# The impact of using different keyboards on free-text keystroke dynamics authentication for Arabic language



Suliman A. Alsuhibany, Muna Almushyti, Noorah Alghasham and Fatimah Alkhudhayr
*College of Computer, Qassim University, Buridah, Saudi Arabia*

Received 2 September 2017
Revised 31 January 2018
29 September 2018
Accepted 29 September 2018

## Abstract

**Purpose** – Nowadays, there is a high demand for online services and applications. However, there is a challenge to keep these applications secured by applying different methods rather than using the traditional approaches such as passwords and usernames. Keystroke dynamics is one of the alternative authentication methods that provide high level of security in which the used keyboard plays an important role in the recognition accuracy. To guarantee the robustness of a system in different practical situations, there is a need to examine how much the performance of the system is affected by changing the keyboard layout. This paper aims to investigate the impact of using different keyboards on the recognition accuracy for Arabic free-text typing.

**Design/methodology/approach** – To evaluate how much the performance of the system is affected by changing the keyboard layout, an experimental study is conducted by using two different keyboards which are a Mac's keyboard and an HP's keyboard.

**Findings** – By using the Mac's keyboard, the results showed that the false rejection rate (FRR) was 0.20, whilst the false acceptance rate (FAR) was 0.44. However, these values have changed when using the HP's keyboard where the FRR was equal to 0.08 and the FAR was equal to 0.60.

**Research limitations/implications** – The number of participants in the experiment, as the authors were targeting much more participants.

**Originality/value** – These results showed for the first time the impact of the keyboards on the system's performance regarding the recognition accuracy when using Arabic free-text.

**Keywords** Information security, Arabic language, Keystroke dynamics, Authentication, Keyboard-layout

**Paper type** Research paper

## 1. Introduction

Authentication is a verification process of the identity of an entity that is requesting access to a system (Karnan *et al.*, 2011). Internet businesses and many other systems require a more stringent authentication process. With requests for trusted authentication ever increasing in many of today's security applications, traditional methods such as username and password fail to meet today's challenges. The drawback of these traditional methods is that they afford easy ways to hack the system, which damages the overall system security (Deng and Zhong, 2013).

The authors would like to thank Qassim University for supporting this research. Also, the authors would like to thank all participants who took part in this study, and to anonymous reviewers for helpful comments.



2056-4961
DOI 10.1108/ICS-09-2017-0062

Biometrics authentication essentially depends on physiological data, such as fingerprint, palm veins, face recognition, DNA, palm print, hand geometry and iris recognition. It can also utilise a person's behavioural data, such as keystroke dynamics, gait analysis, voice ID, mouse use characteristics, signature analysis and cognitive biometrics. Using a person's biometrics provides a significant security improvement to face such challenges, as compared with the traditional methods (Deng and Zhong, 2013). Keystroke dynamics is a behavioural biometric that provides a natural choice for secure access. It is understood in Monrose and Rubin (1997) as "not what you type, but how you type". That is, it authenticates the user based on the keyboard typing pattern. This behavioural biometric, therefore, provides a high level of usability whilst maintaining strong system protection (Alsultan and Warwick, 2013a, 2013b).

Keystroke dynamics is concerned with the way that a user types characters on a keyboard or keypad. Each user has a unique typing manner that differs from others, and it can be measured to generate a distinct pattern that can be used in the authentication process (Chang *et al.*, 2011). The measurements of an individual user from every keystroke can be recorded to determine the timing features. These timing features are: keystroke duration (the time when a key is depressed until it is released), keystroke latencies (the time between two consecutive keys presses or releases), di-graph duration (the time between releasing the first key and pressing the next key) and down-up key time duration (Alsuhibany *et al.*, 2016). After recording, these timing data will be processed through different methods or algorithms to identify users (Chang *et al.*, 2011). This behavioural biometric offers several key advantages. First, capturing data from users is an easy task because computer users are typing often for both work and pleasure. Second, it is not costly because the only hardware needed is a computer with a keyboard (Tappert *et al.*, 2012). Third, the keystroke biometric can provide an effective balance between the aspects of usability and security. Finally, even after verifying a user's identity, keystrokes continue to be entered for potential subsequent checking because they exist as a simple consequence of users making use of their computer (Gunetti and Picardi, 2005).

