS complex of an ECG wave

D. *Classification*

In this stage, ECG data are processed to differentiate individuals using unique and salient features. In the previous study, researchers used various pattern matching method such as Radial Basis Function and Random Forest. In this study, classifiers Multilayer Perceptron (MLP) and Naive Bayes are applied as the techniques that will differentiate between ECG samples to give the best result. At this stage, it involves the decision process which is so important in revealing the identity of individuals. These classifiers will be briefly explained in the next sub-sections.

1. *Multilayer perceptron*

MLP or Neural Network consists of more than one layer. The first layer is the input layer, which does not perform any computational task. Then, there are one or more hidden layers and an output layer, all composed by computational nodes. In MLP network, the nodes from a layer are linked with every node from the earlier and from the subsequent layer. There are no connections between nodes in the same layer or connections between nodes on non-adjacent layers. The non-computational nodes in the input layer uses an identity function, while the computation nodes in the intermediate and the output layers uses a sigmoid function [10]. The MLP network can be presented as in Figure 11.

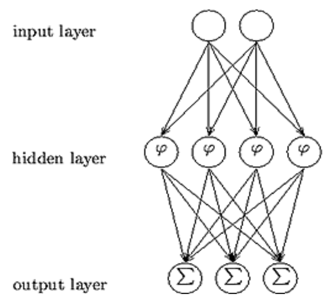

Figure 11: MLP Network

The computational of MLP classifier can be describe as in Equation 1.

$$y = \varphi\left(\sum_{i=1}^{n} \omega_i x_i + b \right) = \varphi(w^T X + b) \tag{1}$$

where:

φ = activation function
w = the vector of weights
x = the vector of inputs
b = the bias

2. *Naive Bayes*

Naive Bayes (NB) uses probability for classification that applies Bayes theorem especially when the dimensionality of the inputs is high. According to Soria et al. (2008), NB classifier has two assumptions which are:

i. The predictive attributes are conditionally independent given the class.
ii. The values of numeric attributes are normally distributed within each class.

The computational equation of NB classifier can be described as follow:

$$p\ (X = x \mid C = c) = \prod_i p(X_i = \ x_i \mid C = c) \qquad (2)$$

where:

C = random variable for the class of an instance
X = variables for the attribute
c = class label
x = attribute value

IV. EXPERIMENTATION AND RESULTS

For this study, the experiment has been performed using five subjects which consist of four females and one male. The mean ages of the subjects are 23 years old. The subjects are under resting condition without any history of heart abnormalities. From each subject, two QRS complexes are extracted where one is used for training and the other is used for testing. After acquiring the ECG samples, preprocessing stage was implemented to remove unwanted noise from the signals. Then, unique features were extracted to represent individuals. Later, classification was performed to recognize the subjects. These processes are further described in the following subsections.

A. Data Acquisition and Preprocessing stages

The first two stages of the proposed ECG based biometric system were experimented in Universiti Teknologi Malaysia (UTM), Johor using their bio-signal data acquisition device. The data acquisition stage was implemented via the KL-75001 ECG Module that works with the KL-72001 Biomedical Measurement System Module, digital storage oscilloscope and the ECG electrodes. Then, data preprocessing which uses a bandpass filter was performed to remove artifacts in the signals. This experiment was under the observation of our supervisor and a qualified lab technician. Figures 12 to 14 shows the ECG signals of the threes subjects which are subject 001, 004, and 005.

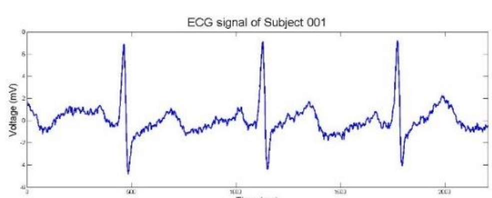

Figure 12: ECG signal of Subject 001

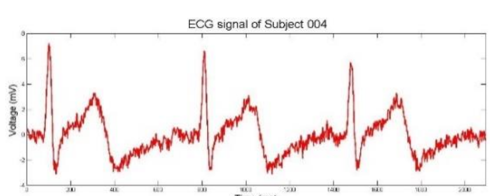

Figure 13: ECG signal of Subject 004

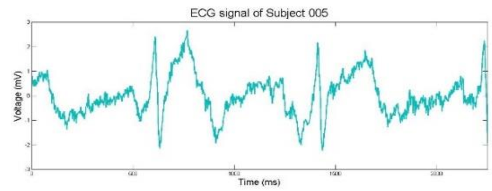

Figure 14: ECG signal of Subject 005

B. Feature Extraction stage

After collecting and filtering the ECG samples, QRS complexes are extracted from the signal which acts as unique features. In this study, a total of 51 fiducial points were used because it is the minimum number of points which forms the QRS complex. Figures 15 to 17 shows the extracted QRS signal from each of the subject.

