The Journal of Systems and Software 85 (2012) 1157–1165

Contents lists available at SciVerse ScienceDirect

The Journal of Systems and Software

journal homepage: www.elsevier.com/locate/jss

A graphical-based password keystroke dynamic authentication system for touch screen handheld mobile devices

Ting-Yi Chang[a,*], Cheng-Jung Tsai[b], Jyun-Hao Lin[a]

[a] Graduate Institute of e-Learning, National Changhua University of Education, No. 1, Jin-De Road, 500 Changhua City, Taiwan, ROC
[b] Department of Mathematics, National Changhua University of Education, No. 1, Jin-De Road, 500 Changhua City, Taiwan, ROC

ARTICLE INFO

Article history:
Received 24 May 2011
Received in revised form
26 December 2011
Accepted 26 December 2011
Available online 11 January 2012

Keywords:
User authentication
Keystroke dynamics authentication
Keystroke feature
Graphical-based password authentication

ABSTRACT

Since touch screen handheld mobile devices have become widely used, people are able to access various data and information anywhere and anytime. Most user authentication methods for these mobile devices use PIN-based (*Personal Identification Number*) authentication, since they do not employ a standard QWERTY keyboard for conveniently entering text-based passwords. However, PINs provide a small password space size, which is vulnerable to attacks. Many studies have employed the KDA (*Keystroke Dynamic-based Authentication*) system, which is based on keystroke time features to enhance the security of PIN-based authentication. Unfortunately, unlike the text-based password KDA systems in QWERTY keyboards, different keypad sizes or layouts of mobile devices affect the PIN-based KDA system utility. This paper proposes a new graphical-based password KDA system for touch screen handheld mobile devices. The graphical password enlarges the password space size and promotes the KDA utility in touch screen handheld mobile devices. In addition, this paper explores a pressure feature, which is easy to use in touch screen handheld mobile devices, and applies it in the proposed system. The experiment results show: (1) EER is 12.2% in the graphical-based password KDA proposed system. Compared with related schemes in mobile devices, this effectively promotes KDA system utility; (2) EER is reduced to 6.9% when the pressure feature is used in the proposed system. The accuracy of authenticating keystroke time and pressure features is not affected by inconsistent keypads since the graphical passwords are entered via an identical size (50 mm × 60 mm) human–computer interface for satisfying the lowest touch screen size and a GUI of this size is displayed on all mobile devices.

1. Introduction

Whenever people use services such as e-bank and e-mail, servers should have the ability to authenticate the users' identities. Otherwise, anyone can easily impersonate a legal user to login to the server. Password-based authentication schemes are simple and practical solutions to user identification because they allow people to choose their own passwords without any device to generate or store them. For personal computers, passwords consist of letters, numbers, and special punctuations on a standard QWERTY keyboard. This is called text-based (alphanumeric-based) password authentication. Handheld mobile devices do not have standard QWERTY keyboard to conveniently enter text-based passwords. Mobile devices often use numeric passwords, which is called PIN-based (*Personal Identification Number*) authentication. Though the password space size of text-based passwords is larger than that of PINs (i.e., the password space size of an 8-character text-based password and an 8-digital PIN are $64^8 \cong 2.8 \times 10^{14}$ and 10^8, respectively), text-based passwords are preferred natural language phrases that people can recognize easily and are therefore susceptible to dictionary attacks. On the other hand, PIN-based authentication is widely used in mobile devices, but it provides a small password space size and therefore compromises security.

The KDA (*Keystroke Dynamic-based Authentication*) system was first proposed by Gaines et al. (1980). It is a biometric measurement method to provide additional security for text-based passwords. Many studies (Araújo et al., 2005; Bergadano et al., 2002; Bleha et al., 1990; Boechat et al., 2007; Chang and Yang, in press; Haider et al., 2000; Harun et al., 2010; Hwang et al., 2009b; Killourhy and Maxion, 2009; Ru and Eloff, 1997; Shih and Lin, 2008; Xi et al., 2011) have been proposed to improve the text-based password KDA system utility. KDA systems confirm the correctness of passwords and also identify users based on individual password keystroke time features. The keystroke time features include the duration of a keystroke (keystroke hold time) and the interval of the keystrokes (keystroke latency time). Even if the password is revealed by dictionary attacks or shoulder surfing

* Corresponding author. Fax: +886 4 7211290.
E-mail addresses: tychang@cc.ncue.edu.tw (T.-Y. Chang), cjtsai@cc.ncue.edu.tw (C.-J. Tsai).

0164-1212/$ – see front matter © 2011 Elsevier Inc. All rights reserved.
doi:10.1016/j.jss.2011.12.044

attacks, the probability of breaking authentication is reduced. The KDA system has the following advantages. It is low-cost with no extra device to obtain the user's features, and does not require complex computations to capture the user's features. Since the process of capturing features is done when the user enters his or her password, it does not create any additional burden on users. Compared with other biometric authentication methods such as fingerprints, eye scan, iris, and signature, the KDA system is simple and useful for providing additional security in identity verification.

As is well known, mobile devices are widely used for accessing various data and information. Campisi et al. (2009) proposed a text-password KDA system that uses a cellular phone keypad. On the other hand, many studies (Clarke and Furnell, 2007a,b; Hwang et al., 2009a) have applied KDA systems to enhance the security of PIN-based authentication in mobile devices. However, the sizes or layouts of keypads of the mobile devices produced by different manufacturers are inconsistent. A user may not get used to entering his or her PIN or text-based password through different mobile devices. This will result in the user's features being entered inconsistently and KDA system verification failing if users enter their PINs or text-based passwords through difference mobile devices. Consequently, the KDA system utility for mobile devices (Campisi et al., 2009; Clarke and Furnell, 2007a,b; Hwang et al., 2009a) is worse than that for QWERTY keyboards (Killourhy and Maxion, 2009; Shih and Lin, 2008).

This paper develops a novel graphical-based password KDA system to improve PIN-based authentication for mobile devices. The password space size of the proposed system is larger than those of PIN-based authentication schemes. Regardless of the size of the user's mobile touch screen, users enter their graphical passwords through clicking or touching an identical human–computer interface. Therefore, the accuracy of users authentication is not affected by inconsistent keypads. In addition, this paper explores the pressure feature, which is a new biometric keystroke feature found in touch screens. The proposed graphical-based password KDA system is implemented in an Android-compatible phone. Serial experiments show the utility of the proposed graphical-based password KDA system is better than the related text-based password and PIN-based KDA systems (Campisi et al., 2009; Clarke and Furnell, 2007a,b). Moreover, when the pressure feature is applied in the proposed system, it further promotes system utility.

