Masaryk University
Faculty of Informatics

# Evaluation of Hand Micromovement Features for Continuous Authentication of Smartphone Users During Typing

Master Thesis

**Zdeňka Sitová**

Brno, January 2015

# Declaration

I declare that this thesis is my own work and has not been submitted in any form for another degree or diploma at any university or other institution of tertiary education. Information derived from the published or unpublished work of others has been acknowledged in the text and a list of references is given.

**Advisor:** prof. RNDr. Václav Matyáš, M.Sc., Ph.D.

# Acknowledgement

To Kiran S. Balagani belong my thanks for endless hours he spent with me during my year at NYIT. Without his support and knowledge, this thesis would not have been possible.

I would also like to thank other co-authors of the paper, upon which this thesis is based. Thank you Jaroslav Šeděnka, Qing Yang, Ge Peng, Gang Zhou and Paolo Gasti, for all the effort you put into our work.

I thank my supervisor Václav Matyáš for his kind guidance and valuable feedback on my writing.

I am also grateful to my lab-mates from NYIT for valuable discussions, fun, and support.

To my friends Pranav Krishnamurthy, Jan Hapala and Martin Vlach belong my acknowledgments for useful comments on the text of this thesis.

Last but not least, I thank my friends, family and my boyfriend Jaroslav for patience and support during my studies.

# Abstract

Biometric authentication is a form of identity verification, which relies on user's distinctive physiological or behavioral characteristics. This thesis presents an evaluation of Hand Movement, Orientation and Grasp (HMOG), a new behavioral biometric for continuous authentication of smartphone users during typing. HMOG unobtrusively captures subtle micromovement and orientation dynamics resulting from how a user grasps, holds, and taps on the smartphone. We evaluated the authentication performance of HMOG features on typing data collected from 100 subjects under two conditions: sitting and walking. Combining HMOG with the state-of-the-art (tap characteristics and keystroke dynamics) improved the performance in walking (the best achieved EER lowered from 10.51% to 6.92%). For sitting users, the performance improved in shorter authentication scans (e.g., from 19.11% to 15.44% for 20-second authentication scan).

# Keywords

behavioral biometrics, continuous authentication, smartphone, typing, virtual keyboard, micromovement

# Contents

1 **Introduction** . . . . . . . . . . . . . . . . . . . . . . . . . . . . 2
- 1.1 *Contributions of This Thesis* . . . . . . . . . . . . . . . . . 3
- 1.2 *Organization* . . . . . . . . . . . . . . . . . . . . . . . . . . 5

2 **Background and Related Work** . . . . . . . . . . . . . . . . . . 6
- 2.1 *Background: Biometrics* . . . . . . . . . . . . . . . . . . . . 6
- 2.2 *Related Work* . . . . . . . . . . . . . . . . . . . . . . . . . 8

3 **Description of HMOG Features** . . . . . . . . . . . . . . . . . 13
- 3.1 *Grasp Resistance Features* . . . . . . . . . . . . . . . . . . 13
- 3.2 *Grasp Stability Features* . . . . . . . . . . . . . . . . . . . 15

4 **Dataset and Evaluation Methodology** . . . . . . . . . . . . . 17
- 4.1 *Dataset* . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 17
- 4.2 *Design of Authentication Experiments* . . . . . . . . . . . 18
- 4.3 *Comparing HMOG to Other Feature Sets* . . . . . . . . 19
- 4.4 *Feature Selection, Preprocessing, and Transformation* . 20

5 **Authentication Results** . . . . . . . . . . . . . . . . . . . . . . 24
- 5.1 *Performance of HMOG Features* . . . . . . . . . . . . . 24
- 5.2 *Comparison of HMOG with Keystroke Dynamics and Tap Features* . . . . . . . . . . . . . . . . . . . . . . . . . . 26
- 5.3 *Fusion of HMOG, Tap, and Keystroke Features* . . . . . 27

6 **Discussion** . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 29

7 **Conclusion** . . . . . . . . . . . . . . . . . . . . . . . . . . . . 32

# Chapter 1

# Introduction

Most methods of smartphone user authentication are based on a password, usually a personal identification number (PIN) or a touchscreen gesture [16]. The password has to be entered for the access to the system to be granted. The main drawbacks of password-based authentication are that

- the password can be forgotten or stolen [29]; and
- the need to enter the password can be so irritating for the user, that she may completely disable the password protection [2].

Both of these problems can be eliminated by authenticating the user *continuously*, using her behavioral patterns (*biometrics*). No password can be taken away or forgotten, as the key to authentication is the user's behavior itself. Furthermore, the authentication process is non-intrusive and proceeds continuously while the user interacts with the phone. Continuous authentication can also be used as an additional line of defense to password-based methods [16].

In this thesis, we present evaluation of Hand Movement, Orientation, and Grasp (HMOG): a new behavioral biometric for continuous authentication of smartphone users during typing. HMOG uses accelerometer, gyroscope, and magnetometer readings to unobtrusively capture subtle hand micromovements and orientation patterns generated when a user taps on the screen.

HMOG biometric is founded upon two core building blocks of human prehension [30]: *stability grasp*, which provides stability to the object being held; and *precision grasp*, which involves precision-demanding tasks like tapping a target. We view the act of *holding a phone* as a stability grasp and the act of *touching targets on the touchscreen* as a precision grasp. We hypothesize that the way in which a user "distributes" or "shares" stability and precision grasps while interacting with the smartphone results in distinctive movement and orientation behavior. The rationale for our hypothesis comes from the following two bodies of research.

First, there is evidence (see [3, 22, 37]) that users have postural preferences for interacting with hand-held devices like smartphones. Depending upon the postural preference, it is possible that a user can have her own way of achieving stability and precision—for example, the user can achieve both stability and precision with one hand if the postural preference involves holding and tapping the phone with the same hand; or distribute stability and precision between both hands, if the posture involves using both hands for holding and tapping; or achieve stability with one hand and precision with the other.

Second, studies in ergonomics, biokinetics, and human-computer interaction have reported that handgrip strength strongly correlates with an individual's physiological and somatic traits like hand length, handedness, age, gender, height, body mass, and musculature (see, e.g., [9, 15, 24]). If the micromovements generated from tapping reflect an individual's handgrip strength, then the distinctiveness of HMOG may have its roots in an individual's physiological and somatic traits.

Motivated by the above, two types of HMOG features were proposed: *resistance features*, which measure the micromovements of the phone in response to the forces exerted by a tap gesture; and *stability features*, which measure how quickly the perturbations in movement and orientation, caused by tap forces, dissipate.

## 1.1 Contributions of This Thesis

This thesis is built upon the paper *HMOG: A New Biometric Modality for Continuous Authentication of Smartphone Users* by Zdeňka Sitová, Jaroslav Šeděnka, Qing Yang, Ge Peng, Gang Zhou, Paolo Gasti, and Kiran S. Balagani, which has been submitted for a journal review [34]. The paper proposes 96 HMOG features and evaluates their performance for continuous authentication and biometric key generation, using typing data collected from 100 subjects under two conditions: sitting and walking. The paper also analyzes energy consumption of HMOG feature computation.

**Table 1.1:** Summary of equal error rates (EERs) achieved with HMOG features alone, and when fused with state-of-the-art features—tap characteristics and keystroke dynamics (KD). The performance improved with longer authentication scans. Here, we present results for the shortest (20 seconds) and longest (140 seconds) authentication scans, with scaled Manhattan—the best performing verifier, and with weighted sum score-level fusion.

| | *Sit* | | *Walk* | |
|---|---|---|---|---|
| | *20-sec* | *140-sec* | *20-sec* | *140-sec* |
| *HMOG* | 21.95% | 17.97% | 18.53% | 12.69% |
| *Tap* | 20.99% | 14.15% | 21.22% | 14.62% |
| *HMOG + Tap* | 15.86% | 11.68% | 14.12% | 8.63% |
| *HMOG + Tap + KD* | **15.44%** | **10.2%** | **12.93%** | **7.07%** |
| *Tap + KD* | 19.11% | 10.86% | 17.93% | 10.51% |

In this thesis, we focus on the evaluation of continuous authentication performance. Our extensive evaluation of HMOG features on a dataset of 100 users[1] who typed on the smartphone led to the following findings (see Table 1.1 for summary of performance):

1. HMOG features extracted from accelerometer and gyroscope signals outperformed HMOG features from magnetometer.
2. HMOG features achieved higher authentication accuracies for walking compared to sitting. We investigated why HMOG had a superior performance during walking by observing the performance of HMOG *during* taps and *between* taps (i.e., the segments of the sensor signal that lie between taps) separately. Our results suggest that HMOG can capture movements caused by walking in addition to micromovements caused by taps.
3. Augmenting HMOG features with tap characteristics (e.g., tap duration and contact size) considerably improved authentication performance.
4. HMOG features complement tap and keystroke dynamics features, especially for short authentication scans at which tap and keystroke dynamics features fare poorly.

