International Journal of Innovative Computing, Information and Control
Volume **8**, Number **11**, November **2012**

ICIC International ©2012 ISSN 1349-4198
pp. 7875–7886

AN APPROACH FOR USER AUTHENTICATION ON NON-KEYBOARD DEVICES USING MOUSE CLICK CHARACTERISTICS AND STATISTICAL-BASED CLASSIFICATION

Cheng-Jung Tsai[1,*], Ting-Yi Chang[2], Yu-Ju Yang[2]
Meng-Sung Wu[3] and Yu-Chiang Li[4]

[1]Graduate Institute of Statistics and Information Science
[2]Graduate Institute of e-Learning
National Changhua University of Education
No. 1, Jin-De Road, Chang-Hua 500, Taiwan
*Corresponding author: cjtsai@cc.ncue.edu.tw; tychang@cc.ncue.edu.tw

[3]Institute of Information Science
Academia Sinica
No. 128, Academia Road, Section 2, Nankang, Taipei 115, Taiwan
wums@iis.sinica.edu.tw

[4]Department of Computer Science and Information Engineering
Southern Taiwan University
No. 1, Nan-Tai Street, Yungkang Dist., Tainan 710, Taiwan
lyc002@mail.stut.edu.tw

Received July 2011; revised December 2011

Abstract. *Internet security has become a serious issue for anyone connected to the Internet. To avoid unauthorized people accessing an information system, keystroke dynamics-based authentication (KDA) systems combine password knowledge with typing characteristics to enhance the security of general password authentication systems. However, some portable computational devices have no computer keyboard, for example, personal digital assistants and mobile phones. That is, KDA systems cannot successfully work while the enrollment phase is implemented based on a standard desktop keyboard. This reduces the portability of the KDA system. This paper adopts rhythms clicked by a mouse as another identifiable factor. Mouse clicks can be replaced by a stylus on non-keyboard device, numeral buttons on mobile phones, or fingers on touch screens to enhance system portability. In our proposed system, the click data are based on the time instances during pressing and releasing the mouse button. Five features based on these time periods are calculated using this data. We invited twenty-five users to participate in our experiment. The experimental results showed that our authentication system can achieve a good accuracy. Our experiments also showed that the rhythm clicked by a mouse can function as the second identifiable factor in general password authentication systems or as the standby identifiable factor in KDA systems.*

Keywords: Classification, Biometric characteristics, User authentication, Keystroke dynamics, Click dynamics

1. Introduction.

As the fast process of computer and network technologies, computers connected over the Internet had become the indispensable electrical appliances of our daily life. When users surf on the Internet, hackers may gain access to their computers or personal digital assistants. Internet security therefore has become a serious issue for anyone connected to the Internet. Password authentication is one of the simplest and most common user authentication mechanisms used to provide basic computer security on demand. In general password authentication systems, the validity of passwords is the main identifiable factor. Since people prefer easily recognizable, natural language phrases, their passwords are usually drawn from a rather limited set of possibilities and susceptible to password guessing attacks [1]. If an impostor obtains a use's password, that impostor is able to masquerade as the user and endanger the system's security.

To guard against unauthorized account access, biometric characteristics are used to identify individuals based on their biological or behavioral characteristics [2]. Keystroke dynamics are one biometric technique that does not require other special devices such as fingerprint, iris, signature scanners, or face scanner [3]. Statistics [4-11], neural networks [12-15], fuzzy logic [16, 17], support vector machines [18, 19], k-nearest neighbor [20-22], and other classifiers [20-24] have all been developed into keystroke dynamics-based authentication (KDA) systems, combined with password knowledge and functional typing characteristics as the second identifiable factor, to achieve higher system accuracy. However, there are potential threat factors that reduce the accuracy and portability of the KDA system. That is, some portable computational devices do not have standard desktop keyboards such as personal digital assistants and mobile phones, resulting in a reduction in system portability if the enrolment phase is implemented based on the desktop keyboard. Differences in computers, for example, desktop versus laptop, may lead to significantly different typing performance [28] and therefore affect the accuracy of KDA systems. Recently, artificial rhythms have been used to improve the keystroke data quality, for instance, a pause between characters [29]. The user in this system should memorize the locations of pauses inserted in his password. Because this is not an innate typing characteristic, a personalized rhythm click dynamics-based authentication system is proposed in this paper.

In previous studies, a common mouse was used as an input device to authenticate users. That is, drawing a circle or other figures with a mouse was used to authenticate the identities of users [30]. This system obtained an average false rate (AFR) of 0.11. [31] utilized a signature written by a mouse to collect the user's biometric data resulting in an AFR of 0.055. Unfortunately, since users are not familiar with writing a signature using a mouse, they have to practice more than eighty times to achieve signing consistency. It is difficult to obtain the same accuracy as an actual signature. We have combined neural network and click dynamics for authentication; however, the proposed system is impractical since it requires the data of impostors [15]. [32] used mouse movements to authenticate users. They obtained a remarkable AFR of 0.0246315. In their scheme, the users installed data collection software on their machines and conducted their usual activities without any restriction. However, the data collection period per user is conducted over a long period of time.

