Hardware Design, Development and Evaluation of a Pressure-based Typing Biometrics Authentication System

Wasil Elsadig Eltahir
Mechatronics Dept,
International Islamic University Malaysia (IIUM),
Jalan Gombak, 53100
Kuala Lumpur. Malaysia
d_jwasel@hotmail.com

W. K. Lai
MIMOS Technology Research Group
lai@mimos.com

Ahmad Faris Ismail
Faculty of Engineering,
International Islamic University Malaysia (IIUM),
Jalan Gombak, 53100
Kuala Lumpur. Malaysia
faris@iiu.edu.my

Momoh Jimoh Salami
Faculty of Engineering,
International Islamic University Malaysia (IIUM), Jalan Gombak,
53100 Kuala Lumpur.
Malaysia
momoh@iiu.edu.my

Abstract

The hardware design of a pressure based typing biometrics authentication system (BAS) is discussed in this paper. The dynamic keystroke is represented by its time duration (Δt) and force (F) applied to constitute a waveform, which when concatenated compose a complete pattern for the entered password. Hardware design is the first part in designing the complete pressure-based typing (BAS) in order to ensure that the best data to represent the keystroke pattern of the user is captured. The system has been designed using LabVIEW software. Several data preprocessing techniques have been used to improve the acquired waveforms. An experiment was conducted to show the validity of the design in representing keystroke dynamics and preliminary results have shown that the designed system can successfully capture password patterns.

1. Introduction

In the strictest sense, biometrics refers to the application of a statistical analysis to biological data and phenomena. The security community, however, widely uses the term to describe technologies for personal identity verification. Biometric devices fall into two categories: those that use physical characteristics, such as fingerprints and hand geometry, and those that use behavioral characteristics, such as signature dynamics and keystroke dynamics. Although a variety of authentication, devices to verify a user identity are in use, password mechanisms have been and probably will remain the preferred method. Password authentication is an inexpensive, familiar paradigm that most operating systems support. Confidence in its ability to provide adequate authentication is, however, waning. This is largely due to the wrongful use of passwords by many users, methods are needed, therefore, to extend and enhance the life of password techniques. One can develop a methodology that improves security by using typing biometrics to reinforce password-authentication mechanisms. Typing biometrics is the analysis of a user keystroke patterns. Each user has a unique way of using the keyboard to enter a password; for example, each user types the characters that constitute the password at different speeds. Fuzzy logic, SVM or Neuro-fuzzy can be implemented to measure the user typing biometrics. This reinforcement is transparent—indiscernible to the users while they are entering the normal authentication information (user ID and password), [1, 4, 5 and 10].

This paper discusses the hardware design of BAS, in such a system it is necessary to determine the dynamics of keystroke typing. One have to identify the parameters of importance that can constitute a convenient pattern or dynamic signature of the user. These parameters can be identified by:

1. The duration of pressing a key, in other words the amount of time a user takes to press and release when typing (Δ t).
2. The amount of force exerted on each button pressed (F).

The application of force (F) over duration of time (Δ t) produces a wave pattern which could be recognized as the typing-template for a sequence of keys pressed.

In [1], a good algorithm for the password authentication is outlined. Here, when a new user requests access to the computer system, or when an existing user password is to expire, the access-control system asks the user to type in the user ID and a new password. The system then asks the user to reenter the user ID and password to verify the previous inputs. Based on the typing patterns displayed on entering and reentering the

information, the typing-biometrics methodology computes a typing template for the user. The access-control system then saves the user identification with the associated template, along with the normal user ID and password pair. On subsequent attempts to access the system, the user goes through the normal password-authentication procedure that is, entering the user ID and password. At the same time, the system monitors the user typing patterns and computes a typing template based on the user ID and password just entered. It then compares this template with the template previously determined for this user. If the new password and typing template match those saved in the database, the system grants access to the user. However, if the password does not match, the normal password-authentication mechanism (without consulting the biometrics component) will reject the user or ask the user to reenter the authentication information. If the password does match, BAS will provide a supporting recommendation that verifies that the user is legitimate.

