Authentication Method through Keystrokes Measurement of Mobile users in Cloud Environment

Mahnoush Babaeizadeh, Majid Bakhtiari, and Mohd Aizaini Maarof

Department of Computer Science, Faculty of Computing, Universiti Teknologi Malaysia, Skudai 81310, Johor, Malaysia
e-mail: mahnoush.b@gmail.com, bakhtiari@utm.my, aizaini@utm.my

Abstract

In order to deal with security of Mobile Cloud Computing (MCC) authentication plays an important role. It is aimed to verify user's identity when they wish to request services via the Internet through Cloud Service Provider (CSP). Personal identification number is the most common mechanism for authentication in mobile devices; however, it is not secure way for authenticating users. This work presents a new behavioral biometric authentication method which enable identify users based on Keystroke Dynamic Authentication (KDA). Furthermore, keystrokes duration is considered as an attribute for measuring keystrokes of mobile. Elliptic Curve Cryptography (ECC) is used as cryptography technique to improve the security of proposed method. Simulation results using JUnit package revealed that, even if an unauthorized person knows the username and password of legal user they cannot gain access rights on 97.33% of efforts. This is due to the keystroke duration of each user which depends on their behavioral characteristic. Therefore, applying this method, it becomes very difficult for an attacker to pretend as the owner. Hence, this method offers the potential to enhance the security of authentication in MCC.

Keywords: Keystroke authentication, Biometric authentication, Mobile cloud computing, Security and privacy.

1 Introduction

Authentication is a main part of every secure communication system especially in wide spread network such as MCC [1-3]. It helps to protect shared information from unauthorized persons, and it is a key technology for information security. AAA is a management module for authentication, authorization, and accounting. When a user tries to access CSP, then AAA checks the user's authentication information. If the user is authenticated, then AAA gets the user's access level,
which has been most recently generated, by inspecting the user's information in the database. In addition, authentication method determine “Who is the legal user” and “Is the user really who he claims himself to be”. In addition, verification of user’s identity is the most important goal behind an authentication.

In other words, an authentication mechanism is a main challenge in security and privacy of cloud users [4-6]. It determines how user identified and verified to access to sensitive information [7]. Verification of user’s identity is the most important goal behind an authentication. PIN is adopted as the only security mechanism for mobile devices. It is obvious that, PIN (something the user knows) is not a very secure mechanism for authenticating users because of its limitation and its difficulty to confirm that the demand is from the rightful owner [8,9].

Strong method of authentication should cover one or several various factors of identification to improve security. These factors are as following:

i. Something we know
ii. Something we have
iii. Something we are

Therefore, biometric authentication [10, 11] is a strong authentication mechanism by providing the factor what we are and what we know [12]. In addition, it is able to identify users based on their unique characteristic [13], and it is more reliable, because it is so difficult for user to pretend as other user by using physical or behavioral biometric authentication.

Keystroke authentication is a type of behavioral biometric authentication. Keystroke based authentication can be categorized in two folds, Keystroke Static Authentication (KSA), as well as Keystroke Dynamic Authentication (KDA). KSA can identify keystroke of users only at particular times, for example the time that user wants to login. This is a huge drawback of KSA; due to system can use by anyone once the user is authenticated at login [14,15]. Static authentication provides more strong and robust user authentication than simple password or PIN; however it cannot keep continuous security. KDA continuously observes the style of typing of the users throughout the whole stage of interaction even after a successful login. In other words, the typing patterns of users are constantly analyzed and when they do not match accessing of users will block [15, 16].

The main goal of KDA is recognizing mobile users by identifying and analyzing their unique feature for authentication such as typing pressure, keystroke duration, typing error, and latency of keystrokes.

KDA has some advantages rather than other types of biometric authentication [17, 18]. These advantages are:
i. Contrasting other biometric methods, KDA does not need any additional tools, therefore it causes to decrease price

ii. High acceptability between mobile users due to it is natural for everybody to type a password for authentication purposes

iii. Preserving privacy and security of users because it is based on behavioral characteristic of users

iv. It could not be forgotten, stolen or lost

This paper presents an android application which used secure type of keystroke dynamic authentication. Furthermore, keystrokes duration is considered as parameter for measuring keystrokes of mobile users. Due to improve the security of communication between mobile user and CSP, TTP is used. It means that, all of security service such as cryptography techniques is done in TTP. ECC is an asymmetric cryptography technique which has rapid computation with the smaller key size, and helps to improve the security of the communication. Moreover, Gaussian Probability Density is used as a function for calculating the similarity of feature (keystroke duration).