The organization of this paper is as follows. Section 2 reviews and discusses studies on graphical-based password authentication and the PIN-based KDA system, respectively. Section 3 proposes the architecture of the methodology, which includes the enrollment phase, the classifier building phase, and the authentication phase. The pressure feature is also introduced in this section. Section 4 presents the experimental results of this paper and compares them with other related studies. At the same time, the performance of the proposed system is presented. It is suitable for low-power mobile devices. Finally, conclusions are given in Section 5.

2. Related works

To improve on the drawbacks of PIN-based KDA systems, this paper first combines a graphical password with the KDA system. This section introduces these two related studies in Sections 2.1 and 2.2, respectively.

2.1. Graphical-based password authentication

The common method for identity authentication is text-based password authentication. In terms of security, a text-based password should consist of a string of eight or more random characters. However, a user is limited by the ability of his or her long-term memory to remember passwords. If the length of the text-based passwords is long, a user's memory load is heavy. Further, people may frequently forget and confuse their text-based passwords (Wiedenbeck et al., 2005a). Users typically cope with text-based password problems as follows. First, they write down their passwords. Second, they use one password for many systems. Third, natural language phrases are preferably used as passwords so they can be recognized easily. However, this leads to some text-based passwords becoming weak passwords that are vulnerable to dictionary and shoulder surfing attacks.

Graphical-based password authentication is an alternative method to withstand dictionary attacks, as originally described by Blonder (1996). Today, graphical-based password authentication can be divided into two categories: recognition-based and recall-based graphical passwords. Recognition-based systems authenticate the users' identities based on a sequence of images selected and remembered by users. These include the Passface system (Brostoff and Sasse, 2000) and Déjà Vu system (Dhamija and Perrig, 2000). However, they are vulnerable to shoulder surfing attacks. As such, Sobrado and Birget (2002) have developed convex-hull click, movable frame, and intersection to withstand such attacks (please see Sobrado and Birget, 2002 for more details). Further, Wiedenbeck et al. (2005a) extended Blonder's idea and developed the PassPoint system. In Weidenback et al.'s method, a user clicks on an image to create his or her graphical password and a tolerance around each chosen pixel is calculated. To authenticate identity, users must click within the tolerance of their chosen pixels and also in the correct sequence. Therefore, the password space size in Weidenback et al.'s method is larger than text-based passwords. On the other hand, the best known recall-based method is DAS (Draw-A-Secret), proposed by Jermyn et al. (1999). In their method, a user is asked to draw a simple picture on a 2D grid, and then the system authenticates the user's identity based on the order of the drawn picture. Moreover, Syukri et al. (1998) used a signature as a substitute for drawing a simple picture. The advantage of using the signature is it is self-given and difficult to copy. Unfortunately, most users require additional devices to finish their signatures since they are not used to writing their signatures with a mouse. Thus, this is inconvenient for users.

Recently, touch screen handle mobile devices have been becoming more widespread. A user inputs his or her password by clicking or touching the touch panel. Because the touch screen size is too small, Sobrado et al.'s and Weidenback et al.'s methods are unsuitable for authentication in mobile devices. Therefore, Jansen (2003) proposed a recognition-based graphical password authentication system for mobile devices. Jansen's system divides an image (e.g. sea, cat, etc.) into thirty thumbnail photos. A user selects several different thumbnail photos and the sequence of these thumbnail photos is the user's graphical password. Further, Jansen (2004) improved his method by allowing the thumbnail photos to be chosen repeatedly. Therefore, the password space size is larger than before. For example, a user chooses three photos and then the password space size is enlarged from $30 \times 29 \times 28$ to 30^3. However, Jansen's methods still cannot withstand shoulder surfing attacks.

The graphical-based password authentication has the following advantages. First, a person's long-term memory need not store all the images, but rather a meaningful interpretation (Mandler and Ritchey, 1977). Psychologists have shown images are remembered more easily than words or sentences (Mandler and Ritchey, 1977; Revett et al., 2005; Wiedenbeck et al., 2005a). In this case, a user is able to remember a complex password and then the password space size will be larger. Second, these graphical-based password authentication methods assume the number of possible images is sufficiently large. Thus, the password space size of the graphical passwords is larger than the text-based passwords. Third, it is able to withstand dictionary attacks. Since the recognition-based

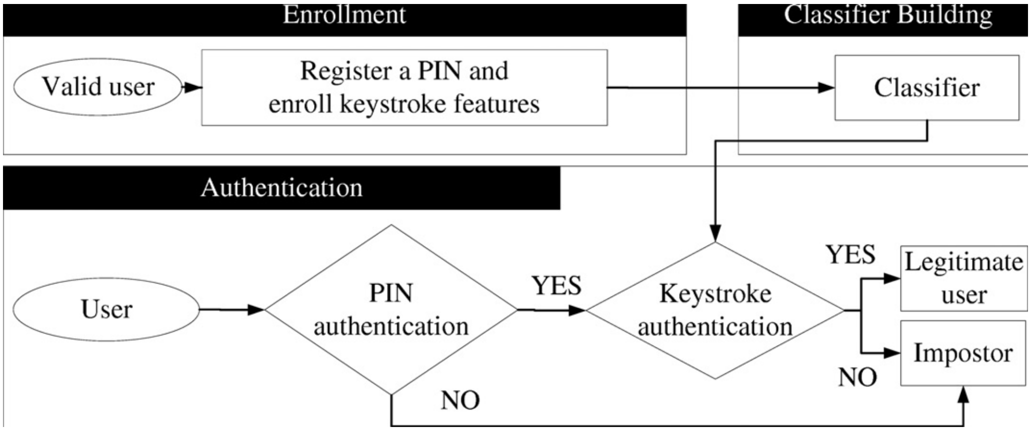


Fig. 1. The flow chart of PIN-based KDA systems (Clarke and Furnell, 2007a,b; Hwang et al., 2009a).

graphical passwords use mouse input instead of keyboard input, there are no pre-existing searchable dictionaries for mounting a dictionary attack. Although some recall-based graphical passwords are susceptible to dictionary attacks, it increases the cost of mounting a dictionary attack because the password space size of the graphical passwords is large. Overall, graphical passwords are less vulnerable to attacks than text-based passwords (Hu et al., 2010).