According to the above descriptions an efficient, practical and feasible system for conveniently collecting user biometric data and precisely authenticating their identity is needed. This paper uses a biometric, the personalized rhythm, as the second identifiable factor for authenticating users. We hypothesize that each person clicks a mouse in a characteristic way. Users are authenticated using biometric click data. The data can be simply and conveniently captured using most devices, and require little time to collect and quickly verify identities. For example, a stylus or fingers on a touch screen and numerical pad can be used to replace the rhythms clicked using a mouse. The methodology of this system is low cost. The click features analyzed are the duration of each click and four click latencies between successive clicks, which are conceptually similar to keystroke features. During the enrolment phase ten samples are collected from each user [4]. Then a statistical-based classifier [33], which does not require the data from impostors, is used to examine the usefulness of the rhythm click-dynamic authentication system based on mouse clicks. Our experiment in Section 3 showed that the combination of rhythms clicked using a mouse and our statistical-based classifier forms a feasible and practical KDA authentication system.

This paper is organized as follows. Section 2 presents the architecture of our authentication system as well as the evaluation metrics and the data collection. Section 3 presents the experimental analysis. Conclusions and the future works are presented in Section 4.

2. The Authentication System.

2.1. The framework of our system.

Figure 1 shows the framework for the main steps of our authentication system. In the enrolment phase the user initializes an account and clicks the target rhythm several times. The click data are captured while the user is clicking and the system checks the number of clicks for the target rhythm. If the number is wrong, the user will be required to click again and the wrong data will be filtered out. A template of the click pattern as well as other related information are computed and stored in a database. In the authentication phase an unknown user tries to access the system and is required to enter the password and click the target rhythm. Here, we assume the unauthorized user knows the password and he/she also clicks the correct target rhythm in the experiment. The system uses only the rhythm clicking characteristic with the designed classifier to determine whether an unknown user is allowed to access the system. Since legitimate users usually fail in the first attempt for authentication [4], a user will be given a second attempt if he or she fails in the first attempt in this scheme.

2.2. A personalized rhythm.

The rhythm used in this paper is well-known as "Encourage with Love" in Taiwan or "Rainbow Claps" in Singapore. Its beats are 2-3-4-2 with a total length of 11 – the first musical note from the left is a quarter note, the third musical note from the right is an eighth note, and the sixth musical note from the left

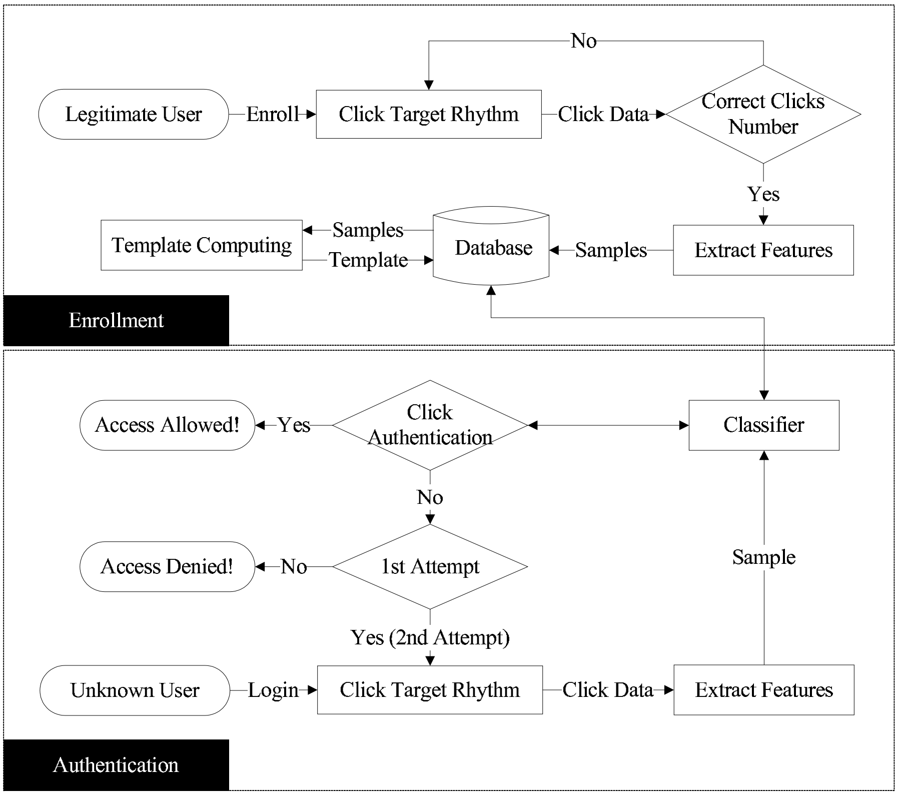

FIGURE 1. The framework of the proposed system [5]

is an eighth rest. This rhythm is used in our experiments since it is very familiar to all participants and provides higher accuracy than unfamiliar rhythms. Our system requires the unknown user to click the same target rhythm ten times in the authentication phase.

2.3. **The mouse click data and features.**

Click dynamics studies the way a user interacts with a mouse or stylus. A mouse event includes the mouse button down and up, producing five features DU, DD, UD, UU, DU2 defined as follows.

(1) *Down-Up* (DU) time: DU time is the interval between the same click being pressed and being released.
(2) *Down-Down* (DD) time: DD time is the interval between the click being pressed and the next click being pressed.
(3) *Up-Down* (UD) time: UD time is the interval between the click being released and the next click being pressed.
(4) *Up-Up* (UU) time: UU time is the interval between the click being released and the next click being released.
(5) *Down-Up*2 (DU2) time: DU2 time is the interval between the click being pressed and the next click being released.