The organization of this paper is as follows: Section 2 discusses the various researches related to biometric authentication, as well as keystroke base authentication. Description about the proposed method is in Section 3. Description results obtained from applying KDA in CSP using Android application development bring in Section 4; finally the conclusions are given in Section 5.

2 Literature Review

Biometric authentication supports the three important factors of information security. These factors are authentication, identification, and non-repudiation. It is an ancient Greek word bios ="life" and metron ="measure". Biometric authentication [11, 15] is an authentication mechanism that identify users base on measuring their unique characteristic. In other words, biometric authentication is based on verifying personal attributes of users [10]. Figure1 shows objectives of biometric authentication are security, cost, computation speed, accuracy, user acceptance, and environment constraints.

Two major vulnerabilities which specially have need more attention in the context of biometric authentication are “spoof attacks” at the user interface along with “template database leakage.” A spoof attack involves presenting an imitation biometric trait not obtained from a live person. Template database leakage implies on the situation which valid user’s biometric template information becomes available to an adversary.
As shown in Figure 2, biometric authentication can categorize in two types. These types are physical biometric and behavioral biometric [10, 18]. Physiological biometric relies on something the users are. It performs authentication base on physical characteristics such as facial features [19-21], palm prints [22-24], retinal patterns [11, 25], finger print [11, 26], iris pattern [21], as well as hand geometry [11, 27]. In other words, physiological biometrics is based on measurements and data derived from direct measurement of a part of the human body. Behavioral biometrics is based on the user’s behavior and authentication may occur perfectly such as signature, keystroke dynamics [28] and voice. Furthermore, voice can considered as physiological biometric. One advantages of behavioral biometrics are that they can be applied in a transparent and continuous authentication system [11, 29].

Biometric authentication has some benefit as compared to other techniques [30]. These include,

i. Biometric characteristics are uniquely individual (something you are)
ii. Non-transferable to others
iii. Impossible to forget or lose
iv. Difficult to reproduce
v. Usable with or without the knowledge
vi. Complicate to alter or modify.

Keystrokes Dynamic Authentication (KDA) is a type of behavioral biometric authentication which developed in the late 19th century. It is based on style of each person’s typing on a keyboard, as well as it can identify the user based on their habitual typing pattern. Keystroke dynamics implies on the process of measuring human’s typing rhythm on digital devices such as mobile devices [9]. Categorization of keystroke dynamic is shown in Fig.3. In this type of
authentication, it is not important what you type; the important point is how you type. Furthermore, keystroke technology can be easily integrated with existing technology environments and processes [31].

Yu et al. [32] proposed the solutions to enhance identity verification by applying an SVM novelty detector, proposing GAe-SVM, as well as an ensemble creation based on characteristic selection. Later, Bartlow and Bojan [33] developed a web application to use keystrokes dynamics by incorporating shift-key patterns. It helps to achieved 14% FRR. Clarke et al. [34] improved biometric authentication by using keystrokes analysis. They utilize telephone number and text message obtained from mobile phone for authenticate users. They compare performance of different type of biometric techniques. It is show that, hand geometry identification with 1.5% EER has lowest performance and vein mechanism has highest performance with 5% EER.

Lee and Sungzoon [35] presented the retraining framework to enhance authentication accuracy. Minetti et al. [36] founded a positional dependence of the relationship between the applied force and the resulting down stroke speed due to the different hammer mass to be accelerated. Kang et al. [37] enhanced quality of data obtained from keystrokes of users by using artificial rhythms and cues. Artificial rhythms cause increasing the uniqueness, as well as raising the stability. Briggs and Olivier [38] recommended creating “biometric daemons” for authenticating users. This method is based on learning user’s behavior.

Hwang et al. [39] utilized artificial rhythms to overwhelm problems resulting from short PIN length. Through the experiment involving human subjects, they decreased the error from 13% to 4%. They measured performance in terms of Equal Error Rate (EER). Giot et al. [17] presented comparative research on various method of keystroke dynamic on keyboard. Moreover, they considered the
operational constraints of use for collaborative systems. Karnan et al. provided an overview on biometric authentication. In fact, their research was on well-known approaches of keystroke dynamics in the last two decades [10]. Chang et al. [40] proposed a method of authentication that is a combination of neural network technique and password keystroke features to produce a long-lived private key dynamically. This method improved security of long-lived private key. In fact, it decreased the likelihood of accessing the private by an unauthorized person.