If the user ID and password are correct, but the new typing template does not match the reference template, the security system has several options, which will be devised accordingly. A typical scenario might be that BAS advises a security or net-work administrator that the typing pattern for a newly entered user ID and password is not what the system expects it to be and that a security breach might be possible. The security administrator then closely monitors the session to ensure that the user does nothing he or she is not authorized to do. A practical situation applies on ATM subscribers where if the user password and pattern don't match restriction is placed on the amount of money that can be withdrawn [1].

The hardware design of BAS has been carried out to achieve desired results with minimal hardware component utilization. The DAQ capture system constitutes of sensors, DAQ hardware, interface card and signal conditioning circuit in addition to the PC unit.

The biometric sensors of the system are force sensors which are needed to give a precise inspection of the amount of force that each user applies while typing. Force sensors can also provide more specific information than just the force applied, hence using data acquisition and analysis one can examine each signal pattern and monitor the time duration (Δt), this will enhance the systems sensitivity and precision.

Flow chart of BAS has been reported in [1]. An important issue to discuss here is the learning stage that the user undergoes when creating a new password.

Initially the system assumes that a user typing undergoes gradual learning curve. The system adjusts the reference typing template for a user when the newly displayed pattern closely resembles the previously saved template. His or her reference typing template has to be updated to reflect the current typing pattern. When the newly typed pattern substantially deviates from the saved template, the system assumes that the user is an impostor, and will not adjust the saved template. However, soon it becomes clear that simultaneously matching patterns and learning new patterns is very difficult, if not impossible, hence it is recommended to separate learning and matching for a prototype system. When a user registers new authentication information by entering the user ID and password, the system monitors when the user pattern stabilizes into a recognizable pattern. The user must therefore repeatedly enter the password until the pattern becomes recognizable. The number of retries before the pattern stabilizes varies from 10 to 20.

The hardware design is divided into two main sections in this paper:

1. The drive and filter circuit, these are external to the main PC unit.
2. The Front panel and block diagram of the LabVIEW VI (virtual instrument).

In the first section the amplifier circuit and the corresponding filter were designed. The sensors were attached to the drive circuit and then from the drive circuit to the A/D interface card.

The overall system block diagram is shown in Figure 1.

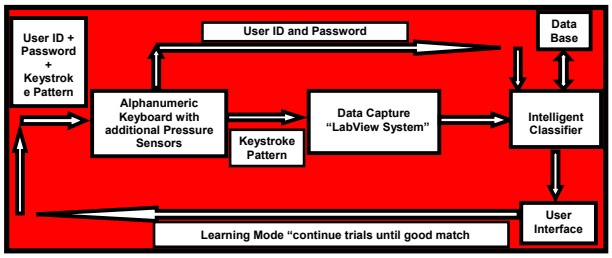

Figure 1. Overall system Block diagram.

Two main authentication frames are emphasized during the overall design of the BAS:

1. The alphanumeric password representing the normal password entered by the user, it consists of alphanumerical combination created by the user or the system administrator and saved by the system.
2. The typing biometrics associated with the user "typing template", this is the second authentication frame needed to study the typing pattern of the user and classify it according to certain parameters associated with typing patterns.

The analysis of BAS is the focus of this paper. However, the overall system as depicted by Figure 1 is currently

being integrated and tested. The overall system will be discussed separately in the future.

2. Hardware design

2.1 Force sensors arrangement

The sensor is an ultra-thin (0.005") flexible printed circuit. It is 0.55" (14 mm) wide and 8" (203 mm) in length. The active sensing area is a 0.375" diameter circle at the end of the sensor. The sensors are attached to the bottom (below) each keyboard key, a plastic spring is fixed on the sensing area of the sensor to insure that the key does not push on the sensing area and cause error pulse. This fabrication allows the template to have a normal distribution for the voltage waveform for each pulse (key pressed), the waveform pulse will hence have a shape similar to half a sine wave cycle.

