Understanding Secure and Usable Gestures for Realtime Motion based Authentication

Yanyan Li and Mengjun Xie
University of Arkansas at Little Rock
Email: {yxli5, mxxie}@ualr.edu

***Abstract*—One promising approach to achieving strong authentication is using gesture based behavioral biometrics, that is, user authentication is based on *how* a gesture is performed. With the advent of smart wearable devices, 3D motion gesture based authentication becomes increasingly appealing. Understanding 3D motion gestures especially their security and usability is fundamentally important but remains to be conducted. Towards this goal, we perform an empirical study on the security and usability of user-created 3D gestures using a realtime free-form motion gesture authentication scheme we have developed. To create a real-world authentication experience, in our experiment, a participant sees the authentication result from his or her wearing smartwatch in real time for each gesture test. Our experiment consists of not only a set of self-tests but also a series of attacks from easily launched random guessing attacks to much more sophisticated and dangerous targeted mimicry attacks. Our experimental results reveal several interesting findings on 3D gestures' security and usability including the correlation between gesture categories and their performance and attack resistance, and the effect of posture and psychological factor, which we believe shed light on the future design of 3D motion gestures for smart wearable devices.**

I. Introduction

Identity authentication is critical to data security and privacy but is lacking in effective techniques that are both secure and usable especially for smart wearable devices. Leveraging the popularity and advancement of mobile sensing technologies, one promising approach to achieving strong authentication for smart wearables is using gesture based behavioral biometrics. Different from conventional physiological biometrics such as fingerprints, behavioral biometrics exploits distinctive traits innate in individuals' behavior such as gestures and gaits. Gesture-based behavioral biometric authentication verifies a user based on *how* a gesture for the claimed identity is performed by the user.

With the advent of smart wearable devices such as smartwatches and smart wristbands, 3D free-form motion gestures become increasingly appealing for behavioral biometric authentication. Several gesture based authentication schemes have been proposed for smart wearable devices (e.g., [1], [2]). However, as an essential piece to making gesture based behavioral biometrics as successful as passwords, a comprehensive understanding of 3D motion gestures especially their security and usability remains missing.

As a first step towards this goal, we perform an empirical study on the security and usability of user-created 3D gestures using the REMOTE (**r**ealtime fre**e**-for**m** m**o**tion ges**t**ure auth**e**ntication) scheme we have designed. We develop a REMOTE mobile app for Android smartwatches and use it to conduct the experiment. Our experiment is designed to create real-world gesture authentication experiences by allowing users to create their personalized 3D motion gestures and showing the authentication result on the smartwatch in real time for each gesture test. Our experiment consists of not only a set of self-tests but also a series of attacks from easily launched random guessing attacks to much more sophisticated and dangerous targeted mimicry attacks. All those self-tests and attacks are performed in two postures: sitting and standing. Fig. 1 shows a participant in our experiment performing his gesture using his left arm in the sitting posture.

Fig. 1: A user performing a gesture in sitting

Our experimental results reveal several interesting findings on 3D motion gestures' security and usability. For example, we find that the user-created gestures can be grouped into three categories and each of them exhibits different characteristics in usability and security. We also find that the motion gestures created in our experiment can effectively defeat both random and content-aware attacks and most of them are highly resistant to targeted mimicry attacks. Our findings also include the effect of posture and psychological state on gestures' security and usability.

The main contributions of this work are as follows:

- We designed and implemented free-form gesture based behavioral biometric authentication for smartwatches.
- We conducted an empirical study on user-created gestures through a series of self-tests and attacks.
- Our experimental results reveal several findings and insights on free-form gestures' security and usability.

The remainder of this paper is organized as follows: We summarize the related work in Section II and present the design and components of REMOTE in Section III. We then describe our experiment design in Section IV. We detail our

evaluation in Section V and conclude the paper in Section VI.

II. Related Work

Recent years have witnessed the increasingly popular application of different physiological biometrics (e.g., fingerprint, face, and retina) to user authentication [3], [4], [5], [6]. However, attacks on those biometrics are also increasingly common (e.g., [7], [8], [9]). Representing humans' unique physiological characteristics, physiological biometrics is not easy to change once being attacked. In addition, physiological biometrics' strong link with people's real identity always raises the privacy concern.

Behavioral biometrics such as gesture and gait is quickly emerging as a promising alternative (and complementary) to existing authentication methods. A number of schemes have been proposed on applying 3D motion gestures to authentication using different types of sensing technologies, e.g., Wii controller [10], Kinect depth-sensing camera [11], smart ring [1] and smartwatch [2]. However, most previous studies on 3D motion gestures do not offer a real-life use experience (i.e., realtime authentication feedback) in their data collection and many focus on predefined gestures including our previous work on smartwatch authentication with four predefined 3D motion gestures [2]. There are numerous studies on 2D gestures that are applied on the touchscreen of a smartphone or tablet (e.g., [12], [13], [14], [15]). Given essential differences between 2D gestures on a touchscreen and 3D free-form gestures in the air, those studies on 2D gestures may not be directly applied to 3D motion gestures performed with smart wearable devices.

Besides 3D motion gestures, other types of behavioral biometrics, e.g., arm movement [16], head movement [17], and touch [18], have been proposed for authentication on smart wearable devices such as smartwatch and Google Glass. In addition, handwaving actions with a smartphone in hand have also been studied for unlocking smartphones [19].