To date, a keystroke dynamic has two main classes: free-text and fixed-text (Alsultan and Warwick, 2013a, 2013b). The free-text keystroke dynamic class uses the typing patterns of the user without entering a predefined text, whilst the fixed-text class uses the typing patterns of the user with entering a predefined text (Karnan *et al.*, 2011). This paper focuses on the free-text class only.

The Arabic and English languages are very different to each other. For example, Arabic is a Semitic language belonging to the Afro-alphabet language family, whilst English is a Germanic language from the Indo-European language family (Miller and Katzner, 2002). Moreover, although Arabic numbers are written from left to right, Arabic letters are written from right to left. An Arabic word includes a set of different letters, and the letters are connected during writing both in printed and handwritten texts. Hence, a study (Alsuhibany *et al.*, 2016) analysed the free-text keystroke dynamic for Arabic language. During the study, three timing features were applied: keystroke duration, di-graph duration and latencies. These features can, thus, be used to construct a unique signature for each user. Furthermore, Euclidean distance was used to find the level of similarity between a user's log-in data and a user's profile. Therefore, the results showed that the approach proposed in Alsuhibany *et al.* (2016) offers better performance than previous results (Alsultan *et al*, 2016) with a false acceptance rate (FAR) of 0.20 and a false rejection rate (FRR) of 0.0.

In this paper, an extension of the work presented previously in Alsuhibany *et al.* (2016) is discussed by investigating how much the system performance would change when using two different keyboards for Arabic language. In particular, the keyboard's layout used during the sign-up phase (i.e. an HP's keyboard) differs from the layout that is used in the testing phase (i.e. Mac's keyboard). By applying the same timing features and classification methods that were used in Alsuhibany *et al.* (2016), we achieved very interesting results. The FRR reached 20 per cent, and the FAR reached 44 per cent when using a Mac's keyboard to collect user profile data. Moreover, the FRR was 0.08, whilst the FAR was 0.60 when using an HP's keyboard in the sign-up phase.

The organisation of the rest of the paper is as follows. Section 2 reviews similar prior works. Section 3 describes the methodology used, whilst the experimental study is explained with more details in Section 4. The results are presented in Section 5 and discussed in Section 6. Section 7 concludes the paper and highlights our research contribution to the future work.

## 2. Background and related work

This section discusses related work in the area of free-text keystroke dynamics in general and those using Arabic free-text in particular. Additionally, a number of studies that sought to investigate the impact of using different keyboards on user authentication are reviewed.

### *2.1 Background*

A recent comprehensive survey of free-text keystroke dynamics methods was compiled in Alsultan and Warwick (2013a, 2013b). There have been many studies undertaken in the area of free-text keystroke dynamics. These studies are differentiated in accordance with variations in selecting time features, as well as classification methods (Alsultan and Warwick, 2013a, 2013b). Additionally, this survey highlights the fact that Euclidean distance has attracted much attention in many works, for instance, Villani *et al.* (2006); Samura and Nishimura (2012) and Buch *et al.* (2008), due to its accuracy. Further, it mentions a number of factors that affect performance in free-text keystroke systems. Some of these factors are text length, word choice and keyboard types. For example, Singh and Arya (2011) used the flight time of key group pairs as a time feature and applied Euclidean distance to classify the users. They achieved a FAR of 0.02 per cent and an FRR of 0.04 per cent. Moreover, a study by E Ahmed *et al.* (2008) selected di-graph latency as a time feature, and a neural network was applied to classify users. The results of this study showed 0.015 per cent for FAR and 4.82 per cent for FRR. The accuracy of the proposed system in Gunetti and Ruffo (1999) reached 90 per cent when the decision tree classification method was used. All of the above studies, including the survey paper, were performed using English language free-texts.