The five features are conceptually similar to the keystroke features. DU time is a feature of any one click. DD, UD, UU and DU2 time are the click latency features relating to any two consecutive clicks. The DU, DD, UD, UU and DU2 time sets of the sample s account a are defined as:

$$\begin{aligned}
\mathrm{DU}_{a,s} &= \{du_1(a,s), du_2(a,s), \ldots, du_n(a,s)\}, \\
\mathrm{DD}_{a,s} &= \{dd_1(a,s), dd_2(a,s), \ldots, dd_{n-1}(a,s)\}, \\
\mathrm{UD}_{a,s} &= \{ud_1(a,s), ud_2(a,s), \ldots, ud_{n-1}(a,s)\}, \\
\mathrm{UU}_{a,s} &= \{uu_1(a,s), uu_2(a,s), \ldots, uu_{n-1}(a,s)\}, \\
\mathrm{DU2}_{a,s} &= \{du2_1(a,s), du2_2(a,s), \ldots, du2_{n-1}(a,s)\},
\end{aligned}$$

where i means the *i-th* mouse click and n is the length of the target rhythm. A target rhythm with 2-3-4-2 beats has a parameter n of 11. That if, there are 11 DU time elements, 10 DD time elements, 10 UD time elements, 10 UU time elements and 10 DU2 time elements in a sample. These features are illustrated in Figure 2, where $C(i)$ means the *i-th* click.

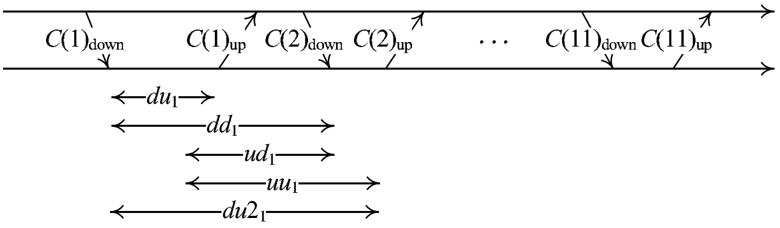

FIGURE 2. The illustration of the DU, DD, UD, UU and DU2 time features

2.4. **The template of a legitimate user.**

The template of a legitimate user is calculated using three statistical-based methods [33] with the enrolment samples. To improve the readability of this paper, here we give a concise description of our statistical-based classifier. In our system, average, maximum, minimum, standard deviation and box plot with median, lower quartile, upper quartile and interquartile ranges are used. The calculations are shown in Equations (1)-(8), respectively, where $M = 10$ is the number of samples needed in the enrollment phase, $feat_j(a)$ and $feat_j(a,k)$ denote the *j-th* element of feature feat (DU, DD, UD, UU, or DU2) from the account a and from the k-th sample of the account a, respectively. Then nine weight scores for each $feat_j(a)$ are derived from the three methods *Method A*, *Method B* and *Method C* in Figure 3 with ten samples of the enrolment phase to model as a template, where N_{feat_a} is the number of elements from the feature feat for the account a. If the feature is the DU time, then $N_{feat_a} = n$; else $N_{feat_a} = n - 1$. The weight scores are calculated based on the three statistical methods to define the individual characteristics. Each weight score is a value between 0 and 1. Finally, the template with the statistical information $\mu_{feat_j(a)}$, $MAX_{feat_j(a)}$, $MIN_{feat_j(a)}$, $\sigma_{feat_j(a)}$, $Q2_{feat_j(a)}$, $Q1_{feat_j(a)}$, $Q3_{feat_j(a)}$ and $IQR_{feat_j(a)}$ is stored in the database.

$$\mu_{feat_j(a)} = \frac{\sum_{k=1}^{M} feat_j(a,k)}{M}, \tag{1}$$

$$MAX_{feat_j(a)} = \max_{\forall k} \{feat_j(a,k)\}, \tag{2}$$

$$MIN_{feat_j(a)} = \min_{\forall k} \{feat_j(a,k)\}, \tag{3}$$

$$\sigma_{feat_j(a)} = \sqrt{\frac{1}{M} \sum_{k=1}^{M} \left(feat_j(a,k) - \mu_{feat_j(a)}\right)}, \tag{4}$$

$$Q2_{feat_j(a)} = \operatorname*{median}_{\forall k} \{feat_j(a,k)\}, \tag{5}$$

$$Q1_{feat_j(a)} = \operatorname*{median}_{\forall k} \left\{x | x \in feat_j(a,k) \cap y < Q2_{feat_j(a)}\right\}, \tag{6}$$

$$Q3_{feat_j(a)} = \operatorname*{median}_{\forall k} \left\{x | x \in feat_j(a,k) \cap y > Q2_{feat_j(a)}\right\}, \tag{7}$$

$$IQR_{feat_j(a)} = Q3_{feat_j(a)} - Q1_{feat_j(a)}. \tag{8}$$

2.5. **The classifier and threshold in our authentication system.** The sample s^* of the account a is evaluated and analyzed by the classifier in the authentication phase. Let $feat_j(a, s^*)$ be the j-th element of feature feat (DU, DD, UD, UU, or DU2) from the sample s^* account a in the authentication phase. The weight scores calculated in Figure 1 are used to assess the degree of similarity between the enrollment samples and the sample s^* in the authentication phase. Through the three reformed methods *Method A**, *Method B** and *Method C** in Figure 2, a score $Score_{feat_j(a,s^*)}$ is initialized with the value zero and obtained for each $feat_j(a, s^*)$. Finally, $Score_{feat_j(a,s^*)}$ for all j of the feature feat, and the sum $SUM_{Score_{feat_{a,s^*}}}$ are obtained.