There exist a few studies on the usability and security of behavioral biometric authentication. Li *et al.* [20] compared the usability and security of PIN and pattern based behavioral authentication on smartphones and tablets and showed that the two schemes can achieve the same level of accuracy. Li *et al.* [21] also proposed that focusing on users' unique segments of their gesture performance can improve security and usability of behavioral biometric authentication. Khan *et al.* [22] studied the usability and security perceptions of human behavior based implicit authentication, which helps understand the barriers to the adoption of implicit authentication. Khan *et al.* [23] evaluated the effectiveness of targeted mimicry attacks on three touch input based implicit continuous authentication schemes and their results show that those schemes failed against shoulder surfing and offline training attacks. Our study uses different types of attacks including targeted mimicry attacks to understand the security of 3D motion gestures and our results show that gesture based behavioral biometrics can be highly resistant to targeted mimicry attacks.

III. Realtime Motion based Authentication

To gain a better understanding of the characteristics of the gestures that are secure and usable for authentication and the important factors that affect gesture authentication, we developed REMOTE, a successor of MotionAuth [2], our prior work that uses simple 3D motion gestures for user authentication on wrist-worn smart devices. Compared to MotionAuth whose primary functionality is data collection for offline analysis, REMOTE was developed for online mobile authentication with features for practical personal use such as profiling (i.e., training and template generation) with only genuine samples and user-based threshold setting in profiling.

A. System Overview

REMOTE uses two motion sensors universally built in smart wearable devices, i.e., accelerometer and gyroscope, to collect motion data for authentication. More sensory data can be incorporated into REMOTE when those sensors become commonly available. REMOTE consists of two phases: enrollment and verification. In enrollment, a user first creates a free-form 3D gesture using the arm wearing the device (e.g., smartwatch or wristband) and then performs the gesture multiple times to register it. The REMOTE system processes the data from genuine gestures and generates the verification templates for the user. In verification, a user is asked to perform the gesture for the claimed identity. REMOTE compares the gesture data with the templates associated with the claimed identity. A decision is made in real time to show whether the identity is verified. In the current design, only timestamp, 3D accelerometer data (a_x, a_y, a_z) and gyroscope data (g_x, g_y, g_z) are collected and used in both enrollment and verification phases.

To measure the similarity between two samples (e.g., two training samples, or one training sample and one testing sample), dynamic time warping (DTW), a time series alignment algorithm, is employed. DTW is a popular method used in gesture authentication [24], [25], [2], [26]. In this study, DTW is applied to measure the similarity between two time series each consisting of a sequence of time points with 6 dimensions (a_x, a_y, a_z, g_x, g_y, g_z). The output of DTW is a distance-like value and it is compared to a predefined threshold to determine whether a gesture attempt is accepted. If the DTW distance is smaller than the given threshold, the gesture is deemed genuine.

B. Template Generation

To represent a user's motion gestures, a given number of the user's training samples (five in the current implementation) are selected as the user's templates. Using multiple templates instead of one is to better characterize the gesture behavior of a user.

In order to create a user's template set, we perform pairwise comparisons on the set of the user's training samples to find those samples with relatively small similarity distances. For a training sample set $S = \{s_1, \cdots, s_n\}$, the similarity distance for sample s_i ($1 \leq i \leq n$), denoted as $D(s_i)$, is defined as the

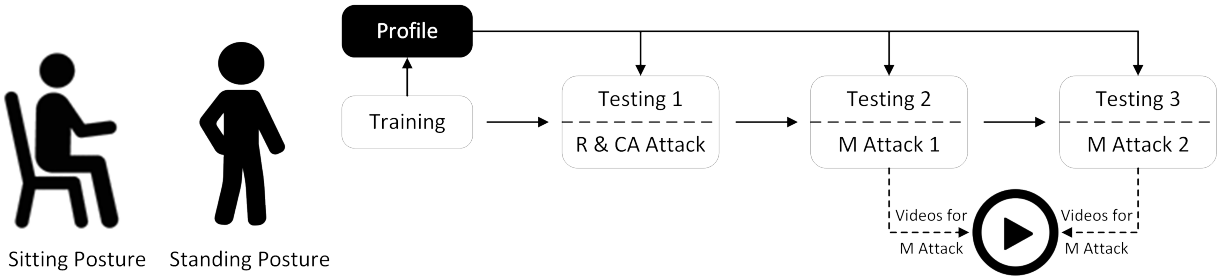


Fig. 2: Experiment process with sitting and standing postures

average of the DTW distances between s_i and the rest training samples, i.e.,

$$D(s_i) = \frac{1}{n-1}\sum_j d(s_i, s_j), \quad j \neq i, \quad j = 1, \cdots, n,$$

where $d(s_i, s_j)$ is the function that calculates the DTW distance between samples s_i and s_j. Those samples that have the five smallest similarity distances are selected as the templates.

C. Threshold Determination

Decision of the threshold can affect authentication accuracy. For REMOTE, a larger threshold can decrease the false rejection rate (accommodate large variations of genuine gestures) but at the same time increase the false acceptance rate, while a smaller threshold can render a lower false acceptance rate but a higher false rejection rate. Since users have their unique way of performing their gestures and each user's gestures can vary, we apply a user-based threshold θ for each user in REMOTE. Let the mean and standard deviation of the similarity distances for all training samples be μ and σ respectively. We define $\theta = \mu + 2\sigma$ based on a number of experiments. Assuming similarity distances follow a normal distribution, 95.5% of possible distances would fall into the range of θ [27].

