

**An Anomaly Detection Model for Signature  
Authentication on Mobile Devices**

نموذج لكشف التباين للتحقق من التوقيع على الاجهزة النقالة

**By**

**Shawq Salman Mahmood Al-Khafaji**

**Supervisor**

**Dr. Mudhafar Al-Jarrah**

**A Thesis Submitted in Partial Fulfillment of the Requirements for the  
Master Degree in Computer Science**

**Department of Computer Science**

**Faculty of Information Technology**

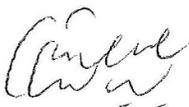
**Middle East University**

**January, 2018**

## Authorization

I, Shawq Al-Khafaji, authorize Middle East University to provide Libraries, Organizations, and Individuals with Copies of my Thesis on Request.

Name: Shawq Al-Khafaji

Signature: 

Date: 30.1.2018

## Thesis Committee Decision

This Thesis "An Anomaly Detection Model for Signature Authentication on Mobile Devices", was discussed and certified on 13 / 1 / 2018.

### Thesis Decision Committee Signature:

Dr. Mudhafar Al-Jarrah (Supervisor / Chairman) .....

Dr. Sharifa Murad Internal Examiner .....

Dr. Mohamad Shkokani External Examiner .....



27/1/2018.

## **Acknowledgment**

I want to thank ALLAH for His blessings and facilitate all necessary to accomplish my work in full.

I am pleased to extend my thanks and gratitude to my supervisor Dr. Mudhafar Al-Jarrah for his efforts in this thesis through his valuable guidance and follow-up to me. Without him I would not have finished this thesis.

I wish to express gratitude to the head of staff of MEU for assistance and efforts.

I must also thank all relatives and friends for their moral support during my academic journey.

Finally, my deep gratitude to my parents, for their patience and encouraging throughout the interval of studying and researching and to achieve my dream.

## **Dedication**

This thesis is dedicated to my Father, Mother.

## Table of Contents

<b>Cover Page</b> .....	<b>I</b>
<b>Authorization</b> .....	<b>II</b>
<b>Thesis Committee Decision</b> .....	<b>III</b>
<b>Acknowledgment</b> .....	<b>IV</b>
<b>Dedication</b> .....	<b>V</b>
<b>Table of Contents</b> .....	<b>VI</b>
<b>List of Tables</b> .....	<b>IX</b>
<b>List of figures</b> .....	<b>X</b>
<b>List of Appendixes</b> .....	<b>XI</b>
<b>List of Abbreviations</b> .....	<b>XII</b>
<b>Abstract</b> .....	<b>XIII</b>
<b>المخلص</b> .....	<b>XV</b>
<b>Chapter One Introduction</b> .....	<b>1</b>
1.1 Research Context .....	2
1.2 Problem Statement .....	4
1.3 Scope of Work.....	4
1.4 Goal and Objectives .....	4
1.5 Motivation .....	5
1.6 Significance of Work.....	5
1.7 Research Questions.....	6
1.8 Thesis Organization.....	6
<b>Chapter Two Background and Literature Review</b> .....	<b>8</b>
2.1 Introduction.....	9
2.2 Classification Methods.....	10
2.3 Biometric Technologies.....	11
2.3.1 Graphic Signature Authentication .....	12
2.3.2 Signature Recognition on Mobile Devices .....	16
2.3.3 One-Class Classification (Anomaly Detection) .....	16
2.3.4 Feature Extraction in One-Class Classification .....	17
2.4 Related Work .....	17
2-5 Summary of Related Work .....	20

<b>Chapter Three Methodology and the Proposed Model.....</b>	<b>22</b>
3.1 Methodology Approach .....	23
3.2 Outline of the Proposed Model .....	23
3.3 Methodology Steps .....	24
3.4 Features Selection.....	25
3.4.1 Feature Sets of Previous Studies .....	26
3.4.2 The Proposed Feature Sets .....	27
3.5 The Anomaly Detector .....	29
3.5.1 The Z-Score Anomaly Detector .....	29
3.5.2 The Average Absolute Deviation (AAD) Anomaly Detector .....	30
3.5.3 The Median Absolute Deviation (MAD) Anomaly Detector .....	31
3.6 The Proposed System.....	32
3.6.1 The Data Collection Module .....	32
3.6.2 The Data Collection Functions.....	35
3.6.3 The Authentication Module .....	36
3.6.4 Feature Extraction Program (Extract – Features).....	38
3.7 Error Metrics.....	39
<b>Chapter Four Experimental Results and Discussion .....</b>	<b>40</b>
4.1 Introduction .....	41
4.2 Objectives of the Experimental Work .....	41
4.3 Limitations of the Proposed Work.....	42
4.4 EER Analysis Steps.....	42
4.5 Feature Sets Selection .....	43
4.6 Analysis of the MOBSIG Dataset .....	44
4.7 Interfaces and Output of the Proposed TDSIG System.....	45
4.7.1 Screen Shots of the Proposed TDSIG System .....	45
4.8 Data Collection Using the Proposed TDSIG System.....	48
4.9 Comparison and Discussion of Results.....	49
4.9.1 Random Forgery Results.....	49
4.9.2 Skilled Forgery Results.....	52
4.9.3 Cross-Validation of the Results.....	53
4.10 Inter-Dataset Analysis .....	54
4.11 Summary of Contributions .....	56

<b>Chapter Five Conclusion and Future Work.....</b>	<b>57</b>
5.1 Conclusion .....	58
5.2 Future Work.....	59
<b>References.....</b>	<b>60</b>
<b>Appendix A .....</b>	<b>65</b>
<b>Appendix B .....</b>	<b>71</b>

## List of Tables

Chapter No. table No	contents	page
2-1	Summary of the Review of Related Study	21
3-1	measured feature set and a sample of the collected data of MOBSIG	26
3-2	The selected measured features	27
3-3	The proposed calculated feature set	28
3-4	The list of functions that are called to measure the raw data	36
4-1	The Proposed Calculated Feature Sets	43
4-2	Random Forgery EER results of the proposed features / models	50
4-3	Random Forgery Training Sample Size Effect on EER, Using Global EER, STD Z-Score and Feature Set B	52
4-4	Skilled Forgery EER results of the proposed features / models	53
4-5	Random Forgery EER results of the proposed features / models Using session 2 data for training and session 1 for positive testing	54
4-6	Inter-Dataset EER Results	55

## List of figures

Chapter No. Figure No	Contents	Page
1-1	Example of signature on touch screen	2
3-1	Methodology steps of the proposed study	24
3-2	Flowchart of the Data collection / Enrollment module	33
3-3	Flow chart of the authentication module	37
4-1	System entry screen	46
4-2	Account creation	46
4-3	Signature enrollment screen	47
4-4	Signature authentication screen	47
4-5	Authentication outcome screen	48

## List of Appendixes

Appendix No	Contents	Page
Appendix A	Samples of raw data features, calculated features and the generated templates of the MOBSIG dataset	65
Appendix B	Samples of raw data features, calculated features and the generated templates of the TDSIG dataset	71

## List of Abbreviations

AA	Active Authentication
AAD	Average Absolute Deviation
CSV	Comma Separated Values
EER	Equal-Error-Rate
FAR	False-Acceptance-Rate
FRR	False-Rejection-Rate
MAD	Median Absolute Deviation
OCC	One Class Classifier
PIN	Personal Identification Number
STD	Standard Deviation

## **An Anomaly Detection Model for Signature Authentication on Mobile Devices**

**By: Shawq Salman Mahmood Al-Khafaji**

**Supervisor: Dr. Mudhafar Al-Jarrah**

### **Abstract**

The use of behavioral biometrics in user authentication has recently moved to new security application areas, one of which is verifying finger-drawn signatures or access codes such as PIN numbers. This thesis investigates the design of anomaly detectors and feature sets for graphic signature authentication on touch mobile devices. The work involved a selection of raw data feature sets that are extracted from modern mobile devices, such as finger area, pressure, velocity, acceleration, gyroscope, timestamp and position coordinates. A set of authentication features have been formulated, which are calculated from the raw features. The proposed anomaly detector is based on the outlier concept, where an input signature's calculated feature element is classified as forgery if it is outside an acceptable zone from a central value such as the mean or median of a set of training values. The Z-Score method is used as the distance function of the anomaly detector, and three versions are investigated; the standard deviation based Z-Score, the modified Z-Score which uses the median-absolute-deviation and the average-absolute deviation Z-Score function. The proposed feature sets and anomaly detectors are implemented as a data collection and dynamic authentication system on a Nexus-9 Android tablet. Experimental work resulted in collecting a signature dataset (TDSIG) from 55 subjects, where the data included genuine and forged signatures. Also, the raw features data from a public dataset (MOBSIG) were converted to the calculated features, for comparison with the collected dataset. The two datasets were analyzed using the Equal-Error-Rate (EER) metric. The results showed that the Z-Score anomaly detector with 3 standard deviations distance from the mean produced the lowest error rates for the two datasets. The TDSIG dataset gave lower EER results compared with the public MOBSIG data, using the same feature sets and anomaly detectors, in both random and skilled forgeries. Variation in training and testing sample sizes indicated that training sample size is more effective than the testing sample size in reducing error rates. Also, skilled forgery error rates were close to random forgery error rates, indicating that behavioral biometrics are the key factors in detecting forgeries, regardless of pre-

knowledge of the signature's shape. The thesis ends with conclusion and suggestion for future work.

**Keywords:** graphic signature, anomaly detector, Z-Score, EER, random forgery, skilled forgery, authentication

## نموذج لكشف التباين للتحقق من التوقيع على الاجهزة النقالة

### إعداد: شوق الخفاجي

### المشرف: د. مظفر الجراح

### الملخص

وقد انتقل استخدام القياسات الحيوية السلوكية في مصادقة المستخدم مؤخراً إلى مجالات تطبيق أمنية جديدة، واحدة منها هي التحقق من التوقيعات المرسومة بالأصابع أو رموز الوصول مثل أرقام (PIN). وفي هذه الأطروحة يحقق تصميم كشف الشذوذ ميزة التحقق من مصادقة توقيع الرسوم البيانية على الأجهزة النقالة التي تعمل باللمس. وشمل العمل مجموعة مختارة من مجموعة البيانات الأولية المستخرجة من الأجهزة النقالة الحديثة، مثل مساحة الاصبع على شاشة اللمس والضغط والسرعة والتسارع والبوصلة (اتجاه المستخدم أثناء التوقيع) والطابع الزمني وإحداثيات الموقع. وقد صيغت مجموعة من خصائص المصادقة (المحسوبة) المستمدة من السمات الخام. وتضمنت مجموعة الخصائص المحسوبة قياسات إحصائية للبيانات الأولية، ومعدلات زمنية ومسافة في المستوى (س, ص).

ويستند كاشف الشذوذ المقترح إلى المفهوم الخارجي، حيث يصنف عنصر توقيع المدخلات على أنه مزيف إذا كان خارج المنطقة المقبولة من القيمة المركزية مثل متوسط أو متوسط مجموعة من قيم التدريب. يتم استخدام طريقة Z-Score كدالة المسافة للكشف عن الشذوذ، ويتم التحقيق في ثلاثة إصدارات. الانحراف المعياري القائم على Z-Score، و Z-Score التي تستخدم average-absolute deviation و Z-median-absolute deviation

يتم تطبيق مجموعات الميزة المقترحة والكشف عن الشذوذ كما جمع البيانات ونظام المصادقة الحيوية على جهاز Nexus-9. وأسفر العمل التجريبي عن جمع بيانات التوقيع من 55 شخصاً، حيث تضمنت البيانات توقيعات حقيقية ومزورة. أيضاً، تم تحويل ميزة البيانات الخام من مجموعة البيانات MOBSIG إلى الميزات المحسوبة، للمقارنة مع مجموعة البيانات التي تم جمعها. تم تحليل مجموعتي البيانات باستخدام مقياس معدل الخطأ (EER) وأظهرت النتائج أن نسخة من مجموعة الخصائص التي استبعدت السرعة والتسارع وشملت القياسات الإحصائية للضغط، مساحة الاصبع والمسافة أنتجت أقل معدل خطأ. وكان أفضل كاشف الشذوذ أداءً للنسخة مع الانحراف المعياري القائم على وظيفة Z-Score، تليها average-absolute-deviation على أساس Z-Score. أعطت مجموعة البيانات الجديدة نتائج (EER) أقل مقارنة مع مجموعة البيانات MOBSIG، وذلك باستخدام نفس مجموعات الميزات والكشف عن الشذوذ. تم جمع مجموعتي البيانات باستخدام نفس الجهاز والتشغيل، وبالتالي يمكن أن تعزى الاختلافات إلى الفرق في حجم عينات التزوير العشوائي. وقد نظر العمل التجريبي في تأثير تغيير نسبة الأحجام الحقيقية (الإيجابية) إلى عينات التزوير (السلبية)، وكانت النتيجة هي تقليل حجم عينة التزوير من 82 إلى 20 لمجموعة بيانات MOBSIG ومن 54 إلى 20 لمجموعة البيانات الجديدة، أدى

إلى انخفاض كبير في قيم معدل (EER) لكل من مجموعتي البيانات، مما يوحي بأن تفسير نتائج معدل (EER) يجب أن يأخذ في الاعتبار نسبة العينات الإيجابية / السلبية، وليس فقط حجم عينة التدريب. وأظهر تحليل نتائج التزوير وجود اختلاف طفيف بين معدلات الخطأ في التزوير العشوائي والماهر، على الرغم من أن التزوير ينبغي أن يؤدي إلى ارتفاع معدلات الخطأ. وكان حجم عينة التزوير 20 لكل مستخدم وهو قريب من حجم الاختبار الحقيقي، ويمكن أن يكون السبب في الفرق الطفيف بين معدلات الخطأ في التزوير والعشوائي أداء وظائف-Z Score أفضل عندما تم استخدام عتبات جديدة بدلاً من 2؛ فإن نسخة الانحراف المعياري كانت أفضل مع عتبة 3، في حين أن الانحراف المتوسط المطلق والانحراف الوسيط المطلق كان أفضل مع عتبة 4. تنتهي الرسالة باستنتاج واقتراحات للعمل المستقبلي.