Each $feat_j(a, s^*)$ through the above three methods will produce a score $Score_{feat_j(a,s^*)}$. $SUM_{Score_{feat_{a,s^*}}}$ derived by summing the $Score_{feat_j(a,s^*)}$ for all j of the features feat. If $SUM_{Score_{feat_{a,s^*}}}$ is greater than or equal to a given decision threshold $Threshold_{feat_a}$, the system considers the sample s^* of feature feat belongs to the account a's owner and allows access. Since $Score_{feat_j(a,s^*)}$ is added with a value between 0 and 1 for each method, the maximum value of $Score_{feat_j(a,s^*)}$ is three. According to the number of elements for each feature set N_{feat_a}, the maximum value of $SUM_{Score_{feat_{a,s^*}}}$ is $N_{feat_a} \times 3$. Finally, the decision threshold $Threshold_{feat_a}$ for the feature feat is $N_{feat_a} \times 3 \times th$, where the parameter th is a value between 0 and 1 and set by [34].

2.6. **Our experimental data.** We invited twenty-five students to participate in our experiment. Two kinds of samples were used to evaluate our system.

(1) *Imitation sample*: the sample collected by the 25 users listened to a sound recording from another user (who is not involved in the 25 users). They can listen to the

```
Three statistical-based methods for calculating template
for $j = 1$ to $N_{feat_a}$
    Method A:
    $Count1 = 0$;
    for $k = 1$ to $M$
        if $(feat_j(a,k) \geq \mu_{feat_j(a)})$ and $(feat_j(a,k) \leq MAX_{feat_j(a)})$
            $Count1++$;
        endif
    endfor
    output $WS^{A1}_{feat_j(a)} = Count1/M$
    output $WS^{A2}_{feat_j(a)} = 1 - WS^{A1}_{feat_j(a)}$;

    Method B:
    $Count2 = 0$; $Count3 = 0$; $Count4 = 0$;
    for $k = 1$ to $M$;
        if $(feat_j(a,k) \geq \mu_{feat_j(a)} - \sigma_{feat_j(a)})$ and $(feat_j(a,k) \leq \mu_{feat_j(a)} + \sigma_{feat_j(a)})$
            $Count2++$
        elseif $(feat_j(a,k) \geq \mu_{feat_j(a)} - 2 \times \sigma_{feat_j(a)})$ and $(feat_j(a,k) \leq \mu_{feat_j(a)} + 2 \times \sigma_{feat_j(a)}$
            $Count3++$
        elseif $(feat_j(a,k) \geq \mu_{feat_j(a)} - 3 \times \sigma_{feat_j(a)})$ and $(feat_j(a,k) \leq \mu_{feat_j(a)} + 3 \times \sigma_{feat_j(a)})$
            $Count4++$
        endif
    endfor
    output $WS^{B1}_{feat_j(a)} = Count2/M$;
    output $WS^{B2}_{feat_j(a)} = Count3/M$;
    output $WS^{B3}_{feat_j(a)} = Count4/M$;
    output $WS^{B4}_{feat_j(a)} = 1 - (WS^{B1}_{feat_j(a)} + WS^{B2}_{feat_j(a)} + WS^{B3}_{feat_j(a)})$;

    Method C:
    $Count5 = 0$; $Count6 = 0$;
    for $k = 1$ to $M$;
        if $(feat_j(a,k) \geq Q1_{feat_j(a)} - 1.5 \times IQR_{feat_j(a)})$ and $(feat_j(a,k) \leq Q3_{feat_j(a)} + 1.5 \times IQR_{feat_j(a)})$
            $Count5++$
        elseif $(feat_j(a,k) \geq Q1_{feat_j(a)} - 3 \times IQR_{feat_j(a)})$ and $(feat_j(a,k) \leq Q3_{feat_j(a)} + 3 \times IQR_{feat_j(a)})$
            $Count6++$
        endif
    endfor
    output $WS^{C1}_{feat_j(a)} = Count5/M$;
    output $WS^{C2}_{feat_j(a)} = Count6/M$;
    output $WS^{C3}_{feat_j(a)} = 1 - (WS^{C1}_{feat_j(a)} + WS^{C2}_{feat_j(a)})$;
endfor
```

FIGURE 3. Three statistical methods used to calculate the weight scores for each $feat_j(a)$

recording many times, and then imitate the sound to click the target rhythm. This collection is designed to make users click the rhythm in a more similar way to heighten the accuracy of the experiment.