2.2 System design in LabVIEW

The DAQ system has been designed with LabVIEW based on the following specifications:

1. The data captured must be saved continuously into a file format that would be transferred to the database and the intelligent classifier.
2. The system must have an independent trigger to start the acquisition

The main objective is to continuously acquire and save data to text file readable by spreadsheet programs. Each row is a scan and each column is a channel; columns are separated by commas and rows by an end-of-line character.

The system uses the circular buffer technique for data acquisition whereby data is continuously acquired into a circular acquisition buffer at the same time that the VI reads the acquired data and processes it. After the program creates the file, it initializes and starts the acquisition. It converts the scan rate to a string ending with an end of line character and writes it as the first line in the file. In each iteration, it reads "number of scans to write at a time" scans from the acquisition buffer, converts the data to a spreadsheet string, and writes it to the file. When one presses the STOP button, the program stops the acquisition and closes the file. The scan backlog indicates how much data remains in the buffer after each retrieval, and is an indication of how well the application is keeping up with the acquisition rate. If the backlog increases with time, that means scanning is too fast and will eventually overwrite the circular buffer [8].

There are inputs for setting the channels, size of the circular buffer, scan rate, and the number of samples to retrieve from the circular buffer each time. For example if a VI has an input buffer size of 2,000 samples and 1,000 number of scans to read at a time, which means the VI reads in half of the buffer's data while the VI fills the second half of the buffer with new data.

Analog trigger signal is connected to one analog input channel. The DAQ device monitors the analog trigger channel until trigger conditions are met. The DAQ hardware has been configured in LabVIEW to begin taking data when the incoming signal is on the rising slope and when the amplitude reaches 1.2V, signal is initiated by a special button on the keyboard.

2.3 Signal preprocessing

The first stage of signal preprocessing is to pass the acquired waveform over a lowpass filter which eliminates any undesirable high frequency components or noise before the sampler and ADC.

After applying the filter one can observe in (Figure 2 and Figure 3) the elimination of noise and the improvement in the waveform signal.

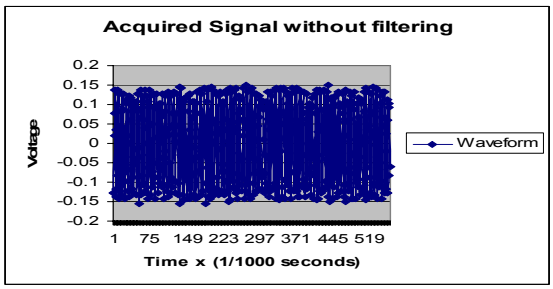

Figure 2. Acquired signal without filtering.

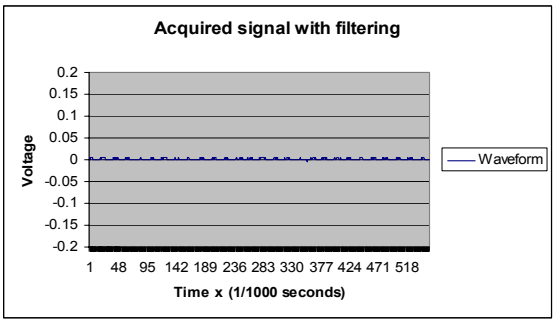

Figure 3. Acquired signal with filtering.

3. Experimental setup

According to the system specifications, the data captured for the password pattern should comply with the type of classifier to be used; the classifier in this case is

SVM (support vector machine). Here are the main criteria for the experiment.