Although there has been a very little research on other languages, such as in Gunetti *et al.* (2005), these languages share the same alphabets with English. However, a recent study by Alsultan *et al.* (2016) investigated the free-text keystroke dynamic using an Arabic language input. This study examined the usefulness of applying an original keyboard's key layout-based method for free-text keystroke dynamics user authentication on Arabic text. The method was originally created to be used with English typing, yet it has crossed over to Arabic input successfully. Additionally, the support vector machines (SVMs) and decision trees (DTs) were utilised to classify individuals based on the timing features. The results of this study showed that the Arabic language input was successfully used for free-text keystroke dynamics user authentication. Moreover, a study in Alsuhibany *et al.* (2016) extends Alsultan *et al.* (2016) by investigating the utilisation of different timing features and applying different classification methods. Both studies Alsuhibany *et al.* (2016) and Alsultan, *et al.* (2016) applied keystroke dynamic authentication for Arabic input; however, different timing features and classification methods were applied. Specifically, in Alsuhibany *et al.* (2016), three timing features were used: keystroke duration, di-graph duration and latencies. Euclidean distance was then applied to find the similarity between user data. The results show that the system performance in Alsuhibany *et al.* (2016) was better than in Alsultan *et al.* (2016), with a FAR of 0.20 and an FRR of 0.0.

The work done in Curtin *et al.* (2006) concentrates on long-text input by conducting three experiments to identify high recognition accuracy that can be acquired if the keystroke system is under the most optimal conditions. The nearest neighbour classification technique using Euclidean distance was implemented in these experiments. The results showed that 100 per cent identification accuracy was achieved when the same keyboard and the same text was used in both the testing and enrolment phases. However, the accuracy was decreased to 97 per cent when different text lengths were used, even if the same keyboard was used.

### *2.2 Related works*

The experiments in Villani *et al.* (2006) aimed to discover the effectiveness of user identification under optimal (the same keyboard type) and non-optimal conditions (the different type of keyboard for enrolment and testing) for English language texts. The keyboard types were desktop and laptop keyboards. When using the same keyboard for enrolment and testing, 99.5 per cent identification accuracy was obtained for laptop keyboards, whilst 98.3 per cent accuracy was achieved for desktop keyboards. However, the accuracy decreased significantly when subjects used different keyboard types for enrolment and testing. That is, when using the laptop for enrolment and the desktop for testing, they obtained 61 per cent identification accuracy. When using the desktop for enrolment and the laptop for testing, 59 per cent identification accuracy was obtained.

A study in Samura and Nishimura (2012) investigated the impact of using same/different keyboards for personal identification by keystroke dynamics in Japanese free-text typing. An experiment was conducted to evaluate the influence of using same/different keyboards. Although the number of participants was limited (i.e. 12 participants), the recognition accuracy was high against the keyboard difference. Likewise, a study in Matsubara *et al.* (2015) investigated the impact of using same/different keyboards for Japanese free-text typing. The results of this study showed that low typing skill users are more easily influenced by the change of keyboards than those wither higher typing skills. As a result, for low typing skill users, the accuracy reached 86 per cent when using different desktop keyboards. For high typing skill users, the recognition accuracy increased to 99 per cent when typing the texts in different desktop keyboards. It is salient to note that the authors used different identification methods for each experiment to maintain high accuracy against the keyboard differences.

This paper, hence, constitutes an extension of work originally presented in Alsuhibany *et al.* (2016) by investigating the impact of using different keyboard layouts on user authentication process for Arabic free-text.

## 3. Methodology

In this section, the time features, the way to extract timing vectors and the Arabic texts that are used in this experiment are explained. Additionally, the methods which are used to classify users based on their collected data are described.

### 3.1 Time features and extracting timing vectors

The users can perform two actions for a key when they are typing on a keyboard, which are pressing a key (*Dn*) and then releasing the key (*Un*). Where *n* indicates the key and the time is recorded in milliseconds.