2.2. Keystroke Dynamic-based Authentication

PIN-based authentication is most widely used in identity verification for mobile devices. However, the security of the system is disintegrated when the PIN is captured by mounting a shoulder surfing attack. Many studies (Campisi et al., 2009; Clarke and Furnell, 2007a,b; Hwang et al., 2009a) have applied the KDA system for the PIN-based authentication scheme in mobile devices. Fig. 1 shows three phases in the KDA systems.

In the enrollment phase, a user is asked to enroll his or her PIN and the corresponding keystroke features are recorded. Because users will probably not endure such a long enrollment procedure, Araújo et al. (2005) suggested the number of training samples should not be more than 10 in this phase. Hwang et al. (2009a) used an artificial rhythm or cue to improve the keystroke data quality when the user enters his or her PIN. A user inserts pauses, staccatos, and legatos while he or she is entering the PIN according to his or her artificial musical rhythm and then a metronome is used to help the user keep tempo. That is, the user must memorize the number of pauses, the lengths of pauses, and the positions of pauses. Although this method improves keystroke data quality and thus promotes KDA system utility, it causes an additional burden on users.

In the classifier building phase, a classifier is built according to collected samples in the enrollment phase. Clarke and Furnell (2007a,b) used various neural networks to build classifiers, such as Feed Forward Multi-Layer Perceptron (FF-MLP), Radial Basis Function (RBF), and Generalized Regression Neural Network (GRNN). However, these neural network classifiers require the impostors' patterns to train the network. That is, the system must provide the legitimate user's password and ask other people to enter the same password so the impostors' patterns can be obtained. Obviously, this is impractical in real applications. In addition, the system computation is complex for training the network and it thus needs more time for building the classifier. Later, Campisi et al. (2009) proposed many statistical methods for classifier normalization, such as the *z*-score, median, and min-max. These methods can be combined for building classifiers and are suitable for low-power cellular phones.

Recently, touch screen mobile devices have become widely used and the most basic equipment in various devices. To apply the KDA system to touch screen mobile devices without keypads, Saevanee and Bhatarakosol (2008) and Chang et al. (2010) used the notebook touch pad and the mouse to simulate users clicking on the touch panel, respectively. Saevanee and Bhatarakosol found the pressure feature on the notebook touch pad and claim it can be utilized on the touch panel of mobile phones. Unfortunately, their experiment used only 10 users to verify system utility and twenty training samples were needed. The simulation result was unable to accurately reflect the touch screen result in handheld mobile phones. Further, Saevanee and Bhatarakosol's classifier is designed by the KNN (K-Nearest Neighbor) algorithm. Although the KNN classifiers require no impostors' patterns, they compare other users' samples to determine which login sample belongs to which user in the system. Thus, the system load will increase based on the number of users in the system and lead to the users spending more time on performing identity authentication. On the other hand, Chang et al. used a neural network to build a classifier. They still required the impostors' patterns to construct the classifier. Under these situations, the simulation results of both Saevanee et al.'s and Chang et al.'s are inapplicable to touch screen handheld mobile devices.

In the authentication phase, the KDA systems confirm the validity of the PIN and also verify the corresponding keystroke features. If the PIN is correct, the system continues to verify the keystroke features. One will reject the PIN as an impostor's pattern if the distance between the prototype and the pattern is greater than a given threshold. Several studies (Campisi et al., 2009; Clarke and Furnell, 2007a,b) used keypads to enter PINs or text-based passwords. If a user uses different mobile devices, then the user's features will be inconsistent.

The performances of the KDA systems are measured based on the recognition error rates, which are defined as follows, and the relationships between these measures are shown as in Fig. 2.

(1) *False Rejection Rate* (FRR): the rate of the system rejecting a legitimate user. FRR is also called *Type* I *error*.
(2) *False Acceptance Rate* (FAR): the rate of the system accepting an impostor. FAR is also called *Type* II *error*.
(3) *Equal Error Rate* (EER): the value at which FAR equals to FRR, which is the most balanced performance index.
(4) *Zero False Rejection Rate* (ZeroFRR): FAR is ZeroFRR when FRR equals zero. It uses for analysis the probability of impostor when legal user all can login.
(5) *Zero False Acceptance Rate* (ZeroFAR): FRR is ZeroFAR when FAR equals zero. It uses for analysis the probability of legal user when impostor all can reject.

When these measures are closer to zero, it means the system of authentication is better. FRR and FAR have a negative correlation,

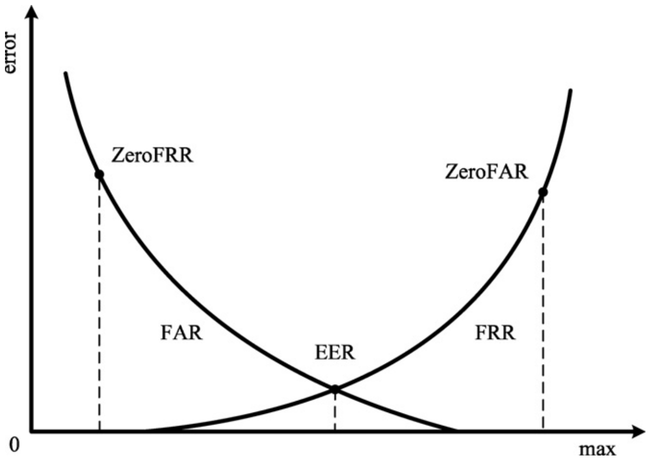


Fig. 2. The relationships of five evaluation criteria.

meaning, when FRR decreases, FAR increases. This means to enhance the security of the system, it would be more difficult for legitimate users to login. Otherwise, if legitimate users login easily then the security of system is low. FRR and FAR only analyze whether the user is legitimate or an impostor. EER can be used to analyze the accuracy of the system. Therefore, this paper uses EER to assay the proposed system.