[1] The dataset is available at `http://www.cs.wm.edu/~qyang/hmog.html`. We also described the data and its release in a poster [38].

## 1.2 Organization

This thesis is organized as follows. We provide background on biometrics and review related research in chapter 2, present the description of HMOG features in chapter 3, describe the details of our dataset and evaluation methodology in chapter 4, present the results of the authentication experiments in chapter 5, discuss the results in chapter 6, and conclude this thesis in chapter 7.

## Chapter 2

# Background and Related Work

## 2.1 Background: Biometrics

The focus of this thesis is continuous authentication with behavioral biometrics—the HMOG modality.

Biometrics are automated methods of authentication based on measurable human *physiological* (e.g. fingerprint or iris pattern) or *behavioral* (e.g. voice sample or signature) characteristics [29]. These characteristics can be referred to by different terms such as: *traits*, *indicators*, *identifiers* or *modalities* [21]. Physiological biometrics are usually more reliable and accurate than behavioral, as the physiological characteristics are easier to repeat, and often are not affected by current (mental) conditions such as stress or illness [28].

*Identity verification* (or *authentication* [21]) occurs when the user claims to be in the system, and the system compares the biometric data to only one record in the database. It is in contrast to *identification*, where the user's identity is unknown, and the user can be anywhere in the database or may not be there at all [29].

There can be one entry point in the authentication process, e.g. the user places her finger on a fingerprint reader; or the authentication process can be *continuous*, e.g. the system can silently authenticate the user each minute, using her previously collected data. The time interval in which user is being authenticated is called *authentication scan* or *window*.

The authentication process has two main phases, *enrollment* and *recognition* [21].

**Enrollment.** During the enrollment phase, user's data is collected, processed and a *template* is stored in a database. The template typically consists of *features*, extracted from raw biometric data.

A user may fail to enroll. The most straightforward example of the *failure to enroll* is when a user without fingers tries to get authenticated with a fingerprint reader. However, the user may also have poor quality ridges—Jain and Ross [20] claim that there is empirical evidence that about 4% of the population's fingers cannot be easily recognized by some of the existing sensors. With behavioral biometrics, failure to enroll can occur when the user does not exhibit the behavior which is used to build the template. A user, who uses voice control to compose messages, cannot use the HMOG modality for her authentication.

**Recognition.** In the recognition phase, biometric data of a user who tries to enter the system is compared to the template. The similarity is expressed by a *score*, which can have different interpretations depending on the comparison technique (*verifier*) used—typically the distance of the data from the template, or the probability that the data come from the template.

*Genuine scores* result from comparison of legitimate (*genuine*) user to the template, *impostor scores* from comparison of *impostor* user to the template. When the impostor is not trying to mimic genuine user's behavior, the impostor scores are called *zero-effort*.

The user can enter the system when her data are similar enough to the template. A *threshold* specifies how similar the data has to be to the template for the verification request to be accepted. Such a threshold is established before the biometric system is implemented in practice, using two performance metrics:

- *False Acceptance Rate (FAR)* = from all impostor users's attempts, percentage of attempts in which the impostor user is incorrectly accepted (= percentage of incorrectly classified impostor scores from all impostor scores); and
- *False Rejection Rate (FRR)* = from all genuine user's attempts, percentage of attempts in which the genuine user is incorrectly rejected (= percentage of incorrectly classified genuine scores from all genuine scores).

In an ideal biometric system, both FAR and FRR are 0%. When the acceptance threshold is changed and FAR decreases, FRR increases, and vice versa. A single number is often reported as an accuracy measure of a biometric system—*Equal Error Rate (EER)*, which equals to FAR and FRR at a threshold, where FAR = FRR. Lower EER means better performance. The EER can be computed from scores of the whole user population (*population EER*), or from each user's data individually (*user-wise EER*).

## 2.2 Related Work

In this section, we review related research and highlight differences between existing work and our work.

Since HMOG features are collected during taps, we review existing work that uses tap activity to authenticate smartphone users. (Other work on continuous smartphone users authentication include swipe-based [16, 32] and gait-based [12, 36] behaviors.)

Trojahn and Ortmeier [35] used digraph, pressure and size of finger features extracted from a 7-digit or 12-letter long phrase and achieved FAR 9.53% at FRR 5.88% on 35 users. The authors used an artificial neural network based verifier. Authentication vector was created by averaging seven samples.

Zheng et al. [39] reported averaged EER 3.65% on 80 users inputting 4- and 8-digit PINs. Features evaluated in this work were pressure, size of finger (both measured at press and release), key hold, key interval, and magnitude of acceleration and angular velocity (at press, release, maximum, minimum and average of each PIN digit). The authors used dissimilarity score verifier, and one authentication vector corresponded to one tap.

Feng et al. [14] performed experiments on 40 users in fixed-text setting with 5- to 60-character long authentication windows, achieving EER 1% with 40-characters long window. The authors used key hold, key interval and pressure features, and performed 2-class verification with decision tree, Bayesian networks and random forest.

Gascon et al. [17] evaluated accelerometer, gyroscope and orientation sensor[1] based features extracted from a time window during user's typing burst. Twelve genuine users typed the same short predefined text ($\approx 160$ characters) more than ten times, while other 303 impostors entered the text only once. There was a group of eight

[1] Orientation sensor is a software-based sensor that derives its data from the accelerometer and the geomagnetic field sensor [1].

users, which were hardly distinguishable (with AUC $< 0.8$); remaining four users could be identified with average FAR 1% and FRR 8%. The verifier was 2-class linear SVM.

Bo et al. [5] used mean magnitude of acceleration and angular velocity during a gesture, as well as touch coordinates, touch pressure, and touch event duration. Taps, scrolls and flings were captured from applications such as mail and social networks. Ten users behaved as phone owners and 90 other users were using owner's phones as guests (with 50 guests in average for each owner). For sitting users, mean FAR is 0% when making judgement after observing 13 taps, FRR almost approaches 0 with only 2 observations. The authors rely on 2-class SVM, which justifies the extremely low EER. In fact, the authors report accuracy 72.36% and FAR 24.99% with 1-class SVM in sitting scenario. EER of gesture-based features in the walking scenario was higher than in sitting (FRR 18% after 4 gestures), and therefore for walking users the authors rely on gait recognition, achieving EER 0% after observing 3 gestures.

In Tables 2.1 and 2.2, we summarize the state-of-the-art in tap-based authentication,[2] and highlight various aspects of each work, such as

1. how the taps were collected—did the user compose free-text or type predefined fixed-text,
2. which body motion conditions (e.g., sitting and walking) were considered,
3. number of subjects (including splitting subjects into owners and attackers, wherever appropriate),
4. how the verifier was trained (training on both victim and impostor data usually allows lower EER compared to training on victim data only),
5. how the authentication vector was created, and
6. which features were used (e.g., motion-sensor-, keystroke-, or touch-based).

[2] We review only studies that used taps from a virtual keyboard. See Clarke and Furnell [10, 11], Buchoux and Clarke [6], Maiorana et al. [26], and Campisi et al. [7] for papers that have performed tap-based authentication with a hardware keyboard.

**Table 2.1:** Comparison of our features with related work on smartphone tap/typing authentication. *ACC* = for acceleration; *AV* = angular velocity; *MF* = magnetic field.

| | *Features* | | | *Best result* | |
|---|---|---|---|---|---|
| | *Motion-based* | *Tap* | *Keystroke* | *FAR* | *FRR* |
| *Trojahn et al. [35]* | ✗ | pressure, contact size | digraph | 9.53% | 5.88 |
| *Zheng et al. [39]* | magnitude of *ACC* and *AV* at press, release, max., min., and avg. of PIN | pressure, contact size at press and release | key interval, key hold | EER = 3.65% | |
| *Feng et al. [14]* | ✗ | pressure | key interval, key hold | EER = 1% | |
| *Gascon et al. [17]* | *ACC*, *AV* and orientation features extracted from typing burst | ✗ | ✗ | 1%[a] | 8%[a] |
| *Bo et al. [5]* | mean magnitude of *ACC* and *AV* during tap | coordinate, pressure, duration | ✗ | 0%[b], 24.99%[c] | 0%[b] |
| *This work* | 60 resistance and 36 stability features extracted during tap from *ACC*, *AV*, and *MF* | contact size (9 f.), duration, velocity between taps | digraph, key hold | EER = 6.92% | |

*a*. Results reported for most distinguishable genuine users only (4 out of 12).
*b*. Trained with impostor data.
*c*. Trained with victim data only.