الكلمات المفتاحية: التواقيع الرسومية، كاشف الشذوذ، نتيجة-Z، معدل الخطأ المتساوي، تزوير

عشوائي، تزوير ماهر

# **Chapter One**

## **Introduction**

## 1.1 Research Context

This thesis deals with the problem of user authentication on mobile devices, using a graphical password of the user on touch screens. The research work considers the use of measurable touch properties devices that can be obtain during the signature, to generate that will enhance authentication accuracy.



**Figure (1-1): Example of signature on touch screen**

Computer security depends largely on passwords to authenticate users. The most common method of authentication is to use usernames and alphanumeric passwords, but this method has many problems including

1. The specific password is easily guessed by the attacker.
2. A password that is difficult to guess is difficult to remember by the user.
3. An over the shoulder attacker can get the password.

To overcome these problems, methods of validation were developed by researchers using images and passwords.

One solution to the password weakness problem is to use two-level authentication where a second limited time passcode is sent to the user via his mobile devices.

This approach also has problems if the mobile devices is infected by a Malware that can re-direct the second passcode to an attacker.

More recently biometrics have been added to the authentication methods on mobile devices, such a physiological and behavior biometrics (Stokes, et. al., 2016). To authenticate users on mobile devise, as are alternative approach to password authentication.

There are three main areas in which interaction between humans and computers is important:

1. Security operations.
2. Develop safe systems.
3. Documentation.

We focus on the authentication issue here. Where user authentication is a primary component in most computer security contexts (Chavan, Gaikwad, Parab & Wakure, 2015).

Studies on passwords show that the user can only remember a limited number of passwords.

Biometrics is one of the various authentication methods used to address problems associated with the user name of traditional passwords. In this research we will deal with another alternative: is to use the graphical password.

Many techniques have been proposed to reduce restrictions on the traditional alphabet password including the proposed use of graphical passwords, which use graphics (images) instead of alphanumeric passwords. This can be achieved by asking the user to select areas of an image rather than typing characters as in the alphanumeric password policy.

Graphical passwords can be easily remembered, since users remember images better than words.

Also, they should be more resistant to brute force attacks, because there is virtually infinite search space (Angeli, Coventry, Johnson, & Renaud, 2005).

Graphical password techniques are classified into two main technologies:

1. based on reminders
2. graphical techniques based on recognition

## **1.2 Problem Statement**

The problem addressed in this study is the increasing reliance on mobile devices by users for the storage of sensitive personal and business data, and the risk of access of such data by unauthorized people. It has become necessary to provide various technical solutions to protect contents of mobile devices, through hardware and/or software.

## **1.3 Scope of Work**

The research work in this thesis includes study of a signature authentication on touch mobile device, using measured and calculated features extracted from the mobile device during the signature process. The work will include the selection and / or development of an authentication model, selection feature set of data collection, and evaluation of the proposed model based on error rates.

## **1.4 Goal and Objectives**

The aim of this research is to improve the authentication of users on touch mobile devices using the graphical password approach. The following objectives are taken into consideration:-

1. Selection of signature features that will be included in the authentication process.
2. Development of an authentication model.
3. Evaluation of the proposed authentication model using an existing dataset.
4. Implementation of the authentication model as a tool on Android operating system.
5. Data collection using the new authentication tool.
6. Evaluation of the proposed authentication model and feature set using the public and the new datasets.

## **1.5 Motivation**

The motivation of this research study is the recent increase in the number of attacks on personal, business and governmental data resources, particularly on important websites such as government or educational sites.

The attacks can access the data for malicious purposes, such as credit card data misuse, and can damage the data or prevent the use of data through ransom demand.

The mobile devices have an additional security problem in that they can be physically stolen with the subsequent risk of its data being exposed to others.

## **1.6 Significance of Work**

It is expected that this study will enhance security of mobile devices by creating a model or application that analyzes the signature of the user according to special features, through which we protect the mobile devices' contents of data and software.

## **1.7 Research Questions**

This thesis attempts to provide answer for the following research questions:

1. Can the proposed graphic signature model improve user authentication on touch mobile device.
2. What will be the error rates' metrics that will be measured in the experimental study?
3. Dose the proposed model produced similar error rates using two independent datasets.
4. Can the fusion of two independent datasets produce consistent result.
5. Will increasing the number of biometric features result in better authentication.

## **1.8 Thesis Organization**

This thesis is divided into five chapters:

- Chapter one: contains general concepts of this thesis which include the topic, background of the study, problem statement, scope of work, limitation of the proposed work, goal and objectives, motivation, significance of work and questions to be answered.
- Chapter two: presents literature review, concepts and definitions which introduced the introduction, classification methods, biometric technologies and related work.

- Chapter three: presents methodology and the proposed model which introduced the methodology approach, outline of the proposed model, methodology steps, features selection, the anomaly detector, the proposed system and error metrics.
- Chapter four: presents experimental results and discussion which introduced the introduction, objectives of the experimental work, EER analysis steps, feature sets selection, analysis of the MOBSIG dataset, the proposed system, data collection using the proposed system and discussion of results.
- Chapter five: contains conclusions and future work.

## **Chapter Two**

### **Background and Literature Review**

## 2.1 Introduction

The area of biometrics for security applications has received considerable interest in the past few years, for various applications, including airport security, banks, education and in public offices. User authentication using biometrics on mobile devices has seen a lot of research effort to investigate techniques, features and modalities that can improve the security of data and software on such devices. Introduction of new features and sensors on smartphones has led to further interest in utilizing the new technologies to enhance user authentication, including the infusion of several modalities and features into an authentication system.

Developments in biometrics technologies have covered two main streams: physiological biometrics and behavioral biometrics. The physiological biometrics has focused on identity checking using features like iris, DNA and finger-print, which require special hardware, and can have hardware related problems, such as hardware malfunction, the need for tuning and maintenance. The other stream of biometrics developments is the use of behavioral biometrics such as typing rhythm, finger movement on touch surfaces, voice and face recognition, mouse dynamics, gait recognition, and device vibrations during strolling. Most behavioral biometrics on mobile devices do not require special hardware, apart from the available built-in features and sensors, (Bubeck & Sanchez, 2003)

An important application of biometrics authentication is the graphic password authentication on touch devices, using stylus or finger-drawn touch input. In this method of authentication, a set of signature data of a user is used in the training phase of an authentication system, from the training data a profile of the signature is extracted. The

extracted profile is used in the authentication phase to classify a new signature as forgery or genuine.

## 2.2 Classification Methods

Authentication of users who are attempting to access a computer resource, based on authentication features, is a classic application of machine learning using the classification methods. The classification methods that are relevant to this research can be divided into two areas, as below:

- 1. Binary classification (two-class classification):** It is a method that classifies data into two categories, based on training of the characteristics of two categories. In authentication applications, the two categories can be genuine or forgery, legitimate or imposter, positive or negative, and the data is divided into two subsets; training subset and testing subset. The training subset contains labeled data from both categories, while the testing subset contains unlabeled data from the two categories. (Koyejo, et. al, 2014).
- 2. Anomaly detection (one-class classification):** It is a way of authenticating a person based on his genuine or correct biometric features in a real application, without having access to negative data samples. This is the case where a security system is trained for user authentication on the basis of the individual's profile of input, without knowledge of how forgers or impostors would input their data. Each person has his own signature profile and his way of signing, which an authentication attempts to capture. The extracted training data is the only data available to the anomaly detector, the one-class classifier. Any input that does not fit the profile of the genuine user will be rejected as negative or in our case a forgery, so the one-class classifier knows only characteristics of the good users,

and any user who doesn't resemble the good user will be rejected. To evaluate the detection performance of a one-class classifier, negative and positive data are needed so as to assess the classifier's capability in distinguishing between genuine and impostor users. The anomaly detector can make mistakes, by false rejecting a genuine person or false accepting an impostor. A template of the user's profile needs to be designed and tuned to avoid two error cases of detection of false acceptance and false rejection (Chandola, et. al, 2009). The performance evaluation of an anomaly detector will measure the Equal-Error-Rate (EER), the point at which the false reject rate (FRR) equals the false acceptance rate (FAR), for a set input from a group of users.

## **2.3 Biometric Technologies**

Biometric systems are able to authenticate or identify people based on physiological or behavioral characteristics which are unique for each person. As biometric systems become increasingly accurate, they will be selected more often as the option of choice for authentication, intrusion detection, or access control within software systems.

One of the most useful applications for biometrics is user authentication. Authentication is a way to prove that a user is who they claim to be.

In most systems, authentication involves asking a person to prove who they are by what they know – such as a username and password combination, (Stokes, et. al., 2016).

Biometric authentication attempts to carry out the verification process based on analysis of characteristics that are unique to a given individual. Physiological biometrics include analysis of characteristics such as fingerprint, iris, or DNA. Behavioral biometrics

focus on the way in which users interact with their computer device. Some examples are mouse movements, keystroke rhythm, and touch screen interaction. The main benefits of biometrics is that they are difficult to mimic and they have an advantage over password authentication in that they are not susceptible to being cracked (via dictionary attacks or brute force attacks), lost, or stolen. An emerging application of biometrics is active authentication (AA).

Active authentication is a way of continuously authenticating or verifying a user's identity during a session. Typically, a user is only authenticated at the beginning of a session. If the user steps away from the computer or if the session is hijacked then the secured assets are vulnerable to exploitation. Active authentication attempts to continually verify that a user's biometric patterns (human to computer interactions) are consistent with those demonstrated during their previous sessions. The goal is to determine whether or not the current user is an imposter or the original authenticated user. (Stokes, et. al., 2016).

### **2.3.1 Graphic Signature Authentication**

One of the earliest methods of verifying user identity is based on his signature. Many official documents require signatures from agreeing parties. Signature recognition can be divided into offline (static) and on-line (dynamic) methods. While on-line systems work with images, therefore only the shape of the signature is available, on-line systems use information related to the behavioral dynamics of the signature. Due to this additional information, on-line systems outperform on-line systems (Impedovo, & Pirlo, 2008). Biometric systems can produce two types of errors: false rejections of genuine signatures (False Rejection Rate - FRR) and false acceptance of forged signatures (False Acceptance Rate - FAR). The overall detection error is usually calculated as EER (Equal Error Rate), which is defined as the detection error rate when FAR and FRR are equal. In signature

dataset evaluations and comparisons two types of forgeries are considered: skilled and random forgeries. Skilled forgery evaluation is based on using the forgery samples available in the dataset, where forgery samples are provided by forgers who know the shape of the imitated signature. Random forgery evaluation is based on using random genuine samples from the dataset, which represents the case when the forger does not know the signature to be forged, therefore is using his own signature.

The rapid development of smart devices and their attractive applications made it desirable and required of all ages, leading to the embrace of smart devices in large numbers in different parts of the world.

In 2014, about 1.75 billion users around the world own and use smartphones, an increase of 25% over the previous year. The security of smart phones and mobile device in general became an important issue in modern times, due to using these devices for storing private information such as contacts, photos, personal documents, business documents, credit card numbers, passport numbers and similar data for access. This has made the mobile devices it vulnerable to many attacks for various malicious purpose (Ranak, Azad, nor & Zamli, 2017) [cited 2017 Aug 30].

Ensuring the security of these devices becomes a burning issue, thus many mobile devices currently employ one or more authentication feature.

One type of authentication is password-based authentication, which is most common because of low implementation complexity, low computational complexity, and low processing requirements.

However, many cryptanalysts have discovered various weaknesses in text-based schemes, such as dictionary attack (Lee, Kim, Kim, Choi, Cho & Lee, 2016)., social engineering attacks (Krombholz, Hobel, Huber & Weippl, 2015)., brute force attack (Saito, Maruhashi, Takenaka & Torii, 2016), guess attack (Reddy, Yoon, Das, Odelu &

Yoo, 2017), etc. Moreover, Smart devices impose some more limitations in text-based schemas, the other type of authentication for mobile devices is graphical schemes. Graphical schemes are used for entry of passwords or personal identification numbers (PIN) using finger-drawn input on a touch screen this scheme has the advantage of adding biometric features in the authentication process.

The graphic password authentication approach is divided into two basic types:

### **1. Recognition Based System**

In recognition-based techniques, authentication is done by challenging the user to identify image or images that the user had selected during the registration stage. Another name for recognition-based systems, is cognometric systems (Angeli, Coventry, Johnson & Renaud, 2005) or search metric systems (Renaud, 2009), generally require that users memorize a number of images during password creation, and then to log in, must identify their images from among decoys. Humans have unique ability to identify images previously seen, even those viewed very briefly (Standing, Conezio & Haber, 1970) and (Nelson, Reed & McEvoy, 1977). From a security point of view, these systems are not acceptable replacements for text password schemes, as they have password spaces which are compared in cardinality to only 4 or 5 digit PINs (considering a set of images whose cardinality remains reasonable, with respect to usability and security).

### **2. Recall Based Systems**

In recall-based techniques, a user is asked to reproduce something that he or she created or selected earlier during the registration stage. Recall-based graphical password systems are occasionally referred to as draw metric systems (Angeli, Coventry, Johnson & Renaud, 2005) since a secret drawing is recalled and reproduced by the user. In these systems, users typically draw their password either on a blank canvas or on a grid (which may arguably act as a mild memory cue). Recall is a difficult memory task (Craik &

McDowd, 1987). Because retrieval is done without memory prompts or cues. Users sometimes devise ways from which the interface could be used as a cue even though it is not intended as such, the task is transformed into one of cued recall, although one where the same cue is available to all users and to attackers. Text passwords can also be categorized as using recall memory. With text passwords, there is evidence that users often include the name of the system as part of their passwords (Vu, Proctor, Bhargav-Spantzel, Tai, Cook & Schultz, 2007) and (Chiasson, Forget, Stobert, Van Oorschot & Biddle, 2009). Although there is currently no evidence of this happening with graphical passwords, it remains a seemingly valid coping strategy if users can devise a way of relating a recall based graphical password to a corresponding account name.