D. Classification

Once the templates and threshold are generated for the gesture created by a user, the user can conduct self-tests and receive instant feedback on whether a test attempt is accepted. A majority voting method is applied in decision making. Specifically, a gesture attempt is compared against the five templates for the claimed identity, which generates five DTW distances. A decision is made based on the majority voting. Acceptance is granted if at least three out of the five DTW distances are smaller than the threshold set for the gesture.

IV. Experiment Design

A. Experiment Overview

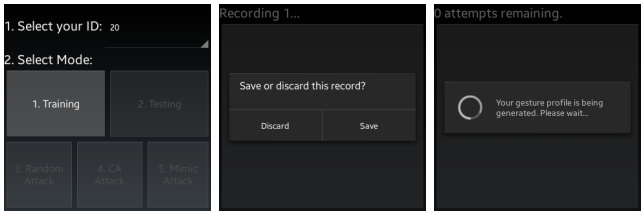
To understand user-created gestures in real world, each experiment participant is asked to create a personalized 3D motion gesture before the experiment. Participants are suggested creating gestures that are easy to perform and recall but difficult to mimic. In the experiment, each participant is given a smartwatch and can choose to wear it on either left or right wrist to perform gestures. Gesture authentication result is shown on the smartwatch instantly.

To study what gestures are secure (or insecure) for authentication, we design the following three types of attacks that cover different levels of sophistication in the experiment.

- **Random Attack**: Attackers have no prior knowledge about the victim's gesture. The attacks are based on random guessing, which is similar to brute force password attack. This type of attacks represents most common attacks that can occur when a lost or stolen device is in the hand of a third party.
- **Content-Aware Attack**: In such an attack, an adversary has the descriptive information about the victim's gesture, e.g., the gesture shape and duration, which can be obtained via social engineering or a third party. This type of attacks represents an escalated level of threat where the victim is targeted and the adversary has already collected certain information about the victim indirectly.
- **Mimicry Attack**: In such an attack, an attacker is able to clearly observe the legitimate user's gesture directly or through a recorded video. The attacker is also able to mimic the gesture skillfully before launching attacks. This type of attacks represents the highest level of threat. The adversary is highly motivated to mimic the legitimate user's gesture and has the highest chance to pass the authentication.

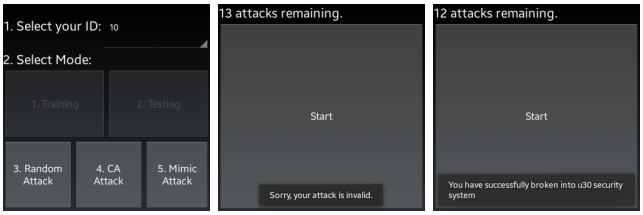
In our experiment, each participant is required to complete a sequence of tasks including training and self-tests in addition to the aforementioned three types of attacks. Multiple self-test tasks are created to examine gestures' usability over time.

In each task, participants are asked to perform their gestures in two postures: sitting and standing. Participants choose their starting posture at their will in the experiment. To minimize the impact on authentication result introduced by the difference of devices or sensors, all the motion gestures performed in the same posture are collected using the same smartwatch. We use two smartwatches of the same model, one for gestures in the standing posture and the other for the sitting posture. We assess self-tests and attacks against the templates generated from training samples in the same posture. In other words, no templates generated in standing are used to assess the self-tests and attacks performed in sitting, or vice versa. The purpose of this design is to achieve a good understanding of motion gestures in a specific posture. Scenarios of combining different postures in gesture authentication, which is much more complex, are left for future study.



(a) (b) (c)

Fig. 3: Screenshots of the REMOTE mobile app



(a) (b) (c)

Fig. 4: Screenshots of attack tests

B. Experiment Process

Before the experiment for a participant starts, the study is explained in detail to the participant and does not proceed without the participant's full consent. Each participant is assigned a unique ID for anonymous data collection and anonymous target selection in attack activities. Participants have no prior knowledge of their attacking targets, which ensures the fairness of random and content-aware attacks.

In the experiment, participants are required to complete a series of tasks with one or two tasks a day and all tasks are completed in four different days (i.e., four visits). Each visit is designed to last about 20 to 40 minutes. The experiment process is illustrated in Figure 2, where we arrange a gesture training task in the first visit, a self-test task (marked as Testing 1) and two attack tasks (random (R) attack task and content-aware (CA) attack task) in the second visit, a self-test task (Testing 2) and a mimicry attack task (M Attack 1) in the third visit, and a self-test task (Testing 3) and another mimicry attack task (M Attack 2) in the last visit. The time interval between two contiguous visits is set to at least two days so that we can measure what gestures can provide good usability over time. The setting of two mimicry attacks is to assess whether there exists an evident increase in attack success rate with more practices on mimicking the victim's motion gestures.

In gesture training sessions, two cameras are set up on site to record users' motion gestures from two directions (one facing the user's front and the other facing the user's left side), respectively. This setting aims to record sufficient movement information for attackers to practice for mimicry attacks by watching the recorded videos. In the training session, each participant is asked to repeat his or her motion gestures 30 times in both sitting and standing postures. Hence, 60 gestures in total are collected for each participant in that session.

In each of the three self-test sessions, participants need to test their gestures at least 30 times in each posture. In the current setting of the REMOTE app, a user has three chances for each test attempt. Either a success or three failures will end a test attempt. This setting imitates the real-world setting where users are always given several chances during a login.