**Table 2.2:** Comparison of details of our authentication experiments with related work on smartphone tap/typing authentication. *COND* = condition; *GEN* = no. of genuine users; *IMP* = avg. no. of impostors.

| | *COND* | *Free text* | *Authentic. vector* | *GEN* | *IMP* | *Verifier* | *Training source* |
|---|---|---|---|---|---|---|---|
| *Trojahn et al. [35]* | sit | ✗ | average of seven samples | 35 | 34 | ANN | unknown |
| *Zheng et al. [39]* | sit | ✗ | each tap | 80 | 79 | dissimilarity score | victim |
| *Feng et al. [14]* | sit | ✗ | 5-60 character window | 40 | 39 | decision tree, random forest, Bayes network | impostor & victim |
| *Gascon et al. [17]* | sit | ✗ | time window | 12 | 303 | linear 2-class SVM | impostor & victim |
| *Bo et al. [5]* | sit, walk | ✓ | each gesture[a] | 10 | 50 | 2-class SVM[b] | impostor & victim[c] |
| *This work* | sit, walk | ✓ | time window | 90 | 98 sit, 92 walk | SM, SE, 1-class SVM | victim |

[a] Judgement is made after 1-13 gestures.
[b] 1-class SVM for initial training.
[c] Victim only for initial training with 1-class SVM.

Among previous papers [5, 17, 39], which have used motion sensors for user authentication, Zheng et al. [39] used fixed pins while Gascon et al. [17] used fixed phrases. The only work that used free-text typing and also the only one to authenticate users under walking condition is the paper of Bo et al. [5]. Therefore, we believe that this is closest work to our paper, and highlight the differences between our paper and [5] as follows:

1. We performed experiments on a large-scale dataset containing 100 users (90 users qualified as genuine, and 92 or more as impostors), while [5] used only 10 genuine users and 50 impostors (on average) from a dataset of 100 subjects. As the genuine population size in [5] is too small, it is difficult to assess how accurately the reported FARs/FRRs represent the achievable authentication error rates with movement-based features, given that the number of users is critical factor in assessing the confidence on empirical error rates.
2. We introduced and evaluated a wide range of micro-movement features, while [5] used only two, which are also used in our work (mean magnitude of acceleration and angular velocity during gesture).
3. Our evaluation is more comprehensive and includes detailed comparison and fusion with additional features such as touch and keystroke. This allowed us to report how fusion with different types of features impacted authentication as well as BKG performance. In contrast, [5] reports only results for movement-based features combined with touch features.
4. HMOG features performed well in both sitting and walking condition, while [5] resorted to gait features for walking.

Chapter 3

# Description of HMOG Features

Recall from the previous chapter that the biometric template typically consist of *features*, extracted from raw biometric data. This is also the case of HMOG biometric. The HMOG features are presented in this chapter.

We define two types of HMOG features: grasp *resistance* and grasp *stability*—that are computed from accelerometer, gyroscope, and magnetometer sensor readings. Because HMOG features aim to capture the subtle micromovements and orientation patterns of a user while tapping on the screen, we extract HMOG features *during* or *close to* tap events.

Let $X, Y$, and $Z$ represent the readings from a sensor (accelerometer, gyroscope, or magnetometer) in $x$, $y$, and $z$ axes respectively. Let $M = \sqrt{X^2 + Y^2 + Z^2}$ represent the magnitude of acceleration, angular speed, or magnetic field. Grasp resistance and stability features are computed as follows.

## 3.1 Grasp Resistance Features

Grasp resistance features measure the resistance of a hand grasp to the forces (or pressures) exerted by touch/gesture events. We quantify resistance as the change (or perturbation) in movement and orientation (i.e., acceleration, angular velocity and magnetic field) caused by a tap event. We define grasp resistance features as follows:

1. **Mean** of $X, Y, Z$, and $M$ during a tap event.
2. **Standard deviation** of $X, Y, Z$, and $M$ during a tap event.

Figure 3.1 illustrates variables used in features 3 through 5. (The features are explained using $Z$-axis as an example.)

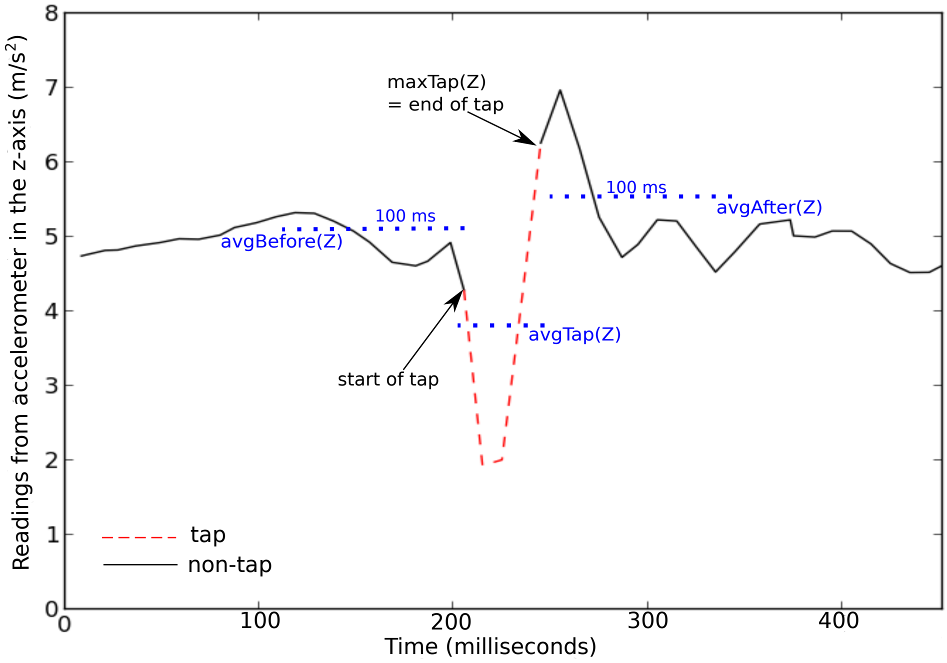


**Figure 3.1:** Variables used for calculating grasp resistance features 3-5, showed on sample data.

3. **Difference** between $X$, $Y$, $Z$, and $M$ values before and after a tap event. We calculate this as

   `avg100msAfter(Z) - avg100msBefore(Z),`

   where `avg100msBefore(Z)` is the mean of $Z$ values in 100 ms window before tap start time and `avg100msAfter(Z)` is the mean of $Z$ values in 100 ms window after tap end time.
4. **Mean change** in $X$, $Y$, $Z$, and $M$ values caused by a tap. We calculate this as

   `avgTap(Z) - avg100msBefore(Z),`

   where `avgTap(Z)` is the mean of $Z$ values during a tap event.
5. **Maximum change** in $X$, $Y$, $Z$, and $M$ values caused by a tap. This is calculated as

   `maxTap(Z) - avg100msBefore(Z),`

   where `maxTap(Z)` is the maximum $Z$ value during a tap event.

Five grasp resistance features from three sensors and four types of readings ($X$, $Y$, $Z$, and $M$) lead to $5 \times 3 \times 4 = 60$ features.

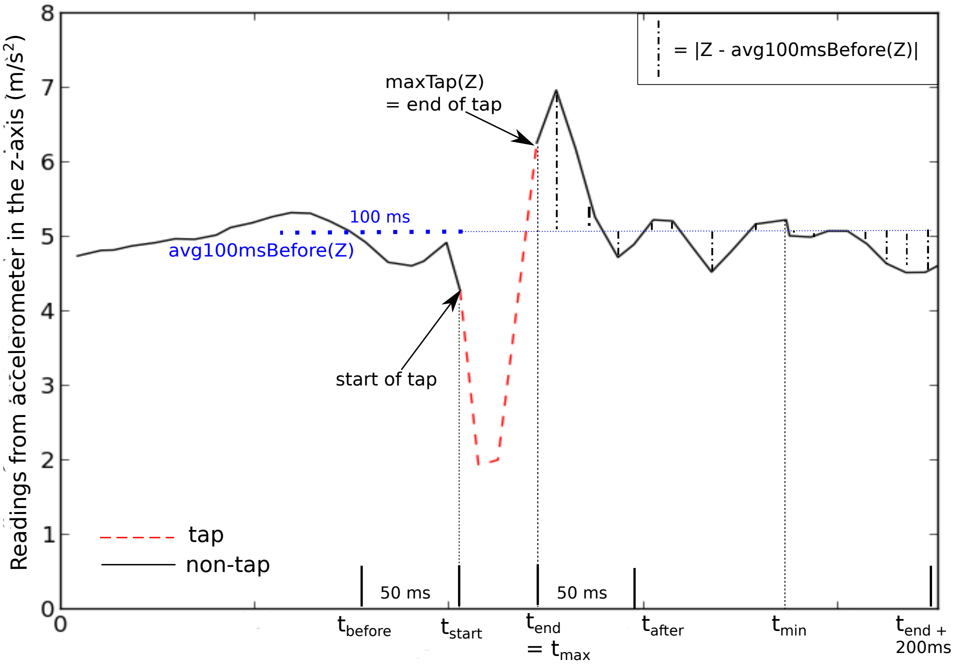


**Figure 3.2:** Illustration of key variables for computing grasp stability features 1-3, showed on sample data.

## 3.2 Grasp Stability Features

Stability features quantify how quickly the perturbations caused by a finger-force from a tap event disappear after the tap event is complete. We compute grasp stability features as follows (Figure 3.2 illustrates variables used in features 1 through 3):

1. **Time duration to achieve movement and orientation stability after a tap event**. This feature is calculated as

   $$t_{min} - t_{end},$$

   where $t_{end}$ can be obtained using the tap event handler, $t_{min}$ is calculated as follows.