To a great extent these systems are generally susceptible to shoulder surfing attack, the entire drawing is visible on the screen as it is being entered, and thus an attacker need to accurately observe or record only one login for the entire password to be revealed. You can secure your password using various techniques in graphical authentication. Here we are proposing a new algorithm of authentication using images. To authenticate, we use a grid based approach by using image as a reference. User will upload the image/set of images along with all his/her details during the time of the registration. Then the image selected by the user will appear on the page with transparent grid layer on it. Then certain grids are selected by the user to set his/her password.

The proposed system was implemented using PHP, CSS, JavaScript and Macromedia flash 2008 (Action Script 2). This Graphical Password can be implemented in authenticating several systems and websites. The implementation has few focuses:

- Password: Contain image as reference & encryption algorithm.
- Grids: Contains unique grid values and grid clicking related methods.
- Login: Contains username, images, Graphical password and related methods.

- SSR shield: Contains shield for Shoulder surfing.

### **2.3.2 Signature Recognition on Mobile Devices**

Recently several research works have been carried out in the field of online signature recognition on mobile devices (Martinez-Diaz, Fierrez & Galbally, 2016), reporting results obtained on signature datasets captured from tablets or smartphones. Most of the studies are concerned with signature recognition results using signature datasets captured on pen tablets. However, touch screen on mobile devices present some drawbacks compared with pen tablets, the most important being the quality of the captured signal. While pen tablets sample the signal uniformly with relatively high frequency, hand-held device sampling is usually event-driven with lower sampling frequency than pen tablets. Moreover, while both touchscreen devices and pen tablets are able to capture trajectory and pressure, the latter can track pen orientation.

Then an average score can be computed from these scores. Both the samples and the users can be evaluated by using only the genuine signatures or using both the genuine and forgery signatures

### **2.3.3 One-Class Classification (Anomaly Detection)**

It is an algorithm whose primary purpose is to build taxonomic models when the negative layer is absent or weak or indefinite by defining the layer boundaries only with the knowledge of the positive layer, where a single layer refers to the positive or exploratory category

An example of one-classification application is the automatic diagnosis of disease, where a patient's data who have disease are considered the positive class.

The negative class is difficult to identify, because it represents the rest of the healthy people.

It appears that Minter (1975) was the first to use the term ‘single-class classification’ four decades ago, in the context of learning Bayes classifier that requires only labelled data from the “class of interest”. Much later, Moya et al. (1993) originate the term One-Class Classification in their research work. Different researchers have used other terms such as Outlier Detection<sup>2</sup> (Ritter and Gallegos, 1997), Novelty Detection<sup>3</sup> (Bishop, 1994), Concept Learning (Japkowicz, 1999) or Single Class Classification (Munroe and Madden, 2005; Yu, 2005; El-Yaniv and Nisenson, 2007). These terms originate as a result of different applications to which one-class classification has been applied. Juszczak (2006) defines One-Class Classifiers as class descriptors that are able to learn restricted domains in a multi-dimensional pattern space using primarily just a positive set of examples (Khan & Madden, 2014).

### **2.3.4 Feature Extraction in One-Class Classification**

Reducing the feature set is often an essential part of solving a classification task. This is done by analyzing key elements of the feature set and eliminating trends of low variance in data and maintaining high contrast trends.

The high-contrast trends are expected to contain information on class differences.

As for the classification of a single class, the task of classification contains one category that is not specified, and which have (almost) no information.

Using a lot of features will increase noise, so the feature set can reduce the detection accuracy, of especially for a sample of limited size.

## **2.4 Related Work**

Authentication of users on mobile devices using graphic signature have been reported in several research papers.

Donato Impedovo and Giuseppe Pirlo (2008) presented the art in automatic signature verification by studying and exploring almost useful and valuable of the more than 300 selected researches to date. The aim is to guide the researchers who are working in the automatic signature verification. The researchers concluded the following points

- (1) Automatic Signature Verification (ASV) is renewable field.
- (2) Several systems based on database and testing protocols to find the accuracy level like the figure print systems.
- (3) Online ASV is very important application because it uses in many fields like banking, driving licence, etc.
- (4) ASV processing is compared the online signature with stored figure signature of handwritten signature in the smart card to verify the rightful owner.

Pascal Bissig, (2011) implemented a signature verification system compatible with touchscreen devices. The author suggested to divide the verification system into two parts Dynamic Time Warping based system and a global feature which based on Vector Machine for classification. the author added the pressure feature to increase the training samples and decreasing the errors of the performance and then integrated the two parts to increase the performance of classification. Finally, the results confirmed that the combination between two features will give a high performance of classification.

Nesma Houmani, Sonia Garcia-Salicetti, Bernadette Dorizzi, and Mounim El-Yacoubi (2012) attempted to demonstrate that the graphic signature is acceptable online signature verification system on a mobile device. The authors used on Hidden Markov Model (HMM) and the output of two results; the first output is from HMM to discover the claimed identity which achieves the arithmetic mean of two results to gain a higher of input signature. The second output came from the segmentation of HMM. The researchers improved their scheme to be executed and verified the given signature when the user's registration but the complexity stayed as is when user's signature verification process is

taken a long time. To improve this case, the authors suggested to enhance the quality of signature itself in order to improve the performance of the scheme.

Ram P. Krish, Julian Fierrez, Javier Galbally and Marcos Martinez-Diaz (2013) focused on dynamic signature verification and the evaluation of smartphone performance for that. The researchers analysed database which was consisted of 25 customers and 500 signatures stored in Samsung Galaxy Notepad. The researchers used a specific verification algorithm which checked the features and functions and then presented the equal error rate as the result of this checking. The researchers achieved the best result of EER which is 0.525%.

Marcos Martinez-Diaz, Julian Fierrez, and Javier Galbally (2015) studied the authentication and free hand sketches and they proposed two models for verification and Gaussian mixture which depended on dynamic signature verification methods. The researchers adapted the sequential forward floating selection algorithm to study the most of features' characteristics. They also used set of training which stored in DooDB database to verify the right person's signature. The results of Equal Error Rates between 3% and 8% are obtained against random forgeries and between 21% and 22% against skilled forgeries. High variability between capture sessions increases the error rates.

A graphic signature database called MOBSIG was collected at Sapiencia University (Ental, M. & Lzsalo, S. (2016, May). The researchers presented data of genuine and forgery signatures using a mobile device. The database contained signatures data obtained in three sessions, resulting in 45 genuine signature per user and 20 skilled forgery signatures against selected users. The user sample of the research consisted of 83 users.

Ental, M. and Lzsalo, S. (2016) analyzed the dataset provided in (Martinez-Diaz, M., Fierrez, J., & Galbally, J., 2015) which contained three raw measurable features (time, x and y coordinates) they suggested two types of equal error rate (EER) to evaluate the

accuracy performance. These types are global threshold and user-specific threshold. The results of evaluation are as the following (i) the skilled forgery produced higher error rate than the random forgery. (ii) The skilled and random forgeries were higher when using global thresholds. The researchers did the same procedures for the doodle dataset. DooDB database to evaluate the result of finger drawn signature. Finally, the result confirmed that the graphic signature approach can be used as a biometric system for user authentication.

Al-Obaidi (2016) investigated the use of Keystroke Dynamics authentication on touch mobile devices. The work presented an authentication model which used measurable features obtained from a mobile device during the typing of a password to build a typing profile of the user. The measured features included pressure, finger area and timing data, and an anomaly detector was based on measuring the distance from the median. The authentication model was implemented on a Nexus-7 tablet, which provided a data collection tool and a dynamic authentication tool based on keystroke stroke dynamics. The experimental work showed a reduction in authentication error rates when the touch features of pressure and finger area were added to the authentication feature set.

## 2-5 Summary of Related Work

Table (2-1) shows a summary of related work and properties of the dataset used in the related experiments.

**Table (2-1): Summary of the Review of Related Study**

Paper	Idea	Users	Device	Input method	#GEN	#FOR	#SESS	Raw Features
Bissig, (2011)	Signature Verification on Finger Operated Touchscreen Devices	NA	HTC Desire 3.7", capacitive	finger	20	NA	NA	x(t),y(t),p(t),fa(t)
Houmani et. al., (2016)	On-line verification of finger drawn signatures	432	PDA HP iPAQ hx2790	pen	30	20	2	x(t),y(t)
Houmani et. al., (2010)	On-line Signature Verification on a Mobile Platform	64	PDA Qtek 2020 ARM	pen	30	20	2	x(t),y(t)

Krish et. al., (2013)	Dynamic Signature Verification on Smart Phones	25	Samsung Galaxy Note	pen	20	0	2	$x(t), y(t), p(t)$
Martinez et. al., (2013)	The DooDB Graphical Password Database: Data Analysis and Benchmark Results	100	HTC Touch HD mobile, resistive	finger	30	20	2	$x(t), y(t)$
Sae-Bae & Meimon (2014)	Online Signature Verification on Mobile Devices	180	user owned iOS devices	finger	30	0	6	$x(t), y(t)$
Antal (submitted 2017)	On-line Signature Verification on MOBISIG Finger Drawn Signature Corpus	83	Nexus 9, Capacitive	finger	45	20	3	$x(t), y(t), p(t), fa(t)$ $vx(t), vy(t),$ $ax(t), ay(t), az(t)$ $gx(t), gy(t), gz(t)$

## **Chapter Three**

### **Methodology and the Proposed Model**

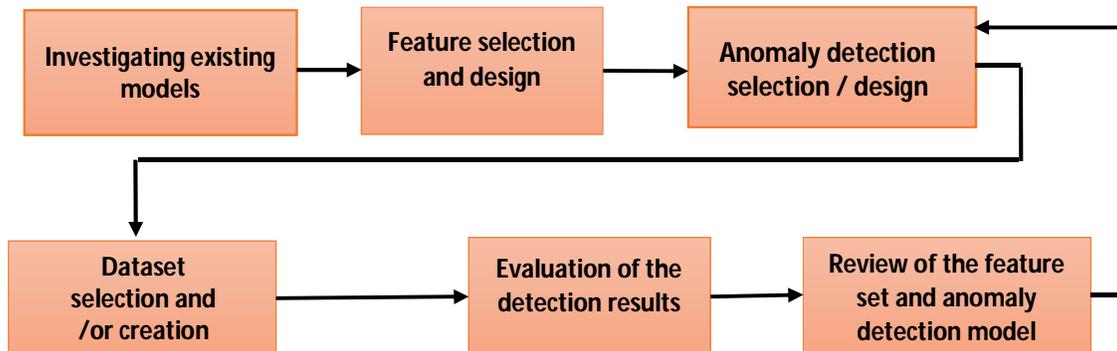
### **3.1 Methodology Approach**

This research follows experimental methodology to achieve its objectives. The proposed model and the related assumptions will be evaluated using a public dataset within the domain of research as well as a dataset collected in this research, to measure the detection accuracy of the proposed model. The study will result in the design of an anomaly detector whose features and structure will be determined by the experimental investigation.

### **3.2 Outline of the Proposed Model**

The aim of the proposed model is to improve detection of forged graphic signatures on mobile devices. The task of detection will be based on the analysis of signatures according to selected measured and calculated features, in order to arrive at an anomaly detection model that will have lower error rate and therefore better detection performance.

### 3.3 Methodology Steps



**Figure (3-1): Methodology steps of the proposed study**

**The proposed study will involve the following steps:-**

- 1. Investigating existing models:** - in this step we will investigate existing models in terms of features, classification methods and the experimental results.
- 2. Feature selection and design:** this will involve identifying possible measurable features that are available on touch mobile devices and the relevance of such features in the authentication process. The features will be in two parts:
  - Raw data features: such as pressure, finger area, timestamp and pixel coordinates.
  - Calculated features: features that are calculated from raw data features, such as total time, total distance average of time between points and 7 average distance of movements.
- 3. Anomaly detector selection / design:** based on the selected and evaluated feature set, an anomaly detector will be selected (or designed) to evaluate signature data of an individual using his genuine data, without the availability of negative data.

The detector will be based on distance measurement from a central point, such the Euclidean distance method (**Barrett, P. (2006)**)

4. **Dataset selection / creation:** public datasets of related study will be used as a source of measured (raw) features data. The datasets will have positive (genuine) and negative (forgery) samples.

Analysis of the datasets will be used to improve the selected feature set and the anomaly detectors.

5. **Experimental evaluation of the proposed model** in this step the model will be evaluated using the evaluation metrics applied to the output result of analyzing the dataset.
6. **Review and update of the feature set and the anomaly detector**, to enhance the anomaly detection performance by reducing the error rates.

### 3.4 Features Selection

There are two feature sets to be considered for this type of research:

1. Measured (raw) feature set, which consist of measurable feature to be collected from the device, such as time stamp.
2. Calculated feature set, which consist of metrics used in the authentication process that are derived from the raw features, such as total time.

The proposed measured feature set is based on existing feature sets (Marcos , Margit) that measure time, location in pixel address, velocity, acceleration, gyro meter reading and other measurement that reflect the behavioral changes during signature, such as switching movement direction.

### 3.4.1 Feature Sets of Previous Studies

**Table (3-1): measured feature set and a sample of the collected data of MOBSIG**

X	y	timestamp	pressure	fingerarea	velocityx	velocityy	accelx	accely	accelz	gyrox	gyroy	gyroz
397.8484	569.407	8984761	0.7375	0.106383	0	0	0.021227	0.013082	-0.01987	0	0	0
399.6107	562.3584	8984788	0.725	0.095745	558.0547	-2232.18	0.021227	0.013082	-0.01987	0	0	0
412.0459	523.0427	8984805	0.7	0.085106	1006.92	-3206.41	0.021227	0.013082	-0.01987	0	0	0
429.6184	469.1492	8984821	0.7125	0.095745	1371.996	-4208.58	0.021227	0.013082	-0.01987	0	0	0
448.3937	423.0226	8984838	0.6875	0.06383	1485.454	-4193.79	0.021227	0.013082	-0.01987	0	0	0
469.453	374.3513	8984855	0.6875	0.074468	1596.782	-4003.08	0.021227	0.013082	-0.01987	0	0	0
485.5256	333.0492	8984872	0.6875	0.085106	1297.522	-2810.83	0.021227	0.013082	-0.01987	0	0	0
496.899	306.4109	8984888	0.7	0.095745	977.2367	-2097.57	0.021227	0.013082	-0.01987	0	0	0
505.3354	288.5942	8984905	0.7	0.06383	529.9851	-1114.52	0.021227	0.013082	-0.01987	0	0	0

Table 3-1 shows the measured (raw) features of the MOBISIG dataset (Martinez-Diaz, Fierrez, & Galbally, 2016) the 1<sup>st</sup> and 2<sup>nd</sup> columns in table are the x and y coordinates of a measurement point.

