An Authentication Framework for Wearable Devices

by

Abdullah Alharbi

A dissertation
submitted to Florida Institute of Technology
in partial fulfillment of the requirements
for the degree of

Doctorate of Philosophy
in
Computer Science

Melbourne, Florida
June, 2018
We the undersigned committee
hereby approve the attached dissertation

An Authentication Framework for Wearable Devices by Abdullah Alharbi

Heather Crawford, Ph.D.
Assistant Professor
School of Computing
Committee Chair

Katrina Ellis, Ph.D.
Assistant Professor
School of Psychology
Outside Committee Member

William Allen, Ph.D.
Associate Professor
School of Computing
Committee Member

Keith Gallagher, Ph.D.
Associate Professor
School of Computing
Committee Member

Philip Bernhard, Ph.D.
Associate Professor and Director
School of Computing
ABSTRACT

Title:
An Authentication Framework for Wearable Devices

Author:
Abdullah Alharbi

Major Advisor:
Heather Crawford, Ph.D.

The popularity of wearable devices (WDs) has been growing rapidly in recent years because of the convenience they add to users lives. Although WDs may be considered an extension of mobile computing devices (e.g., smartphones), their form factor is very different: WDs are always available and expected to be accessible, and they often lack typical input means such as a keyboard. Therefore, they need to be treated, in terms of authentication, differently than other computing devices since we cannot simply apply possibly unsuitable traditional authentication methods such as password that were designed for other modalities. In this dissertation, we introduce the Authentication Framework for Wearable Devices (AFWD) which provides a model to create trustworthy, continuous, transparent, and user-accepted authentication methods for WDs. The AFWD respects the WDs unique form factor by authenticating wearers transparently and continuously. It also respects the WDs limitations such as the lack of input methods by exploiting the current WD wearers use pattern (e.g., behavioral biometrics) to verify that they are the owner of the device. The AFWD is designed to be hardware and software independent, as well as be used with various types of biometric modalities which makes it flexible in light of WDs and biometric future changes.
# Table of Contents

Abstract iii

List of Figures ix

List of Tables xi

Abbreviations xii

1 Introduction 1
   1.1 Research Problem 2
      1.1.1 Problem Statement 3
   1.2 Our Solution: The Authentication Framework for Wearable Devices (AFWD) 4
      1.2.1 Research Question 5
      1.2.2 Research Hypotheses 5
      1.2.3 Null Hypothesis 6
   1.3 Dissertation Structure 6

2 Background 7
   2.1 Overview 7
   2.2 Wearable Devices in History 9
2.3 Wearable Device Classifications ........................................ 14
2.4 Mobile Devices and Wearable Devices Crossover .................... 14
2.5 User Authentication ..................................................... 17
  2.5.1 Textual Passwords and Personal Identification Numbers ..... 19
  2.5.2 Graphical Passwords ............................................ 20
  2.5.3 Biometric Authentication ......................................... 21
    2.5.3.1 Biometric Performance Metrics ............................ 22
    2.5.3.2 Physiological Biometrics .................................. 25
    2.5.3.3 Behavioral Biometrics: ..................................... 28
  2.5.4 Explicit, Transparent, and Continuous Authentication ...... 32
2.6 Authentication in Wearable Devices .................................. 35
  2.6.1 Head-mounted Device Authentication ............................ 36
    2.6.1.1 HMD Authentication Based on Accelerometer and
    Gyroscope Data .................................................. 37
    2.6.1.2 HMD Authentication Based on Multiple Sensors ....... 39
    2.6.1.3 HMD Authentication Based on Camera Data ............ 41
  2.6.2 Hand-worn Device Authentication ................................ 43
    2.6.2.1 Hand-worn Authentication Based on Accelerometer and
    Gyroscope Data .................................................. 43
    2.6.2.2 Hand-worn Authentication Based on Vital Signs ....... 44
  2.6.3 Body-worn Device Authentication ................................ 45
    2.6.3.1 Body-worn WD Based on Accelerometer and Gyrosco-
    pse Data .......................................................... 45
2.7 Pattern Classification .................................................. 46
2.8 Summary ............................................................... 49
3 The Authentication Framework for Wearable Devices

3.1 AFWD Overview .............................................. 50
   3.1.1 Device Classification Method ......................... 53
   3.1.2 AFWD and Security ..................................... 57
   3.1.3 AFWD Assumptions ..................................... 58

3.2 AFWD Design .................................................. 60
   3.2.1 The AFWD Designated Database ......................... 61
   3.2.2 The Enrollment Stage (ES) ............................. 62
   3.2.3 The Continuous Authentication Stage (CAS) ............ 64
      3.2.3.1 The AFWD Input Component ....................... 65
      3.2.3.2 The AFWD Process Component ..................... 65
      3.2.3.3 The AFWD Output Component ..................... 70
   3.2.4 AFWD Data Structure .................................. 72
      3.2.4.1 Score Object (SObj) ............................. 72
      3.2.4.2 Input Object ...................................... 73
      3.2.4.3 Output Buffer .................................... 80
      3.2.4.4 Output Object .................................... 80
      3.2.4.5 SLT Calculation Process ......................... 82

3.3 Summary ....................................................... 87

4 AFWD-based Authentication Method for Hand-worn Devices 88

4.1 Study Goals .................................................... 88
4.2 Equipment ...................................................... 89
4.3 Method ........................................................ 91
   4.3.1 Participants ............................................. 92
   4.3.2 Data Acquisition ....................................... 93
4.3.3 Data Processing .............................................. 95
  4.3.3.1 Authentication Scenarios .............................. 96
  4.3.3.2 Analysis Approaches .................................. 97
  4.3.3.3 Classification using WEKA ............................. 101
4.3.4 Classification .............................................. 102
4.4 Results and Discussion ...................................... 104
  4.4.1 Results Based on the Raw Data Approach ............... 104
  4.4.2 Results Based on the Feature Extraction Approach ...... 106
  4.4.3 Discussions and Observations ............................ 108
  4.4.4 Study Limitations ........................................ 112
4.5 Summary .................................................... 113

5 The AFWD Assessment .......................................... 114
  5.1 Study Goals ................................................ 114
  5.2 Subjective Assessment: User Perceptions of the AFWD .. 115
    5.2.1 Overview ............................................... 115
    5.2.2 Equipment and Methodology ............................ 116
      5.2.2.1 AFWD Prototype .................................... 116
      5.2.2.2 Tasks and Security Levels .......................... 119
      5.2.2.3 Participants ....................................... 121
      5.2.2.4 Procedure ......................................... 122
    5.2.3 Results and Discussion ................................ 124
      5.2.3.1 Per Task Assessment ................................. 124
      5.2.3.2 General AFWD Assessment ........................... 130
  5.3 The AFWD Objective Assessment ........................... 135
    5.3.1 Overview .............................................. 135
List of Figures

2.1 First Wearable Shoes ........................................... 9
2.2 The first Pulsar wrist calculator .............................. 10
2.3 Google Glass ..................................................... 11
2.4 List of smartwatches ............................................ 11
2.5 Structure of the LG Vaunt smart glasses .................. 12
2.6 Biometric system enrollment and verification stages ...... 21
2.7 Accelerometers Structure ........................................ 37
2.8 Gyroscopes Structure ........................................... 38
3.1 Device classification method ................................. 54
3.2 Taxonomy Application ............................................ 56
3.3 AFWD location within the CIA triad ......................... 57
3.4 A high-level overview of AFWD ............................... 61
3.5 The structure of the enrollment stage ....................... 63
3.6 The three components of the continues authentication stage .... 64
3.7 The phases of the biometric path ............................. 69
3.8 The UML object representation of the SObj ............... 73
3.9 The UML object representation of the KObj ............... 75
3.10 The comparison process ....................................... 75
3.11 The UML object representation of the TObj ............... 77
3.12 The UML object representation of the BOBJ .......................... 79
3.13 The UML object representation of the output object ............... 81
3.14 The UML representation of the all objects in the AFWD .......... 82

4.1 The LG-W100 smartwatch .............................................. 90
4.2 Accelerometers Structure .............................................. 90
4.3 Gyroscopes Structure .................................................. 91
4.4 The main screen of the Wear app ..................................... 94
4.5 Screen shots that shows the Wear app status ........................ 95

5.1 The AFWD prototype main interface ................................. 117
5.2 Email app interface ..................................................... 117
5.3 Bank app interface ..................................................... 118
5.4 Alarm app interface .................................................... 118
5.5 The interface of the smartphone’s app ................................. 119
5.6 The security levels available in the AFWD prototype ............. 121
5.7 Low security level alert ................................................. 121
5.8 The participant’s opinions toward alarm task ...................... 126
5.9 The participant’s opinions toward email task ..................... 127
5.10 The participant’s opinions toward banking task ................. 128
5.11 A table that shows the numbers of people who chose each option . 129
5.12 Participants’ current authentication methods vs. the AFWD ...... 132
5.13 How participants describe the AFWD ............................... 134
5.14 The WD users by age .................................................. 146
List of Tables

2.1 The timeline of WD history ........................................ 13
2.2 The confusion matrix of a two-class classification. ............. 23
2.3 The evaluation of several biometrics against the seven factors . . 32

3.1 The used weights for token and biometric objects ................. 83
3.2 Example of the allowed tasks based on the SLT value .......... 86

4.1 All sample element descriptions .................................. 98
4.2 The best achieved result in the raw data approach ............... 105
4.3 The best achieved result in the feature extraction data approach . 108