For attack sessions, a participant is required to attack the same set of targets, who are randomly selected and vary for different attackers, in his or her attack sessions. This design aims to measure the change of attack performance with increasing familiarity with a target's gestures. In each attack session, a participant is asked to launch 15 attacks against each of the three randomly selected targets. Attacks are performed in both sitting and standing postures. Therefore, the total number of attacks a participant generates are 90 ($15 \times 3 \times 2$) in each attack session.

C. Prototype Application for Experiment

We developed a REMOTE mobile app for Android smartwatches. The app has a simple user interface (UI). Screenshots of the app's UI are shown in Fig. 3. Each user is uniquely identified through a preassigned ID number. To start the enrollment (i.e., training), a user has to select his or her assigned ID number and press "Training" button (3 (a)). The user can choose to save or discard each performed gesture in the training session (3 (b)) to control the quality of training samples. After the user performs the personal gesture for a given number of times, the gesture profile will be generated and saved along with the user's ID number (3 (c)).

The app is also used for testing random, content-aware, and mimicry attacks. For attack purpose, the user ID being chosen in an attack session has to be the target victim's ID. Fig. 4 shows some screenshots during attack. In an attack session, a toast message is displayed at the bottom of the screen to show the verification result after each attack. Fig. 4 (b) shows the screenshot of a failed attack while Fig. 4 (c) shows the screenshot of a successful attack. Through instant feedback, the prototype app can give users the sense of real-world authentication.

V. Evaluation

In this section, we present the data acquisition and collected gestures in our experiment, the experimental results, and the analysis and discussion about the results.

A. Data Acquisition

We installed the REMOTE app on two Samsung Galaxy Gear smartwatches and used them for evaluation. Each gesture sample is represented by a time series with seven attributes. The first attribute is the timestamp (t), the second to fourth attributes are the acceleration in 3D (a_x, a_y, a_z), and the last three are the angular acceleration of gyroscope in 3D (g_x, g_y, g_z). The length of a sample varies as it is up to how a user performs a gesture. The sensing frequency was set to 100 Hz (i.e., 100 time series data points per second).

We recruited student volunteers for evaluation through our institutional mail-lists. Each participant would receive $25 as compensation after successfully completing all the required

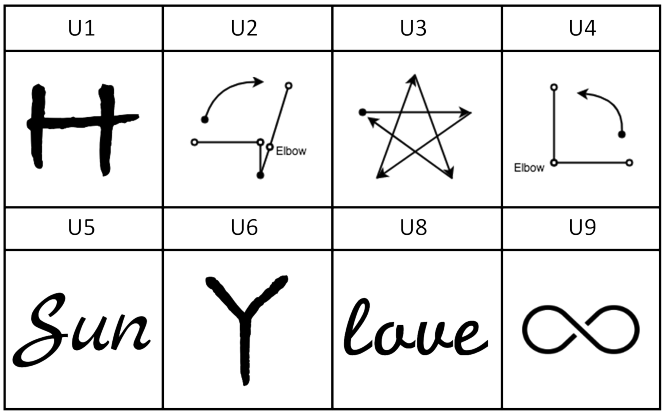


Fig. 5: Gestures of participants (not including U_7) in 2D

tasks. Fig. 1 shows a participant who is performing a gesture with his left arm in sitting. We recruited 11 participants but two of them were excluded from the study as they did not complete. Therefore, our evaluation is based on the data from the left 9 participants. Two of them were college students and the rest were graduate students. Their ages were between 23 and 44 with a mean of 30.3 and a median of 29. Six participants were male and three were female. The entire data collection was carried out in Spring 2017 and spanned over a month. The participants were excited about the experiment especially the attack activities. They were highly motivated to try and see whether they can break the gestures set by others.

B. User-created Gestures

Each participant created a personal 3D motion gesture involving hand and arm movement. Those gestures (except U_7's) are depicted in a 2D form in Fig. 5, which illustrates the composition of those gestures. The brief description of each gesture is given in Table I. Clearly, those user-created gestures vary significantly from simple ones such as forearm movement (U_4) to complex ones such as drawing "Sun" and "love" in the air (U_5 and U_8). Since U_7's gesture is patting on head, chest and thigh once, which is difficult to illustrate in 2D, it is not included in Fig. 5. The nine gestures can be roughly grouped into three categories: I) simple hand/arm movement (U_2, U_4, U_7), II) drawing a geometric shape (U_3, U_9), and III) drawing one or multiple characters (U_1, U_6, U_5, U_8).

C. Experimental Results

1) Self-test Results: The self-test results for the 9 participants are listed in Table II. The numeric values not in percentage format refer to the number of acceptance out of 30 attempts in a self-test session in a specific posture. A successful attempt (i.e., acceptance) means that the participant passes verification within three chances given for each attempt. "S" and "St" in Table II refer to sitting posture and standing posture respectively. From the table, the majority of participants (5 out of 9) have at least 90% success rate (in either sitting or standing) in *all* three self-test sessions. As a self-test session and its prior session (either a training or self-test session) are at least two days apart, those self-test results are encouraging for practical use of 3D motion gestures. Interestingly, the five participants with high success rates belong to categories I (simple hand/arm movement) and II (drawing a geometric shape). In particular, participants U_2, U_4, U_7 and U_9 achieved 100% success rate for all three self-tests in a specific posture. More detailed analysis of those rates is given in the subsequent sections.