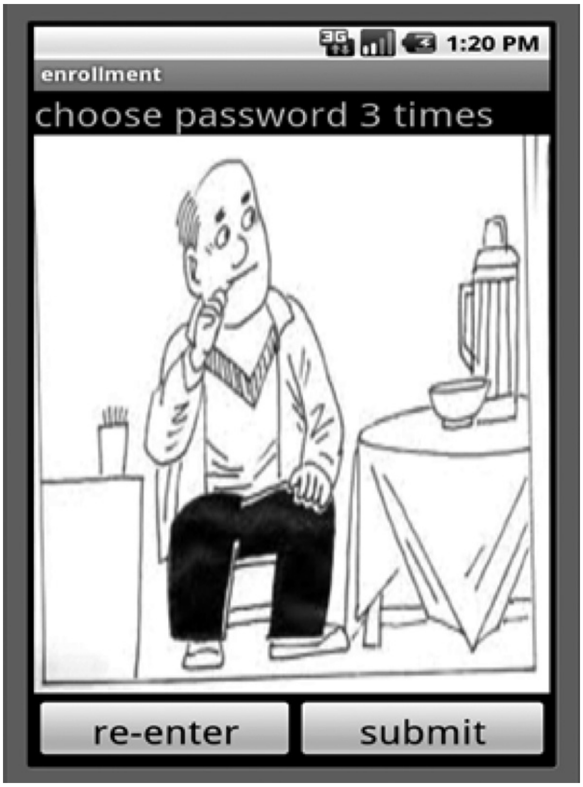
3. Methodology

This paper develops a graphical-based password KDA system for touch screen mobile devices. After observing users using touch screen handheld mobile devices, we found users enter their data through the touch screen in characteristic fashion. The force of each person clicking or touching the touch panel is not necessarily the same when they enter their data, thus, the system captures different pressures from the touch panels on mobile devices. Therefore, this paper assumes the pressure feature can be a new biometric feature and improve data quality. Under this situation, this paper analyzes whether the authentication utility after adding the pressure features is improved. Three phases are the same as in the KDA systems, which are described separately in the following subsections.

3.1. Enrollment phase

The graphical passwords in our system are based on Jansen's method, but there are no edges around each thumbnail photo (because the edges would be ugly and limit the user's freedom of choice for clicking the photos (Wiedenbeck et al., 2005b)). A user was asked to choose his or her favorite image. Regardless of the size of the image, it is transformed into a 50 mm × 60 mm frame and the system cuts it into thirty thumbnail photos each with an identical size of 10 mm × 10 mm. Though the sizes of mobile touch screens are inconsistent, the size of the human–computer interface in our system is identical. The user chooses 3–6 thumbnails in the image through the touch panel on the mobile device and the sequence of these photos is the user's graphical password.

Users enter their graphical passwords through the human–computer interface as shown in Fig. 3. Users enter their graphical passwords by clicking the image in the middle of the view, or clicking the re-enter button to enter their graphical password again, and clicking the submit button to pass their graphical passwords. The message area at the bottom shows the number of times the user needs to enter their graphical passwords.


1:20 PM

enrollment

choose password 3 times

Fig. 3. The human–computer interface for a user to enter his or her graphical password.

This paper provides feedback for users by vibrating so a user knows when he or she enters his or her graphical passwords successfully.

When a user's finger presses the touch screen of the handheld mobile device at thumbnail photo$_j$ the system captures a pressure feature. In the interval between the press and release of the photo$_j$, it will have four kinds of time features. In the ith training sample, the relationships between the above features are shown in Fig. 4, where user clicks 'photo$_1$, photo$_2$, photo$_3$, photo$_4$'.

(1) *Down-Up* (DU) *time*: In the ith training sample, the time duration of press and release of photo$_j$ is called $DU_{i,j}$.
(2) *Up-Down* (UD) *time*: In the ith training sample, the time interval between the release of photo$_j$ and press photo$_{j+1}$ is called $UD_{i,j}$.
(3) *Down-Down* (DD) *time*: In the ith training sample, the time interval between the press of photo$_j$ and press photo$_{j+1}$ is called $DD_{i,j}$.
(4) *Up-Up* (UU) *time*: In the ith training sample, the time interval between the release of photo$_j$ and release photo$_{j+1}$ is called $UU_{i,j}$.
(5) *Pressure*: In the ith training sample, the pressure of the user pressing the screen for photo$_j$ is called $P_{i,j}$.

In Araújo et al.'s research (2005), the best combination of time features for authentication was DU, UD, and DD time. The combination of time features and pressure features was used in this paper. When the user chooses k photos as his or her graphical password, there are $(4k - 2)$-dimensional features. These features of the user's ith training sample include k DU time features, k pressure features, $k - 1$ UD time features, and $k - 1$ DD time features. They are denoted separately as DU_i set, UD_i set, DD_i, and P_i set as follows:

$$\begin{aligned} DU_i &= \{DU_{i,1}, DU_{i,2}, \ldots, DU_{i,k}\} \\ UD_i &= \{UD_{i,1}, UD_{i,2}, \ldots, UD_{i,k-1}\} \\ DD_i &= \{DD_{i,1}, DD_{i,2}, \ldots, DD_{i,k-1}\} \\ P_i &= \{P_{i,1}, P_{i,2}, \ldots, P_{i,k}\} \end{aligned}$$

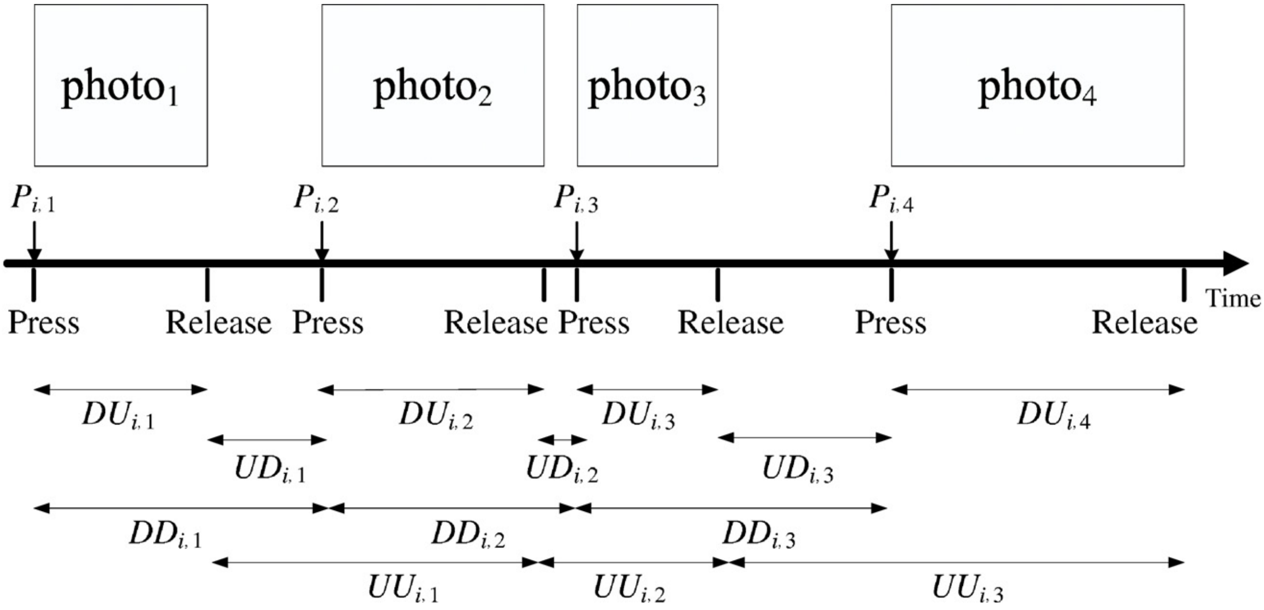


Fig. 4. Keystroke time features and press feature when a user enters a graphical password ‘photo$_1$, photo$_2$, photo$_3$, photo$_4$’.