   Let $T = (t_1, \ldots, t_n)$ represent the series of timestamps collected between $t_{end}$ and $t_{end} + 200$ ms, and $R = (Z_1, \ldots, Z_n)$ the corresponding sensor readings. We define $\texttt{avgDiffs}(t_i)$ as an average of $|Z_j - \texttt{avg100msBefore(Z)}|$ for $j = i \ldots n$.

   $t_{min}$ is such that $\texttt{avgDiffs}(t_{min}) = \min_{t_i \in T}(\texttt{avgDiffs}(t_i))$.

2. **Normalized time duration for mean sensor value to change from before tap to after tap event**, calculated as

$$\frac{t_{after\_center} - t_{before\_center}}{\texttt{avg100msAfter(Z)} - \texttt{avg100msBefore(Z)}},$$

   where $t_{after\_center}$ is the center of the 100 ms window after a tap event, and $t_{before\_center}$ is the center of the 100 ms window before the tap event.

3. **Normalized time duration for mean sensor values to change from maximum sensor value during a tap to after tap event**, calculated as

$$\frac{t_{after} - t_{max\_in\_tap}}{\texttt{avg100msAfter(Z)} - \texttt{maxTap(Z)}},$$

   where `maxTap(Z)` is the maximum sensor value during a tap, and $t_{max\_in\_tap}$ is the time when this value occurred.

We extracted the above three grasp stability features for three sensors and four types of sensor readings ($X$, $Y$, $Z$ and $M$), for a total of $3 \times 3 \times 4 = 36$ features.

## Chapter 4

# Dataset and Evaluation Methodology

## 4.1 Dataset

To evaluate HMOG features, we used sensor data collected from 100 smartphone users (53 male, 47 female) during eight free text typing sessions [38].[1] Users answered three questions, typing at least 250 characters for each answer. In four of the collection sessions, users typed while sitting; in the remaining four, they typed while walking in a controlled environment.[2]

For each user, an average of 1 193 taps per session (standard deviation: 303) and 1 019 key presses (standard deviation: 258) were collected. The average duration of a session was 11.6 minutes, with a standard deviation of 4.6 minutes. All data was collected using a Samsung Galaxy S4 [3] smartphone with Android OS 4.2.2.

The dataset contains accelerometer, gyroscope and magnetometer sensor readings (sampling rate 100 Hz) as well as raw touch data collected from the touchscreen, touch gestures (e.g., tap, scale, scroll, and fling), key press, and key release latencies on the virtual keyboard. Due to security concerns, Android OS forbids third-party applications to access touch and key press data generated on the virtual keyboard. Therefore, a virtual keyboard was designed for data collection, that mimicked the look, feel, and functionality of default Android keyboard, including the autocomplete option, which the users were free to use.

[1] The dataset is available at `http://www.cs.wm.edu/~qyang/hmog.html`.

[2] For one user, sensor data in two sessions were not recorded, and therefore HMOG features were not extracted.

[3] The dataset does not contain pressure information, because Samsung Galaxy S4 cannot record it.

During data collection users were allowed to choose the orientation of the smartphone (i.e., landscape or portrait). Because less than 20 users typed in landscape orientation, we performed all authentication experiments with data collected in portrait mode.

## 4.2 Design of Authentication Experiments

Here, we provide details of the experiments to evaluate the continuous authentication performance of HMOG features. The experiments were performed offline (after data collection) with MATLAB [27].

**1-Class Verifiers.** We performed verification experiments using three verifiers: scaled Manhattan (SM), scaled Euclidian (SE), and 1-class SVM. (Henceforth, we use "SVM" to refer to "1-class SVM".) We chose these verifiers because previous work on behavioral authentication has shown that they perform well. For instance, SM and SVM were top performers in a study on keystroke authentication of desktop users by Killourhy and Maxion [23]. SVM performed well in experiments on touch-based authentication of smartphone users by Serwadda et al. [32]. SE is a popular verifier in biometrics (see for example [4, 18]).

Parameter tuning was not required for SM and SE. However, for SVM, we used RBF kernel and performed a grid search to find the parameters (for $\gamma$, we searched through $2^{-13}$, $2^{-11}$, $2^{-9}$, $\ldots$, $2^{13}$; and for $\nu$, we searched through 0.01, 0.03, 0.05, 0.1, 0.15 and 0.5). We used LIBSVM library [8] for SVM.

We did not include 2-class verifiers in our evaluation. To train a 2-class verifier, in addition to data from smartphone owner, biometric data from a large set of other users (non-owners) is required. Because sharing of biometric information between smartphone users leads to privacy concerns, we believe that 1-class verifiers are more suitable for smartphone authentication. (A similar argument was made in [33].)

**Training and Testing.** For experiments in sitting and walking conditions, we used the first two sessions for training and the remaining two for testing. We extracted HMOG features during each tap. Thus, each training/testing vector corresponded to one tap. With keystroke dynamics features, each vector corresponded to one key press.

For SM and SE, the template consisted of the feature-wise average of all training vectors. We used user-wise standard deviations for each feature for scaling. For SVM, the template was constructed using all training vectors. Users with less than 80 training vectors in their template were discarded from authentication. As a consequence, ten users failed to enroll (and were not included in our experiments).

We created authentication vectors by averaging test vectors sampled during $t$-seconds scan. We report results for authentication scans of $t$ = 20, 40, 60, 80, 100, 120 and 140 seconds. We chose these scan lengths to cover both low and higher authentication latencies. Our preliminary experiments showed that for scans longer than 140 seconds, there is minimal improvement in authentication performance.

**Quantifying Authentication Performance.** We generated two types of scores, genuine (authentication vector was matched against template of the same user) and zero-effort impostor (authentication vector of one user was matched against the template of another). We used population equal error rate (EER) to measure the performance.

## 4.3 Comparing HMOG to Other Feature Sets

We compared the authentication performance of HMOG features with tap characteristics (such as duration and contact size) and keystroke dynamic features (key hold and digraph latencies).

Direct comparison of results reported by the researchers who designed the features is technically impossible due to usage of different datasets and authentication parameters. Therefore, we extracted three feature sets described below (tap, key hold and digraph) from our dataset and performed identical experiments as with HMOG.

**Tap Features.** We chose 11 commonly used (see Section 2.2) tap characteristics:

- duration of the tap;
- contact size features: mean, median, standard deviation, 1st, 2nd and 3rd quartile, first contact size of a tap, minimum and maximum of the contact size during the tap (9 features); and
- velocity (in pixels per second) between two consecutive *press* events belonging to two consecutive taps.

We could not extract pressure features, because of the absence of pressure information in our dataset.

**Key Hold Features.** Key hold latency is the down-up time between press and release of a key. We used 89 key hold features, each corresponding to a key on the virtual keyboard.

**Digraph Features.** Digraph latency is the down-down time between two consecutive key presses. We used digraph features for combinations of the 35 most common keys in our dataset.[4] Thus we have $35^2 = 1\,225$ digraph features.

**Score-level Fusion.** To determine whether HMOG features complement existing feature sets, we combined tap, key hold, digraph and HMOG features using weighted sum score-level fusion. We chose this method because it is simple to implement, and has been shown to perform well in biometrics [19]. We used the technique of Locklear et al. [25] to ensure that weights sum to one and proportion of weights is preserved when scores from some feature sets were missing (e.g. due to lack of accelerometer data). We used grid-search to find the weights which led to the best authentication performance.