TABLE I: Description of user-created gestures

U_{id}	Wrist in Use	Gesture Brief Description
U_1	Right	Draw capitalized 'H' in the air
U_2	Right	Raise up arm like Superman
U_3	Right	Draw a star in the air
U_4	Left	Move forearm up from horizontal to vertical
U_5	Right	Sign with word 'Sun' (S capitalized)
U_6	Right	Draw capitalized 'Y' in the air
U_7	Left	Pat on head, chest and thigh
U_8	Left	Draw 'love' cursively in the air
U_9	Left	Calibration ∞ three times and fist bang

2) Attack Results: The results of all the attacks are listed in Table III. The superscript of a user ID refers to the number of the attacks that user received in a given posture (sitting or standing) in one attack session (Random, CA, M-1, or M-2). For example, U_3^{60} means user U_3 received 60 attacks (from 4 attackers each contributing 15 attacks) in an attack session that were launched in a particular posture. The numeric values not in percentage format refer to the numbers of successful attacks (i.e., false acceptances) in an attack session. The percentage values refer to the false acceptance rates. In "Total" row, the lower a value for a participant, the more secure that participant's gesture is. Clearly, the gestures of U_1, U_2, U_3, U_5, U_8 and U_9 are highly resistant to various types of attacks including targeted mimicry attacks. Almost all their false acceptance rates are 0. The two gestures (U_4's and U_7's) with high false acceptance rates (mainly for mimicry attack sessions) are in category I, indicating that simple hand/arm movement gestures are vulnerable to mimicry attacks. Further analysis is detailed in the next section.

D. Usability and Security Analysis

In this study, we use usability and security as two qualitative concepts to characterize the applicability of a user-created gesture and its resistance to attacks respectively. We measure a gesture's usability and security mainly through its false rejection rate and false acceptance rate, respectively.

1) Usability of Gestures: From Table II, it can be seen that the gestures in categories I and II have much higher true acceptance rates (i.e., much lower false rejection rates) than those in category III. Compared to drawing one or multiple characters, making simple hand/arm movement and drawing a shape (even a bit complex one like calibration gesture) are relatively easy to perform and repeat reliably.

For gestures that draw a character (U_1's 'H' and U_6's 'Y'), low false rejection rates were anticipated as those gestures are

TABLE II: Self-test results (# of attempts that passed (out of 30), S for sitting and St for standing)

Self-test		U_1	U_2	U_3	U_4	U_5	U_6	U_7	U_8	U_9	%
Testing 1	S	25	27	30	30	26	29	28	10	29	87%
	St	29	30	30	29	27	28	30	27	30	96%
Testing 2	S	30	20	27	30	28	10	30	0	29	76%
	St	30	30	28	30	16	30	30	21	30	91%
Testing 3	S	13	30	26	30	30	4	30	0	30	71%
	St	5	30	30	29	10	25	30	16	30	76%
Total (%)	S	76%	86%	92%	100%	93%	48%	98%	11%	98%	78%
	St	71%	100%	98%	98%	59%	92%	100%	71%	100%	88%

TABLE III: Attack results (superscript is # of attacks)

Attack		U_1^{75}	U_2^{45}	U_3^{60}	U_4^{45}	U_5^{45}	U_6^{60}	U_7^{45}	U_8^{15}	U_9^{15}	$\%^{405}$
Rand	S	0	0	0	0	0	0	0	0	0	0%
	St	0	0	0	0	0	0	0	0	0	0%
CA	S	0	0	0	0	0	0	11	0	0	2.7%
	St	0	0	0	0	0	0	0	0	0	0%
M-1	S	0	0	0	28	0	0	15	0	0	10.6%
	St	0	0	0	0	0	1	0	0	0	0.2%
M-2	S	0	0	0	41	0	11	7	0	0	14.6%
	St	0	2	0	1	0	2	1	0	0	1.5%
Total (%)	S	0%	0%	0%	38.3%	0%	4.6%	18.3%	0%	0%	7.0%
	St	0%	1.1%	0%	0.6%	0%	1.2%	0.6%	0%	0%	0.4%

still relatively simple. However, the participants delivered quite different results. U_6 had very poor performance in sitting in the second and third sessions. Based on our observation, U_6 had a minor action when performing the gesture in sitting in the training session, while that action was barely seen in the second and third testing sessions. U_1's passing rate drops dramatically in the third testing session for both sitting and standing postures, which is in sharp contrast to the first two sessions. A possible reason for that change is given in Section V-E. One observation on the two characters is that their gestures involve multiple strokes in the performance. An intuitive explanation for their high false rejection rates is that multi-stroke gestures (e.g., U_1's 'H') in general are more likely to vary in the air and thus more difficult to reliably repeat than single stroke gestures (e.g., U_3's star).

For gestures that draw multiple characters ("Sun" and "love"), the results suggest poor usability. Those gestures often contain multiple strokes, complex curves, and considerable acceleration changes, which makes it difficult to re-draw them in the air in the same manner. As a special case, signatures are often considered in creating such gestures. Our results indicate more comprehensive studies on the usability of signature-based 3D motion gestures are certainly needed for smart wearable devices.