5.1 The usability assessment of the AFWD ......................... 137
5.2 The deployability assessment of the AFWD .................... 139
5.3 The security assessment of the AFWD .......................... 141
### List of Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>AFWD</td>
<td>Authentication Framework for Wearable Devices</td>
</tr>
<tr>
<td>ANN</td>
<td>Artificial Neural Network</td>
</tr>
<tr>
<td>AUC</td>
<td>Area under the ROC curve</td>
</tr>
<tr>
<td>BAC</td>
<td>Balanced accuracy rate</td>
</tr>
<tr>
<td>BObj</td>
<td>Biometric Object</td>
</tr>
<tr>
<td>BSP</td>
<td>biometric-based subprocess</td>
</tr>
<tr>
<td>BW</td>
<td>Body-worn</td>
</tr>
<tr>
<td>CAS</td>
<td>Continuous Authentication Stage</td>
</tr>
<tr>
<td>CIA</td>
<td>Confidentiality, integrity, and availability</td>
</tr>
<tr>
<td>ECG</td>
<td>Electrocardiogram</td>
</tr>
<tr>
<td>EER</td>
<td>Equal error rate</td>
</tr>
<tr>
<td>ES</td>
<td>Enrollment Stage</td>
</tr>
<tr>
<td>FAR</td>
<td>False accept rate</td>
</tr>
<tr>
<td>FIFO</td>
<td>First In First Out</td>
</tr>
<tr>
<td>FN</td>
<td>False negative</td>
</tr>
<tr>
<td>FP</td>
<td>False positive</td>
</tr>
<tr>
<td>FRR</td>
<td>False reject rate</td>
</tr>
<tr>
<td>FTW</td>
<td>first-time wearer</td>
</tr>
<tr>
<td>Glass OTP</td>
<td>Glass One Time Password</td>
</tr>
<tr>
<td>GPS</td>
<td>Global positioning system</td>
</tr>
<tr>
<td>HMD</td>
<td>Head-mounted</td>
</tr>
<tr>
<td>IMEI</td>
<td>International mobile equipment identity</td>
</tr>
<tr>
<td>k-NN</td>
<td>k-Nearest Neighbor</td>
</tr>
<tr>
<td>KObj</td>
<td>Knowledge Object</td>
</tr>
<tr>
<td>KSP</td>
<td>knowledge-based subprocess</td>
</tr>
<tr>
<td>LIFO</td>
<td>Last In First Out</td>
</tr>
<tr>
<td>LR</td>
<td>Logistic Regression</td>
</tr>
<tr>
<td>MAC</td>
<td>media access control address</td>
</tr>
<tr>
<td>MRSs</td>
<td>Motion recording sensors</td>
</tr>
</tbody>
</table>
MSSD  Multiple Data Sources from Single Device
NB    Naive Bayes
NFC   Near field communication
NICA  Non-Intrusive and continuous authentication
NIR   near-infrared
PDA   Personal digital assistant
PIN   Personal identification numbers
QR    Quick Response
ROC   Receiver operating characteristic
RW    returning wearer
SDS   Single Data Source
SLT   Security level tag
SObj  Score Object
SVM   Support Vector Machine
TAR   True acceptance
TBP   touch-based PIN
TN    True negative
TObj  Token Object
TP    True positive
TRR   True rejection rate
TSP   token-based subprocess
VBP   Voice-based PIN
VC    Visual cryptography
WD    Wearable device
WEKA  Waikato Environment for Knowledge Analysis
Acknowledgements

I would like to acknowledge people who have helped to make this dissertation possible. I would like to express my sincere gratitude to my advisor, Dr. Heather Crawford, for her great and invaluable help, support, and guidance throughout this experience. Thanks to my Ph.D. committee members, Dr. William Allen, Dr. Keith Gallagher, and Dr. Katrina Ellis for their comments throughout my journey towards this degree.

I am grateful to all volunteers who participated in the studies that were conducted during this work. I am also grateful for all friends and colleagues for encouragement. Special thanks to Dr. Moahmmed Aljebreen, Dr. Moteeb AlMoteri, Dr. Mehadi Hassen, and Dr. Khaled Slhoub for the thoughtful discussions. Thanks to the government of Saudi Arabia represented by King Saud University for their financial support.

Last, but not least, I owe my deepest gratitude to my mother, for her unconditional love, patience, and unwavering support. Thanks to my brothers, sisters, nephews, nieces, and cousins for their moral support. I am also grateful to my wife Fatimah, who has stood by me and who has been patient while working on this dissertation. To my son Tameem, my daughter Daniah, and my newborn daughter Sama who were a constant reminder that there is life beyond research.
Chapter 1

Introduction

In a 1945 article titled “As We May Think,” Vannevar Bush, an American engineer, introduced his view of an early model of what was called later an augmented-reality device [1]. This device was a head-mounted camera attached to a pair of ordinary glasses that allowed the user to take pictures, record and replay data while their hands are free. Since then, much research has focused on realizing Bush’s view. Augmented-reality is a technology that “overlays a digital interface onto the physical world” [1]. This technology, along with the massive improvement in mobile device technology, has helped to finally have a wearable smart device.

Wearable devices (WDs) are portable devices that extend many smartphone features and that can be worn on and interact with the body. They can gather data and deliver it to their users using several input and output modalities such as screens, heart beat sensor, accelerometer, and gyroscope. It can be worn and taken off since it is not merely attached to the body but is part of the user’s clothing and can be used while its user is moving or busy doing other tasks.

WDs are becoming increasingly popular. There was a 129% increase in WD
sales in 2014, and by end of 2018, there will be more than 250 million WDs in the hands of consumers [2]. Although WDs can be considered an extension of traditional mobile devices (e.g., smartphones), the form factor for each is very different. Mobile devices are used in a “bursty” way, which means people use them frequently but for short periods of time [3]. In contrast, the use pattern of WDs is different: they are always on, always accessible, and always expected to be connected to other devices and services such as smartphones or the Internet [4, 5]. Despite these differences, WDs are often used in conjunction with mobile devices for reasons such as improved processing power as well as additional capabilities such as keyboard and screen and additional memory.

1.1 Research Problem

As WD popularity grows, so does their access to apps and other functionalities that store private information such as email, banking information, and health information. Unauthorized access to the WD may cause interception, interruption, or modification to this private information. Other technologies such as mobile devices often use passwords, biometrics, and personal identification numbers (PIN) to control access to the device and its resources. On WDs, these authentication methods are not viable because the input devices differ; specifically, many WDs do not have keyboards so entering a password or PIN is more difficult. Therefore, we require an authentication method that respects the limitations imposed by the WD’s form factor rather than simply applying possibly unsuitable traditional authentication methods to them.

The problem with WD authentication has several factors that would affect any
proposed solution. The following list summarizes what needs to be considered when designing a WD authentication mechanism:

1. WDs are a new and immature technology that stores sensitive data such as bank and medical information. WDs do not yet have a viable authentication method to protect the confidentiality and integrity of this private data.

2. For WDs to be effective and beneficial, the device needs to be always on and always accessible [5]. While sensitive data needs to be protected, it also has to be always available for authorized users. Therefore, a WD must have an authentication method that respects this unique form factor.

3. WDs have some limitations and diversity compared to any other technologies such as mobile devices. For example, most WDs lack input means such as a keyboard or screen, which limits implementing traditional authentication methods such as passwords or PIN.

4. The users’ willingness to accept the authentication mechanism is a crucial factor. If users are not willing to use an authentication method, they may choose not to activate it and consequently put their sensitive data at risk. Therefore, an authentication method for WDs should be trustworthy and reliable and should require as little effort as possible from users to increase their willingness to use it.

\subsection{Problem Statement}

WDs are a new and immature technology that is known for its uniqueness over other technologies such as mobile devices. They have limitations such as the lack
of keyboard or screen. Also, they have specific form factor as they should be always available and always on. In addition, there is a diversity in WD features and they sometimes deal with very sensitive data. Unfortunately, they do not yet have a viable authentication mechanism that is secure, requires little effort from users, and is respectful of their limitations, diversity, and unique form factor. A framework upon which such an authentication method can be based is needed.

1.2 Our Solution: The Authentication Framework for Wearable Devices (AFWD)

To overcome the issues related to WDs’, we created an Authentication Framework for Wearable Devices (AFWD). The AFWD provides a model to create continuous and transparent authentication methods for WDs. The AFWD exploits the current WD wearer’s biometric pattern to verify that the wearer is also the WD’s owner. The AFWD overcomes the WD’s issues as follows:

- The AFWD supports creating an effective security mechanism for WDs based on sensor data, which can be used in multi-factor and multi-modal biometric authentication.

- The AFWD respects the unique form factor of WDs considering that they should be always available and always on by authenticating wearers transparently as the wearer goes about their regular tasks.

- The AFWD respects the limitations of WDs such as the absences of input means by using factors such as behavioral biometrics.
• The AFDW-based authentication is acceptable by WD’s owners’ methods because it depends on factors such as behavioral biometrics that do not require much effort from the wearer compared to entering a password or a PIN.

1.2.1 Research Question

The following research question defines the research in this dissertation: In terms of error rate, how trustworthy is an AFWD-based authentication method in identifying whether or not the current WD’s wearer is its owner in a transparent and continuous way? Does this method minimize the wearer’s effort and respect the form factor, the diversity, and the limitations of the WD?

1.2.2 Research Hypotheses

The following hypotheses are drawn from the research question:

H1: It is possible to create an authentication method for WDs that is transparent, continuous, trustworthy, and respectful of their unique form factor and limitations.

H2: WD sensors, such as an accelerometer and gyroscope, have low error rates to support determining whether or not the current wearer of the WD is its owner.

H3: Wearers feel that an AFWD-based authenticating method is secure and prefer to use it, if available, over their current authentication method to protect their private data.
1.2.3 Null Hypothesis

The null hypotheses of this research are the following:

**H01:** It is not possible to create an authentication method for WDs that is transparent, continuous, trustworthy, and respectful of their unique form factor and limitations.