The timestamp column shows the timestamp of the measurement event in milliseconds.

The pressure column shows a measurement of the pressure value at the measurement point which represents the amount of the pressure that has resulted from the finger pressing on the touch screen.

The finger area column shows a measurement of the area that finger pressing has occupied on the touch screen at the time of measurement.

The velocity and acceleration column show the X and Y values of these features when moving between two points.

The gyro x, gyro y and gyro z columns shows the gyro meter reading at the touch point the data shows that the gyro results do not change, all zero, because the measurement where taken while the tablet was on a fixed horizontal surface. Also, the z-axis acceleration did not show any changes for the same reason as for the gyro.

The MYCN dataset (Martinez-Diaz, Fierrez, Galbally, 2016) included a limited set of measured features, which includes X and Y coordinates and time duration of the stamp rather time of stamp.

### 3.4.2 The Proposed Feature Sets

The proposed feature sets consist of a measured features set and a calculated features that is derived from the measured feature set. The measured features represent the raw data collected during finger movement between points of the signature. The calculated features represents aggregations of the measured features using various functions, Signature evaluation will be based on the calculated features. In this work we will use the same measured feature of the MOBSIG. Table (3-2) gives description of the (9) measured raw feature that will be used in the proposed work.

**Table (3-2): The selected measured features**

Measured Feature	Description
x	x- coordinate in pixel location
y	y- coordinate in pixel location
timestamp	Time stamp of the current position
Velocity x	Velocity of movement along the x-axis
Velocity y	Velocity of movement along the y-axis
Acceleration x	Acceleration of movement along the x-axis
Acceleration y	Acceleration of movement along the y-axis
Finger area	Area in pixels of the finger touch
Pressure	Pressures during the finger touch

**Table (3-3): The proposed calculated feature set**

Calculated Feature	Description
No of points	Number of signature movements
Total-x	Total absolute distance of all movements along x-axis
Total-y	Total absolute distance of all movements along y-axis
Total-t	Total time duration of the signature
Med-x	Median of absolute distance between two locations on the x- axis
Med-y	Median of absolute distance between two locations on the y- axis
Med-vx	Median of the absolute velocity between two locations on the x- axis
Med-vy	Median of the absolute velocity between two locations on the y- axis
Max-vx	Maximum of the absolute velocity along
Max-vy	Maximum of the absolute velocity along
Med-ax	Median of the absolute acceleration between two locations on the x-axis
Med-ay	Median of the absolute acceleration between two locations on the y-axis
Med-p	Median of the pressure of all measurements
Max-p	Maximum of the pressure of all measurements
Med-fa	Median of the finger area of all measurements
Max-fa	Maximum of the finger area of all measurements
% of x flips	Ratio of reversed movements along x-axis
% of y flips	Ratio of reversed movements along y-axis
Disp-x	Total displacement along x-axis
Disp-y	Total displacement along y-axis
Ratio xy1	Ratio of total traveled distance x over y
Ratio xy2	Ratio of total displacements of x over y

---

Table (3-3) shows the proposed calculated feature set. The features include metrics that are derived from the raw features and from discriminate between signatures. The feature set include total of signature point time, distance and displat, as well as statistical metrics of pressure, finger area, velocity and acceleration. The contribution of these metrics to words discrimination between various signatures will be evaluated experimentally in order to choose the feature set that results in lower authentication errors.

### 3.5 The Anomaly Detector

The selected anomaly detector model is aimed to be used for the detection of outlier anomalous values of signature features, in order to determine whether an unknown signature is genuine or a forgery attempt. Each signature feature is compared with a central value of that feature obtained during the training phase, where the central value can be the mean or the median, depending on the chosen anomaly detection model. For each anomaly detection model, a distance function is used to calculate the distance metric for a feature element value based on its distance from the central value of that feature, and the distance metric will be compared with a threshold.

The following alternative anomaly detection models are used in the proposed system:

#### 3.5.1 The Z-Score Anomaly Detector

The distance function in this model is based on the Z-Score , which is used to detect outliers, (V & Taffler, 2007), and it is calculated for a feature element using the mean and the standard deviation (Wagenmakers & Brown, 2007) of the feature's value that are obtained during training. The Z-Score is calculated below:

$$\text{Z-Score of } X_i = \frac{X_i - \bar{X}}{STD(X)} \dots\dots\dots (1)$$

Where STD is the standard deviation (raw) of the set of values The Z-Score value for a feature element is considered to be genuine if it is within a specified threshold STD (Edjabou, Martín-Fernández, Scheutz & Astrup, 2017). In previous work (Margit / Marcos), the Z-Score threshold was fixed at the value of 2, which means that the acceptable distance for a genuine feature should be within two standard deviations distance from the mean. In the proposed anomaly detector, a variable threshold T is used whose optimum value will be determined based on an empirical investigation. Each feature element is given a feature score (FS) of 1 if it's Z-Score value is within the threshold T.

The total score for a signature attempt is the sum of feature elements' scores:

$$\text{Sig-Score} = \sum_{i=1}^N FS_i \quad \dots\dots\dots (2)$$

Where N is the number of features in the features set.

Due to the nature of behavioral biometrics, a genuine signature is not expected to result in genuine score for all of it feature elements, therefore a signature score threshold is required to determine whether a signature score is within a certain acceptable limit. The signature score threshold is referred to as the pass-mark (Aljarrah), whose value is determined experimentally.

An authentication template is created during the training phase which consists of two sets of reference values: a set of mean values and a set of standard deviation values, where each pair of values corresponds to a feature element of the set of features that will be used in determining the Z-Score of a feature element.

### **3.5.2 The Average Absolute Deviation (AAD) Anomaly Detector**

This model uses a modified version of the Z-Score function, to calculate the acceptable distance metric. This version uses the mean and the Absolute Average Deviation to calculate the modified Z-Score for a given feature element, as below:

$$\text{AAD Z-Score of } X_i = \frac{X_i \times \bar{X}}{\text{AAD}(X)} \dots\dots\dots 3$$

Where the AAD is calculated as below:

$$\text{AAD of } X = \text{Mean of } |X_i - \bar{X}| \dots\dots\dots 4$$

The Z-Score threshold for this version can be different than the STD base Z-Score as the AAD of a range of values covers a smaller area than the Calculating the signature score follows the same steps as with the STD-based Z-Score anomaly detector, using a pass-mark that will be determined experimentally.

**3.5.3 The Median Absolute Deviation (MAD) Anomaly Detector**

This model uses a modified version of the Z-Score function, to calculate the acceptable distance metric. This version uses the median and the Median Absolute Deviation (MAD) (Rousseeuw, & Croux, 1993) to calculate the modified Z-Score for a given feature element F (i), as below:

$$\text{MAD Z-Score of } X_i = \frac{X_i \times \text{Median}(X)}{\text{MAD}(X)} \dots\dots\dots 5$$

Where the MAD is calculated as below:

$$\text{MAD of } X = \text{Median}(|X_i - \text{Median}(X)|) \dots\dots\dots 6$$

The Z-Score threshold for this version can be different than the STD-based Z-Score as the MAD of a range of values covers a smaller area than the STD Calculating the signature score follows the same steps as with the STD-based Z-Score anomaly detector, using a pass-mark that will be determined experimentally.

### **3.6 The Proposed System**

The proposed system hence forth referred as TDSIG (Touch Device Signature) aim to provide two services:-

1. Data collection for experimental analysis.
2. Dynamic user authentication.

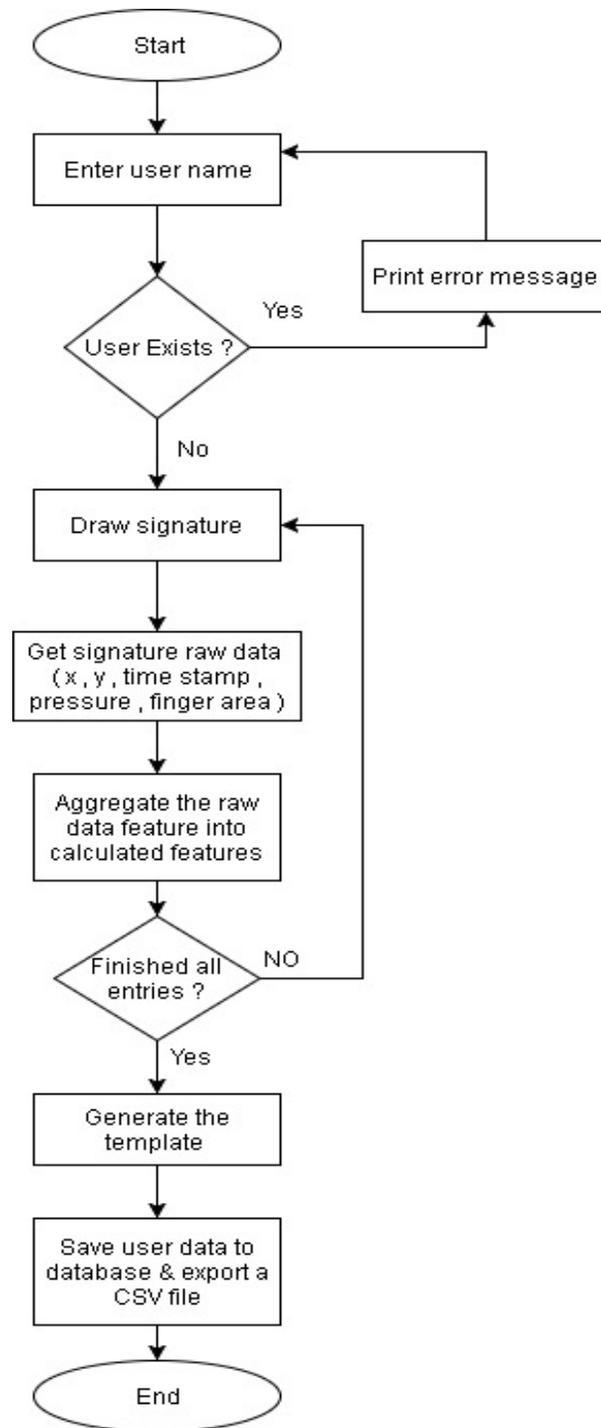
The proposed system is implemented on an Android environment, on a Nuxsus-9 tablet, to be comparable with previous work ( Antal & Szab, 2016)

In addition, a separate data aggregation module is used to aggregate raw data feature into calculated features of a pre-collected dataset.

#### **3.6.1 The Data Collection Module**

This module will provide measurement of the raw feature as in table (3- 2) and calculation of the calculated feature as listed in table (3- 3).

Figure (3-2) shows a flowchart of the steps of this module. The operation of the module is controlled by parameters that define the required number of signature repetitions for particular data collection experiment.



**Figure (3-2): Flowchart of the Data collection / Enrollment module**

The main actions performed in this module are the following:

1. **Get User Name:** the user name or identifier is read, verified that it does not exist in the internal database.
2. **Get Signature Raw Data:** when the user makes contact with the touch surface, the event-listener triggers raw data reading from the built-in functions, at a frequency controlled by the sampling frequency of the device. The collected raw data vector is added to the internal database.
3. **Aggregate Raw Data:** the raw data vectors of a signature attempt that are collected in the Get Signature Data action, are used to calculate the calculated features vector of the signature attempt such as total time, total x-distance and total y-distance, as shown in Figure 3-2.

(Steps 2 and 3 are repeated a number of times, to collect multiple signatures of the user, as determined by the enrollment counter whose value is set using a setting function before the enrollment phase).

4. **Generate Template:** The calculated features vectors that are collected from a sequence of signature attempts are used in calculating the templates that will be used in the authentication module.

The template consists of five vectors:

**Median Vector:** median of each calculated feature column for a group of signatures.

**Mean Vector:** mean of each calculated feature column for a group of signatures.

**STD Vector:** standard deviation of each calculated feature column for a group of signatures.

AAD Vector: average absolute deviation of each calculated feature column for a group of signatures.

MAD Vector: median absolute deviation of each calculated feature column for a group of signatures.

5. Save User Data: the raw data vectors and the calculated features vectors are stored in an internal database and exported at the end of an enrollment session into CSV files, to be used in the empirical study. The template vectors are saved in the database, to be used in authentication mode.

### **3.6.2 The Data Collection Functions**

The data collection module takes measurement of the measured feature set elements using Android – based functions. The measurement is controlled by the event sampling of the device.

Table (3-4) shows the list of Android functions that are used in this work. These functions are available on the Nexus series of touch devices such as Nexus-9.

**Table (3-4): The list of functions that are called to measure the raw data**

Function	Description
GetX()	Get pointer's X position on each Motion Event Occurrence
GetY()	Get pointer's Y position on each Motion Event Occurrence
GetTimeStamp()	Retrieve Current time in timestamp format
GetVX()	Calculate Velocity X using native android VelocityTracker functions of Motion Event
GetVY()	Calculate Velocity Y using native android VelocityTracker functions of Motion Event
FingerArea()	Get finger size using native Android function of pointer (event.getSize)
GetPressure()	Get finger pressure using native android function of pointer (event.getPressure )
getAccX()	Get Linear Acceleration X using Android Sensor Manager (SensorEventListener ) of Type ACCELEROMETER
getAccY()	Get Linear Acceleration Y using Android Sensor Manager (SensorEventListener ) of Type ACCELEROMETER
getAccZ()	Get Linear Acceleration Z using Android Sensor Manager (SensorEventListener ) of Type ACCELEROMETER

The proposed system will export the measured raw features into a CSV file, and generate the calculated features and template for later user for authentication.