(2) *Non-imitation sample*: the sample collected by the 25 users click the target rhythm in accordance with their habit.

```
Three reformed methods for calculating the score for each sample $s^*$
for $j = 1$ to $N_{feat_a}$
    $Score_{feat_j(a,s^*)} = 0;$
    Method A*: Average and Min-Max Method
        if $(feat_j(a,s^*) \geq \mu_{feat_j(a)})$ and $(feat_j(a,s^*) \leq MAX_{feat_j(a)})$
            $Score_{feat_j(a,s^*)} = Score_{feat_j(a,s^*)} + WS^{A1}_{feat_j(a)};$
        elseif $(feat_j(a,s^*) < \mu_{feat_j(a)})$ and $(feat_j(a,s^*) \geq MIN_{feat_j(a)})$
            $Score_{feat_j(a,s^*)} = Score_{feat_j(a,s^*)} + WS^{A2}_{feat_j(a)};$
        else
            $Score_{feat_j(a,s^*)} = Score_{feat_j(a,s^*)};$
        endif

    Method B*: Average and Standard Deviation Method
        if $(feat_j(a,s^*) \geq \mu_{feat_j(a)} - \sigma_{feat_j(a)})$ and $(feat_j(a,s^*) \leq \mu_{feat_j(a)} + \sigma_{feat_j(a)}$
            $Score_{feat_j(a,s^*)} = Score_{feat_j(a,s^*)} + WS^{B1}_{feat_j(a)};$
        elseif $(feat_j(a,s^*) \geq \mu_{feat_j(a)} - 2 \times \sigma_{feat_j(a)})$ and $(feat_j(a,s^*) \leq \mu_{feat_j(a)} + 2 \times \sigma_{feat_j(a)})$
            $Score_{feat_j(a,s^*)} = Score_{feat_j(a,s^*)} + WS^{B2}_{feat_j(a)};$
        elseif $(feat_j(a,s^*) \geq \mu_{feat_j(a)} - 3 \times \sigma_{feat_j(a)})$ and $(feat_j(a,s^*) \leq \mu_{feat_j(a)} + 3 \times \sigma_{feat_j(a)})$
            $Score_{feat_j(a,s^*)} = Score_{feat_j(a,s^*)} + WS^{B3}_{feat_j(a)};$
        else
            $Score_{feat_j(a,s^*)} = Score_{feat_j(a,s^*)} + WS^{B4}_{feat_j(a)};$
        endif

    Method C*: Box Plot Method
        if $(feat_j(a,s^*) \geq Q1_{feat_j(a)} - 1.5 \times IQR_{feat_j(a)})$ and $(feat_j(a,s^*) \leq Q3_{feat_j(a)} + 1.5 \times IQR_{feat_j(a)})$
            $Score_{feat_j(a,s^*)} = Score_{feat_j(a,s^*)} + WS^{C1}_{feat_j(a)};$
        elseif $(feat_j(a,s^*) \geq Q1_{feat_j(a)} - 3 \times IQR_{feat_j(a)})$ and $(feat_j(a,s^*) \leq Q3_{feat_j(a)} + 3 \times IQR_{feat_j(a)})$
            $Score_{feat_j(a,s^*)} = Score_{feat_j(a,s^*)} + WS^{C2}_{feat_j(a)};$
        else
            $Score_{feat_j(a,s^*)} = Score_{feat_j(a,s^*)} + WS^{C3}_{feat_j(a)};$
        endif
endfor
$SUM_{Score_{feat_{a,s^*}}} = 0;$
for $j = 1$ to $E_{feat_a}$
    $SUM_{Score_{feat_{a,s^*}}} = SUM_{Score_{feat_{a,s^*}}} + Score_{feat_j(a,s^*)};$
endfor
output $SUM_{Score_{feat_{a,s^*}}}$
```

FIGURE 4. Three methods used to calculate the score $Score_{feat_j(a,s^*)}$ for each $feat_j(a)$

Each user was asked to provide 60 samples of the target rhythm. Millisecond precision is used to measure the elapsed time between clicks. The first 30 samples were imitational and the 30 remaining samples were clicked in accordance with the users' habit. The samples were collected over a period of six months to avoid the participants becoming impatient.

2.7. **The evaluation metrics.** The following four evaluation metrics were used to evaluate our system.

(1) *False Acceptance Rate* (FAR): the rate the system accepted an impostor.

(2) *False Rejection Rate* (FRR): the rate the system rejected a legitimate user.
(3) *Average False Rate* (AFR): the FAR and FRR simple averages.
(4) *Equal Error Rate* (EER): the value at which FAR equaled FRR, which is the most balanced performance index. EER is defined as the average FAR and FRR when both are at their closest [9].

FAR was calculated based on the results for each user attacking another for one kind of imitation sample and non-imitation sample. Each user took turns as a legitimate user, while other users posed as impostors to attack the legitimate user. Any user had a second attempt regardless of whether they were the legitimate user or impostor. Therefore, the 30 samples for each user were divided into two groups. The first 15 samples were used as the major attempts and the 15 remaining samples were ready for use as the standby attempt if the user failed in the first major attempt. A legitimate user that had been attacked 24 × 15 times with the total number of attacks at 24 × 15 × 25. If the system accepted an attempt from an impostor, the wrong acceptance number was counted as 1. The total number of wrong acceptances was divided by the total number of attacks, producing the FAR value. FRR was calculated by dividing 30 samples into three groups. The first 10 samples were used in the enrolment phase, the second 10 samples and the third 10 samples were used as the major attempts and the standbys for authentication, respectively. In this situation, each user is a legitimate user and has a second attempt. Therefore, the total number of attempts was 10 × 25. If the system rejected an attempt from a legitimate user, the wrong refusal number was counted as 1. The total number of wrong refusals was divided by the total number of legitimate invasions, producing the FRR value.

3. **Experimental Analysis.**

In this paper, five features are analyzed in our classifier to authenticate the identities of users. Several experiments combining the five features were performed. The experimental results were evaluated using two kinds of samples. We found that allowing a second attempt caused a slight increase in FAR and a marked decrease in FRR. This demonstrates the usefulness of the claim from [4]. Based on the reason given above, if the user fails in the first attempt at authentication that user should be given a second attempt. For the sake of brevity, Table 1 lists only the results from the top five EER with FAR and FRR for the non-imitation samples. Table 2 lists the results from the top five AFR that are irrelevant to whether FAR and FRR are closest. Figure 5 shows the charts for FAR and FRR with these six experiments for several possible th values.