**H02:** WD sensors, such as an accelerometer and gyroscope, do not have low error rates to support determining whether or not the current wearer of the WD is its owner.

**H03:** Wearers would not feel that an AFWD-based authentication method is secure and would not prefer it, if available, over their current authentication method to protect their private data.

1.3 Dissertation Structure

This dissertation is organized as follows: Chapter 2 goes into in detail about the background that is needed to better understand the content of this dissertation. In Chapter 3, we present our proposed solution: the Authentication Framework for Wearable Devices (AFWD). Then, in Chapter 4 we present the AFWD-based Authentication Method for Hand-worn Devices study, which explores the feasibility of using WD sensor data in the authentication. In Chapter 5, we evaluate the AFWD from both subjective and objective points of view. This includes a study to assess the AFWD based on user perceptions and an assessment based on existing evaluation metrics that measure usability, deployability, and security. Finally, our research conclusions are drawn in Chapter 6.
Chapter 2

Background

2.1 Overview

The idea of wearable technology goes back decades to when Vannevar Bush discussed his view of an early model of a WD, which was a head-mounted camera connected to glasses. It performed tasks such as taking pictures and playing music [1]. Since then, research has been conducted to characterize and investigate WD-related challenges and opportunities [6], developing a theoretical framework for WDs and designing practical models of WDs [7], and inventing protection methods for their sensitive data [8, 9]. For example, in 1998, Steve Mann, who is recognized as the father of wearable computing, provided what he called operational modes of WearComp (his term to refer to wearable computers): constancy, augmentation, and mediation [7, 10]. Constancy means that the WD should be always on and ready to be used. There should be no turning on or warming up process as with other kinds of computers. Augmentation means that the device should support what a computer does, but not to take its role. Also, it should
be working at the same time its user is doing something else. Mediation means that the WD can “encapsulate” users. This means that the WD should have a comprehensive understanding of what users do and do not want. For example, a WD should be able to filter the data users are getting if it is offensive or simply unwanted.

In 2003, as the technology started to evolve, Bass defined WDs using five characteristics [5]:

- It can be used while its user is moving.
- It can be used even if its user is busy using their hands for another task.
- It should be worn and taken off and not “merely” attached to the body, but it is part of the user’s clothing.
- It can be controlled by its user.
- It should be continuously available and turned on all the time.

Mann was more general in his definition while Bass was more specific such as when he talks about using the WD during movement or how it should be worn. The two major things that both Mann and Bass agreed on are that a WD should be always available to its user and should be working at the same time its user doing something else. In general, just like mobile devices, a WD is a portable computer, but it can be worn on and interact with the user’s body. It may have sensors that mobile devices and laptops do not have, such as the heart rate sensors in some smart watches. It can take many forms, such as glasses, a watch, headset, or wrist band [6].
2.2 Wearable Devices in History

The first WD was used inside a pair of shoes to cheat at roulette in the 1960s and was invented by professor Edward Thorp at Massachusetts Institute of Technology (MIT) [11, 12] (see Figure 2.1). Figure 2.2 shows the first wrist calculator, which was released by Hamilton Watch company in 1975. It was called Pulsar and it showed six digits on its electronic screen and did basic arithmetic operations such as addition, subtraction, division, and multiplication [13].

![Figure 2.1: The shoes Edward Thorp invented and used to cheat at roulette in the 1960s](image)

In 1977, a device with a camera was mounted to a blind person’s head to take pictures [6]. The device converted pictures taken to a tactile grid to allow blind people to feel pictures so they could determine what they look like and also allow them to read text. The device was mounted on a vest that the user wore. In 1981, Steve Mann designed a head-mounted general purpose camera. In the same year, Mann also designed a computer that could take pictures, and text, in addition to other multimedia abilities, and could be worn as a backpack and helmet. The helmet was used as a monitor.

In 1994, Steve Mann invented a wearable wireless webcam that was able to
The first Pulsar wrist calculator released in 1975; it showed six digits on its electronic screen [13].

upload images to the Web [11]. It was also able to stream live video from and to the Web. The attention to wearable technology increased so that in 1997, the first International Symposium on Intelligent Wearable Computers (ISWC) was held in Cambridge, MA, USA.

In 2006, Nike and Apple created the sport kit Nike+iPod to record its users sport activities and sync to an iPod [11]. In 2009, the W200 Wearable computer was released [11]. It was designed for those who are in an emergency situations such as doctors. It was a wrist-worn device that that have fast and large access to large amounts of data such as patient information while doctors are busy.

In 2013, several smartwatches were released including the Pebble which was manufactured by Foxlink Group, and Samsung Galaxy Gear, which was manufactured by Samsung Electronics. In the same year, Google released its smartglasses Google Glass for developers., which is shown in Figure 2.3 [6, 11].

In 2015, the Apple Watch was released the most advanced operating system for watches at that time [15]. After that, the competition between big manufacturers led to yearly releases of new or updated versions of smartwatches and smartglasses.
Figure 2.3: Google Glass was released in 2013 for developers [14].

Figure 2.4: Smartwatches from left to right: Samsung Gear S3 (image from www.samsung.com), Apple Watch Series 3 (image from www.apple.com), and LG Watch Style (image from www.lg.com).

[15]. For example, in 2016, Samsung corporation released its Samsung Gear S3 while Apple released its Watch 2. In 2017, several new WDs were released such as LG Watch Style, Apple Watch Series 3, and Huawei Watch 2, with the faster processor, longer battery life, and bigger storage (Figure 2.4 shows some of these smartwatches). In 2018, a prototype of LG Vaunt smart glasses was released [16]. The LG Vaunt looks different than previous smart glasses, such as Google Glass.
because it does not have a camera nor touchpad (see Figure 2.5). The information is projected directly onto the human retina. It looks more like a traditional pair of non-smart glasses and is meant to be more for simple notification.

Figure 2.5: Two pictures that show the structure of the LG Vaunt smart glasses. The LG Vaunt looks just like normal glasses on the face. It does not have a camera or touchpad. Images source: [16].

The timeline of the WD history is shown in Table 2.1 [6, 11]. The list in the table is not meant to be comprehensive but it is a representation of the WD’s progression throughout history. Throughout the history of WDs, there have been changes in several aspects such as the functionality, the size, the features, and the class (hand-worn, head-mounted, etc.) of the WD. It started with a wearable device worn on the foot that did simple mathematical calculations to cheat at roulette. Then the hand-worn and head-mounted types introduced different functionalities such as texting and taking pictures. Manufacturers exploited new technologies, such as Wi-Fi and Bluetooth, and embedded them in WDs, which allowed wearable devices to interact with other devices and connect to the Internet. Sensors such as an accelerometer and gyroscope made a very important improvement when
<table>
<thead>
<tr>
<th>Year</th>
<th>Event</th>
</tr>
</thead>
<tbody>
<tr>
<td>1966</td>
<td>Edward Thorp invented pair of shoes to cheat at roulette</td>
</tr>
<tr>
<td>1975</td>
<td>First wrist calculator was released</td>
</tr>
<tr>
<td>1977</td>
<td>A vest connected to a camera that took pictures and converted them to a tactile grid for blind people was invented</td>
</tr>
<tr>
<td>1981</td>
<td>Steve Mann designed a head-mounted general purpose camera and a computer with a camera carried in a backpack to take pictures</td>
</tr>
<tr>
<td>1994</td>
<td>Steve Mann invented a wearable wireless webcam that uploads images and streams live video on the Web</td>
</tr>
<tr>
<td>2006</td>
<td>Nike and Apple created the sport kit Nike+iPod</td>
</tr>
<tr>
<td>2009</td>
<td>The W200 Wearable computer was released</td>
</tr>
<tr>
<td>2013</td>
<td>Pebble and Samsung Galaxy smartwatches were released</td>
</tr>
<tr>
<td>2013</td>
<td>Google Glass was released to developers</td>
</tr>
<tr>
<td>2015</td>
<td>Apple Watch was released</td>
</tr>
<tr>
<td>2016</td>
<td>Samsung Gear S3 and Apple Watch 2 were released</td>
</tr>
<tr>
<td>2017</td>
<td>LG Watch Style, Apple Watch Series 3, and Huawei Watch 2, with improved features were released</td>
</tr>
<tr>
<td>2018</td>
<td>LG launched Vaunt smart glasses</td>
</tr>
</tbody>
</table>
they were embedded in WDs as they helped to enable new functionalities such as measuring human activities. Another aspect is the size of wearable devices as they moved from a wearable computer that was carried as a backpack with a helmet to little smart glasses with the same, or perhaps better, abilities.

2.3 Wearable Device Classifications

Jiang et al. provided two classifications for WDs [6]. First, the form-based standard classifies WDs based on the way they are worn such as head-mounted, body-dressed, hand-worn, and foot-worn. An example of the head-mounted is smart glass and example of hand-worn is smart watch. Second, the function-based classifications which groups WDs based on their function or the field they are used in. Examples of this class are wristbands that are used for healthy living, or smartwatches that work as personal digital assistants (PDA).

2.4 Mobile Devices and Wearable Devices Crossover

WD technology may be considered an extension of traditional mobile devices. WD and mobile devices (e.g., smartphones) have several features in common, but they also differ in others. In the following list, some of the characteristics of mobile and WDs are discussed to show the differences and similarities:

- **Portability:** Portability is a main feature in both WDs and mobile devices. However, each one has its own way of being portable. WDs are attached to the body (worn) while mobile devices are carried. For example, mobile devices can be hand-carried or pocket-carried while WDs are portable as
hand-worn or head-worn. WDs can be seen as more portable as they are worn any time and everywhere, sometime even in bed to monitor sleeping habits, while mobile devices are picked up and put down frequently. This means WDs are capable of gathering more real time data compared to mobile devices. In addition, the nature of WD portability suggests it is a single user device given that it is part of the user’s clothing and therefore is unlikely to be shared with others. In contrast, mobile devices can be easily shared with others and may be used by multiple users over time.