1. The experiment aims at proving the consistency of the design in capturing, representing and analyzing the keystroke pattern of a certain password.
2. The keystroke waveform pattern for each individual keyboard button pressed should be appended (attached) to make up one continuous waveform for all the buttons together; this continuous waveform is the dynamic keystroke pattern that will be passed to the classifier. Appending the waveforms can be done externally (before entering the DAQ terminals) or internally (with the LabVIEW software). The results of both external and internal appending were compared, and it was found that external appending is better. Waveforms were appended externally by superimposing them to one channel in a parallel connection, by this way it is possible to save extra processing on the LabVIEW software.
3. The number of data points for each time a user enters his/her password should be consistent, otherwise the classifier will arbitrarily insert small values (or zeros) as replacement for the missing points (to make up the maximum number of points which is 540). It is required to collect only a fixed number of data points for each key pressed or find a way to concise the graph and reduce the number of data points without altering the trend of the data points acquired. This was achieved by applying the moving average algorithm with interpolation in the LabVIEW data processing. The consistency of such approach is shown in the results.

The experiment was done by asking a volunteer to key in a 7 digit password and verifying whether one can constitute a recognizable pattern for the password entered. The experiment showed that password pattern has a very high degree of precision and sensitivity, making it to have the following advantages and disadvantages.
The advantages are:

1. Data is more accurate and more descriptive to keystroke dynamics of the user.
2. Easiness in identifying different patterns of different users.
3. More into fulfilling the real-time dynamic response of password pattern.

While the disadvantages are:

1. Patterns for a certain user can never match completely; this puts a burden on the classifier module.
2. The effect of external factors on the user (his emotional state, his physical position while typing) can greatly affect the keystroke pattern typed

For these reasons one should adjust the sensitivity of DAQ system, the aim is to find the optimal sensitivity for the system. Sensitivity can be adjusted by changing the amplifier gain or by changing the LabVIEW waveform parameters, such as the sampling rate or the numbers of scans to read at a time.

In the experiment results one can notice the effects of using the moving average algorithm in compressing keystroke patterns without altering the trend of the waveforms.

3.1 Data acquisition and processing

The moving average is a simple mathematical technique used primarily to eliminate aberration and reveal the real trend in a collection of data points.

$$y_M[n] = \frac{1}{M}\left(\sum_{k=0}^{M-1} x[n-k]\right) \qquad (1)$$

The algorithm was written so as to make average for (M) number of points and represent them with one point, for convenience an interpolation algorithm has been applied, the interpolation was done to get the average between each two consecutive data points, the scope was to reduce the data points without distorting or losing any significant information in the signal. Simple test on the algorithm have shown that this can improve patterns [9].

1. Data before applying the moving average algorithm

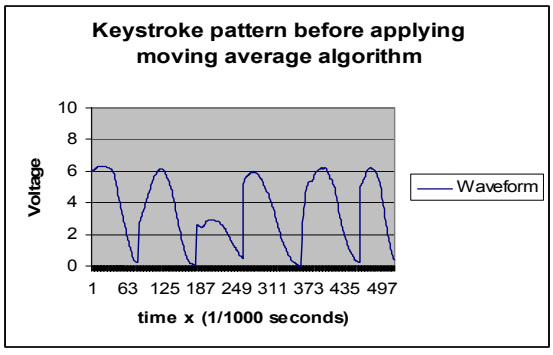

Figure 4. Keystroke pattern without moving average and interpolation.

As one can see the number of points is around 400 to 500 points for each password pattern, the error (difference in number of points for successive password trials) will be high (20-100 points) which is not recommended for the SVM classifier. Therefore it is mandatory to reduce the number of points representing each pattern in order to comply with design specifications

2. Data after applying the moving average algorithm.

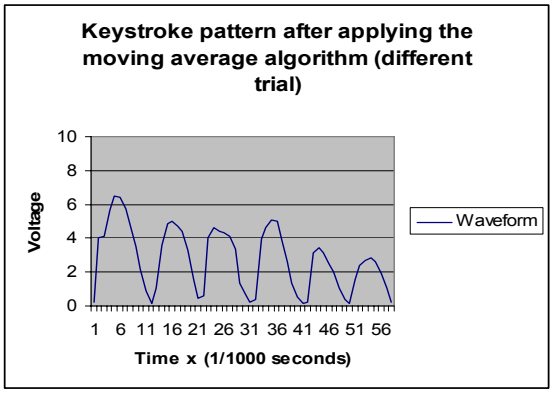

Figure 5. Keystroke pattern with Moving average and interpolation.