These sets of the *i*th training sample are denoted as:

$$feat_i = \{DU_i, UD_i, DD_i, P_i\} = \{X_{i,1}, X_{i,2}, \ldots, X_{i,4k-2}\}$$

Note, every user in the system only needs to provide five training samples (i = 1–5) in the enrollment phase, which is smaller than that in Araújo et al.’s suggestion.

3.2. Classifier building phase

The classifier is built to verify the user’s identity after obtaining the personal features. This paper employs a computation-efficient statistical classifier (Boechat et al., 2007). The mean μ_f and the standard deviation σ_f are calculated for each element in $feat_i$ by Eqs. (1) and (2), respectively.

$$\mu_f = \frac{1}{5}\sum_{i=1}^{5} X_{i,f}, \quad \text{where} \quad f = 1 \text{ to } 4k-2. \tag{1}$$

$$\sigma_f = \frac{1}{5-1}\sum_{i=1}^{5} |X_{i,f} - \mu_f|, \quad \text{where} \quad f = 1 \text{ to } 4k-2. \tag{2}$$

The classifier does not require the impostors’ patterns to build. Eqs. (1) and (2) can be calculated efficiently and are suitable for low-power mobile devices.

3.3. Authentication phase

In this phase, the classifier is used to verify the user’s identity. After the user enters his or her graphical password, the system compares the sequence of it with the registered one in the enrollment phase. If it is inconsistent, the system rejects the user’s login request. Otherwise, the corresponding features are examined. An unknown user’s features are denoted as $feat_v = \{DU_v, UD_v, DD_v, P_v\} = \{X_{v,1}, X_{v,2}, \ldots, X_{v,4k-2}\}$ and the system calculates the average distance D between each element in $feat_v$ and $feat_i$ by Eq. (3). The system then accepts or rejects the user’s login based on a threshold t. If $D <= t$, then the user is legitimate. Otherwise, the system rejects the user’s login. That is, a user is able to login to a system only if he or she can be successfully authenticated via the graphical password and the KDA authentications. Thus, even if the password is revealed by shoulder surfing attacks, the probability of breaking the authentication is reduced.

$$D = \frac{1}{4k-2}\sum_{f=1}^{4k-2} \frac{X_{v,f} - \mu_f}{\sigma_f} \tag{3}$$

This phase only analyzes the distance between the user’s login sample and the training samples. Even if the number of users increases, the time for finishing the authentication will not be affected. Similarly, computation in this phase is efficient.

4. Experimental results

This paper provides a graphical-based password KDA system developed by Java language and implemented in Android-compatible devices. The handheld mobile devices used in the experiment were a Motorola Milestone (with an ARM Cortex A8 550 MHz CPU and 256 MB memory), an HTC Desire HD (with a Qualcomm 8255 Snapdragon 1 GHz CPU and 768 MB memory), and a Viewsonic Viewpad (with an Intel Atom N455 1.66 GHz CPU and 1 GB memory). The features in our system were obtained by Android API MotionEvent function library. The press time and the release time were obtained by the getDownTime() and getEventTime() methods, respectively. The system calculated DU, UD, and DD time based on the press time and the release time. The keystroke time features were measured in ms (milliseconds). Android also provided the pressure value via the getPressure() method and it exerted on the device in kilopascals (Meier, 2010) (please see Android function library for more details).

Fawcett (2006) pointed out the KDA system utility can be measured by a ROC (Receiver Operation Characteristic) curve. This paper uses Fawcett’s method to calculate FRR and FAR to build the ROC curve and calculates EER to obtain the most balanced optimum threshold for the system. The EER is the system utility when the optimum threshold is obtained. The ROC curve is better than others if its southwestern point is close to the southwest (FAR is lower, FRR is lower, or both). Fig. 5 shows two ROC curves in the experimental results. The solid line denotes the result that uses time features and the dashed line denotes the result that uses time + pressure features. Fig. 5 shows the dashed line is close to the southwest; in

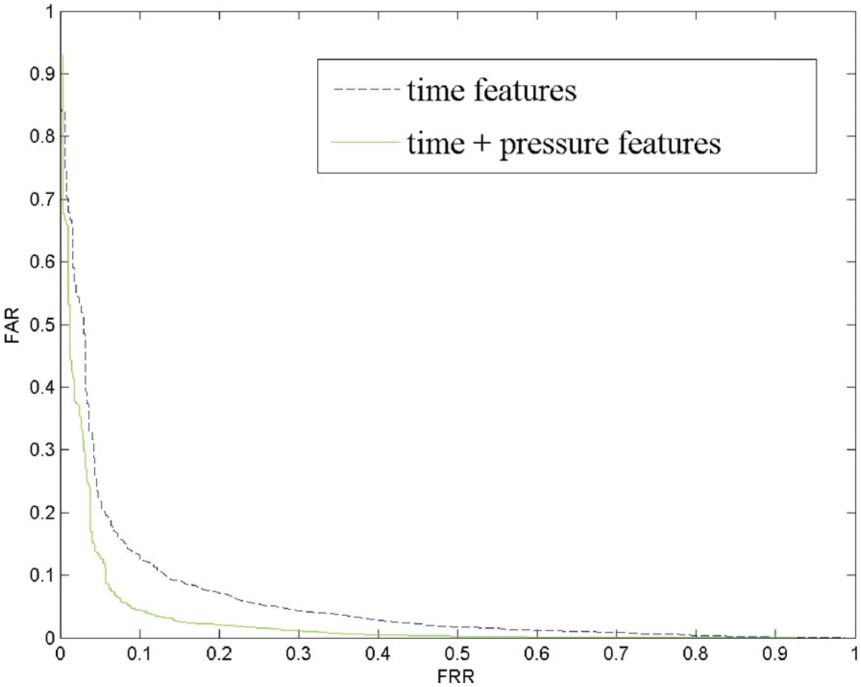


Fig. 5. The experiment results of this paper by ROC.

other words, the pressure feature is useful for promoting system utility.