### 3.6.3 The Authentication Module

This module uses the calculated features template that are obtained during the training phase to authenticate a new signature attempt .The operation of this module is controlled by the selected anomaly detection model's thresholds such as pass-mark and Z-Score thresholds .

A new signature is classified as genuine or forgery depending on the total score for all the feature elements and in reference to the thresholds whose values are calculated during training .As shown as Figure (3-3)

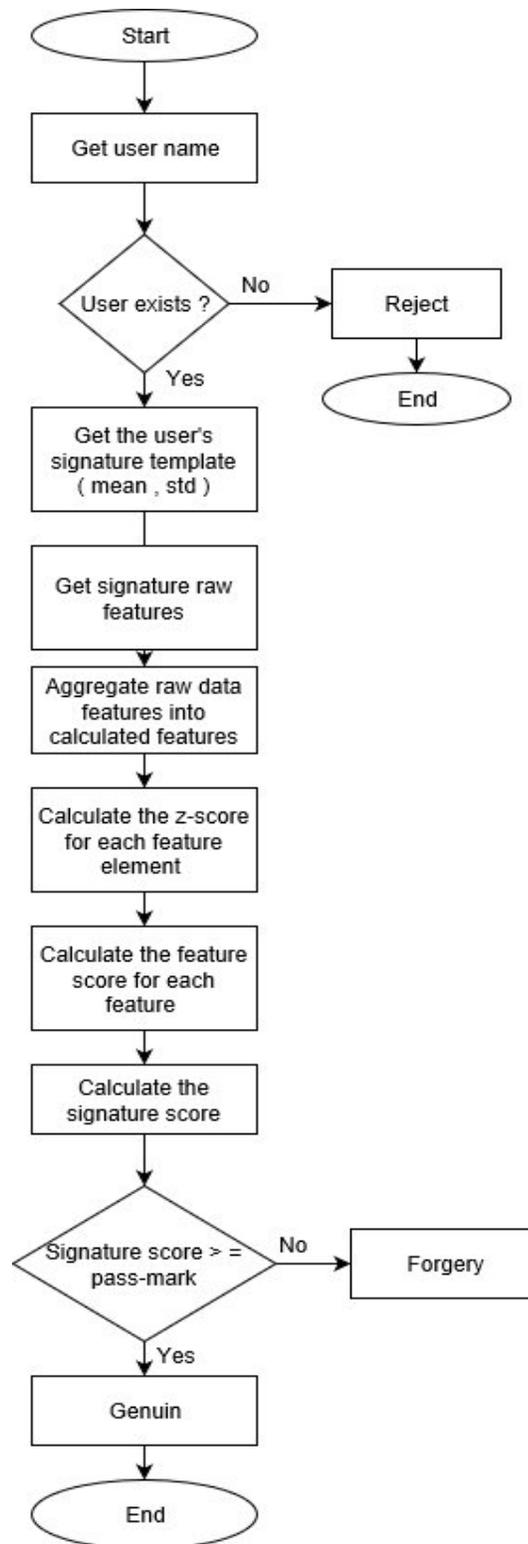


Figure (3-3): Flow chart of the authentication module

The main actions performed in this module are the following:

1. Log-in User: Get user name or identifier that was used in the enrollment phase.
2. Get User Template: Load the template vectors of the logged-in user that were generated in the data collection and enrollment module.
3. Get Signature Raw Data: perform the same task as in the enrollment phase, but only once, for the authentication purpose.
4. Aggregate Raw Data: perform the same task as in the enrollment phase, to generate a calculated features vector for the signature to be authenticated.
5. Generated Z-Score: calculate the Z-Score for each feature element of the calculated features vector of the signature.
6. Determine Features Scores: for each feature element, determine a score of 1 if the feature's value is within the Z-Score threshold, otherwise 0.
7. Determine Signature Score: Calculate the signature score which is the sum of features scores.
8. Determine Outcome: if the signature score is greater than or equal to the pass-mark threshold then the authentication outcome is genuine, otherwise forgery.

#### **3.6.4 Feature Extraction Program (Extract – Features)**

This program aggregates raw data features from an existing dataset into a set of calculated feature vectors, where each vector represents one signature attempt.

The program calculates the template for a set of signature attempts for the evaluation of the EER metric of the given dataset.

### 3.7 Error Metrics

The following error metrics will be used in this model:

**False Acceptance (FA):** Number of forgery signature attempts that are detected as genuine.

**False Rejection (FR):** Number of genuine signature attempts that are detected as forgery.

**False Acceptation Rate (FAR):** ratio of the number of false acceptance to the total number of attempts.

**False Rejection Rate (FRR):** ratio of the number of false rejection to the total number of attempts.

**Equal Error Rate (EER):** the average of FAR and FRR when they are closet to each other.

The EER metric is used in the evaluation of the detection performance of an anomaly detector over a certain experimental data. Two versions of the EER metrics are used in the literature

Antal & Szab´ (2016) and Al-Obaidi (2016) .For comparison of experimental results, these are:-

1. Global EER ( $EER_g$ ): This is average of EER for a set of user data using a common fixed pass-mark threshold.
2. User EER ( $EER_u$ ): This is the average of EER for a set of user data using a variable pass-mark threshold to obtain the lowest EER for each user individually.

## **Chapter Four**

### **Experimental Results and Discussion**

## **4.1 Introduction**

This chapter presents the practical side of the research work: an implementation of the models discussed in chapter three, the data sources used in the experiments, the data collection and authentication modules, and analysis and discussion of the results. The data sources consist of a public dataset of related research, and a dataset collected using the developed data collection tool.

## **4.2 Objectives of the Experimental Work**

The experimental work is designed to fulfill the following tasks:

1. Evaluating the proposed anomaly detectors and feature sets using a public dataset.
2. Implementing a data collection tool based on the selected feature set and an authentication tool using a selected anomaly detector.
3. Data collection of user signatures.
4. Evaluating the feature sets and anomaly detectors using the new dataset.
5. Providing answers to the research question that are related to the research hypotheses.

### **4.3 Limitations of the Proposed Work**

The proposed work has the following limitations:-

1. The software platform for the proposed work is the Android operating system, therefore it will need adaptation to work on iOS and other operating systems.
2. The selected features set includes touch features that are measurable in modern touch devices, but might not be available on previous platforms.

### **4.4 EER Analysis Steps**

To measure the EER value for a set of mixed genuine and forgery samples for a group of users, the EER for each user is calculated separately using either a global pass-mark for all or a separate user pass-mark. The EER for a user is the average of False-Acceptance-Rate (FAR) and False-Rejection Rate of his signature attempts. The EER analysis will be performed using Excel and it consists of two analyses:

1. Random forgery analysis: for each subject, a set of signature samples are used for training to obtain the authentication template, and a similar number of genuine samples are used for positive testing. The random forgery signature samples are taken from all-other subjects, one feature vector selected randomly from signature data of the other subjects. The random forgery samples are used as the negative testing samples.
2. Skilled forgery analysis: for each subject that is the target of forgery, the same number of training and genuine testing samples are used as in the random forgery, while using the skilled forgery samples that are collected for forgeries against the targeted subject.

## 4.5 Feature Sets Selection

In chapter three a set of 26 authentication (calculated) features was proposed, to be used in the anomaly detection process. The authentication features are calculated from raw data features collected from the touch device during the signature process. The calculated features were chosen based on a preliminary analysis of the public MOBSIG dataset, on which we applied various features to test their contribution to reducing error rates. In order to evaluate the effectiveness of the proposed features, three alternative feature sets will be considered, as shown in Table 4-1.

**Table (4-1): The Proposed Calculated Feature Sets**

Feature set	Number of Features	Calculated Feature Set Elements
A	26	#Points, TotX, TotY, TotT, MedX, MedY, MedVX, MedVY, MedAccX, MedAccY, MedP, MedFA MaxVX, MaxVY, MaxAccX, MaxAccY, MaxP, , MaxFA, %XFlips, %YFlips, DispX, DispY, RatioXY1, RatioXY2, StdX , StdY
B	18	Set A, excluding velocity and acceleration related features
C	14	Set B, excluding pressure and finger area related features

Set A is the complete set of 26 elements, set B excludes velocity and acceleration and set C excludes velocity, acceleration, and pressure and finger area. Based on the data analysis results, the feature set which is associated with the lowest error rate, the EER, will be included in the authentication module.

## 4.6 Analysis of the MOBSIG Dataset

The MOBSIG dataset provides several categories of finger drawn signature measurements of movement over the touch surface. The dataset contains signatures' raw data of 83 subjects stored in comma-separated-value (CSV) files, where each subject has made 45 signatures over three sessions of 15 entries each. Also, the dataset contains skilled forgery data, where a skilled forgery is a signature attempt by a forger who knows the target signature, and there are 20 signature attempts against 77 of the subjects. The number of raw features vectors per signature varies from 40 to 300, where each row represents a point of measurement during the signature. Table (A-1) in appendix A shows the raw data features and a sample of the measurements. The data was collected on a 9-inch Nexus-9 tablet under Android 6.0.

The MOBSIG public dataset contained raw data features from which we extracted the proposed calculated features using the Extract-Features MATLAB program, where each signature's raw data vectors were aggregated into one vector of calculated features and stored in an Excel file. Table (A-2) shows a sample of the proposed calculated features extracted from the MOBSIG dataset.

A sample of the template that was generated using feature set B and the STD Z-Score model is shown in Table (A-3).

The EER, FAR and FRR results of the entire MOBSIG subjects are shown in table (A-4), which shows the individual subjects values as well as the average for the population. The results were calculated using feature set B, the STD Z-Score anomaly detector with threshold value of 3, and a global pass-mark of 14. The z-score threshold and the pass-mark value were chosen as they gave the lowest average EER results for the population.

## **4.7 Interfaces and Output of the Proposed TDSIG System**

The proposed system implementation consists of two parts: the data collection module and the authentication module, implemented in Java for Android. The data collection module collects the full signature raw data features shown in Table (B-1) in appendix B.

The authentication module implements the proposed feature set B and the STD Z-Score using the new threshold.

### **4.7.1 Screen Shots of the Proposed TDSIG System**

The proposed system provides the following interface screens:

1. System entry screen shown in Figure (4-1). Apart from registration, this screen provides settings change function, to update the pass-mark, Z-Score threshold and number of enrollment repetitions. The user can decide on the pass-mark based on his experience in using the system. The screen provides options for creating an account (registration) and enrollment, and for login for authentication of registered users, as well as the change settings function.
2. Account creation as shown in Figure (4-2).
3. Signature enrollment screen as shown in Figure (4-3). The user enters his signature a number of times as determined in the setting.
4. Signature authentication screen as shown in Figure (4-4). The logged-in user is allowed to enter his signature once, for authentication using the anomaly detector.
5. Authentication outcome screen as shown in Figure (4-5). The user receives the outcome of the authentication.

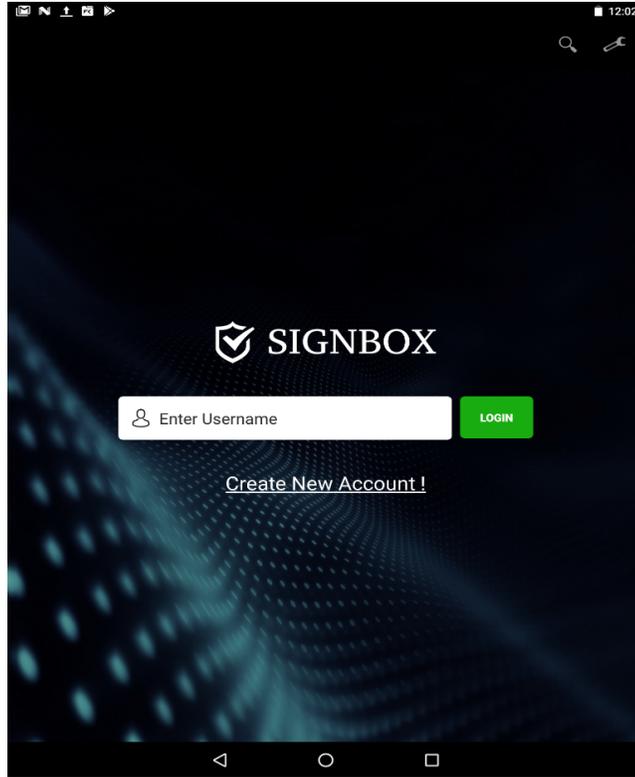


Figure (4-1): System entry screen

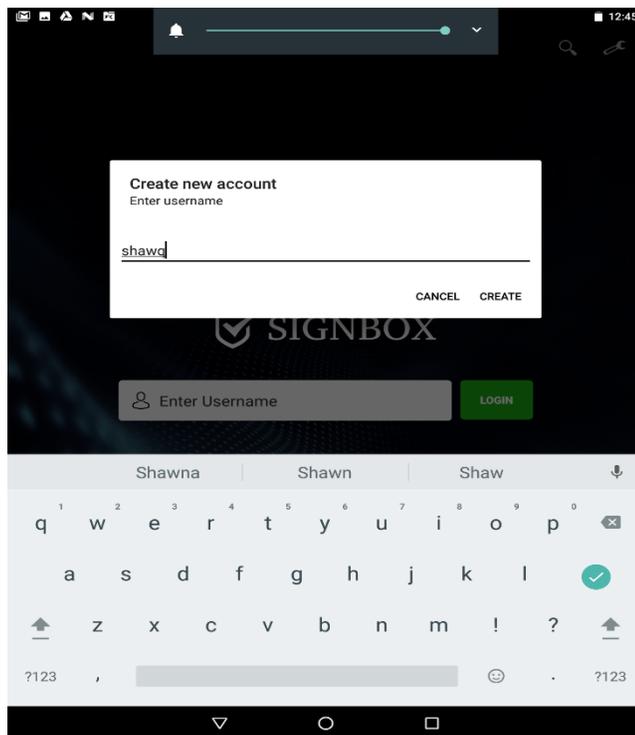
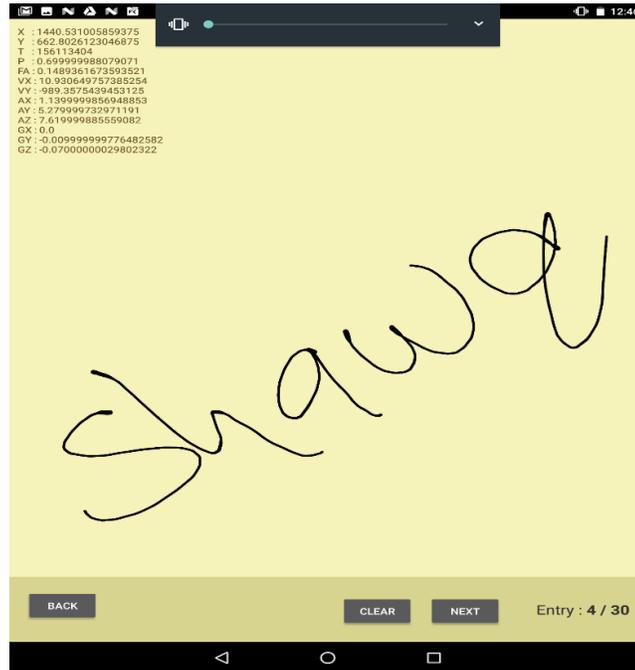