## 4.4 Feature Selection, Preprocessing, and Transformation

To improve authentication performance, we performed preprocessing (e.g., outlier detection), feature selection, and feature transformation with Principal Component Analysis. All procedures described in this section were performed on training data, separately for each verifier, feature set and for sitting and walking conditions.

**Feature Selection.** We evaluated two methods for feature selection: Fisher score [13], and minimum-Redundancy Maximum-Relevance (mRMR) [31]. Fisher score ranking was computed independently for each feature as the ratio of *between-user* to *within-user* variance. The higher the Fisher score, the higher the discriminability of the corre-

[4] All 26 alphabetic keys, 5 keyboard switches (shift, switch between numerical and alphabetical keyboard, delete, done, return) and 4 special characters (space, dot, comma and apostrophe). The availability of other keys in our training data was extremely low ($<1$ in average per user).

sponding feature. We experimented with feature subsets whose sum of Fisher scores accounted for 85, 90, 95, 98 and 100% of the sum of Fisher scores of all features. Our experiments showed that Fisher score performed better for HMOG features, while mRMR performed well for tap features. With key hold and digraph features, the best performing feature set contained all the features.

For HMOG, feature selection was performed on (1) all HMOG features, and (2) features extracted from best performing sensors only. Using the best performing sensors only led to better results. The best performing sensors were accelerometer combined with gyroscope for SM and SVM in all scans and conditions, and for SE 20-second scans in sitting and 20- and 40-second scans in walking. For other scans with SE, using only gyroscope features led to best performance.

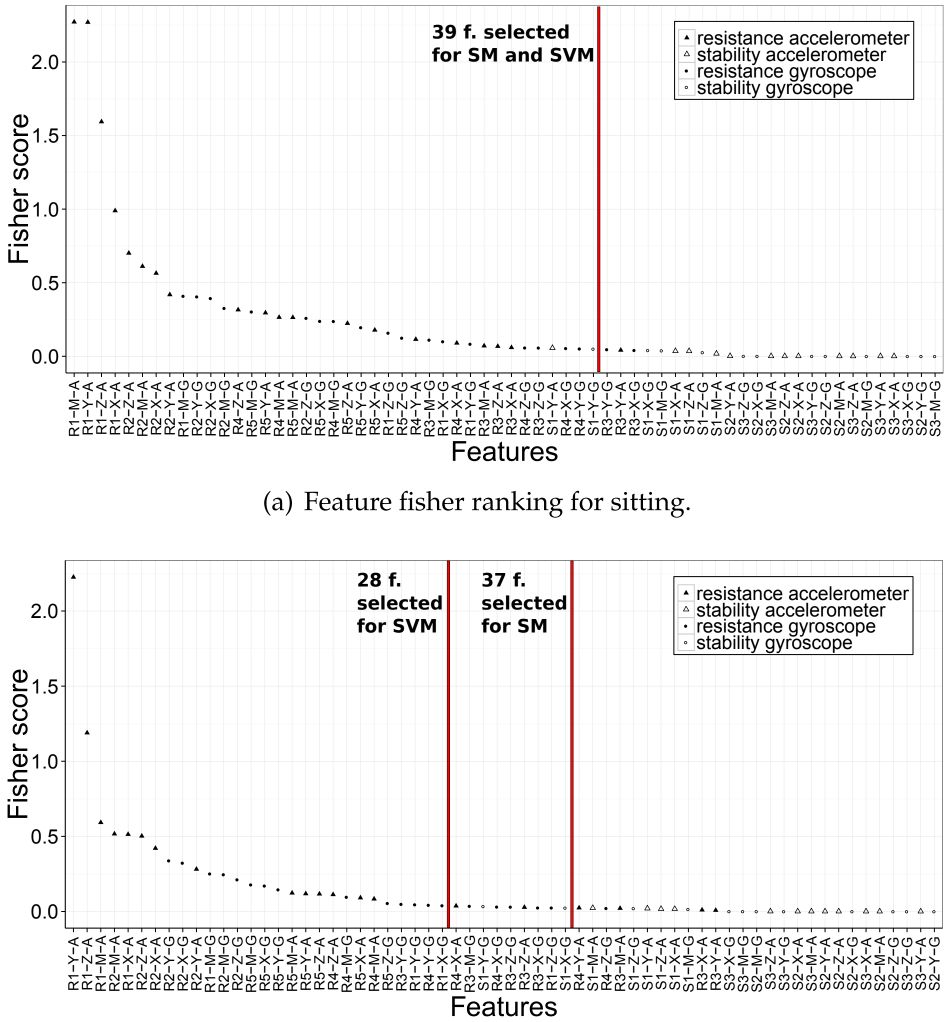
The following parameters for Fisher score ranking provided the best authentication results: 98% (39 features) for SM and SVM during sitting; 98% (37 features) for SM during walking; and 95% (28 features) for SVM during walking. For SE, we achieved lowest EER by including resistance features only, than with features selected by feature selection. Figure 4.1 reports the ranking of features during sitting (4.1(a)) and walking (4.1(b)), and shows the selected features.

For tap features, with SM verifier we achieved the best results with 3 features chosen by mRMR (threshold 0) and for SE and SVM with 2 features (threshold 0.1). The best three features according to mRMR are (in this order): duration of the tap; mean of contact size; and velocity between two consecutive down events.

**Outlier Removal.** For HMOG and tap templates, we evaluated the interquartile outlier removal, i.e. kept only values in interval $[firstQ - k * iqr, thirdQ + k * iqr]$, where $firstQ$ is the first quartile, $thirdQ$ is the third quartile, $iqr = thirdQ - firstQ$ is the interquartile range, and $k$ is the parameter. The higher $k$, the less values are considered as outliers—we experimented with $k$ = 3, 10, 20, 50, 100, 200 and 500. Experiments with SM verifier showed that outlier removal does not improve authentication accuracy, so we did not perform it in our experiments.

For key hold and digraph, using only outlier removal and no feature selection or transformation led to the best authentication results. Outlier removal was done by restricting two parameters: (1) maximal length (values longer than $l$ ms were discarded), and (2) minimum number of occurrences of a feature in template (if the feature occurs less than $m$ times in user's template, then this feature is discarded). The values evaluated for $l$ were 100, 200, 300, 400, 500 and 1000 for key hold and 200, 350, 500, 650 and 800 for digraph. For $m$ value we experimented with 2, 5, 10, 15, 20, 40 and 60. The optimal $l$ value was 200 for key hold and ranged between 350–500 for digraph; the optimal $m$ value ranged between 2–60 for key hold and between 2–5 for digraph.

**Feature Transformation.** We evaluated Principal Component Analysis (PCA), which transforms original features into principal components that are subsequently used in authentication experiments. Our motivations for using PCA are (1) to remove correlation between features to meet the assumptions in SE and SM, and (2) to reduce dimensionality by using only those principal components, which explain most of variance in data. We used PCA on (1) all features (except magnetometer features), and (2) features selected by feature selection. We experimented with components explaining 90%, 95%, 98% and 100% of total variance in order to find the threshold for dimensionality reduction. PCA improved EER for HMOG features with SE when performed on resistance features and with SVM in sitting when performed on features selected by feature selection. PCA performed on all tap features improved results with SM and SE.



(a) Feature fisher ranking for sitting.

(b) Feature fisher ranking for walking.

**Figure 4.1:** HMOG accelerometer and gyroscope features sorted by Fisher score computed from training data. Higher scores correspond to features with higher discriminative power. Overall, resistance features perform better than stability. We show selected features for SM and SVM. For SE, only resistance features led to better performance than feature selection. Feature names (on x-axis) are encoded as follows:
[**R**esistance|**S**tability][ID][-][**X**|**Y**|**Z**|**M**agnitude][-][**A**cc|**G**yro]

Chapter 5

# Authentication Results

In this chapter, we report authentication performance of HMOG features. We compare the performance of HMOG with keystroke and tap features and also report results with fusion.

## 5.1 Performance of HMOG Features

With SM and SVM verifiers, HMOG features extracted from both accelerometer and gyroscope outperformed those extracted from individual sensors (see Table 5.1 for EERs). With SE, HMOG features from gyroscope outperformed all other sensors and sensor combinations, except for 20- and 40-second scan lengths in walking condition and 20-second scan in sitting, where accelerometer and gyroscope features performed best. HMOG features from magnetometer performed consistently worse than accelerometer and gyroscope features with all verifiers, in both sitting and walking conditions. Combining magnetometer features with features from accelerometer and gyroscope did not improve performance.