One hundred participants who were frequent mobile users took part in the experiments. Table 1 lists the age, sex, and which finger was used to input a graphical password.

The one hundred users could freely choose their favorite photos to construct their graphical passwords and provide 10 samples. Five samples were collected at the same time through the same mobile phone (Motorola Milestone 3.7 in. screen) and used in the enrollment phase to build the classifier. The other five samples were collected over a period of five weeks through two mobile devices (HTC Desire HD 4.3 in. screen and Viewsonic Viewpad 10.1 in. screen). These had different screen sizes in the enrollment phase provided for users and for the legitimate user's login test. Similarly, Giot et al.'s (2011) used two keyboards (the original laptop keyboard and a USB keyboard plugged into the laptop) to analyze the system utility. The purpose here was to show screen size and time will not affect system accuracy. The total number of legitimate user samples was 100 × 5 = 500. The total number of impostor samples was 10 × 100 × 5 = 5000, which was obtained by 10 people who were given the graphical passwords of the one hundred users and told to act as an impostor five times (the feature data of our experiment has been linked at http://ty.ncue.edu.tw/N27/data.html). Dividing the number of times the system rejected a legitimate user by the number of legitimate users obtains the FRR; dividing the number of times the system accepted an impostor by the number of impostors obtains the FAR. Table 2 shows the time + pressure feature (t = 2.69, EER = 6.9%) is superior to others. Since the number of legitimate samples and the number of impostor samples differ, the confidence intervals for specific FAR and FRR are provided in the time feature, pressure feature, and time + pressure features. The reader has a 95% level of confidence with the results in Table 2.

Since the related studies of graphical-based password authentication do not combine with the KDA system, Table 3 directly compares the performance with related works (Campisi et al., 2009; Clarke and Furnell, 2007a,b; Hwang et al., 2009a; Saevanee and Bhatarakosol, 2008) that apply the KDA system in mobile devices. In Clarke and Furnell (2007a,b), the EER is a range because they calculate the EER for every person. Here, the EER in this paper and in Campisi et al. (2009) and Hwang et al. (2009a) is a value for the system. That is, the EER is able to analyze the system utility rather than each user utility. Compared with Clarke and Furnell (2007a,b) and Campisi et al. (2009), this paper requires only five training samples for constructing the classifier and the EER is superior. Further, the classifiers building in Clarke and Furnell (2007a,b) require the impostors' patterns and complex computations to train the neural network. Compared with Saevanee and Bhatarakosol (2008), although their EER is better, only 10 users participated in the experiment and more than five training samples were needed. The KNN classifier requires other users' samples to determine which login sample belongs to which user in the system. The EER of Hwang et al. (2009a) is better than our system, but the users have to memorize the number of pauses, the lengths of the pauses, and the positions of the pauses. Without causing any additional burden on users, our system adds the pressure features for improving data quality and the EER of our system is better than all of the above related works. Like the keystroke time features, the keystroke pressure features

Table 1
The data of one hundred users.

Age	
Less than 20	72
21–25	25
More than 26	3
Sex	
Male	62
Female	38
3–6 Thumbnails	
3	66
4	24
5	8
6	2
Which finger is used to input a graphical password	
Right thumb	7
Right index finger	87
Right middle finger	4
Left thumb	1
Left index finger	1

Table 2
The EER with the most balanced optimum threshold of our system.

	Time feature	Pressure feature	Time + pressure features
Threshold t	2.33	2.30	2.69
EER (%)	12.2	14.6	6.9
Legitimate user samples	500	500	500
# of times to reject a legitimate user	61	73	34
FRR (%)	12.2	14.6	6.8
Confidence interval[a]	[0.0927, 0.1513]	[0.1144, 0.1776]	[0.0455, 0.0905]
Impostor samples	5000	5000	5000
# of times to accept an impostor	561	727	346
FAR (%)	11.22	14.54	6.92
Confidence interval[a]	[0.1033, 0.1211]	[0.1354, 0.1554]	[0.0620, 0.0764]

[a] Confident level = 95%, sampling error = $\sqrt{\widehat{p}(1-\widehat{p})/n}$, where $\widehat{p}$ is FRR or FAR, and n is # of legitimate user samples or impostor samples (Chernick, 2007).

Table 3
Comparison of the related works.

	Password	# of training participants	Password space size	# of training samples	Classifier	EER (%)
Clarke and Furnell (2007a,b)	4-digit PIN	30	1×10^4	30	Neural network	9–16
Clarke and Furnell (2007a,b)	11-digit telephone number	30	1×10^{11}	30	Neural network	5–13
Clarke and Furnell (2007a)	6-character text	30	$26^6 \cong 3 \times 10^8$	30	Neural network	15–21
Saevanee and Bhatarakosol (2008)	10-digit number	10	1×10^{10}	20	KNN	1
Campisi et al. (2009)	10-character text	30	$26^{10} \cong 3 \times 10^{14}$	6	Statistical	13
Hwang et al. (2009a)	4-digit PIN	25	1×10^4	5	Statistical	4
This paper (time feature)	3–6 Thumbnails	100	30^3–30^6 = 2.7×10^4 to 7.29×10^8	5	Statistical	12.2
This paper (pressure feature)	3–6 Thumbnails	100	30^3–30^6 = 2.7×10^4 to 7.29×10^8	5	Statistical	14.6
This paper (time + pressure features)	3–6 Thumbnails	100	30^3–30^6 = 2.7×10^4 to 7.29×10^8	5	Statistical	6.9

captured will not cause any additional burden on users. Note, the design of our experiment is via different mobile devices. The experimental results precisely show our system is able to promote the security of PIN-based authentication in mobile devices and system accuracy will be not affected by screen size.