![Keystroke Dynamic Categorization](image)

**Fig. 3** Categorization of keystroke dynamic

Bours [41] experimented with the new idea of continuous keystroke dynamics. This authentication mechanism can continuously monitor the typing behavior of a user and verify that the request is from the legal user. Furthermore, Wang et al. [42] introduced a new user authentication approach by using keystroke dynamic method. This method includes two parts which are training and authentication. Training contains a set of feature vector generated from keystroke characteristics of valid user that successfully authentications. Authentication consists of, current keystroke feature vector of the set orthogonal bases. It has better performance in terms of False Acceptance Rate (FAR) and False Rejection Rate (FRR).

Chang *et al.* [43] suggested a new graphical based password KDA system for identifying users in touch screen mobile devices. In addition, they applied pressure feature in their system. The results are shown, this method cause to develop EER to 12.2%. Moreover, using pressure and time features in this system helps to decrease EER to 6.9%. Therefore, this system is suitable for low-power mobile devices. Teh et al. [9] provides a survey on keystroke dynamics biometrics authentication. It covers research performed during the last three decades, as well as proposing some future works in this approach. Nauman et al. [44] proposed a protocol for keystroke dynamics analysis which allows web-based applications to
make use of remote attestation and delegated keystroke analysis. Moreover, they presented a prototype implementation of their protocol using Android operating system. Bhatt and Santhanam have presented a survey paper that explained about the researches work on keystroke dynamics, as well as they discussed advantages and disadvantages of these researches [16].

The aforementioned contributions have mainly been focused on performance metrics while attributes measures have been rarely considered. Therefore, there is a gap in literature regarding keystrokes method. Furthermore, to the best of authors' knowledge, the keystroke duration has not been considered as parameters for measuring keystrokes of mobile users. This paper is aimed to address this gap. In addition, the existing authentication methods employ RSA, DES or AES [45] as cryptography technique in MCC, however in this work ECC method is employed to enhance the security of the proposed method.

3 The Proposed Method

This section describes the proposed method for MCC. MCC is an infrastructure where both data processing and data storage occur outside of mobile set [5]. Therefore, the mobile device does not require powerful Control Processing Unit (CPU) and memory capacity. In this paper, Google drive (CSP) is considered as data storage. As mentioned before the proposed method is based on keystroke authentication. It is a type of behavioral biometric authentication. There are different parameters for measuring keystrokes of mobile users which can be categories in two main parts as attributes measured and performance metrics [10].

Attributes Measured are

i. Latencies between successive keystrokes (flight time)
ii. Keystroke durations (dwell time)
iii. Finger placement and applied pressure on the keys
iv. Typing speed
v. Use of additional keys in the keyboard
vi. The order in which the keys are pressed

Performance Metrics are

i. False Rejection Rate (FRR): The percentage of attempts of wrongly recognizing a legitimate user as an imposter
ii. False Acceptance Rate (FAR): The percentage of attempts of wrongly allowing an impostor to access the system
iii. Equal Error Rate (EER): The point at which FRR equals FAR
In this method, keystroke duration is considered as an attribute to measure keystrokes of users. As shown in Figure 4, this method is multi-factor authentication. It uses username/password as well as keystrokes authentication to identify users. The process consists of the following steps. In first time login, user has to insert username/password to login to the application, after inserting password the application measures keystrokes duration. The calculated keystroke value will then be encrypted. The security service is neither deployed in mobile user nor CSP, but in the Trusted Third Partner (TTP). It provides the basic security service, such as the data encryption and decryption, and the process of calculating similarity of features. Finally, all the values (username, password, keystrokes duration) will send to CSP.

There are various methods of cryptography which can be categorized as follow,

i. Symmetric (Secret key encipherment or secret key cryptography)

ii. Asymmetric (Public key encipherment or public key cryptography)

iii. Hashing

Symmetric key encipherment uses a single secret key for both encryption and decryption. It is based on sharing secret key between two sides and symbols are permuted or substituted. These technics are Data Encryption Standard (DES), Advanced Encryption Standard (AES), Electronic Code Book (ECB), Cipher Feedback (CFB), and Output Feedback (OFB).

Asymmetric key encipherment uses two keys instead of one, public key and private key. Public key is available to anyone who might send a message, and private key is kept secret. It is based on personal secrecy (each person creates and keeps own secret). These techniques include ElGamal, Rivest Shamir Adleman (RSA), and Elliptic Curve Cryptography (ECC).
Hashing is capable to generate a fixed length message digest from variable length message. The digest is much smaller than the message. These techniques include SHA-512, whirlpool and so on.