Resistance features outperformed stability features in both walking and sitting conditions (and also had a higher Fisher score, see Figure 4.1). This suggests that the ability of resistance features to discriminate between users is higher than that of stability features. For SM and SVM, we achieved the lowest EERs when we used both resistance and stability features and performed Fisher score-based feature selection (see Figure 4.1 for the selected features). For SE, using only resistance features led to best EER. In some cases, using PCA after feature selection further lowered EERs. In Table 5.1, we indicate when feature selection and PCA led to the lowest EERs.

In Figure 5.1, we show the EERs of all verifiers under sitting and walking conditions, when the authentication scans varied between 20 and 140 seconds. Among the three verifiers, SM overall had lower EERs for both sitting and walking conditions and therefore we present the results only with SM hereafter.

**Table 5.1:** Summary of lowest EERs achieved with individual HMOG features. FS = Fisher score-based feature selection.

| | *Best performing motion sensors* | *Best performing features* | *Sitting* | *Walking* |
|---|---|---|---|---|
| *Scaled Manhattan* | accelerometer + gyroscope | resistance + stability (w/ FS) | 17.88% | 12.69% |
| *Scaled Euclidean* | gyroscope, or accelerometer + gyroscope | resistance | 19.15% (w/ PCA) | 13.88% (w/ PCA) |
| *1-class SVM* | accelerometer + gyroscope | resistance + stability (w/ FS) | 22.43% (w/ PCA) | 14.65% |

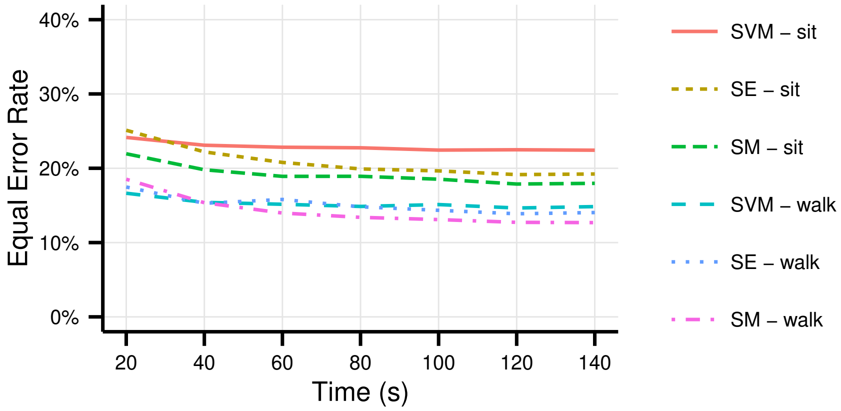


**Figure 5.1:** Comparison of HMOG features in sitting and walking conditions for three verifiers. The reported EERs are with PCA for SVM in sitting condition and for SE, and without PCA for SVM in walking condition and for SM.

## 5.2 Comparison of HMOG with Keystroke Dynamics and Tap Features

HMOG and tap features performed better than keystroke dynamics features in both sitting and walking conditions (see Figure 5.2 for authentication performance with SM—the relative performance of HMOG, tap, and keystroke dynamics features with SE and SVM verifiers was virtually the same as with SM). HMOG features outperformed tap features in walking condition, while tap outperformed HMOG in sitting. The performance of tap and keystroke dynamics features did not change significantly between sitting and walking. However, the performance of HMOG improved considerably (up to 5.52%) during walking.

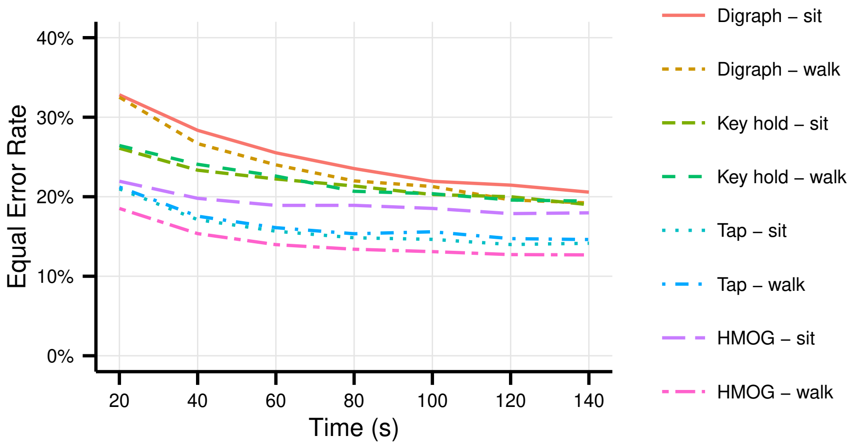


**Figure 5.2:** Comparison of EERs of HMOG with keystroke dynamics (i.e., key hold and digraph) and tap features with SM verifier.

## 5.3 Fusion of HMOG, Tap, and Keystroke Features

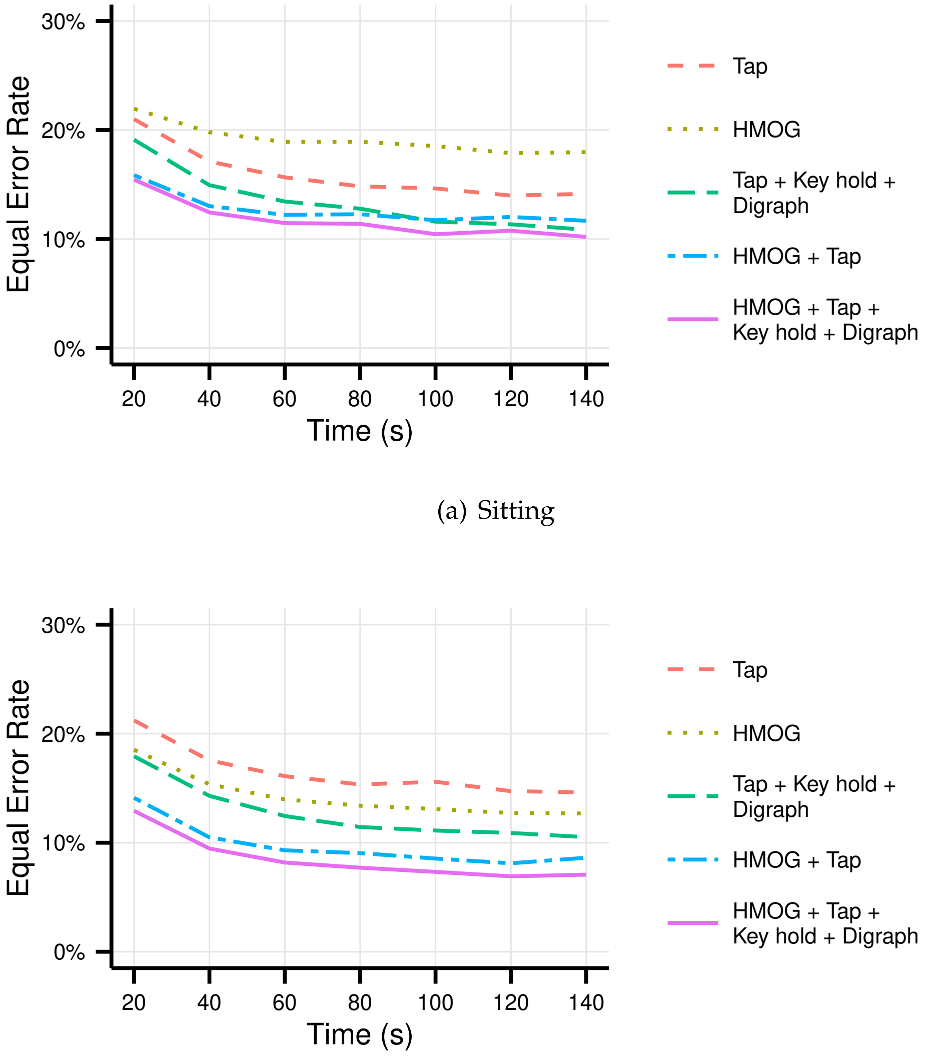
We used SM verifier and performed score-level fusion with the following feature combinations: {HMOG, tap, keystroke dynamics}; {tap, keystroke dynamics}; and {tap, HMOG}. Detailed fusion results for sitting and walking conditions are presented in figures 5.3(a) and 5.3(b), respectively. The lowest EERs achieved with fusion are summarized in Table 5.2.

Our results show that: (1) for both walking and sitting conditions, score-level fusion of all signals led to the lowest EER; and (2) fusing HMOG with tap features led to a decrease in EERs and either outperformed (in the case of walking and shorter scans in sitting) or was comparable (in the case of longer scans in sitting) to fusion of tap and keystroke dynamics (see figures 5.3(a)) and 5.3(b)).

Both (1) and (2) indicate that HMOG provides additional distinctiveness to tap and keystroke dynamics, especially in walking condition. (2) shows that keystroke dynamics can be replaced by HMOG features to obtain better performance in walking and in shorter scans in sitting condition.