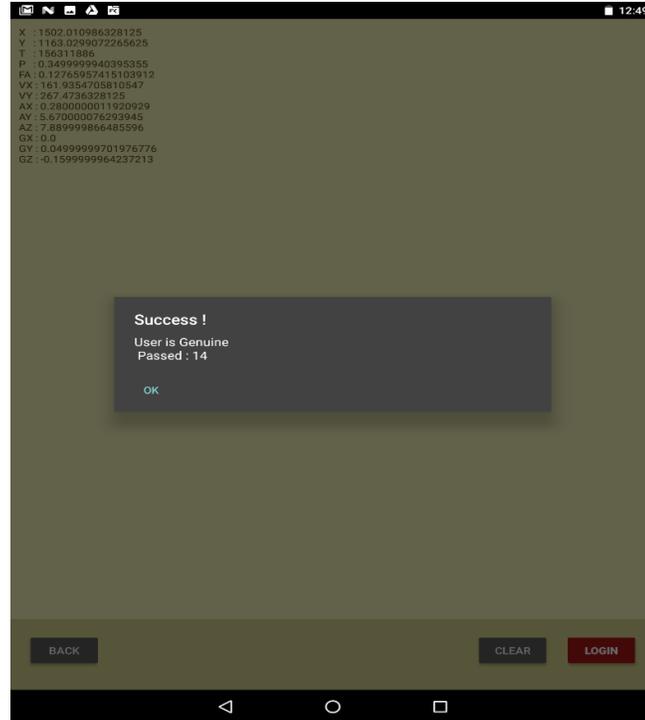
Figure (4-2) Account creation



**Figure (4-3) Signature enrollment screen**



**Figure (4-4) Signature authentication screen**



**Figure (4-5) Authentication outcome screen**

#### **4.8 Data Collection Using the Proposed TDSIG System**

The proposed system was implemented on a Nexus-9 tablet under Android 7.1 to provide two functions: data collection and signature authentication. The data collection module performed the tasks of collecting the signature raw data features vectors, aggregating the raw data into the calculated features vectors, one vector for each signature attempt, as shown in table (B-2) and then generating the authentication template which will be used by the authentication module, as shown in table (B-3).

The data collection module was used in collecting signature data of 55 subjects, 30 genuine signature attempts each, and 20 forgery signature attempts against each of the subjects. The forgery signature attempts were collected from 4 subjects who made 5 forgery signature attempts each, against the selected target subject. The collected data was partitioned for random and skilled forgery as follows:

1. **Random forgery:** 15 signature attempts are used for training, 15 signature attempts are used for genuine signature testing, and 54 signature attempts are used from all-others, one vector from each other user, randomly selected.
2. **Skilled forgery:** 15 genuine signature attempts are used for training samples, 15 genuine signature attempts are used for genuine testing samples, and 20 forgery signature attempts are used for forgery testing samples.

The EER, FAR and FRR results of the entire TDSIG subjects are shown in table (B-4), which shows the individual subjects values as well as the average for the population.

The results were calculated using feature set B, the STD Z-Score anomaly detector with threshold value of 3, and a global pass-mark of 14. The z-score threshold and the pass-mark value were chosen as they gave the lowest average EER results for the population.

## **4.9 Comparison and Discussion of Results**

This section presents experimental results of analyzing the new TDSIG dataset and the public MOBSIG dataset, both analyzed using the proposed feature sets and anomaly detectors. The EER metric is calculated using two scenarios; a global EER (EERg) where the pass-mark threshold is fixed for all subjects, and a user-based EER (EERu) where the pass-mark for each user is tuned to get to the point of equal FAR and FRR for the particular user.

### **4.9.1 Random Forgery Results**

Table 4-2 shows the random forgery EER results obtained by analyzing the two datasets using combinations of the proposed feature sets and anomaly detectors.

**Table (4-2): Random Forgery EER results of the proposed features / models**

Feature Set	Anomaly Detector	Z-Score Threshold	TDSIG Dataset		MOBSIG Dataset	
			EER <sub>g</sub>	EER <sub>u</sub>	EER <sub>g</sub>	EER <sub>u</sub>
A	STD Z-Score	3	3.49	0.71	7.43	4.23
	MAD Z-Score	4	2.24	0.64	10.14	7.25
	AAD Z-Score	4	3.26	0.52	7.75	4.10
B	STD Z-Score	3	2.89	0.90	6.50	3.14
	MAD Z-Score	4	3.02	1.07	7.65	4.39
	AAD Z-Score	4	2.76	0.81	6.55	3.05
C	STD Z-Score	3	3.48	1.45	7.57	3.96
	MAD Z-Score	4	4.29	1.71	8.31	4.93
	AAD Z-Score	4	3.30	1.21	7.39	4.05

The best Z-Score threshold for each combination was determined experimentally, as being the value that gave the lowest EER for that combination. The shown EER values are the average of individual EER values for the 83 subjects for the MOBSIG dataset and 55 subjects for the TDSIG dataset. The results show that feature set B gave the lowest EER with all anomaly detectors and for the two datasets. This suggests that the velocity and acceleration features did not have a positive contribution in improving the detection accuracy. Also, the STD Z-Score based anomaly detector gave the lowest EER among the other models, using a Z-Score threshold of 3, i.e. the acceptable distance from the mean is 3 standard deviations. Comparison of results of the two datasets show that both results have the same pattern in terms of the better anomaly detector, which is the STD Z-Score model, and the better feature set which is set B. However, by comparison between results of the two datasets, the new dataset results show lower EER in all combinations, which can be attributed to the difference in datasets size; the MOBSIG

dataset has 82 forgery attempts while the TDSIG dataset has 54 forgery attempts. The two datasets were collected using the same tablet (Nexus-9), but the data collection software were different, which might have contributed to the differences.

To investigate the effect of changing the training and forgery testing sample sizes on the error rates, the EER results of the two datasets were calculated using three different training samples (5, 10, 15), and two random forgery sample sizes (15 and 82 for MOBSIG and 15 and 54 for the new dataset) while keeping the number of genuine testing samples the same (15) for all cases. Table (4-3) shows that the lowest training sample size (5 each) produced the highest EER results for both datasets. However, for training sample sizes of 10 and 15 the MOBSIG results produced near equal EER values, while the TDSIG dataset results produced lower EER for the higher training sample size. These results can be used as a guideline in determining the number of training samples for a signature authentication application. The effect of reducing the random forgery sample size to be equal to the genuine sample size to showed less than 1% difference in both datasets, which indicates that increasing the negative sample size does not lead to significant improvement in error rates.

**Table (4-3): Random Forgery Training Sample Size Effect on EER,  
Using Global EER, STD Z-Score and Feature Set B**

MOBSIG Results			
Training Sample Size			EERg
#Training Samples	#Genuine Samples	#Forgery Samples	
5	15	82	12.55%
10	15	82	9.34%
15	15	82	9.43%
5	15	15	11.85%
10	15	15	8.39%
15	15	15	8.96%

TDSIG Results			
Training Sample Size			EERg
#Training Samples	#Genuine Samples	#Forgery Samples	
5	15	54	8.57%
10	15	54	4.19%
15	15	54	2.55%
5	15	15	8.24%
10	15	15	3.70%
15	15	15	2.79%

#### 4.9.2 Skilled Forgery Results

A skilled forgery signature attempt is based on knowledge by the forger of the shape of the target's signature. Generally, it is assumed that skilled forgery attempts would lead to higher authentication error rates because a skillfully forged signature is more likely to pass as a case of false acceptance than the random signature. The two datasets provide skilled forgery signature data collected from entries of some subjects attempting to forge signatures of others. Both datasets have 20 skilled forgery signatures per target subject. The new dataset provides skilled forgery data against all subjects of the dataset, while MOBSIG dataset has skilled forgery data against 77 subjects.

Table (4.4) shows skilled forgery EER results for the two datasets using the proposed feature sets and anomaly detectors. It can be seen that the skilled forgery results are slightly higher than the random forgery results, but there is no significant difference. The small gap between the skilled and random forgeries can be the result of using equal negative and positive samples.

**Table (4-4): Skilled Forgery EER results of the proposed features / models**

Feature Set	Anomaly Detector	Z-Score Threshold	TDSIG Dataset		MOBSIG Dataset	
			EER <sub>g</sub>	EER <sub>u</sub>	EER <sub>g</sub>	EER <sub>u</sub>
A	STD Z-Score	3	4.55	0.48	9.81	4.97
	MAD Z-Score	4	5.36	1.32	12.92	7.05
	AAD Z-Score	4	4.92	0.58	10.39	5.68
B	STD Z-Score	3	4.21	0.97	7.76	3.18
	MAD Z-Score	4	5.26	1.68	10.68	5.32
	AAD Z-Score	4	4.35	0.86	7.92	3.70
C	STD Z-Score	3	4.97	1.62	6.79	2.66
	MAD Z-Score	4	7.64	3.59	9.16	3.90
	AAD Z-Score	4	4.80	1.62	6.95	2.66

#### 4.9.3 Cross-Validation of the Results

To cross validate the experimental results, we switched the training and positive testing signature data, hence to have a 2-fold cross validation. Table (4-5) shows EER results of the switched training / testing samples for random forgery of the global EER metrics. There is no significant difference between the first and second folds for both datasets using the feature sets and anomaly detections combinations.

**Table (4-5): Random Forgery EER results of the proposed features / models Using session 2 data for training and session 1 for positive testing**

Feature Set	Anomaly Detector	Z-Score Threshold	TDSIG Dataset	MOBSIG Dataset
			EER <sub>g</sub>	EER <sub>g</sub>
A	STD Z-Score	3	4.29	9.10
	MAD Z-Score	4	5.77	10.82
	AAD Z-Score	4	4.42	9.34
B	STD Z-Score	3	4.97	7.59
	MAD Z-Score	4	6.34	8.78
	AAD Z-Score	4	4.89	7.65
C	STD Z-Score	3	6.64	8.50
	MAD Z-Score	4	7.45	8.78
	AAD Z-Score	4	6.47	8.40

#### 4.10 Inter-Dataset Analysis

To compare the effectiveness of the two datasets as a source of training samples that can be used in detecting forgeries from an independent source, an Inter-Dataset approach was applied, in which the training samples were from one dataset and the forgery testing samples were from another dataset. In this experiment, the MOBSIG and the TDSIG datasets were used interchangeably as training and random forgery testing sources. The genuine testing samples were from the same dataset that was used for training. Table 4-6 shows the EER results for two inter-dataset testing cases: training with MOBSIG and random forgery testing with the TDSIG dataset and vice versa, and in both cases the number of training samples were 5, 10 and 15 samples, while random forgery sample size was 55 in both cases. The STD Z-Score anomaly detector and feature set A were used in this analysis.

**Table (4-6): Inter-Dataset EER Results**

<b>Training Dataset</b>	<b># Training Samples</b>	<b>Testing Dataset</b>	<b># Forgery Testing Samples</b>	<b>EER</b>
MOBSIG	5	TDSIG	55	5.67%
MOBSIG	10	TDSIG	55	3.43%
MOBSIG	15	TDSIG	55	3.48%
TDSIG	5	MOBSIG	55	2.94%
TDSIG	10	MOBSIG	55	1.15%
TDSIG	15	MOBSIG	55	1.29%

The results show similar pattern of EER variability versus training sample size, with the 10-sample case providing the lowest error rate. This suggests that for an authentication application of this type, the choice of training sample size should be based on experimental results, in order to achieve lower authentication errors. In terms of comparison between the two datasets, the TDSIG EER results are consistently lower than the MOGSIG results, which indicates that the TDSIG training samples can lead to more effective rejection of forgeries. Moreover, the inter-mixing of the two datasets and the obtained results, confirm that the proposed anomaly detector and feature set produce similar pattern of results despite the fact that the training and testing samples are from independent sources.

## 4.11 Summary of Contributions

Contributions of the work in this thesis, as presented in chapters 3 and 4, can be summarized as follows:

1. Formulating a new anomaly detector based on the Z-Score Outlier distance function, using the Average Absolute Deviation metric.
2. Enhancing the Z Score distance functions through experimental work to determine a better value for the thresholds which resulted in more accurate authentication (the classical threshold is 2 for STD Z-Score, our thresholds are 3 for STD, and 4 for AAD and MAD Z-Scores).
3. Designing three feature sets, and experimentally selecting the best set that improved authentication and reduced error rates.
4. Evaluating the proposed anomaly detectors and feature sets using a public dataset, and a dataset collected in this research.
5. Highlighting the effect of training sample size or authentication accuracy.