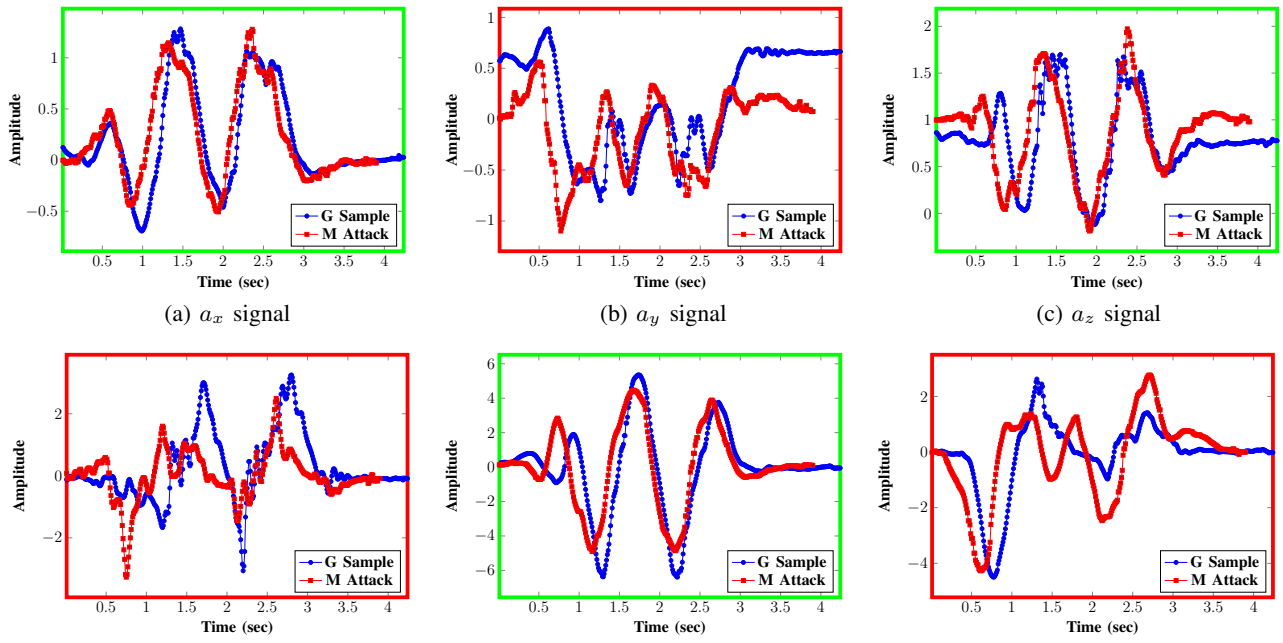
2) Security of Gestures: From Table III, we can see that random attacks in general are not effective to gesture-based behavioral biometric authentication. More importantly, our results show that the additional descriptive information given in the content-aware attacks only helps attackers gain very limited advantage over random attacks. As our motion gesture authentication is based on behavioral biometrics, descriptive information offers little help to perform the gesture in the same manner as the gesture creator does. Mimicry attacks are more effective than random and content-aware attacks although they achieve no success in most cases. For the results of mimicry attacks (M-1 and M-2 in the table), the overall success rates in sitting and standing in the second session (M-2) are both higher than those in the first session (M-1), indicating that more practices can help increase success chance for mimicry attacks.

There exist some users (e.g., U_4 and U_7) whose gestures in the sitting posture were successfully mimicked. Those gestures are relatively simple (in gesture category I). No turning points or changes in action strength/speed, which are important in differentiating genuine gestures from mimicry ones, can be observed from those gestures. However, it is interesting that the same gestures in the standing posture are much more resistant to mimicry attacks, which suggests that posture may need to be considered in creating secure gestures.

U_7's gesture (pat on head, chest and thigh) is the only one that was successfully attacked in content-aware attacks. Interestingly, only one participant (out of three) was able to successfully mimic it in mimicry attacks. By examining U_7's gesture, we find that it is actually not easy for other people to have the same or similar arm movement distance in performing that gesture as U_7 does, due to the physical difference between individuals. The only participant who had successful attacks is very similar to U_7 in height and arm length. Therefore, physical body features such as body shape and size should also be considered and exploited in creating secure gestures, in addition to personalized variation in action strength and rhythm.

Many participants were surprised by the fact that the vast majority of their mimicry attacks failed. To help understand why most mimicry attacks fail, we pick a sample from U_3's training data and compare it with a mimicry attack sample against U_3 in 3D acceleration (a_x, a_y, a_z) and angular acceleration (g_x, g_y, g_z). The comparison result is illustrated in Fig. 6. From the figure, although the time series of a_x, a_z, and g_y signals for the mimicry attack sample (curves in red) are similar to those for the genuine sample (curves in blue), evident differences between the attack sample and the genuine sample manifest in the time series of a_y, g_x, and g_z signals and those differences enable the system to reject the attack sample. To differentiate, the border lines for a_x, a_z, and g_y figures are in light green while those for a_y, g_x, and g_z figures are in bold red.

In fact, the mimicry attack sample used in the comparison is a special case, which is quite close to being successfully authenticated. In our experiment, most mimicry attack samples have more or less obvious differences with genuine samples. Duration difference (e.g., mimicry gestures finish half second earlier or later) is a common factor that defeat mimicry attacks. Action strength and rhythm are also difficult to mimic, even if the gesture trajectory can be well imitated. By examining raw



(a) a_x signal (b) a_y signal (c) a_z signal

(d) g_x signal (e) g_y signal (f) g_z signal

Fig. 6: Comparison between a genuine sample and a mimic attack sample using accelerometer and gyroscope data

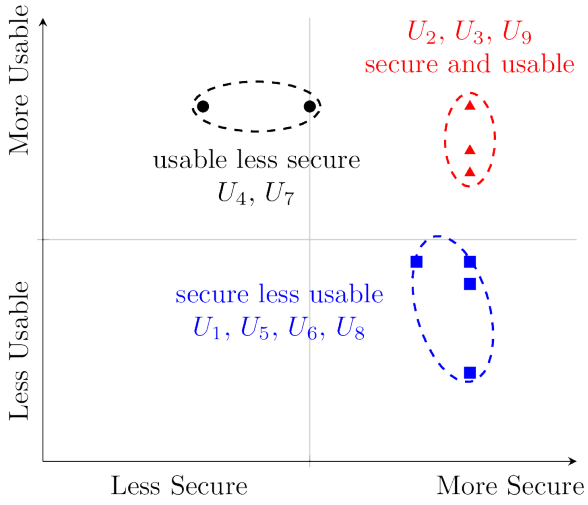


Fig. 7: Gestures' security and usability

sensor data, we find that a mimicry attack sample is rarely able to match a genuine sample in all aspects.