In this study, as suggested in Alsuhibany *et al.* (2016), the following three timing features were extracted, as shown in Figure 1:

(1) *Hold time for a key (H)*: This is the time that a user presses a key before releasing it. It can be calculated by the following formula $H_{k1} = U_{k1} - D_{k1}$.

(2) *Keystroke latencies*: This can be called down-down (DD), which is the elapsed time between two consecutive keys presses. It can be calculated as: $DD = D_{k2} - D_{k1}$.

(3) *Di-graph duration*: This is the time between releasing the first key and pressing the next key, e.g. $UD = D_{k2} - U_{k1}$.

Figure 1 illustrates the actions that can be performed on the keys and the features that can be extracted for each key during the time period, which is recorded in milliseconds.

After defining and extracting the timing features, we noticed a number of outliers and noisy data, such as when a user presses two keys together by mistake. These outlying data may lead to incorrect results. Therefore, these data have been deleted during the pre-processing phase. The timing vector in our study is the same as in Alsuhibany *et al.* (2016) which is the means of keystroke duration, flight time and di-graph duration. This time vector was extracted and stored in the database during enrolment and testing phases to be used in the classification process.

### 3.2 Typing data

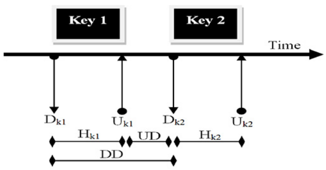
In this section, the Arabic text that participants were asked to type using two different keyboard layouts is presented. For the usability aspect, we selected common words and letters that can be typed using only one key for each letter (e.g. we avoid words that have letters such as "أ" which needs to press H and shift key at the same time on a Mac keyboard).

(1) *Arabic text for the sign-up phase*: This Arabic text is made of 199 characters (with spaces) as follows:

"حسن الظن يوحد قوه الروح وقوة الجسد ومن استقرار الروح تزدهر الصحة النفسية التي ترتبط غاية الارتباط بقدرة الشخصية للتوافق مع نفسها ومجتمعها الذي نعيش فيه وهذا يقود للتمتع بحياة ساكنه سوية مفعمة بالحماس".

(2) *Arabic text for the log-in phase*: This Arabic text has 179 characters (with spaces) as follows:

"الاسلام منهج شامل لجميع جوانب الحياة وسلوك الانسان وهذا الشمول لا يقبل الاستبعاد ولا التخصيص بل هو كامل تام بكل ما تحمله الشمولية ومن البديهيات في الاسلام انه لعموم البشر وكافه الخلق".




**Source:** Alsuhibany *et al.* (2016)

**Figure 1.**
Keystroke features

### 3.3 Classification methods

We applied Euclidean distance due to its accuracy, as discussed in Section 2, to check whether the user's data during a sign-up phase are close to the user's data during the enrolment phase.

As the timing vectors consist of three dimensions, including the means of keystroke duration, flight time and up-down keys duration, we have used a Euclidean distance equation in three-dimensional space (Waterman, 2016):

$$d(x,y) = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + (x_3 - y_3)^2} \quad (1)$$

where $x$ *and* $y$ are two vectors. In this study, $x$ indicates the user's testing vector, and $y$ indicates the user's training vector. Additionally, $d(x,y)$ should not be less than zero (i.e. $d(x,y) \geq 0$) (Waterman, 2016).

To identify users, we needed to classify the data by determining a threshold for Euclidean distance. Based on the work done in Alsuhibany *et al.* (2016), we determined the standard deviation (SD) of a user's data vector during sign-up phase as the threshold between two user vectors.

After collecting data, we used MATLAB, which provides a statistical toolbox that can assist us to calculate the distances between two vectors to analyse the data and extract timing vectors for each user.

## 4. Experiment

We carried out a controlled laboratory experiment in such a way that all participants were asked to type texts in Arabic language using two different keyboard layouts. In this section, we describe the setup and procedure of the experiment.