**Using** an inter-data approach to evaluate quality of the two datasets as a training samples source.

## **Chapter Five**

### **Conclusion and Future Work**

## 5.1 Conclusion

The work in this thesis presented the design of a graphic signature authentication system based on the empirical study of anomaly detectors and feature sets, with the aim of improving authentication accuracy. The work involved analysis of a public graphic signature dataset, collection of a new dataset using the implemented system, and comparison of the analysis of results of the two datasets. The following points summarize the conclusion of this thesis:

- ❖ Experimental evaluation of the proposed anomaly detector(s) and feature sets for signature authentication have shown that it is possible to reduce error rates by choosing better models and features without the need for additional sensors or hardware.
- ❖ Variation of training sample size for both datasets showed a significant change in the EER values, with better results obtained for training sample sizes of 10-15. However, there were insignificant difference in the EER values when the random forger sample size were reduced to the same number as the genuine testing samples. The EER rate for both datasets using the proposed models and features showed that
- ❖ The proposed model performed almost equally well in detecting random and skilled forgeries. The skilled forgery error rate should have been much higher. The difference in negative sample size between random and skilled forgery (82 vs. 20 for MOBSIG and 54 vs. 20 for TDSIG dataset) did cause a significant difference between the two cases. This indicates that knowing the shape of the signature is not as important as the behavioral biometrics of the signature, in the process of detecting a forgery attempt.

- ❖ When the pass-mark threshold is tuned per user, the average EER is much lower than the case of using a global (fixed) pass-mark, this suggests that the authentication application would do better if the pass-mark is tuned to the user's signature behavior.

## 5.2 Future Work

The graphic signature authentication field is still in its early stages and further improvements are needed. Based on results of the present work, the following suggestions are put forward:

1. Investigating other sensor features that become available on new mobile devices
2. Collecting a larger graphic signature dataset with higher number of skilled forgery signature samples
3. Investigating other statistical features that can reduce authentication errors.
4. Comparing the detection performance of the enhanced Z-Score anomaly detectors with other distance-based model.

## References

- Agarwal, V., & Taffler, R. J. (2007). Twenty-Five Years Of The Taffler Z-Score Model: Does It Really Have Predictive Ability?. *Accounting and Business Research*, 37(4), 285-300..
- Barrett, P. (2006). Euclidean Distance: Raw, Normalised, And Double-Scaled Coefficients. Unpublished Paper Retrieved From [http://www.pbmetrix.com/techpapers / Euclidean\\_Distance. pdf](http://www.pbmetrix.com/techpapers/Euclidean_Distance.pdf).
- Bissig, P. (2011). Signature verification on finger operated touchscreen devices. *ETH Zürich, Distributed Computer Group*.
- Bubeck, D. S. U., & Sanchez, D. (2003). Biometric Authentication. *Universidade Estadual de San Diego*.
- Chandola, V., Banerjee, A., & Kumar, V. (2009). Anomaly detection: A survey. *ACM computing surveys (CSUR)*, 41(3), 15.
- Chiasson, S., Forget, A., Stobert, E., van Oorschot, P. C., & Biddle, R. (2009, November). Multiple password interference in text passwords and click-based graphical passwords. *In Proceedings of the 16th ACM conference on Computer and communications security* (pp. 500-511). ACM.
- Craik, F. I., & McDowd, J. M. (1987). Age differences in recall and recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 13(3), 474.
- De Angeli, A., Coventry, L., Johnson, G., & Renaud, K. (2005). Is a picture really worth a thousand words? Exploring the feasibility of graphical authentication systems. *International journal of human-computer studies*, 63(1-2), 128-152.

- Durgesh, K. S., & Lekha, B. (2010). Data classification using support vector machine. *Journal of Theoretical and Applied Information Technology*, 12(1), 1-7.
- Edjabou, M. E., Martín-Fernández, J. A., Scheutz, C., & Astrup, T. F. (2017). Statistical analysis of solid waste composition data: Arithmetic mean, standard deviation and correlation coefficients. *Waste Management*, 69, 13-23.
- Antal, M., & Szabó, L. Z. (2016, May). On-line verification of finger drawn signatures. In *Applied Computational Intelligence and Informatics (SACI), 2016 IEEE 11th International Symposium on* (pp. 419-424). IEEE.
- Ental, M. & Lzsalo, S. (2016). The MOBSIG database, <http://www.ms.sapientia.ro/~manyi/mobisig/MOBISIG.ZIP> (viewed on 10/7/2017).
- Hall, M., Frank, E., Holmes, G., Pfahringer, B., Reutemann, P., & Witten, I. H. (2009). The WEKA data mining software: an update. *ACM SIGKDD explorations newsletter*, 11(1), 10-18.
- Houmani, N., Garcia-Salicetti, S., Dorizzi, B., & El-Yacoubi, M. (2010, October). On-line signature verification on a mobile platform. *In International Conference on Mobile Computing, Applications, and Services* (pp. 396-400). Springer, Berlin, Heidelberg.
- Impedovo, D., & Pirlo, G. (2008). Automatic signature verification: The state of the art. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, 38(5), 609-635.

- Natarajan, N., Koyejo, O., Ravikumar, P. K., & Dhillon, I. S. (2014). Consistent binary classification with generalized performance metrics. *In Neural Information Processing Systems (NIPS)*.
- Krish, R. P., Fierrez, J., Galbally, J., & Martinez-Diaz, M. (2013, May). Dynamic signature verification on smart phones. *In International Conference on Practical Applications of Agents and Multi-Agent Systems* (pp. 213-222). Springer, Berlin, Heidelberg.
- Lopes, H., & Chatterjee, M. A Survey of User Authentication Schemes for Mobile Device.
- Martinez-Diaz, M., Fierrez, J., & Galbally, J. (2013). The doodb graphical password database: Data analysis and benchmark results. *IEEE Access*, 1, 596-605.
- Martinez-Diaz, M., Fierrez, J., & Galbally, J. (2016). Graphical password-based user authentication with free-form doodles. *IEEE Transactions on Human-Machine Systems*, 46(4), 607-614.
- Nelson, D. L., Reed, V. S., & McEvoy, C. L. (1977). Learning to order pictures and words: A model of sensory and semantic encoding. *Journal of Experimental Psychology: human learning and memory*, 3(5), 485.
- Al-Obaidi, Noor Mahmood Shakir (2016). **A New Statistical Anomaly Detector Model for Keystroke Dynamics on Touch Mobile Devices**, Master Thesis, Department of Computer Science, Faculty of Information Technology, Middle East University.

- Renaud, K. V. (2009). Guidelines for designing graphical authentication mechanism interfaces. *International Journal of Information and Computer Security*, 3(1), 60-85.
- Richiardi, J., Ketabdar, H., & Drygajlo, A. (2005). Local and Global Feature Selection For On-Line Signature Verification. In Document Analysis And Recognition, 2005. Proceedings. *Eighth International Conference on IEEE* (pp. 625-629).
- Rousseeuw, P. J., & Croux, C. (1993). Alternatives to the median absolute deviation. *Journal of the American Statistical association*, 88(424), 1273-1283.
- Sae-Bae, N., & Memon, N. (2014). Online signature verification on mobile devices. *IEEE Transactions on Information Forensics and Security*, 9(6), 933-947.
- Standing, L., Conezio, J., & Haber, R. N. (1970). Perception and memory for pictures: Single-trial learning of 2500 visual stimuli. *Psychonomic Science*, 19(2), 73-74.
- Stokes, R., Willis, A., Bryant, K. S., Tyler, Z., & Dobson, A. (2016). Comparison of Biometric Authentication Software Techniques: **GEFE** vs. Angle Based Metrics. In *MAICS* (pp. 75-89).
- Vu, K. P. L., Proctor, R. W., Bhargav-Spantzel, A., Tai, B. L. B., Cook, J., & Schultz, E. (2007). Improving password security and memorability to protect personal and organizational information. *International Journal of Human-Computer Studies*, 65(8), 744-757.

- Wagenmakers, E. J., & Brown, S. (2007). On the linear relation between the mean and the standard deviation of a response time distribution. *Psychological review*, 114(3), 830.
- Bissig, P. (2011). Signature verification on finger operated touchscreen devices. *ETH Zürich, Distributed Computer Group*.
- Donato Impedovo, a. G. (2008). Automatic Signature Verification: The State of the Art. *IEEE*, 609-626.
- Martinez-Diaz, M., Fierrez, J., & Galbally, J. (2016). Graphical password-based user authentication with free-form doodles. *IEEE Transactions on Human-Machine Systems*, 46(4), 607-614.
- Zabó, L. Z., & Tordai, T (2016). On-line Signature Verification on MOBISIG Finger Drawn Signature Corpus.
- Houmani, N., Garcia-Salicetti, S., Dorizzi, B., & El-Yacoubi, M. (2010, October). On-line signature verification on a mobile platform. *In International Conference on Mobile Computing, Applications, and Services* (pp. 396-400). Springer, Berlin, Heidelberg.
- Ram P. Krish, J. F.-D. (2012). Dynamic Signature Verification on Smart Phones. *ACM*, 1-10.

## **Appendix A**

**Samples of raw data features, calculated features ,the  
generated templates and summery of the results of the  
MOBSIG dataset**

**Table (A-1) the raw data features and a sample of the measurements from the MOBSIG dataset**

<b>x</b>	<b>y</b>	<b>timestamp</b>	<b>pressure</b>	<b>fingerarea</b>	<b>velocityx</b>	<b>velocityy</b>	<b>accelx</b>	<b>accely</b>	<b>accelz</b>	<b>gyrox</b>	<b>gyroy</b>	<b>gyroz</b>
329.3706	500.1036	1.23E+08	0.5625	0.08511	0	0	0.00676	-0.00646	-0.00987	0	0	0
329.3706	500.1036	1.23E+08	0.625	0.08511	-2.8639	-4.33915	0.00676	-0.00646	-0.00987	0	0	0
329.3706	500.1036	1.23E+08	0.6125	0.06383	-0.01615	-0.02076	0.00676	-0.00646	-0.00987	0	0	0
338.2527	482.3416	1.23E+08	0.625	0.07447	564.4932	-1128.97	0.00676	-0.00646	-0.00987	0	0	0
356.1869	437.1559	1.23E+08	0.625	0.06383	1157.399	-2718.01	0.00676	-0.00646	-0.00987	0	0	0
376.026	375.8759	1.23E+08	0.625	0.07447	1486.332	-4081.46	0.00676	-0.00646	-0.00987	0	0	0
395.8948	325.0769	1.23E+08	0.625	0.06383	1625.73	-4574.83	0.00676	-0.00646	-0.00987	0	0	0
409.969	294.1569	1.23E+08	0.625	0.10638	1473.892	-3954.92	0.00676	-0.00646	-0.00987	0	0	0
417.8289	279.778	1.23E+08	0.6125	0.08511	649.6567	-1206.68	0.00676	-0.00646	-0.00987	0	0	0

**Table (A-2) sample of the proposed calculated feature set B extracted from the MOBSIG dataset**

Subject	#Points	TotX	TotY	TotT	MedX	MedY	MedVx	MedVY	MaxVX	MaxVY	MedAccX	MedAccY	MaxAccx	MaxAccy	MedP	MaxP	MedFA	MaxFA	%X Flips	%Y Flips	DispX	DispY
G01-01.csv	146	1317.635	2393.759	2966	7.5946	10.09314	474.5528	794.6798	2222.28	6027.167	0.006756	0.015913	0.019738	0.028731	0.575	0.65	0.085106	0.12766	0.219178	0.506849	812.7355	410.736
G01-02.csv	142	1111.886	2409.851	3450	7.26593	12.56594	535.8574	868.9414	1715.978	4367.561	0.011525	0.010133	0.033268	0.019909	0.5	0.6375	0.085106	0.12766	0.197183	0.471831	751.7552	424.704
G01-03.csv	135	1096.392	2243.358	2962	7.59725	10.78632	565.5355	800.8658	1658.667	4637.368	0.005317	0.016871	0.023354	0.018911	0.5125	0.6375	0.085106	0.12766	0.222222	0.459259	705.6499	391.829
G01-04.csv	145	1136.578	2100.193	2819	6.977335	8.473395	479.0134	783.9153	2096.098	4501.585	0.006706	0.001031	0.018635	0.036737	0.475	0.575	0.085106	0.12766	0.255172	0.462069	744.7576	332.300
G01-05.csv	152	1151.215	1984.596	3068	6.44561	9.21351	446.5247	605.5988	1872.689	3495.715	0.009773	0.017466	0.023078	0.018028	0.5375	0.6375	0.085106	0.12766	0.190789	0.506579	745.2574	355.401
G01-06.csv	156	1135.246	2109.508	3223	6.1078	9.86036	462.9098	646.6266	1962.199	4054.042	0.017858	0.01724	0.032116	0.023173	0.55	0.625	0.085106	0.12766	0.230769	0.480769	775.7475	316.307
G01-07.csv	162	1078.602	2028.659	3496	5.59903	7.3944	363.1393	537.5905	1392.014	3904.885	0.009598	0.018831	0.014143	0.027077	0.55	0.65	0.085106	0.12766	0.216049	0.469136	715.2671	377.936
G01-08.csv	172	1045.881	2143.972	3351	4.689	6.76758	344.2302	549.1402	1331.82	3902.885	0.014768	0.022114	0.036185	0.033478	0.3875	0.525	0.074468	0.12766	0.197674	0.476744	717.7664	332.300
G01-09.csv	148	929.7366	2195.247	3256	4.65027	8.59593	330.463	613.8983	1775.56	4800.913	0.014715	0.012346	0.020228	0.024567	0.4875	0.6	0.085106	0.12766	0.263514	0.452703	644.4225	326.080
G01-10.csv	141	853.399	2129.596	3802	4.75685	8.810785	328.805	672.8036	1661.369	5331.528	0.01779	0.017561	0.030232	0.021805	0.50625	0.6625	0.074468	0.12766	0.262411	0.503546	634.5482	330.241