**Table 5.2:** Summary of lowest EERs achieved with score-level fusion of HMOG, Tap, and Keystroke Dynamics (KD) Features. With Scaled Manhattan verifier.

| *Score-level fusion (with SM verifier)* | *Sitting* | *Walking* |
|---|---:|---:|
| HMOG, Tap, and KD | **10.20**% | **6.92**% |
| HMOG and Tap | 11.68% | 8.12% |
| Tap and KD | 10.86% | 10.51% |



(a) Sitting

(b) Walking

**Figure 5.3:** Score-level fusion of combinations of feature types with SM verifier.

## Chapter 6

# Discussion

Our results in Chapter 5 show that HMOG outperform tap features in walking condition, and that fusing HMOG features with other features improves the performance.

With all three verifiers, HMOG features achieve lower error rates for walking compared to sitting. This is not the case for the touch features and keystroke dynamics.

Here we investigate why HMOG features performed better during walking. Specifically, we investigate whether the high authentication accuracies of HMOG features during walking were due to hand movements caused by taps, or due to movements caused by walking, or a combination of both.

**Experiment setup.** We extracted 64 HMOG features from two segments of an accelerometer/gyroscope signal: (1) *during tap*, as discussed in previous sections; and (2) *between taps*, in which HMOG features were extracted when the user was *not* tapping the screen (see Figure 6.1). In (2), the signal between taps was segmented into non-overlapping blocks of 91 ms; one HMOG feature vector was extracted from each block. We selected 91 ms as the block size because it was the median duration of a tap in our training data. This ensured that the number of sensor readings used to extract a HMOG feature vector *between* and *during* tap remained same.

HMOG features extracted *during* taps use sensor readings from 100 ms before and 200 ms after a tap event (see Section 3). We extracted HMOG features *between* taps starting 300 ms after a tap until 300 ms before the next tap, to avoid any overlap between *during* and *between* HMOG features.

The average number of the training vectors per user for HMOG *during* taps was 1 122 for sitting, and 1 186 for walking. For *between*

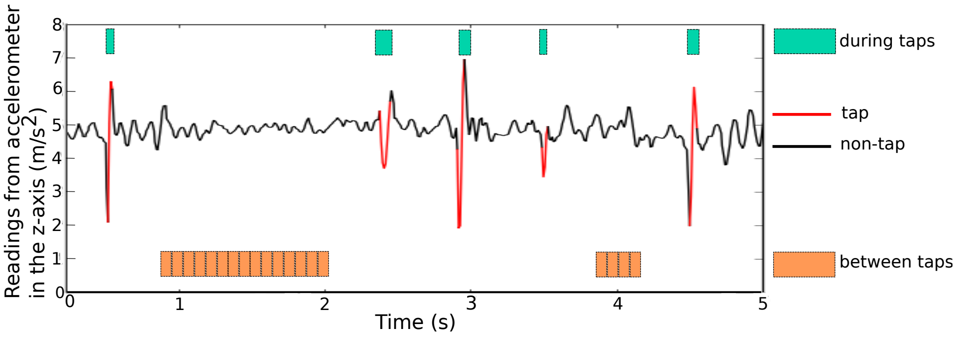


**Figure 6.1:** Area from which HMOG features were extracted. The figure shows a sample of readings from the z-axis of accelerometer in sitting condition.

taps, it was 7 692 for sitting and 7 462 for walking. The average number of testing vectors per user for HMOG features *during* taps was 897 for sitting and 972 for walking. For *between* taps, it was 5 885 for sitting and 5 768 for walking. User population was the same for both settings. Verification experiments were performed using SM.

**HMOG Features Extracted During vs. Between Taps.** We compared HMOG features extracted during taps with the same features extracted between taps for sitting and walking conditions. For sitting, HMOG features extracted *during* taps performed consistently better than those extracted *between* taps (see EERs in Figure 6.2). This indicates that HMOG features were able to capture distinctive hand micromovement patterns when the users tapped on the phone. Similarly, for walking, HMOG features extracted *during* taps performed better than those extracted *between* taps (see EERs in Figure 6.2). This again indicates that HMOG features capture users distinctive hand micromovement patterns when the user is tapping, regardless of the motion condition.

**Impact of Walking on HMOG Features Extracted Between Taps.** HMOG features extracted *between* taps during walking outperformed the same when extracted during sitting. In Figure 6.2, compare *between* tap EERs for sitting and walking. This indicates that HMOG features capture distinctive movements induced by walking, even in the *absence* of tap activity.

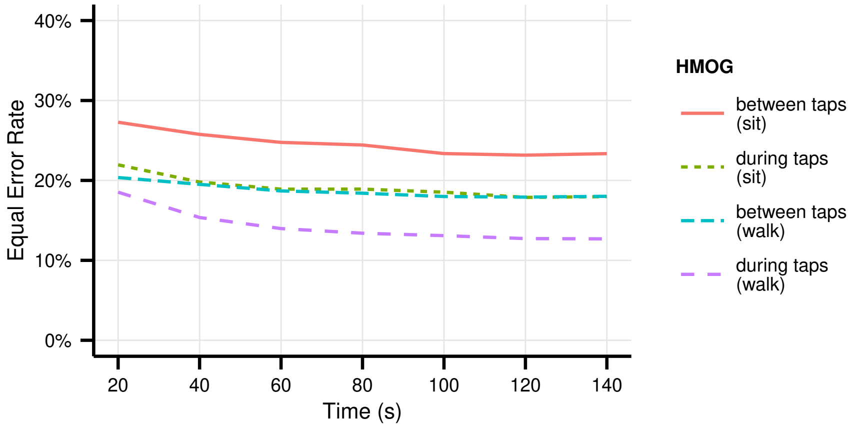


**Figure 6.2:** Performance of HMOG features when extracted from the tap and between the taps.

Supported by the above results, the high authentication accuracies achieved by HMOG features during walking can be jointly attributed to: (a) the distinctiveness in hand movements caused by tap activity and (b) the distinctiveness in movements caused by walking.

**Note.** Our data collection was not designed to collect gait related information (e.g., ground truth via video captures) and therefore, we could not perform experiments to directly compare our HMOG results with authentication results of gait based features.

Nevertheless, HMOG features do not overlap with features traditionally used for gait recognition. All our features are extracted from a time windows with median size 91 ms and the time window location is determined by taps. Traditional gait recognition methods detect users' steps from accelerometer and than extract features from the steps, or compare the signals of steps. See for example [12], achieved EER of 20.1% for 51 volunteers with phone in pocket attached to a belt or [36], used data of 31 users and reported 14.3% EER when the phone is held in hand and 13.7% EER when it is in a breast pocket. Work of Bo et al. [5] (already mentioned in Section 2.2) evaluates their gesture-based accelerometer and gyroscope features as sufficient only for the sitting condition; relying on gait-recognition-based features in walking.

## Chapter 7

# Conclusion

In this thesis we presented evaluation of HMOG, a novel behavioral biometric modality that harnesses hand micromovements during taps for continuous user authentication.

We evaluated the authentication performance of HMOG features on data from 100 users under two motion conditions (sitting, and walking). Our results show that HMOG features (best achieved EER 12.69%) outperform tap features (14.62%) in authentication of walking users. Moreover, HMOG combined with tap and keystroke dynamics features provide the best combined performance—as low as 10.2% in sitting and 6.92% in walking.

Surprisingly, HMOG features collected when a user is walking perform better than the same features collected while the user is sitting. Our analysis suggests that this is due to the ability of HMOG to capture body movements in addition to hand micromovements.

This thesis was built upon the paper *HMOG: A New Biometric Modality for Continuous Authentication of Smartphone Users* by Zdeňka Sitová, Jaroslav Šeděnka, Qing Yang, Ge Peng, Gang Zhou, Paolo Gasti, and Kiran S. Balagani, submitted for a journal review [34]. The paper proposes 96 HMOG features for continuous authentication of typing smartphone users, and our contribution is the offline evaluation of the continuous authentication performance. The paper moreover evaluates biometric key generation (BKG) performance of HMOG, and energy consumption of HMOG feature computation.

In future studies, HMOG during swipe and pinch gestures should be investigated. Although our preliminary experiments showed that our features do not perform well for swipe and pinch, we believe that by designing new HMOG features specifically for these touchscreen interactions, the performance of our biometrics can be improved.

# Bibliography

[1] Android developers: Position sensors. `http://developer.android.com/guide/topics/sensors/sensors_position.html`. Accessed: 2015-01-03.

[2] Byod insights 2013: A cisco partner network study. Cisco mConcierge, March 2013.