### 4.1 Experiment setup

In this part, the system, participants and materials of the experiment are detailed:

- *System*: The developed system in Alsuhibany *et al.* (2016) is used in this experiment to collect data from users. This system then has been installed on two different laptops with different keyboard layouts. The first laptop was MacBook Pro, and the layout of the keyboard is shown in Figure 2. The second laptop was HP, and Figure 3 shows its keyboard layout. In both laptops, the data are stored in a MySQL database. Once the sign-up phase is successfully completed, the users are directed to

**Figure 2.**
Arabic-English keyboard of the *MacBook Pro* laptop

the log-in interface, as illustrated in Figures 4 and 5. There is an email box in both phases of the interface where the user's email operates as a user's ID in the database.

- *Participants*: In total, 34 users participated in the experiment. The participants had different levels of typing skills and were aged between 19 and 60 years. Furthermore, they were native Arabic speakers with different educational backgrounds (21 undergraduate students and 13 graduates). A proportion of participants had prior experience with one of the keyboards that we used in this experiment.
- *Materials*: The stimulus material presented to participants consisted of two texts written in Arabic: sign-up text and log-in text. The sign-up text in Arabic was 199 characters (with spaces) in length, and 179 characters (with spaces) for the log-in text. Although several studies have shown an interest in using short free-text, for instance, Singh and Arya (2011); Monrose and Rubin (1997) and Chantan *et al.* (2012), it seems insufficient to use short texts to analyse keystrokes, as it does not provide an adequate amount of information to discriminate between users (Dowland *et al.*, 2001; Gunetti and Picardi, 2005; Sim and Janakiraman, 2007). Therefore, we decided to use these lengths of characters for both languages.


**Figure 3.** Arabic-English keyboard of the *HP* laptop


**Figure 4.** Sign-up interface in Arabic


**Figure 5.**
Log-in interface in Arabic

We have used the same text for sign-up in both laptops, and users type the same text for log-in using two different keyboards. The details of these texts have been given in Section 3.

Furthermore, a study in Villani *et al.* (2006) concluded that the system has a good chance of accurately identifying a user when a participant uses the same type of keyboard for both training and testing. However, investigating the recognition accuracy for Arabic language input whilst using different keyboards constitutes an important contribution to this field of research.

### *4.2 Experiment procedure*

In this section, we clarify how we conducted the experiment, i.e. instructions that were given to the subjects, procedures and data collection.

First, we explained the main goal of our experiment to all participants, which is examining the effect of using different keyboards on the user authentication process. To facilitate concentration for the subjects during the experiment, we carried out this experiment in a laboratory environment. The subjects were requested to put their phones off or switched to silent to avoid any type of interruption during the experiment. Each user was asked to type the displayed text in two laptops which have different keyboard layouts. The participants were asked to complete all phases (training and testing) in one of the laptops and then switch to the second laptop and follow the same procedure (i.e. sign-up then log-in). The participants were permitted to take a break between sign-up and log-in phases. Moreover, the use of the backspace and spacebar keys was permitted during this experiment. Finally, a welcome message was displayed at the end of the sign-up phase, whilst a confirmation message was shown at the end of the log-in phase, which indicates the end of the experiment. The participants were asked to use the same email address during the experiment whilst using MacBook and HP laptops. For each participant, all typing attempts are recorded in milliseconds and stored in the database. For each laptop, we collected the timing vector for the participants.

## 5. Results

All 34 participants successfully completed their given task. To evaluate the system, two performance measures are used: FRR and FAR (Alsultan and Warwick, 2013a, 2013b). FRR is the ratio of the authorised users who are incorrectly prohibited access to the system. On the other hand, FAR is the ratio of the unauthorised users who gain access to the system. In this experiment, we calculated the FAR and FRR in two cases as follows. For the FAR, we compared users' test data that were stored in the Mac laptop with two users randomly based on their profile data collected in the HP laptop. Also, we compared users' test data in the HP laptop with two users randomly based on their training data collected in the Mac laptop. We used an automated function in Java to pick a random user. We calculated the FRR by comparing users' test data collected in the HP laptop with users' profile data, which were gathered by the Mac laptop. We compared users' log-in data collected in the Mac laptop with users' training data collected in the HP laptop.