As the figures and tables indicate, an experiment combining DU, DD, UD, and DU2 time has a better EER of 0.0697 and a better AFR of 0.0628 while the th value are

Table 1. The results of the top five EER with FAR and FRR for non-imitation samples

Experiments	Non-imitation sample			th
	FAR	FRR	EER	
(a)+(b)+(c)+(e)	0.0754	0.0640	0.0697	0.57
(a)+(b)+(c)+(d)+(e)	0.0797	0.0640	0.0719	0.56
(a)+(b)+(c)	0.0846	0.0640	0.0743	0.58
(a)+(b)+(d)	0.0811	0.0680	0.0746	0.58
(a)+(c)+(e)	0.0728	0.0800	0.0764	0.58

(a) DU time; (b) DD time; (c) UD time; (d) UU time and (e) DU2 time

Table 2. The results of the top five AFR with FAR and FRR for non-imitation samples

Experiments	Non-imitation sample			th
	FAR	FRR	AFR	
(a)+(b)+(c)+(e)	0.0896	0.0360	0.0628	0.56
(a)+(c)+(e)	0.1023	0.0320	0.0672	0.56
(a)+(b)+(c)	0.0854	0.0560	0.0707	0.57
(a)+(b)+(c)+(d)+(e)	0.0797	0.0640	0.0719	0.56
(a)+(c)+(d)+(e)	0.0882	0.0600	0.0741	0.56

(a) DU time; (b) DD time; (c) UD time; (d) UU time and (e) DU2 time

Table 3. FAR and FRR with the same experiments and th values in Table 1 for imitation samples

Experiments	Imitation sample			th
	FAR	FRR	AFR	
(a)+(b)+(c)+(e)	0.1798	0.0720	0.1259	0.57
(a)+(b)+(c)+(d)+(e)	0.1858	0.0560	0.1209	0.56
(a)+(b)+(c)	0.1971	0.0600	0.1286	0.58
(a)+(b)+(d)	0.1824	0.0720	0.1272	0.58
(a)+(c)+(e)	0.1750	0.0800	0.1275	0.58

(a) DU time; (b) DD time; (c) UD time; (d) UU time and (e) DU2 time

Table 4. FAR and FRR with the same experiments and th values in Table 2 for imitation samples

Experiments	Imitation sample			th
	FAR	FRR	AFR	
(a)+(b)+(c)+(e)	0.2083	0.0440	0.1262	0.56
(a)+(c)+(e)	0.2269	0.0400	0.1335	0.56
(a)+(b)+(c)	0.2262	0.0360	0.1311	0.57
(a)+(b)+(c)+(d)+(e)	0.1858	0.0560	0.1209	0.56
(a)+(c)+(d)+(e)	0.1998	0.0520	0.1259	0.56

(a) DU time; (b) DD time; (c) UD time; (d) UU time and (e) DU2 time

0.57 and 0.56, respectively. There are a reasonable number of results showing that a rhythm click-dynamics authentication system based on mouse clicks with our statistical-based classifier is feasible and practical. To express more precise results, the results for the imitation samples with the same experiments and th values in Tables 1 and 2 are presented in Tables 3 and 4, respectively.

Tables 3 and 4 show an experiment combining all five features has a better AFR of 0.1209. Under the same condition for the non-imitation sample, this has an AFR of 0.0719. The best EER and AFR in this study were 0.0697 and 0.0628, respectively. With the non-imitation sample and the experiment combined DU, DD, UD and DU2 time, accompanying the two AFR with imitation samples are 0.1259 and 0.1262 in Tables 3 and 4, respectively. However, these error rates are not different from the two experiments. These results also show that the error rates are reasonable, even in the imitation samples,