3) Combination of Security and Usability: To gain a comprehensive understanding of user-created gestures' security and usability, we assign two numeric scores to each participant's gesture–one for usability and the other for security–based on the gesture's self-test and attack results, and we place the nine gestures on a 2D coordinate system ($x-$axis for security and $y-$axis for usability) using gestures' scores as their coordinates, which is illustrated in Fig. 7. Given the emphasis on understanding gestures' security and usability via visualization, the scoring detail is omitted.

Interestingly, the 9 gestures exhibit 3 clusters based on their relative locations in Fig. 7. The gestures created by U_4 and U_7, i.e., in gesture category I, are in "usable but less secure" cluster in that they are simple and easy to perform but meanwhile are also easy to imitate. The gestures that draw one or multiple characters (gesture category III) are in "secure but less usable" cluster since those gestures are difficult to repeat in the same manner. The gestures by U_2, U_3, and U_9, which are the mix of categories I and II, are "secure and usable." The correlation between gesture categories and their security and usability characteristics suggests the following features for a secure and usable gesture: 1) single stroke with sufficient variation in movement, 2) smooth and natural change of moving direction and speed, and 3) consistent performance under various circumstances.

E. Discussion

1) Effect of Posture: According to Table II, for self-tests, the average success rates (last column in the table) of gestures in standing are consistent higher than those in sitting. The overall success rate of gestures in standing is 10% higher than that in sitting. This difference is reflected in most of the individuals' self-test results. Interestingly, from Table III, the attack success rates for the gestures in standing are consistently lower than those in sitting. Although in our study gestures in standing generally show an edge over those in sitting, the conclusive impact of posture on gesture's usability and security remains unclear and demands further comprehensive studies.

Three gestures (U_5, U_6, U_8) exhibit outstanding disparities between sitting and standing. U_5's gesture is an exception in that its two self-test results in sitting are much better than those in standing. Part of its possible reason is given in the next discussion section. U_8's gesture is drawing word "love," which requires more horizontal space. Performing this gesture

in sitting is more constrained in movement range and speed due to limited space and thus less stable compared to its performance in standing, which may explain in part why U_8's gesture has the worst performance in sitting. U_6's gesture also exhibits similar disparities as U_8's. The talk with U_6 reveals that the participant felt more natural and comfortable to draw 'Y' in standing than in sitting.

2) Psychological Factor: In Table II, the passing rate for some users (e.g., U_1 and U_5) drops substantially in certain tests. Through our observation and talks with the participants, we believe that performance of a gesture may also be affected by the performer's psychological state. U_1 had a dramatic drop in the third self-test session (from 100% passing rate in the second session to 30% in the third session). Coincidentally, U_1 had a doctor's visit right before the test, which might have a negative psychological and/or physiological effect on gesture performance. For U_5, we observed that she took a phone call in the middle of her third self-test session in the standing posture and that she failed for all the following attempts after the call. This dramatic change of behavior might be attributed to her psychological change caused by the call. This study demonstrates that systematic and in-depth research is needed to understand and evaluate psychological factors to gesture performance and their impact on behavioral biometric authentication.

VI. Conclusion

In this paper we presented our empirical study on realtime motion gesture based authentication for smart wearable devices as a first step towards a comprehensive understanding of 3D free-form motion gestures for behavioral biometric authentication. We designed a realtime motion gesture authentication scheme named REMOTE and developed a REMOTE mobile app for Android smartwatches. We conducted a multi-stage experiment, which consists of a series of self-tests and attacks of different levels to understand gestures' security and usability in practice. Our experimental results and analysis suggest the following features for a secure and usable gesture: 1) single stroke with sufficient variation in movement, 2) smooth and natural change of moving direction and speed, and 3) consistent performance under various circumstances. Our results also suggest that posture and psychological state may affect gesture performance. Our future work includes a more comprehensive user study to gain a deeper understanding of secure and usable gestures.

References

[1] M. Roshandel, A. Munjal, P. Moghadam, S. Tajik, and H. Ketabdar, "Multi-sensor finger ring for authentication based on 3d signatures," in *Proc. HCI International*. Springer, 2014, pp. 131–138.

[2] J. Yang, Y. Li, and M. Xie, "MotionAuth: Motion-based authentication for wrist worn smart devices," in *Proc. the 1st Workshop on Sensing Systems and Applications Using Wrist Worn Smart Devices*, 2015, pp. 550–555.

[3] K. Xi, T. Ahmad, F. Han, and J. Hu, "A fingerprint based bio-cryptographic security protocol designed for client/server authentication in mobile computing environment," *Security and Communication Networks*, vol. 4, no. 5, pp. 487–499, 2011.