- **Display methods:** Information is displayed differently in both devices. In mobile devices, information is displayed on a screen that is an integral part of the structure of the device. WDs may or may not have a screen. For example, a smartwatches and smart glasses have a screen while some fitness wristbands rely on other devices (maybe another mobile device) to display their content. Mobile devices and WDs are also different in terms of the screen size. In mobile devices, the screen is bigger, which consequently gives more space to show more information. In contrast, screens in WDs were designed to display a limited amount of information. Data that is shown in WDs is real-time data (because they should be always turned on and they should be always gathering data), while with mobile devices, the user sees data only while using the device. The difference in displaying data might also mean that users will be notified differently. With a smartphone, users might be notified in multiple ways such as by an banner that appears on the top of the screen or by an alert. Given the bursty use pattern of smartphones [3], the user probably will not notice this notification until picking up the smartphone and looking at it. In contrast, a WD’s users are notified almost
instantly given its use pattern such as being always on. For example, a notification might pop up on the screen which is always visible to the user.

- **Input methods:** In mobile devices, there are several ways to enter data to the device such as keyboards of all kinds (touch and physical keyboards). The microphone is used to input audio data and the camera is used to input still images and video. Recently, some mobile devices have had biometric sensors, such as a fingerprint readers, which can also be considered an input method. However, many WDs lack a keyboard which, makes entering text to the device limited. Some WDs such as smart glasses have a touchpad that can sense gestures such as tap and swipe that can be considered an input method. Like mobile devices, the microphone is a good source of audio input in WDs as well as the camera to input still images and video. In general, WDs sensors can be considered as input methods. For example, the hand-worn ones have a built-in heart rate sensor that can be considered an input source.

- **Sensors:** Mobile and WDs share sensors such as an accelerometer and gyroscope. However, WDs, like some smartwatches, have built-in heart rate sensors that mobile devices do not have. Also, some sensors, such as fingerprint sensors, are currently available only in mobile devices.

- **Authentication methods:** Mobile devices have multiple authentication methods such as PIN, password, graphical password, or biometrics. In contrast, the nature of WDs narrows down the authentication method options. For example, the lack of a keyboard will hinder having some authentication methods like password and personal identification number (PIN).
- **Connectivity:** WDs and mobile devices share the same connectivity technologies such as Wi-Fi, Bluetooth, and Near Field Communication (NFC). However, most WDs lack mobile network connectivity, such as 3G or 4G networks, but they can get them by being tethered to a smartphone.

### 2.5 User Authentication

User authentication is an important mechanism that protects the system by differentiating between authorized and unauthorized users. Authentication is used to verify the identity of a user who wants to access a system.

There are two different types of authentication, identification and verification:

**Identification:** “is the act of asserting who a person is.” [17, pp. 38]. Identification is one-to-many matching in which there is no claim of identity where we compare one person is identity against several people in the database to tell who that person is.

**Verification:** “is the act of proving that asserted identity: that the person is who she says she is.” [17, pp. 38]. When a user wants to be granted access to the protected resources, this process verifies whether this user is the one who they claim to be. Verification is considered one-to-one matching in which users claim an identity by providing evidence for that identity via a user name, ID card, etc. The system compares what the user provided to the stored one in the database.

Authentication methods are broken into three factors [18]:

17
**Something they know:** which is also called a knowledge-based authentication method in which the system authenticates users based on shared secrets between them and the system, such as a passcode or a password.

**Something they have:** which is also called token-based authentication in which users need to possess a token in order to authenticate themselves. Examples of tokens are an ATM card or smartcard token.

**Something they are:** or a biometric-based authentication method, which authenticates the user based on behavioral or physiological characteristics such as a fingerprint, iris, or voice.

These factors can be combined to make two different types of authentications: multi-factor authentication and multi-modal biometric authentication. Multi-factor is based on a combination of two or more of the three authentication factors. An example of this authentication is using a credit card, which represents something the users have, and the PIN which represents something the users know. Multi-modal biometric authentication is based on a combination of two or more biometrics, such combining fingerprint with retina. These multi-factor and multi-modal biometric authentications are known to provide improved security level [19].
2.5.1 Textual Passwords and Personal Identification Numbers

A password is a sequence of characters that is used to grant or deny individual access to a system. A PIN is an authentication method that uses four or more digits for authentication. Passwords and PINs are knowledge-based authentication methods that depend on the user’s memory. Both PINs and passwords are easy to create and remember, which make them common methods for authentication [20]. However, with the improvement of technology, PIN and simple password have some drawbacks. They can be easily guessed using a brute force attack, which is where an attacker tries all combinations of letters, numbers, and special characters until the password is found [17, pp. 42]. For example, a 4-digit PIN that consists of numbers only (0-9), the total number of possible PINs is $10^4$ which is 10000. Assuming that we can try 1000 PINs per second, a 4-digit PIN can be cracked in 10 seconds using a brute force attack.

Also, passwords suffer from several problems such as difficulty in memorizing a complicated password or multiple passwords [21]. This becomes complicated if average user is required to memorize about 25 passwords for all accounts they have [15]. Therefore, users tend to choose short and easy to remember passwords or even popular phrases [21, 22]. Additionally, passwords and PINs cannot assure the user’s identity because anyone who knows the secret gains access. The possibility of implementing this kind of method is low since most WDs do not have the traditional inputs means such as a keyboard that is used in mobile devices.
2.5.2 Graphical Passwords

A graphical password is an authentication method that depends on selecting images or drawing sketches instead of entering characters. The user is then required to recognize that selection during authentication [23]. Graphical passwords is an alternative knowledge-based authentication method designed to overcome the human factor effects in the password method such as memorizing a complicated password [24]. Although a graphical password can mitigate some password weaknesses, it is at risk of smudge attacks where an attacker analyzes the residual oils (smudges) left by the device owner’s fingers to determine the password [25]. There are different types of graphical password authentication: recognition-based, pure recall-based, and cued recall-based [18, 26, 27]. A recognition-based depends on whether or not the user is able to recognize an image they have seen before. When a user first enrolls into the system, a set of images are chosen. During the authentication stage, the selected images are shown along with other images that is called “detractor” images. The user will not gain access to the device until choosing exactly the same images that were chosen during the enrollment stage. In a pure recall-based, users draw a sketch or a picture on a grid during the enrollment stage. In the authentication stage, users are asked to redraw that secret sketch. The cued recall-based mechanism is based on the same idea in the pure recall-based mechanism but in this case a user is given a hint (like a hot spot or gestures on an image background) that will help in remembering the password. This will help to overcome the memory weakness issue for some people.
2.5.3 Biometric Authentication

In a biometric-based authentication methods, a human’s unique biological characteristics or traits (both physiological, such as fingerprint or behavioral, such as gait) are used to identify users and accept or reject their access to a specific system [28, pp. 2-3][29]. Biometrics are considered to be an effective authentication method compared to the other methods such as passwords [17, pp. 53] because it is hard for biometric traits to be shared, lost, stolen, or forgotten. However, they often require additional hardware and user collusion. The general biometric system includes two processes: enrollment and the verification processes.

![Figure 2.6: The enrollment and verification stages in a general biometric system. Adapted from [28, pp. 6]](image)

The flow of the enrollment process includes three phases: data acquisition, data processing, and data storage, as shown in Figure 2.6. During the enrollment process, users are asked to provide raw data, e.g., biometric samples. The raw biometrics are first processed in the processing stage where salient and distinctive features are extracted and then represented digitally to make an enrolled template or biometric template. The biometric template is then stored in the system database.
The verification process includes three stages: data acquisition, data processing, and data comparison, as shown in Figure 2.6. During this process, users are asked to scan a trait and then the distinctive features are extracted and represented digitally to create what is called a biometric query. During the comparison stage, the biometric query is compared to the biometric template that was stored during enrollment, to make the authentication decision. The result of the comparison process is the matching score, which represents the amount of similarity between the biometric template and the biometric query. The higher the similarity value, the closer the match between the biometric query and template.

2.5.3.1 Biometric Performance Metrics

The accuracy and performance of biometric systems can be measured and evaluated with methods such as confusion matrix, false acceptance rate (FAR), false rejection rate (FRR), equal error rate (EER), and Area Under the ROC Curve (AUC) [28, pp. 24][30].

- **Confusion Matrix:** is a method that shows pattern classifier performance and is often used for binary classification. A confusion matrix shows the class predictions (correct and incorrect) compared to the actual class [31]. The results of binary classification can be: True positive (TP), False positive (FP), True Negative (TN), and False Negative (FN). TP represents the number of positive instances that were correctly predicted. FP represents the number of negative instances that were incorrectly predicted as positive. TN represents the number of negative instances that were correctly predicted. FN represents the number of positive instances that were incorrectly predicted as negative. For example, in Table 2.2, we have a 2x2 matrix which represents
two classes: positive and negative. The diagonal values from the left to right, top to bottom of the matrix represent the correctly classified instances while the other values represent the misclassified instances.

- **FAR**: the probability of providing access to an unauthorized person. A high value of FAR means the system is at higher risk since it means allowing illegitimate users to access protected resources when it is supposed to keep them out. FAR is calculated as follows [32]:

\[
FAR = \frac{FP}{FN + TN}
\]  

(2.1)

where FP is number of false positives or the number of unauthorized users accepted, FN is the total number of false negative samples, and TN is the total number of true negative samples.