In this paper, ECC technique is used as a cryptography technique. As it is a type of asymmetric encipherment, and it capable to generate more efficient and secure cryptographic keys. In compare with RSA and ElGamal, ECC has the same level of security with smaller key size, and rapid computation. The general equation for an elliptic curve is in Equation (1).

$$y^2 + b_2xy + b_3y = x^3 + a_1x^2 + a_2x + a_3$$  \hspace{1cm} (1)

Elliptic curve over real numbers use a special class of elliptic curves of the form Equation (2).

$$y^2 = x^3 + ax + b$$  \hspace{1cm} (2)

The process of ECC technique can be divided into three parts as follows,

i. Key generation: CSP (the receiver) chooses an elliptic curve, a base point and private key (some number or mathematical procedure that can be applied to data), and then it generates and publishes public key.

ii. ECC encryption: Mobile user (the sender) receives CSP's public key, and encrypts the plaintext (value of keystroke duration) with this key. After that she sends the cipher text to CSP.

iii. ECC decryption: CSP decrypts the cipher text with its private key to obtain mobile user’s plaintext.

As shown in Figure 5, after storing the values obtained from mobile device inserted values have to be compared with the stored values for the next login. Calculating similarity of keystrokes is based on Gaussian Probability Density (GPD) function. Moreover, there are various cloud servers such as Dropbox, Amazon, Sky Drive, Box, as well as Google drive which provide different types of cloud computing services (Software as a Service (SaaS), Platform as a Service (PaaS), and Infrastructure as a Service (IaaS)). In this paper Google Drive is used as an IaaS.

There are different way for calculating feature matching or similarity of keystrokes, such as GPD [46], and Direction Similarity Measure (DSM) function. In this paper, similarity between keystroke duration of mobile's user and stored value in database is calculated using GPD. Algorithm (1) and Algorithm (2) show
pseudo code of the proposed method. If this is the first time applying the method, Algorithm (1) is invoked, otherwise Algorithm (2) is used.

In proposed method, fixed username and password are defined for the first time login. Moreover, there is a constant parameter which is related to the limitation of login attempt. At most user can try three times attempt to login to the application otherwise they will be blocked. Furthermore, after successfully login, pseudo-random session ID will create and store in database. It means that, user is login until session expires. It helps to continuously check validity of user.

![Fig. 5 Process of calculating similarity of values in mobile device and CSP](image)

When user wants to run an application for the first time login button is inactive. They have to change fix password and may username to allow login to the application. In the time users change their password keystroke duration can measure. As mentioned before values of username, password, as well as keystroke duration will store in to the CSP.

![Fig. 6 Process of changing password](image)
Figure 6 shows the process of changing password. For changing password, user should insert old password. If old password is match, user can insert new password. After that, they should insert confirm password. In the time user inserts confirmation of password, keystroke duration of them will measure. An important point is that, this method is more accurate because it capable to calculate the keystrokes duration up to milliseconds. At last, new password and value of keystroke duration send to CSP via TTP.

---

**Algorithm 1**

//Change password
Insert old password
   if password is incorrect then
      Output (“Password is incorrect”)  
      Exit
   else
      Insert new password, confirm password,
      Measuring keystrokes duration
      if password match then
         Output (“Password is changed”)
         Store new password and keystrokes
duration in CSP
      else
         Output (“Password is not match”)  
      end if
   end if

---

Fig. 7 login to the Application
As mentioned before, for improving the security of communication between mobile device and CSP (Google Drive) and protecting shared data against attackers a cryptography technique is used. In addition, user should insert their password to allow changing username. However, it is not compulsory to change user name. When password and may be username changed, mobile user is allowed to login to the application. As shown in Figure 7, after inserting username and password by user, the application can compare inserted username and password as well as keystrokes duration with stored values. If they match user can login successfully.

```
Algorithm 2

if like to change username then  //change username
  Insert old password
  if password is correct then
    Insert new username
    Output: (“Username is changed”) Store new username in CSP
  else
    Output (“password is incorrect”)
  end if
else    //login to the application
  Insert username
  if username matches with the stored value then
    Insert password and Measuring keystrokes duration
    Calculate the similarity using Equation (3) // GPD function
    if keystrokes duration is similar with stored value then
      Output (“Successfully login”)
      Create pseudo-random session ID and Store it
      if session exists on database then
        User is logged in        Count= Count+1
      else
        Output (“login again”)
      end if
    else
      Output ("You cannot login")
    end if
  else
    Output ("You cannot login")
  end if
else
  Output ("You cannot login")
end if
```
In addition, session expiration time helps to expire the session of user and force them to login again. It helps to continuously monitor keystrokes duration of user through the whole stage of interaction even after successful login, as well as identifying unauthorized person after verification.