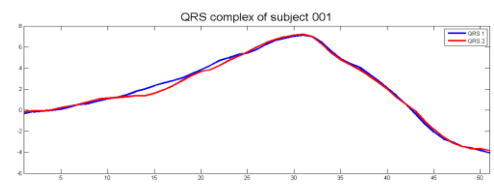

Figure 15: QRS complex of Subject 001

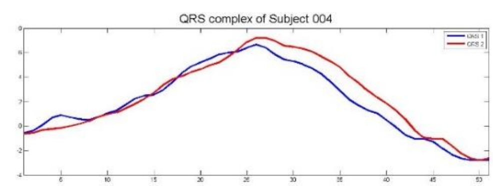

Figure 16: QRS complex of Subject 004

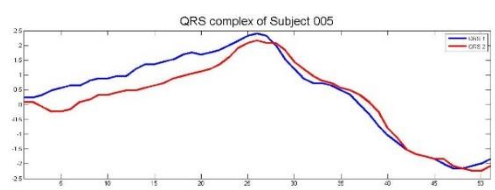

Figure 87: QRS complex of Subject 005

From the observation of the study, each ECG signal shows a different pattern between one and another which suggests ECG biometric recognition is possible. For example, the QRS complexes of Subject 001 are different from Subject 005. Besides the pattern, each ECG signal also has different values of Q, R and S which proves that every individual has a unique feature of an ECG. For instance, the values of Q wave for Subject 002 are 4.56 and 3.28 whereas for Subject 004, the Q wave values are 7.2 and 6.64.

C. Classification Stage

Then, the final stage of ECG based biometric system is classification. This step uses data mining software called Weka. It is an application that consists of a collection of algorithms. Data pre-processing, classification, clustering, association rules and visualization are the applications that can be found in Weka. As stated in the previous section, MLP and

NB are used in the study. The result of the classification techniques is shown as in Table 1.

Table 1: The classification accuracies of MLP and NB

Classifier	Accuracy (%)
MLP	90
NB	80

From the results, classification accuracy of 90% and 80% were achieved when using MLP and NB respectively. The outcome using MLP gives better output because it has these advantages [11]:

i. Dominant for non-linear classifier
ii. Sophisticated solutions built of continuous basis-functions
iii. Handles noisy data well
iv. Run faster

In addition, the result suggest that ECG based biometric identification system is robust across a neural network classifier and a non-neural network classifier attaining high accuracy rates.

Moreover, the high amplitude value of the QRS complexes in female subjects as compared to the male subject contributes to a high identification rate. In other words, biometric features for females are more unique and discriminant.

Therefore, as a proof of concept, it is proven that ECG based biometric system in the university scenario is feasible to be used in higher institution which suggest high identification rate.

V. Conclusion

As a conclusion, the study achieved the research objectives. Generally, it can be said that ECG based biometric system for university scenario is a possible way to substitute the traditional method. All four stages have been completed and the processes of an ECG based biometric system have been understood. Based on the experimentation results, classification accuracies of 90% and 80% were achieved using Multilayer Perceptron and Naïve Bayes that suggest the capability of our proposed system to identify individuals. The result provides an alternative mechanism to detect a person besides using the traditional methods in higher learning institutions.

References

[1] CIFAS The UK’s Fraud Prevention Service, Fraudscape Depicting The UK’s Fraud Landscape. (Access on September 2014). https://www.cifas.org.uk/research

[2] Harrell. E, Langton. L, “Victims of Identity Theft, 2012,” U.S. Department of Justice, Office of Justice Programs, Bureau of Justice Statistics, pp.1-26, December 2013

[3] John Rondoni, (2002) *Paroxysmal Atrrial Fibrillation and Electrocardiogram Predictors*. MAS.622J Fall 2002. http://courses.media.mit.edu/2002fall/mas622j/proj/students/rondoni/

[4] Kyoso. M and Uchiyama. A. (2001) *Development of an ECG identification system*. Engineering in Medicine and Biology Society, 2001. Proceedings of the 23rd Annual International Conference of the IEEE, 4, pp. 3721-3723

[5] Low, Y.F., Mustaffa I.B., N.B.M., and Hamidon, A.H. (27-28, June 2006) Development of PC-Based ECG Monitoring System. Research and Development, 2006. SCOReD 2006. 4th Student Conference pp.66-69.

[6] Azim.M.A., Abou-Chadi, F>E>Z, Bakr, H.M. and Soliman,H.H., (2001) Design and development of an annotated ECG database. Radio Science Conference, 2001. NRSC 2001. Proceedings of the Eighteenth National , 2, pp.661-665.

[7] Belgacem, N., Ali, N., Fournier, R., and Reguig, F.B. (June, 2012) *ECG Based Human Authentication Using Wavelet and Random Forest.* International Journal on Cryptography and Information Security (IJCIS), vol. 2 (2).

[8] Belgacem, N., and Reguig, F, B. (6, June 2012) *Person Identification System Based on Electrocardiogram Signal Using LabVIEW*. International Journal on Computer Science and Engineering (IJCSE), vol. 4 (6).

[9] Lawrence J.C., and Terry. J.F. (July 2011). The Band Pass Filter. pp. 1-28.

[10] Bishop, C. M. (1995) Neural Networks for Pattern Recognition, Oxford, New York.

[11] Crnkovi, L. (July 2010). Classifier Showdown. Retrieved May, 2015 from http://blog.peltarion.com/2006/07/10/classifier-showdown/.