- **FRR**: the probability that the biometric system will deny access to an authorized user. The FRR is more related to user experience than security as it is not pleasant to frequently ask authorized users to try authenticating again. FRR is calculated as follows [32]:

\[
FRR = \frac{FN}{TP + FN}
\]  

(2.2)
where FN is number of false negative, TP is the total number of true positive samples, and FN is the total number of false negative samples.

- **EER**: the point where FAR and FRR are equal, and the lower the value of EER (near 0% is better), the more accurate the biometric authentication system is. EER is valuable as it helps when comparing multiple systems under the same conditions such as experimental and acquisition conditions. It is also used to compare the same system under several conditions.

- **Receiver Operating Characteristic (ROC) curve**: a graphical representation based on the relationship between the FAR and true positive rate (TPR) which is the probability of positive samples that are correctly identified as such. It is created by plotting TPR against FAR at various threshold settings. The ROC curve helps to evaluate and compare the biometric systems’ implementation. In addition, it also used to evaluate multiple biometric systems that have the same experimental conditions.

- **Area Under the ROC Curve (AUC)**: the area under the receiver operating characteristic (ROC) curve. It represents the probability of a classifier ranking the random positive samples higher than the random negative samples [33]. The value of AUC ranges from 0.5 to 1 where 0.5 means a random classifier and 1 means the classifier is at perfect performance.

The European Standard for Access Control provides recommended acceptable FAR and FRR values for authentication purposes: FAR is recommended to be less than 0.001% while FRR should be less than 1% [34]. The reason that FAR is much lower than FRR because the FAR represents an attacker gaining access to the protected system. On the other hand, FRR is more related to usability
as authorized users are required to reauthenticate when misclassification happens. This does not generally put protected data at risk.

2.5.3.2 Physiological Biometrics

A physiological biometrics depend on “an anatomical or physiological characteristic rather than a learned behavior” [29]. Examples of these characteristics or traits are fingerprint, face, hand geometry, ear shape, and iris. The applicability of these physiological biometrics in WD authentication depends on the availability of the sensors or scanners that can collect them. For example, a WD cannot implement a fingerprint-based authentication method if it does not have a fingerprint scanner.

- **Fingerprint**: The human finger has a “a flow-like pattern of ridges and valleys”, which are also called friction ridges, that are unique in every finger [28, pp. 51]. These friction ridges are processed into a mathematical representation that is based on the extracted features and the set of points that describe the minutiae. Minutiae are the points of ridge bifurcations and endings [35]. The feature extraction process could be affected by the resolution of the raw fingerprint image and the clarity of the friction ridges themselves [28, pp. 51]. Fingerprint-based authentication has been implemented in several smartphones such as Apple products and Samsung products [36, 37]. In WDs, however, this technology is not widely implemented but there has been increasing interest in integrating fingerprint sensors into smartwatches. For example, Samsung has recently filed patents that show their interest in adding new features, one of which is a fingerprint sensor [38].

It is possible to manipulate the recognition system by making a fake fingerprint. For example, the iPhone fingerprint sensor was attacked by using a
cloned fingerprint lifted from shiny surface and regenerated using a glue to gain unauthorized access to the device [39]. Having a fingerprint cloned or copied makes it dangerous to use this kind of biometric trait because the chance of a fingerprint being changed over time is very low, and once it is copied, it is vulnerable [35].

- **Face**: The face is the area from the left ear to the right ear and from the forehead to the chin. The nose, mouth, eyes, and cheeks are the main parts of the face. The idea of face recognition in biometric systems depends on detecting facial characteristics such as shape, skin color, nose, eyes, marks, or moles [28, pp. 97]. Based on the sensor used in the face recognition biometric system, different format images can be acquired such as 2D, 3D, and video. The 3D images reveal more depth information compared to 2D images and therefore improved facial recognition. The video acquisition happens by taking sequence of images which enables choosing a good quality image.

- **Iris**: Iris recognition systems work by analyzing the pattern of the area between the pupil and the white (sclera) of the eye. The uniqueness of the iris structure among people is very high, even between identical twins [28, pp. 141]. As a biometric trait, the iris is acquired by taking its image in the near-infrared (NIR) spectrum to illuminate the iris and to provide more textural details from the iris surface. After taking the image of the eye, the iris is segmented and localized from the rest of the eye. Then, the segmented iris region is transformed, normalized, and encoded to create a binary template of the distinctive feature that is used for comparison. An Iris is a good source for authentication because of its uniqueness and it is hard to copy it. Iris
has not been commercially implemented in mobile or WDs. In 2015, Fujitsu revealed an Iris-based authentication method in a prototype of a smartphone under development. It uses an infrared LED to illuminate the iris and an infrared camera to capture the iris [40].

- **Veins**: Yanagawa et al. [41] have shown that the human vein patterns, which are the blood vessels in the finger and hands, patterns are unique among individuals, even twins. This suggests that vein patterns can be used to identify people. A camera with near-infrared light is used to take an image, which is then normalized. Then, the distinctive features such as the vein’s location, starting point, distribution, bifurcation points, and ending points are extracted and then encoded to be stored digitally in the database [42]. Like some other biometrics, veins are also liable to change due to aging or health-related matters. The possibility of using veins is high in WDs, especially hand-worn devices, such as smartwatches or wrist bands because they are attach to the hand where the veins are located.

- **Retina**: Retina-based authentication methods depend on identifying individuals based on the unique structure of the network of blood vessels in the human retina. Retina recognition requires a user’s cooperation as they need to concentrate on a particular spot in the visual field to take the image [28, pp. 34]. Also, careful scanning is required in which an alignment is needed. All this may affect user acceptance. Retina scanning may show some underlying medical issues which may also affect user acceptance [43].
2.5.3.3 Behavioral Biometrics:

Behavioral traits are “learned and acquired over time rather than based upon biology.” [29]. Examples of behavioral traits are signature, gait, voice, and keystroke dynamics. In this case, there is a need for some sort of activity, such as keyboard typing from the user in order to gather some behavioral features. Behavioral biometrics are helpful in transparent authentication because they can be gathered in the background and no explicit user participation is needed.

- **Voice**: after acquiring the sound from the user in the enrollment stage, noise needs to be removed from the voice first; then the voice is converted to a digital format. After that, the voice features are extracted and the biometric template is created [29]. Voice can be classified as both a physiological and behavioral biometric [28, pp. 33][29]. The physiological characteristics of the voice depend on some aspects such as the shape and size of the mouth and vocal cords [29]. For example, the voice of a person will change if the shape of the mouth has changed. The behavioral aspect also exists in the voice biometric trait. For example, the tone, the volume of the voice, and the rhythm [29].

Many modern mobile devices and WDs have a microphone, which gives the ability to implement the voice authentication method. However, voice biometric is not widely used as an authentication method because voice is not a very unique trait for each individual [28, pp. 33][29]. In addition, voice is affected by human conditions such sicknesses and age.

- **Contextual biometrics**: Dey and Abowd defined context as “any information that can be used to characterize the situation of an entity [44]. An
entity is a person, place, or object that is considered relevant to the interaction between a user and an application, including the user and applications themselves.” [44]. Contextual biometrics are the side information that are related to the interaction between the smart devices and users. For example, knowing that a specific user should be at a particular location at a particular time can be exploited to improve the authentication process if we knew that no one else could be at that location at the same time. Other examples of contextual information are: conversation, gestures, and ambient sound.

Although contextual biometrics are not unique by nature, they would be very helpful and very applicable in WD authentication. WDs are not like any other technologies since they are attached to people wherever they go. This gives the ability to collect a large amounts of data about the environment surrounding users that can be used in authentication.

- **Heartbeat**: Based on some studies, the heartbeat of a human is considered to be unique once the individual reaches the age of adolescence [45, 46]. Heartbeat data is acquired based on an electrocardiogram (ECG) which captures the heart’s electrical signals. The ECG is different among people based on several factors, such as the size and anatomy of the heart and the chest size. Feature extraction in heartbeat-based authentication depends mainly on depolarization and repolarization phases. Depolarization and repolarization represent the waves reported from the heart parts, such as atrial and ventricular through the ECG.

- **Gait** Gait-based authentication is the method of identifying individuals based on the way they walk. Gait recognition requires users to be walk-
ing in order to authenticate them, which may have some limitations, for example, they may not like to stand up and walk every time they need to access their devices. Thus, it is preferable to have gait recognition used with other authentication methods [47]. According to [47] and [48], there are three different ways of gathering data for gait recognition:

**Machine Vision Based:** This method uses a camera to capture individuals while walking. The relevant features, such as gait silhouettes, are extracted from the captured raw biometrics after applying some segmentation, machine learning, and image processing techniques [49]. In this case, individuals do not need to have direct and conscious interaction with any sensor like in fingerprint recognition, for example. The person’s walking style is captured from a distance. This approach is also called motion based [50].

**Floor Sensor Based:** The raw gait biometric is captured by sensors that are placed on the floor. The authentication when using these sensors is based on features like the force and the pressure on the sensors. Individuals are authenticated just by walking on these sensors.

**Wearable Sensor Based:** the sensors that are used in this approach have to be carried or worn by individuals in order to collect gait recognition data. Sensors like an accelerometer, gyroscope, and force sensors can be used to captured gait biometric data. Examples of the features that can be extracted from gait data are gait cycles and maximum amplitudes. Many WDs are embedded with types of sensors that are able to collect gait data, which makes it a potential WD authentication
To have a valid biometric trait, the following requirements should be met [28, pp. 29][51]:

- **Universality**: This means most of the biometric system users should have the trait. If the user does not have the trait or if the system is not able to extract the trait feature, the non-universality problem will happen, which means that not all users are able to enroll into the system.