In comparisons of password space size, if the graphical password consists of six photos, the password space size of our system is larger than the 6-character text-based password of Clarke and Furnell (2007a). However, the 11-digit telephone number of Clarke and Furnell (2007a,b) and the 10-character text-based password of Campisi et al. (2009) are larger than ours. On the other hand, the keystroke features are affected special in text-based passwords while the sizes or layouts of the touch panels on mobile devices are inconsistent. These passwords are too long and put a heavy memory burden on users, while the text-based passwords are susceptible to dictionary attacks. Compared with the PIN-based authentications of Hwang et al. (2009a) and Clarke and Furnell (2007a), the password space size is $30^3 = 2.7 \times 10^4$ while a user clicks three photos as his or her graphical password, which is larger than those methods 1×10^4 when a user chooses a 4-digit PIN. In summary, our system is able to enlarge the password space size and improve the security of PIN-based authentication without causing any additional burden on users. In addition, the features captured are not affected by the inconsistent sizes or layouts of keypads on mobile devices.

Since the EER = 6.9% at the threshold t = 2.69 is for the system rather than each user, Table 4 lists the FRR and FAR on different numbers of thumbnails at the threshold t = 2.69. A greater the number of thumbnails to construct a graphical password leads to a larger password space size. According to Wiedenbeck et al.'s research (2005a), graphical password users agreed more strongly than alphanumeric users that the rules for creating a password made it easy to remember the password.

On the other hand, Araújo et al. (2005) suggested the training sample size should not be more than ten; otherwise, the user will feel uncomfortable. The number of training samples in our system is five which satisfies Araùjo et al.'s suggestion and is less than other systems. Fig. 6 analyzes the system utility on different numbers of training samples. The experiment randomly chooses the samples (from 2 to 5, when the number of training samples = 1, Eq. (3) cannot be determined) in the enrollment phase to build the classifier. Obviously, when the sample size increases, the system utility will increase.

Table 5 shows the average time for performing the classifier building phase and the authentication phase in the proposed scheme for different numbers of photos and different combinations of features. The average time for the classifier building phase is between 1 ms and 4 ms and the average time of the authentication phase is between 1 ms and 2 ms. Even for the proposed system uses time and pressure features; the worst times for the classifier

Table 4
Comparison of different numbers of thumbnails.

# of thumbnails	FRR (%)	FAR (%)
3	7.27	5.73
4	8.33	10.67
5	0	7.25
6	0	0

Table 5
The average time (ms) of the building classifier phase and the authentication phase for different numbers of thumbnails for graphical-based passwords and different combinations of features.

	Time feature	Time and pressure features
3 Photos		
Classifier building phase	1	2
Authentication phase	1	1
4 Photos		
Classifier building phase	2	2
Authentication phase	1	1
5 Photos		
Classifier building phase	2	3
Authentication phase	1	1
6 Photos		
Classifier building phase	3	4
Authentication phase	1	2

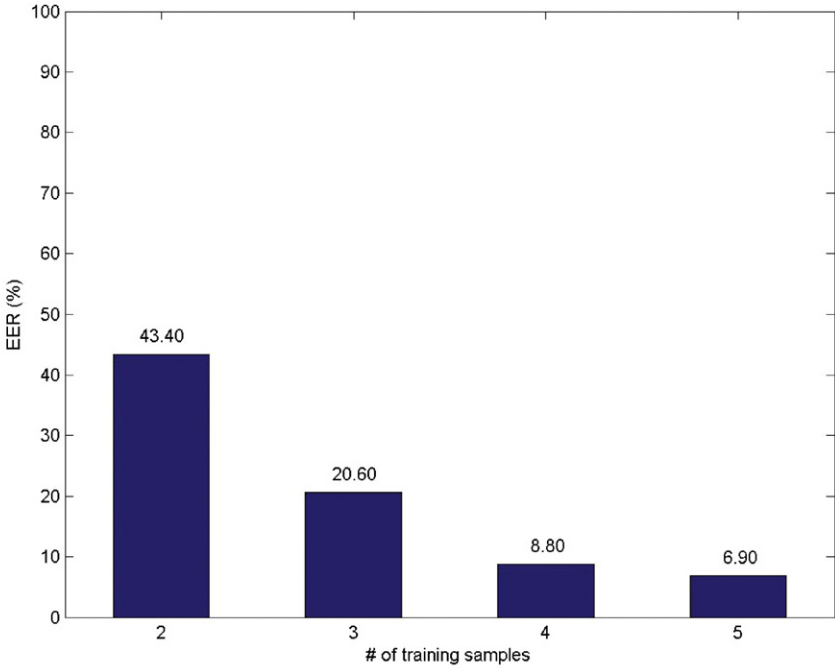


Fig. 6. Comparison of different numbers of training samples.

building phase and the authentication phase were 4 ms and 2 ms, respectively. The pressure features used in the statistical classifier are still suitable for low-power mobile devices.

5. Conclusion

In this paper, we proposed a graphical-based password KDA system for touch screen handheld mobile devices. A user enters his or her graphical password through an identical human–computer interface and therefore the user's keystroke features will not be affected if the user uses different devices. In the experiment, the novel pressure feature applied to touch screens could improve data quality and further promote KDA system utility. The time and pressure features are obtained when a user enters his or her graphical password so this system does not cause any extra burden on users. In our system, a user is able to login to the system when he or she can successfully authenticate via the graphical password and KDA authentications. Namely, even if the graphical password is revealed by a shoulder surfing attack, the probability of breaking the authentication is reduced. Finally, the performance of the proposed system is excellent, and is suitable for low-power mobile devices.

Acknowledgments

I would like to thank the referees for many valuable comments and suggestions which have resulted in several improvements of the presentation of the paper. This research was partially supported by the National Science Council, Taiwan, ROC, under contract nos.: NSC100-2221-E-018-025, NSC100-2221-E-018-034 and NSC100-2622-E-018-004-CC3.

References

Araújo, L.C.F., Sucupira, L.H.R., Lizárraga, M.G., Ling, L.L., Yabu-Uti, J.B.T., 2005. User authentication through typing biometrics features. IEEE Transactions on Signal Processing 53, 851–855.

Bergadano, F., Gunetti, D., Picardi, C., 2002. User authentication through keystroke dynamics. ACM Transactions on Information and System Security 5, 367–397.

Bleha, S.A., Slivinsky, C., Hussien, B., 1990. Computer-access security systems using keystroke dynamics. IEEE Transactions on Pattern Analysis and Machine Intelligence 12, 1217–1222.

Blonder, G.E., 1996. Graphical Passwords. United States Patent 5,559,961.

Boechat, G.C., Ferreira, J.C., Filho, E.C.B.C., 2007. Authentication personal. In: IEEE International Conference on Intelligent and Advanced Systems, pp. 254–256.