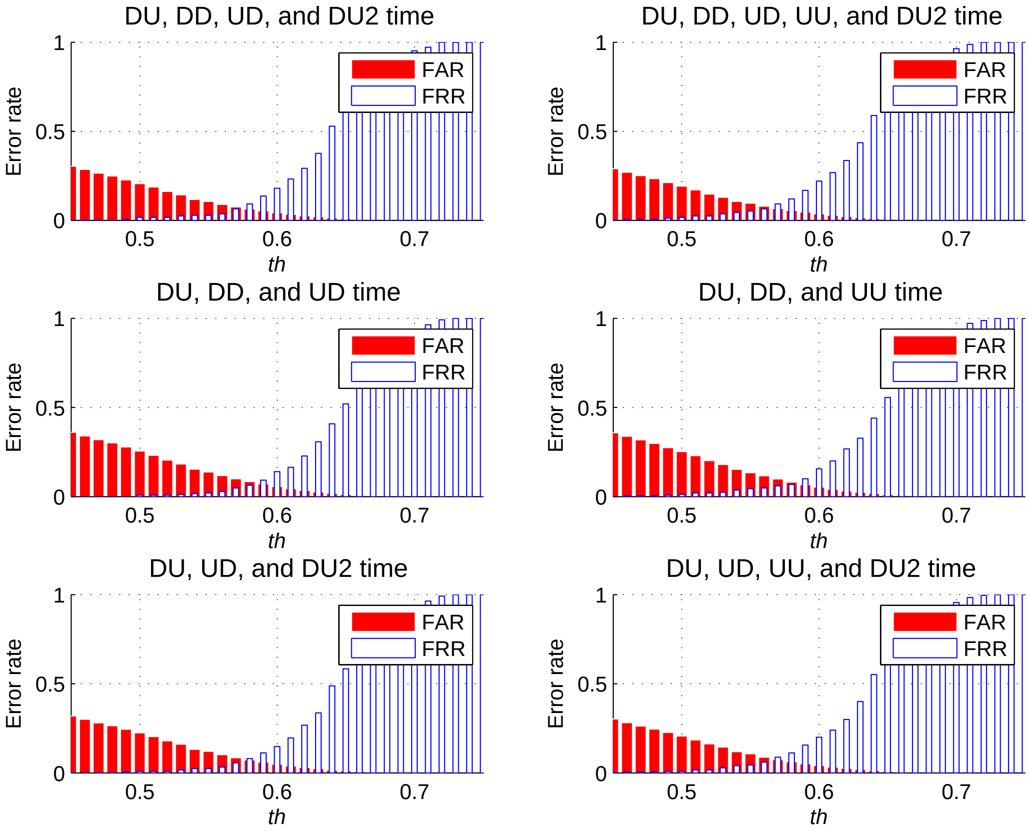

Figure 5. FAR and FRR for several possible *th* values with six experiments

and demonstrate the usefulness of the rhythm click-dynamics authentication system based on mouse clicks.

4. **Conclusions and Future Works.** As the fast process of computer and network technologies, computers connected over the Internet had become the indispensable electrical appliances of our daily life. Internet security therefore has become a serious issue for anyone connected to the Internet. This paper examined the usefulness of a rhythm click-dynamics authentication system based on mouse clicks and a statistical-based classifier. A fixed common rhythm which is very familiar for all participants was used to facilitate the system accuracy. If each legitimate user adopts a personal rhythm as the target rhythm, it will significantly reduce the error rate since any impostor has to guess the target rhythm. Our experiment also has a reasonable amount of results with the classifier used in this paper, showing that the combination of rhythms clicked using a mouse and our statistical-based classifier forms a feasible and practical KDA authentication system. Our system increases portability and can be applied to electronic devices with touch or numerical input pads. The proposed system can also be used as the standby identifiable factor in KDA systems to improve the system security. Clicking in rhythm will create noise through the mouse, which could allow other people to easily observe and listen to a user's clicking rhythm and subsequently imitate the speed and tempo to impersonate the user. Therefore, we designed an imitation sample to conform to this situation. The results are still reasonable and show that the rhythm click-dynamics authentication system based on mouse clicks has a certain degree of usefulness even in this situation.

However, the error rate of our authentication system is still high and does not reached an ideal level in real world. Saevanee and Bhatarakosol [35] found the pressure characteristics on the notebook touch pad and show it can improve the KDA system utility. In the future, we will combine the new feature in the proposed authentication system to reduce the error rate. On the other hand, some useful classifiers can be examined and analyzed to find the most suitable classifier.

Acknowledgment. This work is partially supported by the National Science Council of Taiwan under Grant No. NSC 100-2221-E018-025 and NSC 100-2221-E-018-034.

REFERENCES

[1] T. Y. Chang, W. P. Yang and M. S. Hwang, Simple authenticated key agreement and protected password change protocol, *Computers & Mathematics with Applications*, vol.49, no.5-6, pp.703-714, 2005.

[2] A. K. Jain, R. Bolle and S. Pankanti, *Biometrics: Personal Identification in Networked Society*, Springer, 2002.

[3] S. Hocquet, J. Y. Ramel and H. Cardot, User classification for keystroke dynamics authentication, *Advances in Biometrics, LNCS*, vol.4642, pp.531-539, 2007.

[4] L. C. F. Araujo, L. H. R. Sucupira Jr., M. G. Lizarraga, L. L. Ling and J. B. T. Yabu-Uti, User authentication through typing biometrics features, *IEEE Transactions on Signal Processing*, vol.53, no.2, pp.851-855, 2005.

[5] S. Bleha, C. Slivinsky and B. Hussien, Computer-access security systems using keystroke dynamics, *EEE Transactions on Pattern Analysis and Machine Intelligence*, vol.12, no.12, pp.1217-1222, 1990.

[6] D. Hosseinzadeh and S. Krishnan, Gaussian mixture modeling of keystroke patterns for biometric applications, *IEEE Transactions on Systems, Man, and Cybernetics – Part C: Applications and Reviews*, vol.38, no.6, pp.816-826, 2008.

[7] K. Revett, S. T. de Magalhaes and H. M. D. Santos, Enhancing login security through the use of keystroke input dynamics, *Advances in Biometrics, LNCS*, vol.3832, pp.661-667, 2005.

[8] Y. Sheng, V. V. Phoha and S. M. Rovnyak, A parallel decision tree based method for user authentication based on keystroke patterns, *IEEE Transactions on System, Man, and Cybernetics – Part B: Cybernetics*, vol.35, no.4, pp.826-833, 2005.

[9] P. S. Teh, A. B. J. Teoh, T. S. Ong and H. F. Neo, Statistical fusion approach on keystroke dynamics, *Proc. of IEEE Conf. on Signal-Image Technologies and Internet-Based System*, pp.918-923, 2008.

[10] T. Y. Chang, C. J. Tsai, Y. J. Yang and P. C. Cheng, User authentication using rhythm click characteristics for non-keyboard devices, *Proc. of the International Conf. on Remote Sensing and Data*, pp.167-17, 2011.