In practical situation keystrokes duration does not exactly equal with the stored value of keystrokes duration. Therefore, parameter $\sigma$ is considered as tolerance. It means that keystrokes duration should be in range of defined tolerance to successfully login to the application, otherwise application does not allow user to login. Parameter $\sigma$ is set based on keystrokes duration of user which is calculated on average after several times login to the application.

4 Experimental Results

Previous section was explained the proposed method of authentication in MCC. The purpose of this section is to discuss about the performance of this method. Results obtained from communication of mobile user with CSP (Google drive) through TTP (Security server). In order to test the performance of proposed method JUnit was used. It is a unit testing framework of Eclipse which is used for test and analysis of obtained results. In addition, the keystrokes duration of legal user was set at 4.048 second, the tolerance parameter $\sigma$ was set at 529 milliseconds namely user may successfully login to the application, if their keystrokes duration is between 3.548 second and 4.548 second.

As mentioned above there are different method for calculating the similarity of features in keystrokes authentication and in this paper, GPD is used. Generally Gaussian function is defined as Equation (3).

$$\text{Similarity}_{GPD} = \frac{2^{\frac{-(T-\mu)^2}{2\sigma^2}}}{\sqrt{2\pi}\sigma}$$  \hspace{1cm} (3)

Where the variable in Equation (3) are as follow,

- $T$: test data’s latency of a particular character
- $\mu$: the mean of keystroke duration
- $k$: Total number of keystroke in password
- $\sigma$: Standard deviation of a keystroke duration
When user want to log in to the application, the obtained value of keystrokes duration will compare with the stored value in CSP using Equation (3).

An important point is that username and password of the mobile’s owner were known in the time of testing. In this experimental result, there are three important factors, username, password, as well as keystrokes duration. In addition, zero is considered for incorrect value and one for correct value of each factor. For example “000” means username and password and keystrokes duration are inserted incorrectly, “110” means username and password are inserted correctly however keystrokes duration is not in the range of keystrokes duration of mobile’s owner.

In our experimental result keystrokes duration of owner was KD=4.679 seconds and $\sigma$ is equal to 529 milliseconds (these variables are based on repeated login by mobile user). Therefore, acceptable range of keystrokes duration is between KD-$\sigma$ to KD+$\sigma$. It is obvious that, the main reasons of preventing unauthorized access relate to the situation that username and password are correct but keystrokes duration is not match with the stored value in CSP (110). It means that keystrokes duration can prevent unauthorized person to login to the application and access to the shared information in to the CSP. This is due to the fact that it is difficult for unauthorized person or attackers to pretend as an owner.

The reasons of unsuccessfully login are as follow,

i. Username and password are correct; however keystrokes duration is not correct

ii. Username and keystrokes duration are correct, but password is incorrect

iii. Username is correct, but password and keystrokes duration are incorrect

iv. Username, password, and keystrokes duration are correct

v. Username, password, as well as keystrokes duration are incorrect
As shown in Figure 8, 9 the main reason of preventing unauthorized users to pretend as a legal user is keystrokes duration.

![Pie chart of different options mentioned upper](image)

**Fig. 9** Pie chart of different options mentioned upper

![Keystrokes duration](image)

**Fig. 10** Keystrokes duration

Figure 10 shows acceptable and unacceptable range of keystrokes duration. There are four black points that put in the acceptable range. It means that in these tests the method works incorrectly. Three of them occurred, when username was correct, password was incorrect, keystrokes duration was correct. Another one related to the situation that username, password as well as keystrokes duration was correct.

Based on experimental results, proposed method of authentication has 97.33% successful performance when performing 150 tests. This suggests that the proposed method is reliable enough to prevent unauthorized user from login and access to shared information in cloud server.


5 Conclusion

This paper addresses using keystroke authentication in MCC. It has described a secure biometric method based on measuring keystrokes duration. It helps to identify users based on their unique behavioral biometric and unlike the other biometric methods, Keystroke analysis does not require the aid of extra special tools. Therefore it is cheaper than other type of biometric authentication methods. Experimental results show that this method can work 97.33% correctly in authenticating mobile’s users, and it helps to improve the security and privacy of authentication in mobile communication. Some ideas of potential future investigation exist in this area; there are some other parameters for measuring keystrokes like as finger placement and applied pressure on the keys, and it is suggested that performance metric can use for measuring keystrokes. This naturally brings the idea of improving KDA by measuring False Rejection Rate (FRR) which is percentage of attempts of wrongly recognizing a legitimate user as an imposter.

References


111

Authentication Method through Keystrokes