[4] W. Wang, Y. Yuan, and N. Archer, "A contextual framework for combating identity theft," *IEEE Security & Privacy*, vol. 4, no. 2, pp. 30–38, 2006.

[5] Á. Serrano, I. M. de Diego, C. Conde, and E. Cabello, "Recent advances in face biometrics with gabor wavelets: A review," *Pattern Recognition Letters*, vol. 31, no. 5, pp. 372–381, 2010.

[6] M. Ortega, M. G. Penedo, J. Rouco, N. Barreira, and M. J. Carreira, "Retinal verification using a feature points-based biometric pattern," *EURASIP Journal on Advances in Signal Processing*, vol. 2009, no. 1, p. 235746, 2009.

[7] Z. Kleinman, "Politician's fingerprint 'cloned from photos' by hacker," http://www.bbc.com/news/technology-30623611, December 2014.

[8] C. McGoogan and D. Demetriou, "Peace sign selfies could let hackers copy your fingerprints," http://www.telegraph.co.uk/technology/2017/01/12/peace-sign-selfies-could-let-hackers-copy-fingerprints, Jan. 2017.

[9] T. Warren, "Windows 10's face authentication defeated with a picture," https://www.theverge.com/2017/12/21/16804992/microsoft-windows-10-windows-hello-bypass-security, Dec. 2017.

[10] J. Liu, L. Zhong, J. Wickramasuriya, and V. Vasudevan, "uWave: Accelerometer-based personalized gesture recognition and its applications," *Pervasive & Mobile Computing*, vol. 5, no. 6, pp. 657–675, 2009.

[11] J. Tian, C. Qu, W. Xu, and S. Wang, "KinWrite: Handwriting-based authentication using kinect," in *Proc. the 20th NDSS*, 2013.

[12] A. De Luca, A. Hang, F. Brudy, C. Lindner, and H. Hussmann, "Touch me once and i know it's you!: Implicit authentication based on touch screen patterns," in *Proc. CHI'12*, 2012, pp. 987–996.

[13] M. Shahzad, A. X. Liu, and A. Samuel, "Secure unlocking of mobile touch screen devices by simple gestures: You can see it but you can not do it," in *Proc. MobiCom'13*, 2013, pp. 39–50.

[14] M. Sherman, G. Clark, Y. Yang, S. Sugrim, A. Modig, J. Lindqvist, A. Oulasvirta, and T. Roos, "User-generated free-form gestures for authentication: Security and memorability," in *Proc. ACM MobiSys'14*, 2014, pp. 176–189.

[15] P. Saravanan, S. Clarke, D. H. P. Chau, and H. Zha, "Latentgesture: Active user authentication through background touch analysis," in *Proc. 2nd Intl. Symposium of Chinese CHI*. ACM, 2014, pp. 110–113.

[16] R. Kumar, V. V. Phoha, and R. Raina, "Authenticating users through their arm movement patterns," *arXiv preprint arXiv:1603.02211*, 2016.

[17] S. Li, A. Ashok, Y. Zhang, C. Xu, J. Lindqvist, and M. Gruteser, "Whose move is it anyway? authenticating smart wearable devices using unique head movement patterns," in *Proc. IEEE PerCom'16*, 2016, pp. 1–9.

[18] G. Peng, G. Zhou, D. T. Nguyen, X. Qi, Q. Yang, and S. Wang, "Continuous authentication with touch behavioral biometrics and voice on wearable glasses," *IEEE Transactions on Human-Machine Systems*, vol. 47, no. 3, pp. 404–416, 2017.

[19] L. Yang, Y. Guo, X. Ding, J. Han, Y. Liu, C. Wang, and C. Hu, "Unlocking smart phone through handwaving biometrics," *IEEE Transactions on Mobile Computing*, vol. 14, no. 5, pp. 1044–1055, 2015.

[20] Y. Li, J. Yang, M. Xie, D. Carlson, H. G. Jang, and J. Bian, "Comparison of pin- and pattern-based behavioral biometric authentication on mobile devices," in *Proc. IEEE MILCOM'15*, 2015, pp. 1317–1322.

[21] Y. Li, M. Xie, and J. Bian, "SegAuth: A segment-based approach to behavioral biometric authentication," in *Proc. IEEE CNS'16*, Philadelphia, PA, USA, 2016.

[22] H. Khan, U. Hengartner, and D. Vogel, "Usability and security perceptions of implicit authentication: Convenient, secure, sometimes annoying," in *Proc. USENIX SOUPS'15*, 2015, pp. 225–239.

[23] ——, "Targeted mimicry attacks on touch input based implicit authentication schemes," in *Proc. ACM MobiSys'16*, 2016, pp. 387–398.

[24] N. Sae-Bae, K. Ahmed, K. Isbister, and N. Memon, "Biometric-rich gestures: A novel approach to authentication on multi-touch devices," in *Proc. ACM CHI'12*, 2012, pp. 977–986.

[25] A. De Luca, E. Von Zezschwitz, N. D. H. Nguyen, M.-E. Maurer, E. Rubegni, M. P. Scipioni, and M. Langheinrich, "Back-of-device authentication on smartphones," in *Proc. CHI'13*. ACM, 2013, pp. 2389–2398.

[26] M. T. I. Aumi and S. Kratz, "AirAuth: evaluating in-air hand gestures for authentication," in *Proc. MobileHCI '14*, 2014, pp. 309–318.

[27] T. Gjedrem and I. Olesen, "Basic statistical parameters," in *Selection and Breeding Programs in Aquaculture*. Springer, 2005, pp. 45–72.