[11] T. Y. Chang, C. J. Tsai and J. H. Lin, A graphical-based password keystroke dynamic authentication system for touch screen handheld mobile devices, *Journal of Systems and Software*, 2012.

[12] H. Lee and S. Cho, Retraining a keystroke dynamics-based authenticator with impostor patterns, *Computers & Security*, vol.26, no.4, pp.300-310, 2006.

[13] M. S. Obaidat and S. Member, A multilayer neural network system for computer access security, *IEEE Transactions on System, Man, and Cybernetics*, vol.24, no.5, pp.806-813, 1994.

[14] M. S. Obaidat and B. Sadoun, Verification of computer users using keystroke dynamics, *IEEE Transactions on System, Man, and Cybernetics – Part B: Cybernetics*, vol.27, no.2, pp.261-269, 1997.

[15] T. Y. Chang, Y. J. Yang and C. C. Peng, A personalized rhythm click-based authentication system, *Information Management and Computer Security*, vol.18, no.2, pp.72-85, 2010.

[16] W. G. de Ru and J. H. P. Eloff, Enhanced password authentication through fuzzy logic, *IEEE Expert: Intelligent Systems and Their Applications*, vol.12, no.6, pp.38-45, 1997.

[17] S. Haider, A. Abbas and A. K. Zaidi, A multi-technique approach for user identification through keystroke dynamics, *Proc. of the the IEEE International Conf. on System, Man, and Cybernetics*, pp.1336-1341, 2000.

[18] G. L. F. Azevedo, G. D. C. Cavalcanti and E. C. B. Carvalho Filho, Hybrid solution for the feature selection in personal identification problems through keystroke dynamics, *Proc. of the IEEE International Joint Conf. on Neural Networks*, pp.1947-1952, 2007.

[19] W. Martono, H. Ali and M. J. E. Salami, Keystroke pressure-based typing biometrics authentication system using support vector machines, *Computational Science and Its Applications, LNCS*, vol.4706, pp.85-93, 2007.

[20] J. Hu, D. Gingrich and A. Sentosa, A k-nearest neighbor approach for user authentication through biometric keystroke dynamics, *Proc. of the IEEE International Conf. on Communications*, pp.1556-1560, 2008.

[21] F. Monrose and A. D. Rubin, Keystroke dynamics as a biometric for authentication, *Future Generation Computer Systems*, vol.16, no.4, pp.351-359, 2000.

[22] F. Bergadano, D. Gunetti and C. Picardi, User authentication through keystroke dynamics, *ACM Transactions on Information and System Security*, vol.5, no.4, pp.367-397, 2002.

[23] J. A. Robinson, V. M. Liang, J. A. M. Chambers and C. L. MacKenzie, Computer users verification using login string keystroke dynamics, *IEEE Transactions on System, Man, and Cybernetics – Part A: Systems and Humans*, vol.28, no.2, pp.236-241, 1998.

[24] M. Choras and P. Mroczkowski, Keystroke dynamics for biometrics identification, *Adaptive and Natural Computing Algorithm, LNCS*, vol.4432, pp.424-431, 2007.

[25] D. Gunetti, C. Picardi and G. Ruffo, Keystroke analysis of different languages: A case study, *Advances in Intelligent Data Analysis VI, LNCS*, vol.3646, pp.133-144, 2005.

[26] K. Revett, S. T. de Magalhaes and H. Santos, Data mining a keystroke dynamics based biometrics database using rough sets, *Proc. of the IEEE Portuguese Conf. on Artificial Intelligence*, pp.188-191, 2005.

[27] K. Revett, S. T. de Magalhaes and H. M. D. Santos, On the use of rough sets for user authentication via keystroke dynamics, *Progress in Artificial Intelligence, LNCS*, vol.4874, pp.145-159, 2007.

[28] G. P. Szeto and R. Lee, An ergonomic evaluation comparing desktop, notebook, and subnotebook computers, *Archives of Physical Medicine and Rehabilitation*, vol.83, no.4, pp.527-532, 2002.

[29] P. Kang, S. Park, S. S. Hwang, H. J. Lee and S. Cho, Improvement of keystroke data quality through artificial rhythms and cues, *Computer & Security*, vol.27, no.1-2, pp.3-11, 2008.

[30] K. Hayashi, E. Okamoto and M. Mambo, Proposal of user identification scheme using mouse, *Proc. of the 1st International Conf. on Information and Communication Security, LNCS*, vol.1334, pp.144-148, 1997.

[31] A. F. Syukri, E. Okamoto and M. Mambo, A user identification system using signature written with mouse, *Information Security and Privacy, LNCS*, vol.1438, pp.403-414, 1998.

[32] A. A. E. Ahmed and I. Traore, A new biometric technology based on mouse dynamics, *IEEE Transactions on Dependable and Secure Computing*, vol.4, no.3, pp.165-179, 2007.

[33] C. C. Peng, T. Y. Chang, C. J. Tsai, J. W. Li and C. S. Wu, A novel and simple statistical fusion method for user authentication through keystroke features, *Journal of Convergence Information Technology*, vol.6, no.2, pp.347-356, 2011.

[34] T. Fawcett, An introduction to ROC analysis, *Pattern Recognition Letters*, vol.27, no.8, pp.861-874, 2006.

[35] H. Saevanee and P. Bhatarakosol, User authentication using combination of behavioral biometrics over the touchpad acting like touch screen of mobile device, *Proc. of the International Conference on Computer and Electrical Engineering*, pp.82-86, 2008.