It is observed that the 5 moving average has greatly reduced the number points without altering the overall shape and trend of the keystroke pattern, in the 7 digit experiment it will shown that the error in that case was only 5 data points for all trials in the experiment.

3.2 Experiment results and discussion

The experimental results for keying a 7 digit password for one user is shown in figure 6, 7 and 8 respectively.

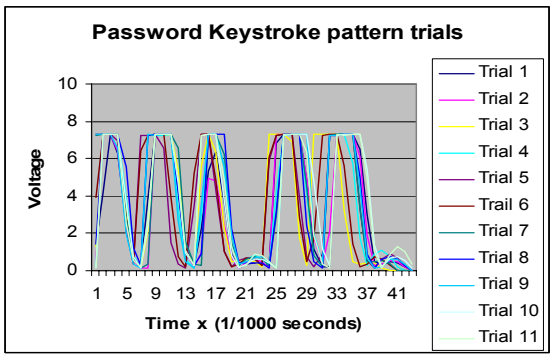

Figure 6. Keystroke pattern trials for single user and 7 digit password.

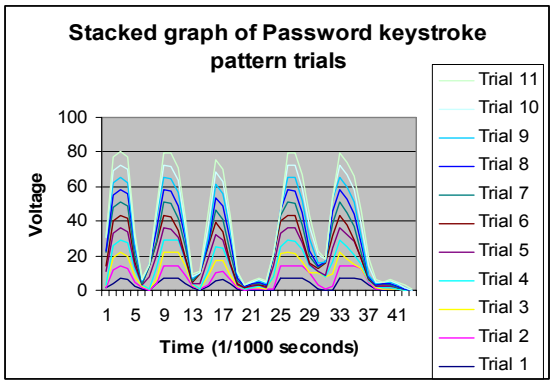

Figure 7. Keystroke pattern trials for single user and 7 digit passwords (stacked graph).

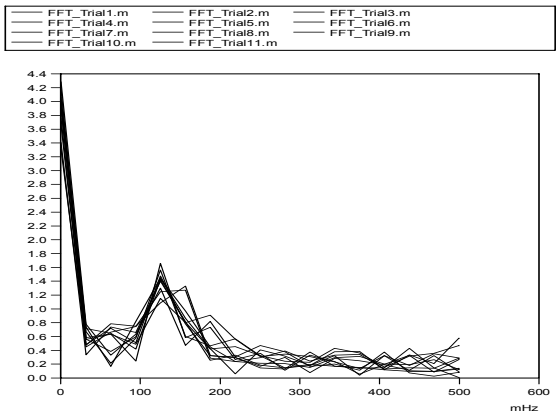

Figure 8. FFT of keystroke pattern trials for single user and 7 digit password.

In Figure 6, it is evident that the user is following distinct stroking rhythms, a general pattern is easily recognizable from the graph. This proves that the hardware has been successful in capturing a recognizable

pattern for the password, justifying that the approach of designing the dynamic keystroke pattern is correct.

The difference in the number of data points for each password pattern is only 5 for all 11 trials, this is confirmation with the SVM requirements of the system.

Figure 7 shows the stacked graph for the 7 digit password trials, this graphs is more descriptive to the degree of compatibility between each password trial, it is clear that there is a lot of similarity between keystroke patterns.

It is also possible to examine the frequency analysis of the concatenated waveform. This can be achieved by computing the discrete Fourier transform (DFT) of $y[n]$ using the FFT algorithm. This would produce the energy spectral density (ESD) of $y[n]$ which is directly related to the pattern of keystroke typing (it is expected that different users should produce different ESD). However, information about the latency is lost due to the procedure of acquiring the data.