As mentioned in Section 3, the SD of the users' training data has been chosen as a threshold. Consequently, we have achieved an interesting FRR and FAR when using two different keyboard layouts. In the first case, when we use login data from the Mac laptop and compare it with HP's training data, we found that the FFR is equal to 0.08 and the FAR is equal to 0.60. In the second case, when comparing HP's testing data with Mac's profile data, the value of the FRR equals 0.20 and the FAR equals 0.44. These results are outlined below in Table I.

Interestingly, the system performance is influenced by utilising the two keyboard layouts in the sign-in and log-in phases. For example, Figures 6 and 7 show the system performance when using Mac and HP keyboards, respectively.

## 6. Discussion

The results of this experiment show that using two different keyboard layouts during the log-in and sign out phases affects the performance of the system. That is, when the SD is used as a threshold for Euclidean distance, we found interesting results regarding the FAR

| Testing data | Mac's keyboard | | HP's keyboard | |
|---|---|---|---|---|
| Training data | HP's keyboard | | Mac's keyboard | |
| Threshold | SD<br>Euclidean distance | | | |
| | FAR | FRR | FAR | FRR |
| | 0.60 | 0.08 | 0.44 | 0.20 |

**Table I.** Error rates of using Euclidean distance

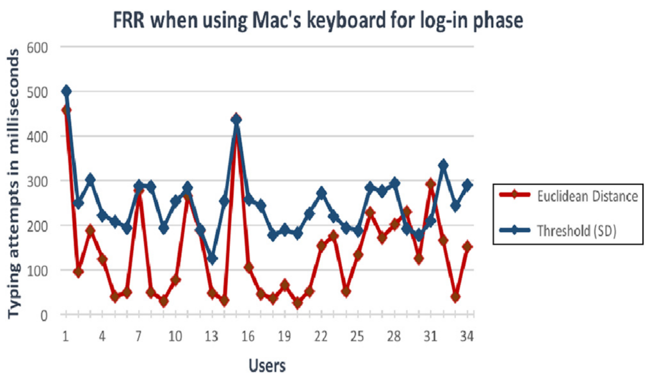



**Figure 6.** The system performance when using a Mac's keyboard

and FRR whilst using two different keyboard layouts, as they reflect the accuracy of the system. It is important to note that the achieved results may confirm a known fact, i.e. that keystroke layout has an influence on typing patterns. However, the aim of our study was to evaluate and validate the system performance whilst using Arabic keystroke layout, as Arabic language has different features comparing to previous studies (Villani *et al.*, 2006; Matsubara *et al.*, 2015). Additionally, we argue that our findings have important implications for supporting proven studies for using English keystroke layout (Villani *et al.*, 2006) and using Japanese keystroke layout (Matsubara *et al.*, 2015).

Moreover, the features used in our experiment are the same as in Alsuhibany *et al.* (2016), as well as the typed text. However, the accuracy in Alsuhibany *et al.* (2016) was much better than the achieved accuracy in this experiment. In particular, the FRR in Alsuhibany *et al.* (2016) was equal to 0.0 for Arabic text, whereas in this experiment, the ratio reaches 0.08 and 0.20. The possible explanation behind this may be that the same keyboard layout was used in both the enrolment and testing phases.

Moreover, using different keyboard layouts not only affects the FRR ratio but also the FAR. That is, in the previous study (Alsuhibany *et al.*, 2016), if 20 per cent FAR was achieved, whilst we achieved 44 per cent, which could be increased to reach 60 per cent. It is observable that the FAR is deteriorating when a keyboard layout is used in the sign-in that differs from the keyboard that was used in the sign-up phase.