**Table (A-3) sample of the template that was generated using feature set B and the STD Z-Score model**

	#Points	TotX	TotY	TotT	MedX	MedY	MedVx	MedVY	MaxVX	MaxVY	MedAccX	MedAccY	MaxAccx	MaxAccy	MedP	MaxP	MedFA	MaxFA	%X Flips	%Y Flips	DispX	DispY	RatioXY1	RatioXY2
<b>MED</b>	151.000	1000.975	2119.552	3308.500	5.160	9.216	373.143	670.911	1583.960	3934.121	0.011	0.011	0.024	0.024	0.475	0.600	0.085	0.128	0.228	0.485	684.785	338.455	0.463	1.907
<b>MEAN</b>	151.900	1005.891	2126.168	3322.650	5.600	9.211	398.236	687.077	1578.091	4186.694	0.011	0.012	0.029	0.027	0.481	0.601	0.083	0.128	0.233	0.488	681.225	347.530	0.472	1.971
<b>MAD</b>	6.000	126.873	96.160	191.500	0.549	0.632	43.510	78.961	314.634	436.664	0.004	0.005	0.006	0.004	0.050	0.038	0.000	0.000	0.027	0.015	55.105	19.476	0.044	0.104
<b>AAD</b>	7.590	126.627	115.337	220.815	0.955	0.882	61.664	82.022	329.002	575.718	0.004	0.005	0.010	0.008	0.049	0.038	0.003	0.000	0.024	0.017	62.327	29.841	0.047	0.165
<b>STD</b>	9.754	149.563	150.276	285.664	1.113	1.239	75.036	98.906	385.351	719.832	0.005	0.006	0.014	0.014	0.061	0.045	0.004	0.000	0.028	0.022	75.750	39.360	0.055	0.217

**Table (A-4) Random Forgery EER Results of MOBSIG Dataset**  
**Using STD Z-Score Anomaly Detector and Feature Set B**

**Z-Score Threshold: 3, Global Pass-Mark: 15**

<b>Subject</b>	<b>FRR</b>	<b>FAR</b>	<b>EER</b>
user 1	0.00%	0.00%	0.00%
user 2	0.00%	3.66%	1.83%
user 3	10.00%	1.22%	5.61%
user 4	30.00%	0.00%	15.00%
user 5	5.00%	4.88%	4.94%
user6	0.00%	3.66%	1.83%
user 7	25.00%	0.00%	12.50%
user8	0.00%	2.44%	1.22%
user9	0.00%	6.10%	3.05%
user 10	0.00%	1.22%	0.61%
user 11	5.00%	4.88%	4.94%
user 12	45.00%	0.00%	22.50%
user 13	0.00%	8.54%	4.27%
user 14	0.00%	1.22%	0.61%
user 15	0.00%	1.22%	0.61%
user 16	10.00%	0.00%	5.00%
user 17	0.00%	3.66%	1.83%
user 18	10.00%	2.44%	6.22%
user 19	0.00%	4.88%	2.44%
user 20	5.00%	0.00%	2.50%
user 21	0.00%	0.00%	0.00%
user 22	10.00%	2.44%	6.22%
user 23	0.00%	20.73%	10.37%
user 24	0.00%	46.34%	23.17%
user 25	0.00%	0.00%	0.00%
user 26	0.00%	3.66%	1.83%
user 27	0.00%	36.59%	18.29%
user 28	0.00%	19.51%	9.76%
user 29	0.00%	0.00%	0.00%
user 30	15.00%	2.44%	8.72%
user31	30.00%	1.22%	15.61%
user 32	10.00%	1.22%	5.61%
user33	0.00%	7.32%	3.66%
user34	0.00%	1.22%	0.61%
user35	25.00%	0.00%	12.50%
user36	35.00%	0.00%	17.50%
user37	5.00%	0.00%	2.50%
user38	0.00%	9.76%	4.88%
user39	5.00%	0.00%	2.50%
user40	20.00%	0.00%	10.00%

user41	5.00%	3.66%	4.33%
user42	0.00%	3.66%	1.83%
user43	5.00%	0.00%	2.50%
user44	0.00%	0.00%	0.00%
user45	30.00%	0.00%	15.00%
user46	0.00%	1.22%	0.61%
user47	0.00%	2.44%	1.22%
user48	0.00%	3.66%	1.83%
user49	0.00%	2.44%	1.22%
user50	0.00%	1.22%	0.61%
user51	0.00%	3.66%	1.83%
user52	5.00%	0.00%	2.50%
user53	0.00%	4.88%	2.44%
user54	10.00%	10.98%	10.49%
user55	15.00%	4.88%	9.94%
user56	5.00%	1.22%	3.11%
user57	0.00%	20.73%	10.37%
user58	0.00%	0.00%	0.00%
user59	0.00%	6.10%	3.05%
user60	0.00%	0.00%	0.00%
user61	15.00%	0.00%	7.50%
user62	0.00%	3.66%	1.83%
user63	0.00%	1.22%	0.61%
user64	15.00%	2.44%	8.72%
user65	0.00%	10.98%	5.49%
user66	0.00%	9.76%	4.88%
user67	40.00%	1.22%	20.61%
user68	35.00%	1.22%	18.11%
user69	0.00%	30.49%	15.24%
user70	0.00%	36.59%	18.29%
user71	0.00%	4.88%	2.44%
user72	15.00%	2.44%	8.72%
user73	10.00%	0.00%	5.00%
user74	5.00%	1.22%	3.11%
user75	0.00%	2.44%	1.22%
user76	0.00%	17.07%	8.54%
user77	50.00%	6.10%	28.05%
user78	0.00%	31.71%	15.85%
user79	0.00%	15.85%	7.93%
user80	10.00%	14.63%	12.32%
user81	0.00%	1.22%	0.61%
user82	30.00%	8.54%	19.27%
user83	0.00%	2.44%	1.22%
<b>Average</b>	<b>7.23%</b>	<b>5.77%</b>	<b>6.50%</b>

## **Appendix B**

**Samples of raw data features, calculated features , the  
generated templates and summary of results of the  
TDSIG dataset**

**Table (B-1) the raw data features and a sample of the data collected by proposed system**

X	Y	Timestamp	Pressure	Finger Area	VelocityX	VelocityY	AccelX	AccelY	AccelZ	GyroX	GyroY	GyroZ
191.5129	951.8472	111874415	1.375	0.12766	234.8681	426.7452	-0.05	0.2	9.63	0	0	0
208.2494	984.376	111874431	1.375	0.12766	277.7388	505.9216	0.01	0.1	9.59	0	0	0
223.1955	1025.562	111874448	1.3625	0.12766	249.0601	525.1476	-0.02	0.2	9.59	0	0	0
241.5047	1058.85	111874465	1.325	0.12766	188.4229	438.3649	-0.02	0.2	9.59	0	0	0
263.266	1093.965	111874481	1.2875	0.117021	154.8212	316.152	-0.02	0.12	9.63	0	0	0
287.0883	1124.65	111874498	1.2625	0.12766	199.1228	271.2251	-0.02	0.16	9.57	0	0	0
307.9219	1143.947	111874515	1.2375	0.117021	222.3889	196.6022	-0.02	0.14	9.59	0	0	0
319.0198	1153.228	111874532	1.25	0.117021	164.301	93.21954	-0.02	0.18	9.57	0	0	0
325.2705	1155.922	111874548	1.2875	0.117021	63.56973	-14.0668	-0.02	0.18	9.57	0	0	0
331.3954	1153.609	111874565	1.275	0.12766	1.88467	-91.1636	-0.02	0.1	9.59	0	0	0

Table (B-2) sample of the proposed calculated feature set B

Subject	#Points	TotX	TotY	TotT	MedX	MedY	MedVx	MedVY	MaxVX	MaxVY	MedAccX	MedAccY	MaxAccx	MaxAccy	MedP	MaxP	MedFA	MaxFA	%X Flips	%Y Flips	DispX	DispY	RatioXY1	RatioXY2
hind-01	164	1731.44	2912.403	8643	6.437408	11.64026	42.06781	103.8559	284.3928	610.4698	0.02	0.14	0.11	0.41	0.9125	1.025	0.12766	0.148936	0.926829	0.329268	904.7055	481.5687	0.594505	1.878663
hind-02	131	1640.986	2275.185	7299	6.989624	11.12372	45.90464	99.14857	317.0632	470.3076	0.03	0.12	0.19	0.29	0.9625	1.0375	0.12766	0.148936	1.175573	0.312977	955.6778	392.761	0.721254	2.43323
hind-03	135	1692.986	2380.734	6998	7.984863	13.09897	66.12941	109.9301	286.3073	534.9914	0.03	0.14	0.11	0.33	1.075	1.2625	0.12766	0.148936	1.118519	0.325926	994.6761	401.6034	0.711119	2.476762
hind-04	145	2329.488	2591.96	8305	8.201782	9.911194	61.52888	104.052	337.1725	581.2111	0.03	0.12	0.23	0.31	1.1	1.275	0.12766	0.12766	1.096552	0.37931	1000.174	512.6664	0.898736	1.950927
hind-05	127	1713.915	2602.956	7694	9.176331	14.90771	82.05565	134.7946	305.8917	630.6149	0.03	0.12	0.11	0.29	1.1125	1.2625	0.12766	0.148936	1.228346	0.346457	1062.694	451.3596	0.65845	2.354428
hind-06	122	1763.772	2510.97	7780	10.37366	14.16235	93.11111	135.0598	307.8489	509.5622	0.03	0.14	0.07	0.31	1.1	1.3	0.117021	0.12766	1.311475	0.336066	996.1756	498.4503	0.702426	1.998546
hind-07	96	1711.937	1939.724	6944	13.75516	14.4165	121.6994	152.1004	410.7907	635.0549	0.03	0.14	0.13	0.25	1.15	1.3625	0.12766	0.148936	1.708333	0.375	965.1857	443.3631	0.882567	2.176965
hind-08	90	1826.665	2294.864	6939	14.81604	17.83301	126.2783	166.1281	548.9266	828.2991	0.03	0.14	0.36	0.35	1.1625	1.375	0.12766	0.12766	1.844444	0.333333	968.1848	438.0707	0.795979	2.210111

**Table (B-3) sample of the template that was generated using feature set B and the STD Z-Score model**

	#Points	TotX	TotY	TotT	MedX	MedY	MedVx	MedVY	MaxVX	MaxVY	MedAccX	MedAccY	MaxAccx	MaxAccy	MedP	MaxP	MedFA	MaxFA	%X Flips	%Y Flips	DispX	DispY	RatioXY1	RatioXY2
Med	127.000	1763.772	2275.185	7128.000	10.294	13.037	93.111	122.448	337.173	534.991	0.030	0.140	0.170	0.330	1.150	1.325	0.128	0.149	1.272	0.333	968.185	438.071	0.796	2.210
Mean	124.600	1801.691	2306.489	7335.733	10.102	13.019	88.775	125.153	354.712	556.694	0.030	0.135	0.173	0.338	1.113	1.293	0.124	0.143	1.314	0.340	966.208	437.978	0.788	2.231
MAD	5.000	70.786	158.197	184.000	1.146	1.379	13.597	13.713	33.659	46.220	0.000	0.000	0.060	0.020	0.013	0.050	0.000	0.000	0.096	0.006	27.991	39.564	0.085	0.259
AAD	12.667	114.539	195.544	410.542	1.681	1.537	20.081	16.237	49.288	68.084	0.003	0.008	0.057	0.054	0.056	0.081	0.005	0.008	0.160	0.014	42.809	39.665	0.074	0.229
STD	18.259244	170.96535	256.74871	535.3807	2.2761443	1.9927883	25.6856	19.643653	68.515076	96.417552	0.0065465	0.0091548	0.0742262	0.0919783	0.0782719	0.1171461	0.005191	0.0097391	0.2312918	0.0186556	55.889339	49.433739	0.0928969	0.2655215

**Table (B-4) Random Forgery EER Results of TDSIG Dataset**  
**Using STD Z-Score Anomaly Detector and Feature Set B**  
**Z-Score Threshold: 3, Global Pass-Mark: 15**

Subject	FRR	FAR	EER
User1	0.00%	0.00%	0.00%
User2	0.00%	0.00%	0.00%
User3	6.67%	0.00%	3.33%
User4	0.00%	3.70%	1.85%
User5	0.00%	1.85%	0.93%
User6	0.00%	0.00%	0.00%
User7	0.00%	1.85%	0.93%
User8	0.00%	1.85%	0.93%
User9	0.00%	0.00%	0.00%
User10	0.00%	40.74%	20.37%
User11	0.00%	1.85%	0.93%
User12	0.00%	0.00%	0.00%
User13	0.00%	0.00%	0.00%
User14	0.00%	7.41%	3.70%
User15	0.00%	0.00%	0.00%
User16	0.00%	3.70%	1.85%
User17	0.00%	0.00%	0.00%
User18	20.00%	0.00%	10.00%
User19	0.00%	1.85%	0.93%
User20	6.67%	0.00%	3.33%
User21	46.67%	0.00%	23.33%
User22	0.00%	18.52%	9.26%
User23	0.00%	0.00%	0.00%
User24	0.00%	3.70%	1.85%
User25	0.00%	11.11%	5.56%
User26	0.00%	0.00%	0.00%
User27	0.00%	5.56%	2.78%
User28	0.00%	3.70%	1.85%
User29	6.67%	0.00%	3.33%
User30	0.00%	1.85%	0.93%
User31	6.67%	0.00%	3.33%
User32	0.00%	0.00%	0.00%
User33	0.00%	1.85%	0.93%
User34	6.67%	1.85%	4.26%
User35	0.00%	1.85%	0.93%
User36	0.00%	0.00%	0.00%
User37	13.33%	0.00%	6.67%
User38	0.00%	3.70%	1.85%
User39	0.00%	1.85%	0.93%
User40	0.00%	1.85%	0.93%
User41	0.00%	0.00%	0.00%

User42	13.33%	0.00%	6.67%
User43	0.00%	1.85%	0.93%
User44	0.00%	0.00%	0.00%
User45	0.00%	9.26%	4.63%
User46	0.00%	0.00%	0.00%
User47	6.67%	1.85%	4.26%
User48	6.67%	0.00%	3.33%
User49	6.67%	1.85%	4.26%
User50	0.00%	0.00%	0.00%
User51	6.67%	5.56%	6.11%
User52	20.00%	0.00%	10.00%
User53	0.00%	0.00%	0.00%
User54	0.00%	1.85%	0.93%
User55	0.00%	0.00%	0.00%
<b>Average</b>	<b>3.15%</b>	<b>2.63%</b>	<b>2.89%</b>