Figure 8 shows the ESD graph for the 7 digit password trials. FFT algorithm calculates the discrete spectrum of a particular waveform. from the graph it is noticed that the ESD for the patterns are very similar, it is also noticed that the FFT estimates of the keystroke patterns are very similar too.

4. Conclusions and recommendations

This paper is aimed at proving the consistency of the hardware design for BAS. The approach was to represent the keystroke by its force and duration so as to constitute a waveform pulse, the addition of several waveforms for several keystrokes (appending) makes the password pattern of the user.

The experiment was conducted to prove that this approach is reliable and that the overall design for the BAS is feasible and valid. This result gives the green light to proceed to the second stage of the design which involves the classifier testing and the construction of the keystroke pattern database.

The second stage will emphasize on the inference options that can be implemented to classify the keystroke pattern templates. Currently SVM classifier (support vector machines) is being implemented and tested for; other classifier techniques such as Neuro-fuzzy and ARMA (auto regressive moving average) are going to be tested as well.

For several password trials, the degree of matching between each keystroke pattern entered can vary. However, the general keystroke harmonics preserves this variation within a limited range. In the design one has to emphasize on ways to increase the matching degree of patterns for a distinct user, while passwords for different users must not match. This is somehow related to the sensitivity of the hardware when measuring the voltage waveform; it was experimented that if one increases the sensitivity the waveforms of keystroke patterns will reach the threshold voltage for most of the keys pressed when typing a certain password, this is not desired as it increases the possibility of match between patterns of different users. To solve this problem one have to attain the optimal settings for the sensitivity which can only be achieved by conducting more random experiments for many users and passwords.

The FFT was used to prove the analogy between the ESD of keystroke patterns; this is a good criteria to include in further analysis of pattern classification.

References

[1] Willem G. de Ru, Jan H. P. Eloff, "Enhanced Password Authentication through Fuzzy Logic", IEEE Expert 12(6): 38-45 (1997).

[2] P. Conn, J.H. Parodi and M. Taylor, "The Place of Biometrics in a User authentication Taxonomy," Proc. 13th Nat'l Computer Security Conf., Nat'l Inst. Standards and Technology/Nat'l Computer Security Center, Gaithersburg, Md.,1990.

[3] Sajjad Haider, Ahmed Abbas and Abbas K.Zaidi, "A multi-technique Approach for User Identification through Keystroke Dynamics", Proc.IEEE SMC Society Conference 2000. ,8 Oct 2000.

[4] D.L. Jobusch and A.E. Oldehoeft, "A Survey of Password Mechanisms, Weaknesses and Potential Improvements, Part 1," Computers & Security, Vol. 8, 1989, pp. 587–604.

[5] W.G. de Ru and J.H.P. Eloff, "Improved Password Mechanisms through Expert System Technology," Proc. Ninth Ann. Computer Security Applications Conf., IEEE Computer Society Press, Los Alamitos, Calif., 1993, pp. 272–280.

[6] D. Russell and G.T. Gangemi, Sr., "Computer Security Basics", O'Reilly and Associates, Sebastopol, Calif., 1991.

[7] O. Coltell, G. Torres, and J. M. Badía, "Biometric Identification System Based in Keyboard Filtering". Proc. of 33rd Annual 1999 International Carnahan Conference on Security Technology, IEEE Publisihng Services, pp. 203-209, Piscataway (USA), 1999.

[8] "LabVIEW measurement manual", National Instrument, July 2000 edition.

[9] James H.Mcllellan, Ronald W.Schafer and Mark A. Yoder, "DSP First a multimedia approach", Prentice-Hall, international edition 1998, PP 119-155.

[10] Brigitte Wirtz, "Biometric systems 101 and beyond, an introduction to and evaluation of the technology and an overview on current issues", Infineon Technologies AG.