It is stated in Matsubara *et al.* (2015) that using different keyboards for typing Japanese text affects the accuracy of the system, except for those who have very high typing skills, where the accuracy reached 99 per cent. Our results prove that the change of keyboard's layout from phase to phase (e.g. sign-up phase to sign-in phase) affects the accuracy of the system for Arabic text also, as illustrated in Figures 6 and 7. Furthermore, as shown in Table I, the ratio of FAR is increased by at least 20 per cent compared to previous work (Alsuhibany *et al.*, 2016) even when the same text and timing features were used.

Additionally, during this experiment, we found that most of the participants were familiar with HP's keyboard layouts. This might be the reason underlying the increased FAR ratio or decrease FRR ratio, compared to the results of the Mac keyboard. Additionally, even though all of the participants are Arabic native speakers, the level of typing skills varies across participants, which affects the accuracy of the system as indicated in Matsubara *et al.* (2015).

A limitation of this study is that the sample size of participants was relatively small. However, we consider this work to validate the impact of keyboards on the system's performance regarding the recognition accuracy when using Arabic free-text. A future study investigating this impact with a larger sample size would further corroborate these results.

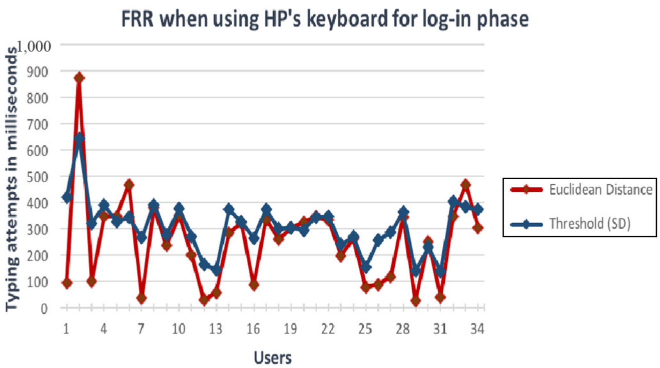



**Figure 7.**
The system performance when using an HP's keyboard

## 7. Conclusion and future work

This study has investigated the impact of using different keyboard layouts on the user authentication process when typing Arabic free-text. Therefore, an experimental study was conducted. The results of this investigation showed that FAR was 0.44 using the first keyboard and 0.60 using the second, whilst the FRRs were 0.2 and 0.08 for the first and second keyboards, respectively. Hence, a considerable change in the accuracy of the system is observed when using different keyboards even under the same time features and Arabic text inputs.

This work can be extended by applying some methods to achieve more accurate results in light of the keyboard differences. Moreover, more timing features might be used to increase the accuracy of the system. These results could be further enhanced by using different classification methods. Finally, this study could be extended by investigating the effect of using different keyboard layouts in other languages.

## References

Alsuhibany, S., Almushyti, M., Alghasham, N. and Alkhudier, F. (2016), "Analysis of free-text keystroke dynamics for Arabic language using Euclidean distance", in *12th International Conference on Innovations in Information Technology*, *IEEE*, pp. 1-6.

Alsultan, A. and Warwick, K. (2013a), "User-friendly free-text keystroke dynamics authentication for practical applications", in *IEEE International Conference on Systems, Man, and Cybernetics*, *IEEE*, pp. 4659-4662.

Alsultan, A. and Warwick, K. (2013b), "Keystroke dynamics authentication: a survey of free-text methods", *International Journal of Computer Science Issues*, Vol. 10 No. 4, pp. 1-10.

Alsultan, A., Warwick, K. and Wei, H. (2016), "Free-text keystroke dynamics authentication for Arabic language", *IET Biometrics*, Vol. 5 No. 3, pp. 164-169.

Buch, T., Cotoranu, A., Jeskey, E., Tihon, F. and Villani, M. (2008), "An enhanced keystroke biometric system and associated studies", in *Proceedings of Student-Faculty Research Day*, *Pace University*.

Chang, T., Tsai, C., Yang, Y. and Cheng, P. (2011), "User authentication using rhythm click characteristics for non-keyboard devices", in *International Conference on Asia Agriculture and Animal*, *IACSIT Press*, *Singapore*, pp. 167-171.