- **Uniqueness** or distinctiveness: Every trait should be unique enough among all other people. Two persons should not have the same trait. In fact, they should be different enough to be distinguished based on their traits.

- **Permanence** The trait should not change, or change slowly over time. If it changes over time, the same person using the same trait will be rejected by the system after some time, causing a false reject. This may require re-enrollment.

- **Measurability**: This means that the features extracted from the trait can be measured in order to be able to use them in comparison during the authentication mechanism.

- **Acceptability**: This refers to how willing people are to accept a biometric as an identifier and are willing to have their traits acquired and measured.

- **Circumvention**: This represents how easy an authentication method that is based on a particular biometric can be fooled so it grants access to an unauthorized user.
Table 2.3: The evaluation of several biometrics against the seven factors. H refers to high, M refers to Medium, and L refers to low. Modified from [52].

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Face</td>
<td>H</td>
<td>L</td>
<td>M</td>
<td>H</td>
<td>L</td>
<td>H</td>
<td>H</td>
</tr>
<tr>
<td>Fingerprint</td>
<td>M</td>
<td>H</td>
<td>H</td>
<td>M</td>
<td>H</td>
<td>M</td>
<td>M</td>
</tr>
<tr>
<td>Gait</td>
<td>M</td>
<td>L</td>
<td>L</td>
<td>H</td>
<td>L</td>
<td>H</td>
<td>M</td>
</tr>
<tr>
<td>Hand Vein</td>
<td>M</td>
<td>M</td>
<td>M</td>
<td>M</td>
<td>M</td>
<td>M</td>
<td>L</td>
</tr>
<tr>
<td>Iris</td>
<td>H</td>
<td>H</td>
<td>H</td>
<td>M</td>
<td>H</td>
<td>L</td>
<td>L</td>
</tr>
<tr>
<td>Retina</td>
<td>H</td>
<td>H</td>
<td>M</td>
<td>L</td>
<td>H</td>
<td>L</td>
<td>L</td>
</tr>
<tr>
<td>Voice</td>
<td>M</td>
<td>L</td>
<td>L</td>
<td>M</td>
<td>L</td>
<td>H</td>
<td>H</td>
</tr>
<tr>
<td>Heartbeat</td>
<td>H</td>
<td>M</td>
<td>H</td>
<td>M</td>
<td>M</td>
<td>H</td>
<td>L</td>
</tr>
</tbody>
</table>

In Table 2.3, an evaluation of the biometrics discussed above against the seven factors is presented [52].

2.5.4 Explicit, Transparent, and Continuous Authentication

Some authentication methods currently implemented in mobile devices and laptops are considered static or explicit authentication, in which users are required to explicitly interact with the system to authenticate themselves. Examples of explicit authentication include but are not limited to password and PIN. In addition, explicit authentication is considered one-time proof of identity in a session or point-of-entry authentication [53, 54]. After this point-of-entry, the resources are considered equal in terms of sensitivity as further authentication is not required [54].

In contrast, in transparent and continuous authentication, the user’s efforts are minimized and are not explicitly required. For example, users can be authenticated
based on their voice, walking, or heartbeats, which smart devices can gather implicitly without user interaction. Unlike explicit authentication, transparent and continuous authentication does not happen only once a session but throughout the whole session. In the transparent and continuous authentication scheme, the system continuously and transparently makes sure that the user is still the same one who logged into the system initially [55]. Some research has shown that behavioral biometrics such as gait [47] and keystroke dynamics [56] can be used in transparent authentication. Behavioral biometrics are efficient for transparent authentication because they can be gathered while the user goes about other tasks and do not require explicit user participation.

As discussed in Section 2.5.3, the authentication decision in a biometric-based authentication is not a binary yes or no answer like in knowledge-based authentication. Rather, it depends on the level of similarity between the biometric template and the biometric query. In transparent and continuous authentication, the similarity level represents how much trust or confidence a system has that the identity claimers were who they claimed to be. The higher the similarity level, the higher the confidence that the user is legitimate. Therefore, in continuous and transparent authentication systems, a confidence level threshold is set where the authentication decision is made based upon it. For example, if the confidence level is above or equal to this threshold, a user who claimed the identity is considered legitimate. Similarly, the confidence level is below this threshold, a user who claimed the identity becomes illegitimate [47, 56].

The idea of confidence level has been seen in the transparent and continuous authentication research for years [47, 54, 56]. Derawi [47] and Crawford [56] discussed the confidence level and how it can be implemented in transparent biometric-based
authentication. They both suggested that this level initially starts at a low level and then increases or decreases every time new biometric traits are collected, analyzed, and compared against the biometric template. If the similarity is high, the confidence or trust level increases and vice versa. This confidence level becomes at its highest (e.g., 100) after an explicit authentication is performed. The confidence or trust level gives more flexibility to transparent and continuous authentication compared to the explicit authentication. This is because, unlike explicit authentication, tasks such as texting, taking pictures, and viewing email can be associated with thresholds that reflect how sensitive they are and how much security they require. For example, highly sensitive resources should require a high threshold value while less sensitive resources will require a low threshold value.

To the best of our knowledge, there is no authentication framework specifically for WDs yet. However, some researchers have created similar frameworks for mobile devices. Clarke et al. [57] described an authentication framework called the Non-Intrusive and Continuous Authentication (NICA). NICA is an biometric-based authentication framework for mobile devices that verifies the user identity in a transparent and continuous way. The NICA is meant to be more user friendly as it reduces user annoyance found in knowledge-based authentication. In addition, it is intended to provide a security to the mobile device throughout its use time as compared to the point of entry protection that the knowledge-based system provides. The NICA uses the confidence levels idea in which there is no access provided to the protected sensitive data if the level is low. In contrast, an access to the sensitive data is provided without explicit user interaction, such as entering a password, if the confidence level is high. Clarke et al. evaluated NICA in a study that involved 27 people. From the perspective of 92% of the participants, NICA
provides a more secure environment compared to what they already have [57].

Crawford, Renaud, and Storer developed a continuous and transparent authentication framework for mobile devices [56]. The framework aims to minimize users’ effort during the authentication process by using behavioral biometrics to identify the device owner. In addition, it aims to address some of the issues related to traditional authentication methods, such as memorability. The authentication in this framework is transparent since it does not require explicit interaction from the users and it is continuous since it is not performed on-demand but dynamically in the background. The framework uses device confidence level that represents how confident the system is that the current user is the device owner. The higher this level is, the more confident the device is that it is being used by its owner. Users of this framework are able to use explicit authentication, such as a password as a backup authentication in order to, for example, elevate the confidence level temporarily.

Crawford et al. also evaluated the framework in terms of security and usability [56]. The evaluation shows that, with transparent authentication, users were able to do all tasks while having about 67% less explicit authentication requests compared to when they do not have the transparent authentication. Additionally, the evaluation shows that the confidence level decreased relatively fast once attackers start using the device, which consequently led to denying their access.

2.6 Authentication in Wearable Devices

In mobile devices, the resources and the features of the device provides multiple options of authentication methods. Features like touchscreen, keyboard, Global
Positioning System (GPS), accelerometer, and fingerprint sensor are very helpful in providing a variety of authentication methods. The keyboard, for example, can be used with passwords, keystroke dynamics, or PINs as an authentication method.

The nature of WDs makes authentication resources limited compared to mobile devices. Many WDs do not have a keyboard, which makes it hard to implement authentication methods like passwords or PINs [8]. However, wearable devices have sensors such as accelerometer and gyroscope which can collect data about the unique features of human movement. This kind of sensors can be used to improve not only security but usability as well. This is because the information needed for authentication is collected transparently without having the user manually enter a password, for example.

Despite WDs limitations, they have the potential to adopt the three most common authentication factors: knowledge-based, token-based, and biometric-based as we will see in the next section. However, the applicability of authentication factors depends on the available sensors in the WD. In the following sections, we explore a cross section of the current WD authentication research. The research that has been conducted ranges from theoretical proposals of authentication methods to experimentation with proof-of-concept implementations of proposed methods. The applicability of authentication factors depends on the sensors available in the WD. Therefore, we present the state-of-the-art based on the device type such as head-mounted and hand-worn that were discussed in Section 2.3.

2.6.1 Head-mounted Device Authentication

Head-mounted (HMD) WDs are typically a type of glasses; for example, Google Glass and Epson Moverio [58]. The research on HMD devices has focused on all
three authentication factors: something you know [59, 60], something you have [61], and something you are [19, 62].

2.6.1.1 HMD Authentication Based on Accelerometer and Gyroscope Data

The accelerometer and gyroscope are hardware sensors that measure movement. An accelerometer measures the acceleration based on three directions: the x, y, and z directions as shown in Figure 4.2. Similarly, a gyroscope measures the rotation and orientation for the three directions: x, y, and z based on gravity, as shown in Figure 4.3 [63].

![Figure 2.7: Accelerometers measure changes in velocity along the x, y, and z axes.](image)

Gait, or how a person moves while walking, is a biometric that can be based on accelerometer and gyroscope data [47, 66, 67, 68, 69]. Gait recognition data can be collected by placing the sensor (accelerometer) on different parts of the body such as foot, arm, hip, and leg. The authentication accuracy varies based on the sensor location [67]. For example, Gafurov and Snekkenes [67] studied gait recognition data collected from Motion Recording Sensors (MRSs) that were placed on the
Participant’s foot, hip, pocket, and arm. The values of the EER were 5% for foot, 7% for pocket, 10% for arm, and 13% for hip. Although the foot-worn sensors study by Gafurov and Snekkenes [67] achieved a relatively good EER (5%), the authors, unfortunately, did not report the number of participants in their study. The number of participants in a user study affects the results’ applicability; it is difficult to generalize the result of study with a very small number of participants to a larger population. In addition, few participants many not have as much reliability or statistical power compared to a study with many participants, even if small and large studies report similar overall results.