Brostoff, S., Sasse, M.A., 2000. Are passfaces more usable than passwords? a field trial investigation. In: People and Computers XIV—Usability or Else: Proceedings of HCI, pp. 405–424.

Campisi, P., Maiorana, E., Bosco, M.L., Neri, A., 2009. User authentication using keystroke dynamics for cellular phones. IET Signal Processing 3, 333–341.

Chang, T.Y., Yang, Y.J. A simple keystroke dynamics-based authentication system using means and standard deviation. Journal of Internet Technology, in press.

Chang, T.Y., Yang, Y.J., Peng, C.C., 2010. A personalized rhythm click-based authentication system. Information Management and Computer Security 18, 72–85.

Chernick, M.R., 2007. Bootstrap Methods: A Guide for Practitioners and Researchers. Wiley Series in Probability and Statistics, 2nd edition. John Wiley and Sons.

Clarke, N.L., Furnell, S.M., 2007a. Advanced user authentication for mobile devices. Computers & Security 26, 109–119.

Clarke, N.L., Furnell, S.M., 2007b. Authenticating mobile phone users using keystroke analysis. International Journal of Information Security 6, 1–14.

Dhamija, R., Perrig, A., 2000. Déjà vu: A user study using images for authentication. In: 9th USENIX Security Symposium.

Fawcett, T., 2006. An introduction to ROC analysis. Pattern Recognition Letters 27, 861–874.

Gaines, R.S., Lisowski, W., Press, S., Shapiro, N.,1980. Authentication by keystroke timing: some preliminary results. In: Rand Report R-256-NSF. Rand Corporation.

Giot, R., El-Abed, M., Hemery, B., Rosenberger, C., 2011. Unconstrained keystroke dynamics authentication with shared secret. Computers & Security 30, 427–445.

Haider, S., Abbas, A., Zaidi, A.K., 2000. A multi-technique approach for user identification through keystroke dynamics. In: IEEE International Conference on System, Man, and Cybernetics, pp. 1336–1341.

Harun, N., Woo, W.L., Dlay, S.S., 2010. Performance of keystroke biometrics authentication system using artificial neural network (ann) and distance classifier method. In: International Conference on Computer and Communication Engineering (ICCCE), pp. 1–6.

Hu, W., Wu, X.P., Wei, G.H., 2010. The security analysis of graphical passwords. In: 2010 International Conference on Communications and Intelligence Information Security, pp. 200–203.

Hwang, S., Cho, S., Park, S., 2009a. Keystroke dynamics-based authentication for mobile devices. Computers & Security 28, 85–93.

Hwang, S., Lee, H., Cho, S., 2009b. Improving authentication accuracy using artificial rhythms and cues for keystroke dynamics-based authentication. Expert Systems with Applications 36, 10649–10656.

Jansen, W.A., 2003. Authenticating users on handheld devices. In: Canadian Information Technology Security Symposium.

Jansen, W.A., 2004. Authenticating mobile device users through image selection. In: Data Security.

Jermyn, I., Mayer, A., Monrose, F., Reiter, M.K., Rubin, A.D., 1999. The design and analysis of graphical passwords. In: 8th USENIX Security Symposium.

Killourhy, K.S., Maxion, R.A., 2009. Comparing anomaly-detection algorithms for keystroke dynamics. In: IEEE/IFIP International Conference on Dependable Systems & Networks, pp. 125–134.

Mandler, J.M., Ritchey, G.H., 1977. Long-term memory for pictures. Journal of Experimental Psychology: Human Learning and Memory 3, 386–396.

Meier, R., 2010. Professional Android 2 Application Development. John Wiley & Sons Inc.

Revett, K., de Magalháes, S.T., Santos, H.M.D., 2005. Enhancing login security through the use of keystroke input dynamics. In: Advances in Biometrics, Lecture Notes in Computer Science 3832, pp. 661–667.

Ru, W.G.D., Eloff, J.H.P., 1997. Enhanced password authentication through fuzzy logic. IEEE Expert: Intelligent Systems and Their Applications 12, 38–45.

Saevanee, H., Bhatarakosol, P., 2008. User authentication using combination of behavioral biometrics over the touchpad acting like touch screen of mobile device. In: Computer and Electrical Engineering, pp. 82–86.

Shih, D.H., Lin, T.C., 2008. User authentication system by using keystroke dynamics. In: International Conference on Pacific Rim Management 18th Annual Meeting, pp. 102–115.

Sobrado, L., Birget, J.C., 2002. Graphical passwords. In: The Rutgers Scholar, An Electronic Bulletin for Undergraduate Research.

Syukri, A.F., Okamoto, E., Mambo, M., 1998. A user identification system using signature written with mouse. In: 3rd Australasian Conference on Information Security and Privacy, pp. 403–441.

Wiedenbeck, S., Waters, J., Birget, J.C., Brodskiy, A., Memon, N., 2005a. Authentication using graphical passwords: basic results. In: 11th Human–Computer Interaction International (HCII).

Wiedenbeck, S., Waters, J., Birget, J.C., Brodskiy, A., Memon, N., 2005b. Passpoints: design and longitudinal evaluation of a graphical password system. International Journal of Human–Computer Studies 63, 102–127.

Xi, K., Tang, Y., Hu, J., 2011. Correlation keystroke verification scheme for user access control in cloud computing environment. The Computer Journal 54, 1632–1644.

Ting-Yi Chang received his M.S. from the Graduate Institute of Computer Science and Information Engineering at Chaoyang University of Technology, and his Ph.D in the Department of Computer Science at National Chiao Tung University, Taiwan. Currently, he is an Associate Professor with the Graduate Institute of e-Learning, National Changhua University, Taiwan. His current research interests include artificial intelligence, e-Learning, information security, cryptography, and mobile communications.

Cheng-Jung Tsai was born in Tainan, Taiwan, Republic of China, 1973. He received the B.S. degree in mathematics and science education from National Ping Tung University of Education in 1995, the M.S. degree in information education from National University of Tainan in 2000, and the Ph.D degree in computer science and information engineering from National Chiao Tung University in 2008. Currently, he is an Assistant Professor in the Graduate Institute of Statistics and Information Sciencethe, National Changhua University, Taiwan. His current research interests include data mining, information security, and e-learning.

Jyun-Hao Lin received his M.S. degree from the Graduate Institute of e-Learning at National Changhua University, Taiwan. His research interests include keystroke dynamic systems.