Chantan, C., Sinthupinyo, S. and Rungkasiri, T. (2012), "Improving accuracy of authentication process via short free text using Bayesian network", *International Journal of Computer Science Issues*, Vol. 9 No. 2, pp. 10-16.

Curtin, M., Tappert, C., Villani, M., Ngo, G., Simone, J., Fort, H. and Cha, S. (2006), "Keystroke biometric recognition on long-text input: a feasibility study", in *International Multi Conference Engineers and Computer Scientist*, *Pace University*.

Deng, Y. and Zhong, Y. (2013), "Keystroke dynamics user authentication based on Gaussian mixture model and deep belief nets", ISRN Signal Processing, 2013, pp. 1-7.

Dowland, P., Singh, H. and Furnell, S. (2001), "A preliminary investigation of user authentication using continuous keystroke analysis", in *The International Conference on Information Security Management and Small Systems Security*, pp. 215-226.

E Ahmed, A., Traore, I. and Almulhem, A. (2008), "Digital fingerprinting based on keystroke dynamics", in *The Second International Symposium on Human Aspects of Information Security and Assurance*, *University of Plymouth*, pp. 94-104.

Gunetti, D. and Picardi, C. (2005), "Keystroke analysis of free text", *ACM Transactions on Information and System Security*, Vol. 8 No. 3, pp. 312-347.

Gunetti, D. and Ruffo, G. (1999), "Intrusion detection through behavioral data", in *International Symposium on Intelligent Data Analysis*, *Springer*, *Berlin, Heidelberg*, pp. 383-394.

Gunetti, D., Picardi, C. and Ruffo, G. (2005), "Keystroke analysis of different languages: a case study", in *International Symposium on Intelligent Data Analysis*, *Springer*, *Berlin, Heidelberg*, pp. 133-144.

Karnan, M., Akila, M. and Krishnaraj, N. (2011), "Biometric personal authentication using keystroke dynamics: a review", *Applied Soft Computing*, Vol. 11 No. 2, pp. 1565-1573.

Matsubara, Y., Samura, T. and Nishimura, H. (2015), "Keyboard dependency of personal identification performance by keystroke dynamics in free text typing", *Journal of Information Security*, Vol. 6 No. 3, pp. 229-240.

Miller, K. and Katzner, K. (2002), *The Languages of the World*, Routledge.

Monrose, F. and Rubin, A. (1997), "Authentication via keystroke dynamics", in *The 4th ACM Conference on Computer and Communications Security, ACM Digital Library*, pp. 48-56.

Samura, T. and Nishimura, H. (2012), "Influence of keyboard difference on personal identification by keystroke dynamics in Japanese free text typing", in *Fifth International Conference on Emerging Trends in Engineering and Technology*, *IEEE*, pp. 30-35.

Sim, T. and Janakiraman, R. (2007), "Are digraphs good for free-text keystroke dynamics?", in *IEEE Conference on Computer Vision and Pattern Recognition*, *IEEE*, pp. 1-6.

Singh, S. and Arya, K. (2011), "Key classification: a new approach in free text keystroke authentication system", in *Third Pacific-Asia Conference on Circuits, Communications and System*, *IEEE*, pp. 1-5.

Tappert, C., Cha, S., Villani, M. and Zack, R. (2012), "A keystroke biometric system for long-text input", IGI Global, pp. 32-57.

Villani, M., Tappert, C., Ngo, G., Simone, J., Fort, H. and Cha, S. (2006), "Keystroke biometric recognition studies on long-text input under ideal and application-oriented conditions", in *Conference on Computer Vision and Pattern Recognition Workshop*, *IEEE*, pp. 39-39.

Waterman, G. (2016), *Linear Algebra I Skills, Concepts and Applications*, 1st ed, OR Institute of Technology, pp. 36-43.

**Corresponding author**
Suliman A. Alsuhibany can be contacted at: salsuhibany@qu.edu.sa