Li et al. built a system called Headbanger to authenticate users based on head movement when listening to an audio recording such as music [70]. They gathered accelerometer data from 95 people and found that the system achieved an average FAR of 4.42% and an average true acceptance rate (TAR) of 95.57% (TAR is the probability of the system identifying the legitimate user as an authorized one). The authors also showed that the more data gathered (i.e., more head movements), the more robust the system becomes against attacks such as imitating head movement. This is because the system learns more about how users move their heads and

Figure 2.8: Gyroscopes measure the rotation rate around the x, y, and z axes [65].
therefore, the EER is improved (Headbanger achieved EER 6.65% during 5 seconds of data gathering and 4.43% during 10 seconds) [70]. This means that imitation attacks are less plausible since accurately imitating a longer series of movements is less likely to be successful.

2.6.1.2 HMD Authentication Based on Multiple Sensors

A combination of two biometrics produces a multi-modal biometric that is intended to improve FAR and FRR values. For example, Ishimaru et al. [19] were able to extract eye blink features from the eye-facing infrared sensor in Google Glass and combine them with information about the wearer’s movement from the Google Glass’s accelerometer. Their goal was to recognize and identify what activity the 8 participants were performing: predefined activities such as reading, having a conversation with someone, and doing math calculations. Combining eye blinks with head movement increased the accuracy to 82% while with eye blinks alone the accuracy was 67% [19].

Similarly, Rogers et al. [62] exploited the data collected from Google Glass’s infrared, accelerometer, and gyroscope sensors to identify the WD wearer without depending on predefined tasks. The study was conducted on 20 participants and collected their blinking pattern using the infrared sensor, and their head movement using an accelerometer and gyroscope sensors. The authors evaluated their work using FAR, FRR, and balanced accuracy rate (BAC). BAC is the sum of half of the true acceptance rate (TAR) and half of the true rejection rate (TRR). The system achieved 11.3% FAR, 0.5% FRR, and 94.4% BAC [62]. The FAR value is not promising since it means giving a relatively high number of individuals unauthorized access, putting both the confidentiality and integrity at risk.
Bailey et al. [59] proposed knowledge-based authentication methods for Google Glass. Their theoretical methods depended on a challenge-response authentication protocol where the system asks the user questions and the user has to respond with the correct answer in order to gain access. The response is delivered to the device through voice, gestures, blinking, or head movements. Although these recommendations help to shed light on how to approach WD authentication, the authors did not implement any of their proposed solutions. Implementation will help to determine the feasibility of these suggestions in terms of security as well as usability. As an example, using voice with a shared secret is not secure due to the risk of eavesdropping and subsequent secret reuse. Therefore, implementation might help to overcome such an issue by, for example, obfuscating the secret key.

The implementation issue was addressed by Yadav et al. [60] who presented an authentication method for Google Glass called voice-based PIN (VBP) in which the PIN was obfuscated. This method used a keypad where each digit is mapped to another random digit. Users enter their PIN by speaking the mapped digits that correspond their actual PIN. Yadav et al. [60] also presented another knowledge-based authentication method for Google Glass called touch-based PIN (TBP). TBP uses the touchpad located on the side of the Glass to enter the PIN using a randomized keypad shown on the screen. The users enter their PIN by creating the same saved pattern on the touchpad. Both VBP and TBP were developed to overcome the shoulder surfing attack seen with Google Glass’ built-in authentication method that uses any “combination of swipe, hook swipe, or tap with one or two fingers” [71]. Both TBP and VPB are based on the idea that the PIN on the screen can be seen by the wearer only and cannot (theoretically) be seen by others. Therefore, observation attacks are less likely to succeed.
Evaluating against the built-in PIN in Google Glass, VBP and TBP are better in authentication accuracy, security level, and usability. Both VBP and TBP achieved 80% accuracy compared to the built-in method, which achieved 68% accuracy [60]. Comparing the work of Yadav et al. [60] to that of Bailey et al. [59], the former contained a user study to evaluate their work on 30 participants where the latter was simply a thought experiment and had no implementation. Therefore, comparison between the results or efficacy of the two proposed methods is not possible.

2.6.1.3 HMD Authentication Based on Camera Data

Sajid and Cheung [72] tried to identify Google Glass users based on writing their signature in mid air, which was captured by the Google Glass’s camera as they signed. They achieved 97% accuracy with 10 study participants, although they did not explicitly study whether users would be willing to use their system in their everyday lives; signing in mid air in order to authenticate may not be acceptable to all users [72].

So far, the research presented has depended only on the WD itself and not on additional hardware or software. Chan et al.’s [61] work, however, depends on the fact that some WDs are companion devices to smartphones. They implemented Glass One Time Password (Glass OTP), which is a token-based authentication method that depends on having the Glass’s built-in camera scan a QR code generated and displayed on a designated Android smartphone. Chan et al. [61] built two Android applications: one for the smartphone to generate the Quick Response (QR) code and one to allow the Google Glass to scan and read the QR, which will then unlock it. To evaluate their work, Chan et al. [61] used Bonneau et. al.’s
[73] framework that was developed to evaluate authentication schemes in terms of security, usability, and deployability; the authentication scheme either “offers” the benefit, “almost offers the benefit,” or “does not offer the benefit” [73]. Their overall assessment shows that Glass OTP is a good alternative that is generally better than the official lock in Google Glass in terms of security. In terms of usability, Glass OTP fails to offer some of Bonneau et al.’s framework usability benefits such as nothing-to-carry, physically-effortless, and efficient-to-use, which the authors explain is due to the setup process needed, such as launching the companion application on the smartphone [61]. Similarly to Yadav et al. [60], they were attempting to address the issue of the built-in authentication method in Google Glass being at risk of shoulder surfing attacks. However, Yadav et al.’s VBP method may outperform Glass OTP in terms of usability as it meets Bonneau et al.’s requirements of nothing-to-carry and physically-effortless.

Visual cryptography (VC) is the process of splitting an object picture or characters into two or more components that consist of black and white sub-pixels. Splitting these components is considered the encryption stage. During the authentication (decryption) process, the two components must be overlaid in order to reveal the original object [74, 75]. This object represents the secret for knowledge-based authentication; the WD wearer is the only one who can see it. Andrabi et al. performed a study to evaluate the effectiveness of VC in Epson Moverio smartglasses [74]. In this study, 26 out of 30 participants were able to successfully decrypt the encrypted object. According to the authors, the risk of the glasses being compromised is not large because the decryption is not actually done by the glasses “but only in the eye of the recipient” [74, 76]. However, participants took an average of 8.88 seconds to decode the encrypted object, which is quite long for
authentication on WDs and, particularly, on a mobile device where authentication is frequent given the bursty use pattern [3].

2.6.2 Hand-worn Device Authentication

2.6.2.1 Hand-worn Authentication Based on Accelerometer and Gyroscope Data

Hand movements and gestures are other examples of the behavioral biometrics that were used for identification purposes. Xu et al. [77] studied how hand-worn device accelerometer and gyroscope data can be leveraged to recognize arm, hand, and finger gestures as well as mid-air finger-writing. The purpose of this work was to find a new input method for smartwatches. They were able to identify 37 gestures with an accuracy of 98% and accuracy value of 95% in recognizing mid-air writing using the index finger [77]. This may help in developing a virtual keyboard for WDs, as the author suggested, which allows for authentication methods such as passwords and PINs. However, the authors did not conduct a usability study to determine whether users are willing to choose to use such a method or not as well as providing estimates of time to enroll and time to authenticate.

Yang et al. [78] proposed the MotionAuth authentication method that uses hand-worn WDs to gather participant arm movement data using the smartwatch’s accelerometer and gyroscope sensors. This data is used to create a user profile, which is later used to verify the user’s identity. The profile is built based on four human gestures, three of which are natural (arm up, arm down, and forearm rotation) and one of which is specific (drawing a circle). An Android smart watch was used to collect the data from 30 participants. MotionAuth achieved relatively
good result with 2.6% EER [78].

Al-Naffakh et al. explored the applicability of implementing transparent authentication by exploiting smartwatch accelerometer and gyroscope sensors [79]. The authors conducted a study to collect two sessions of 5-minute of walking from 10 users where each session took place on two different days. The data was collected using the Microsoft Band 2 accelerometer and gyroscope sensors. The best Euclidean distance achieved was 5.5% [79]. The authors considered this result encouraging because it shows that accelerometer and gyroscope data is distinctive enough to distinguish between individuals. However, the low number of participants (10 users in this study) affects the general applicability of the results: a user study with few participants cannot be generalized easily to a larger population because it is unlikely that participants are a representative cross section of the target population. In addition, the measurement method is not widely used among the research we reported in this document and, therefore, it is hard to do a comparison.

2.6.2.2 Hand-worn Authentication Based on Vital Signs

Vital signs are the clinical measurements of the body’s activities, such as heartbeats, body temperature, and blood pressure. Some of these measurements are considered distinctive from one individual to another. For example, human heartbeats are considered to be unique once the individual reaches the age of the adolescence [46, 45]. Kang et al. [80] attached an electrode to a traditional (non-smart) hand-worn watch and implemented an ECG-based authentication system. The system achieved 5.2% FAR and 1.9% FRR [80]. Overall, heartbeats are a potential authentication type that can be further explored in WDs as heart rate sensors.
are becoming part of hand-worn devices.