[3] S. Azenkot and S. Zhai. Touch behavior with different postures on soft smartphone keyboards. In *Proceedings of the 14th International Conference on Human-computer Interaction with Mobile Devices and Services (MobileHCI 2012)*, pages 251–260. ACM, 2012.

[4] M. Blanton and P. Gasti. Secure and efficient protocols for iris and fingerprint identification. In *Proceedings of the 16th European Symposium on Research in Computer Security (ESORICS 2011)*, pages 190–209. Springer-Verlag, 2011.

[5] C. Bo, L. Zhang, X. Li, Q. Huang, and Y. Wang. Silentsense: Silent user identification via touch and movement behavioral biometrics. In *Proceedings of the 19th Annual International Conference on Mobile Computing & Networking (MobiCom 2013)*, pages 187–190. ACM, 2013.

[6] A. Buchoux and N. L. Clarke. Deployment of keystroke analysis on a smartphone. In *Proceedings of the 6th Australian Information Security Management Conference*, pages 48:1–48:7. Edith Cowan University, 2008.

[7] P. Campisi, E. Maiorana, M. Lo Bosco, and A. Neri. User authentication using keystroke dynamics for cellular phones. *Signal Processing, IET*, 3(4):333–341, 2009.

[8] C. Chang and C. Lin. Libsvm: A library for support vector machines. *ACM Transactions on Intelligent Systems and Technology*, 2(3):27:1–27:27, 2011.

[9] S. Chatterjee and J. Chowdhuri. Comparison of grip strength and isomeric endurance between the right and left hands of men and their relationship with age and other physical parameters. *Journal of human ergology*, 20(1):41–50, 1991.

[10] N. L. Clarke and S. M. Furnell. Authenticating mobile phone users using keystroke analysis. *International Journal of Information Security*, 6(1):1–14, 2007.

[11] N. L. Clarke and S. M. Furnell. Advanced user authentication for mobile devices. *Computers & Security*, 26(2):109–119, 2007.

[12] M. Derawi, C. Nickel, P. Bours, and C. Busch. Unobtrusive user-authentication on mobile phones using biometric gait recognition. In *Proceedings of the 6th International Conference on Intelligent Information Hiding and Multimedia Signal Processing (IIH-MSP 2010)*, pages 306–311. IEEE, 2010.

[13] R. O. Duda, P. E. Hart, and D. G. Stork. *Pattern Classification. 2nd Edition*. Wiley-Interscience, 2001.

[14] T. Feng, X. Zhao, B. Carbunar, and W. Shi. Continuous mobile authentication using virtual key typing biometrics. In *Proceedings of the 12th IEEE International Conference onTrust, Security and Privacy in Computing and Communications (TrustCom 2013)*, pages 1547–1552. IEEE Computer Society, July 2013.

[15] I. M. Fiebert, K. E. Roach, J. W. Fromdahl, J. D. Moyer, and F. F. Pfeiffer. Relationship between hand size, grip strength and dynamometer position in women. *Journal of Back and Musculoskeletal Rehabilitation*, 10(3):137–142, 1998.

[16] M. Frank, R. Biedert, E. Ma, I. Martinovic, and D. Song. Touchalytics: On the applicability of touchscreen input as a behavioral biometric for continuous authentication. *IEEE Transactions on Information Forensics and Security*, 8(1):136–148, 2013.

[17] H. Gascon, S. Uellenbeck, C. Wolf, and K. Rieck. Continuous authentication on mobile devices by analysis of typing motion behavior. In *Sicherheit, Schutz und Zuverlässigkeit (Sicherheit 2014)*, pages 1–12. Gesellschaft für Informatik, 2014.

[18] S. Govindarajan, P. Gasti, and K. S. Balagani. Secure privacy-preserving protocols for outsourcing continuous authentication of smartphone users with touch data. In *Proceedings of the IEEE 6th International Conference on Biometrics: Theory, Applications and Systems (BTAS 2013)*, pages 1–8. IEEE, September 2013.

[19] A. K. Jain and A. Ross. Information fusion in biometrics. *Pattern Recognition Letters*, 24(13):2115–2125, 2003.

[20] A. K. Jain and A. Ross. Multibiometric systems. *Communications of the ACM*, 47(1):34–40, January 2004.

[21] A. K. Jain, R. Bolle, and S. Pankanti. *Introduction to Biometrics*. Springer US, 1996.

[22] A. Karlson, B. Bederson, and J. L. Contreras-Vidal. Studies in one-handed mobile design: Habit, desire and agility. Technical report, Computer Science Department, University of Maryland, 2006.

[23] K. S. Killourhy and R. A. Maxion. Comparing anomaly-detection algorithms for keystroke dynamics. In *Proceedings of the 39th IEEE/IFIP International Conference on Dependable Systems Networks (DSN 2009)*, pages 125–134. IEEE Computer Society, June 2009.

[24] K. Kim, W. Chang, S. Cho, J. Shim, H. Lee, J. Park, Y. Lee, and S. Kim. Hand grip pattern recognition for mobile user interfaces. In *Proceedings of the 18th Conference on Innovative Applications of Artificial Intelligence (IAAI 2006)*, volume 2, pages 1789–1794. AAAI Press, 2006.

[25] H. Locklear, S. Govindarajan, Z. Sitová, A. Goodkind, D. Brizan, A. Rosenberg, V. V. Phoha, P. Gasti, and K. S. Balagani. Continuous authentication with cognition-centric text production and

revision features. In *Proceedings of the International Joint Conference on Biometrics (IJCB 2014)*, pages 1–8. IEEE, 2014.

[26] E. Maiorana, P. Campisi, N. González-Carballo, and A. Neri. Keystroke dynamics authentication for mobile phones. In *Proceedings of the ACM Symposium on Applied Computing (SAC 2011)*, pages 21–26. ACM, 2011.

[27] MATLAB. *Version 8.2 (R2013b)*. The MathWorks Inc., 2013.

[28] V. Matyáš and Z. Říha. Biometric authentication systems. Technical report, Faculty of Informatics, Masaryk University, 2000.

[29] V. Matyáš and Z. Říha. Toward reliable user authentication through biometrics. *Security & Privacy, IEEE*, 1(3):45–49, May 2003.

[30] J. Napier. The prehensile movements of the human hand. *Journal of Bone and Joint Surgery*, 38(4):902–913, 1956.

[31] H. Peng, F. Long, and C. Ding. Feature selection based on mutual information: criteria of max-dependency, max-relevance, and min-redundancy. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 27(8):1226–1238, August 2005.

[32] A. Serwadda, V. V. Phoha, and Z. Wang. Which verifiers work?: A benchmark evaluation of touch-based authentication algorithms. In *Proceedings of the IEEE 6th International Conference on Biometrics: Theory, Applications and Systems (BTAS 2013)*, pages 1–8. IEEE, September 2013.

[33] C. Shen, Z. Cai, X. Guan, Y. Du, and R. Maxion. User authentication through mouse dynamics. *IEEE Transactions on Information Forensics and Security*, 8(1):16–30, January 2013.

[34] Z. Sitová, J. Šeděnka, Q. Yang, G. Peng, G. Zhou, P. Gasti, and K. S. Balagani. Hmog: A new biometric modality for continuous authentication of smartphone users. Submitted to IEEE Transactions on Information Forensics and Security.

[35] M. Trojahn and F. Ortmeier. Biometric authentication through a virtual keyboard for smartphones. *International Journal of Computer Science & Information Technology*, 4(5):1–12, October 2012.

[36] E. Vildjiounaite, S. Mäkelä, M. Lindholm, R. Riihimäki, V. Kyllönen, J. Mäntyjärvi, and H. Ailisto. Unobtrusive multimodal biometrics for ensuring privacy and information security with personal devices. In *Proceedings of the 4th International Conference on Pervasive Computing(PERVASIVE 2006)*, pages 187–201. Springer-Verlag, 2006.

[37] J. O. Wobbrock, B. A. Myers, and H. H. Aung. The performance of hand postures in front- and back-of-device interaction for mobile computing. *International Journal of Human-Computer Studies*, 66(12):857–875, December 2008.

[38] Q. Yang, G. Peng, D. Nguyen, X. Qi, G. Zhou, Z. Sitová, P. Gasti, and K. S. Balagani. A multimodal data set for evaluating continuous authentication performance in smartphones. In *Proceedings of the 12th ACM Conference on Embedded Network Sensor Systems (SenSys 2014)*, pages 358–359. ACM, 2014.

[39] Nan Zheng, Kun Bai, Hai Huang, and Haining Wang. You are how you touch: User verification on smartphones via tapping behaviors. Technical report, Department of Computer Science, College of William & Mary, 2012.