Enamamu et al. introduced a unique authentication mechanism for smartphones called BT-Authen that is based on body temperature [81]. Enamamu et al. collected body temperature from 30 people over a period of four hours using a Microsoft Band 2 smartwatch. Each participant performed the experiment on three different days. During the 4 hours period, participants were asked to act normally and do what they usually do in their daily lives. The EER value started at 1.46% on the first day, 2.18% on the second day, and finally achieved 3.4% on the third day [81]. However, the authors concluded that this does not mean that the EER is going to get worse over time because a good number of participants recorded better EER on the third day. This study shows the viability of using body temperature in WD authentication.

2.6.3 Body-worn Device Authentication

2.6.3.1 Body-worn WD Based on Accelerometer and Gyroscope Data

The body-worn (BW) WD authentication studies we have covered here are not only implemented on an actual WD (some are based on devices that were not meant to be worn, such as a smartphone). However, the way the sensors work is the same so we include them as examples of BW authentication. Derawi et al. [47] exploited the accelerometer data for authentication purposes. The authors placed a Google G1 smartphone on 51 participants’ hips in a gait-based authentication method and achieved 20% EER [47]. The authors state that this is a high error rate because of the low sampling rate, which was 4050 samples per second, in the accelerometer of the Google G1 phone. Studies on gait may be affected by factors
other than what part of the body is used. For example, the *sampling rate*, which is the number of times the accelerometer reads data per second, could affect the result since a lower rate means fewer samples upon which to base a model of the wearer’s movements, and therefore is a possibly less distinctive biometric.

### 2.7 Pattern Classification

In order to use a biometric as an identifier, we need to have a model by which we can identify the owner. This is generally accomplished using a pattern classification processes in which we create a model for biometrics. To create this model, we choose a classifier that is trained with part of data called the “training set.” Then, different part of the data called the “test set” is compared to this model using machine learning methods to identify the owner. Some examples of different machine learning methods are listed below:

1. **Decision Tree**: In a decision tree, classification happens by structuring data like a tree. The tree is built based on a sequence of questions that are asked about the data [82, 83]. The answers to these questions build up the tree nodes which are: the root node, internal nodes that have input and output, and leaf nodes that have only one input and no output and represent the final decision. This algorithm has some features that make it commonly used in the classification field. For example, it deals very well with missing data and numeric values. For example, J48, which is a decision tree-based algorithm, does pruning by removing nodes or leaves from the tree without affecting the result, which helps to simplify the result and makes the processing faster [55, 82].
2. **Naive Bayes**: is a probabilistic classifier algorithm that assumes independence in all of the data’s features, which means the existence of a specific feature in a class is not associated with the existence of other features in the same class. It is fast and simple to build because it requires only one pass on the data, which also makes suitable for large amount of data [83]. In addition, it does not require a lot of memory for training and classification. Since it is fast and simple to build, it is a good choice to be implemented in resource-constrained environments like mobile devices and WDs.

3. **k-Nearest Neighbor (k-NN)**: classifies unlabeled/unknown instances by assigning them to the majority label of their k-nearest neighbors where k is the predefined parameter [83]. In other words, a test instance will have the same label as its nearest neighbors of the training set. The neighbors are found by calculating the distance between the training instances and the neighbor test instances. The algorithm depends on voting, and the k parameter should always be an odd number as a tie-breaking technique [83]. Increasing the k parameter might increase the accuracy rate until a certain point at which the accuracy rate starts to decrease. This is because having a large k leads to an underfitting issue where the classifier may lean toward the most frequent class (the majority class) [84].

4. **Artificial Neural Network (ANN)**: ANN processes data in way that simulates how a human brain’s biological neural network works [83]. The ANN algorithm consists of a group artificial “neurons” that are connected with links that can be weighted. Each neuron receives inputs, process them, and then the output is sent to other nodes. All neurons work in a parallel
way to solve problems. A known disadvantage of ANN is that it is very slow [83].

5. **Support Vector Machine (SVM):** SVM is a supervised machine learning algorithm that predicts to which class a given subject belongs to [85]. SVM plots data as points on n-dimensional space. These points represent data categories or classes while n represents the number of feature. Then it uses a hyperplane that best separates the data into two classes. The best hyperplane placement is the one that maximizes the margins between the classes of training set [86].

6. **Logistic Regression (LR):** LG is used to predict the probability of occurrence of an event [85]. It explains the association between a dependent variable and one or more independent variables. Typically, the dependent variable is a binary type and the independent variables can be any type such as nominal, ordinal, and interval. LR fits the data on a logistic curve which is usually limited to two values that can be numerical or categorical such as 0 and 1 or male and female.

Based on the type of data we are dealing with in this dissertation (i.e., accelerometer and gyroscope data), researchers tend to gravitate toward common classification algorithms, such as Decision Tree, Naive Bayes, k-Nearest Neighbor (k-NN), and Logistic Regression because of some factors such as speed in making the decision, how quickly they train, and the accuracy of the decision [55, 82]. The other options such as SVM and ANN might provide good accuracy but they take a long time and much data to train, which make them not preferred in the WDs as they have limitations in memory, battery, and processing power.
2.8 Summary

In this Chapter, we introduced and discussed the background that is needed to understand the components of this research. We started by talking about wearable technologies, what they are, how they started, how they evolved throughout history, how they are classified, and how they relate to mobile devices. Also, because this research is about authentication in WDs, we discussed it by presenting its type, such as transparent authentication and explicit authentication. In addition, we talked about the three factors used for authentication: something you know, something you are, and something you have. This Chapter also went into a discussion of the current research about WD authentication which was presented based on the type of WD (i.e., hand-worn and head-worn). Finally, multiple pattern classification algorithms that are used to create a model for biometrics are discussed.
Chapter 3

The Authentication Framework for Wearable Devices

3.1 AFWD Overview

The AFWD is a framework that provides a basis to create a transparent and continuous authentication method to verify that the user who is currently wearing a WD is its owner. Authentication methods can be built based upon this framework, and they can be evaluated and compared to each other, which is one of the main contributions of this research. AFWD authentication is transparent, meaning authentication data is gathered in the background and users do not need to explicitly provide information to authenticate. They need less effort and less frequent interaction with the WD. The authentication is also continuous, meaning the identity of the user is repeatedly checked in an instantaneous way. It happens with a frequency that may change if needed. Continuously does not mean the authentication is happening every single moment but instead periodically.
The AFWD’s target audience includes, but is not limited to, WD manufacturers, developers, and security researchers. We will use the word developer to refer to the framework user who may develop an authentication method based on the AFWD. We will use the word wearer to refer to the person who wears and uses the WD.

WDs store sensitive data such as bank information, but they have some uniqueness and limitations that prevent them from implementing traditional authentication methods such as password or PIN. They do not have input means such as a keyboard and they have to be always available and always on while mobile devices are used in a bursty way [3]. Therefore, these issues make it challenging to build an authentication method to protect sensitive data. The AFWD provides a good basis for WD authentication because:

- It provides a reliable and trustworthy security mechanism for WDs. This is demonstrated by:

  - using multi-factor and multi-modal biometric authentications where the security improved. This is shown by the results of a user study, presented in Chapter 4, which we conducted using a hand-worn device in which multi-modal biometric data provided a very promising outcome in terms of biometric performance metrics such as having FAR value of 5.08% and an AUC value of 91.12%.
  
  - the results of a user study that assesses the capability of AFWD to provide a sufficient security from the user’s perspective, as presented in Chapter 5.

- It respects the unique form factor of WDs considering that they should be
always available and always on by authenticating users transparently and continuously. In addition, it respects the limitations of WDs such as the lack of keyboard or screen by using behavioral biometrics.

- Using factors that do not require explicit user interaction, such as behavioral biometrics, the AFWD minimizes the wearer’s interaction and effort to authenticate them. This is satisfied by the ability to collect the authentication data transparently as demonstrated by a user study we conducted to collect WD’s sensor data, which is presented in Chapter 4.

- AFWD is hardware and software independent. This means the AFWD-based authentication method should be able to work in any WD no matter what kind of hardware or software is installed. This is shown by the provided assumptions and suggested data structure.

The output of the AFWD is a security level tag (SLT) that represents the current security level in the WD. The level of security (represented by the SLT) depends on how certain the AFWD is that the current wearer of the WD is also its owner. The SLT is calculated based on the data that is gathered during the authentication process, and it can have three different values: Low, Medium, and High.

The SLT is meant to provide an authentication level to other processes that might require authentication so that authentication is centralized and does not have to be reinvented for each process. The AFWD reports the SLT to the WD’s operating system periodically. The operating system uses the SLT to apply restrictions on access the WD resources based on the data’s sensitivity and how certain the AFWD is that the current wearer is the owner of the WD. The operating sys-
tem assigns the WD tasks or functionalities to one of SLT values, Low, Medium, and High and allows/disallows access based on these levels. For example, if the SLT is Low, accessing to a high-ranked sensitive task, such as bank information is not permitted. However, a low-ranked sensitive task such as news feed or weather information would still be accessible.

The AFWD is designed to accommodate the three authentication factors: something you know, something you have, and something you are [17, 18]. In other words, the possible AFWD-based authentication methods are knowledge-based, token-based, and biometrics-based methods. This gives the AFWD the ability to include different types of authentication such as multimodal and the multifactor authentication schemes. For example, one, two, or all three common factors can be represented, or more than one input for one or more of the factors can be represented. The candidate methods depend on the WD in which the AFWD-based authentication will be implemented. For example, we can implement a knowledge-based method (e.g., PIN) in a WD that has a numeric keyboard such as a smartwatch, some of which have soft keyboards as add-ons. In addition, a WD with accelerometer, gyroscope, touchpad, and microphone can have candidate biometrics such as gait, voice, and gestures (including tapping, swiping, etc.).

3.1.1 Device Classification Method

The work described here requires a method of determining what type of device we are working with, whether it is wearable, mobile, or portable. A review of the relevant literature has not revealed a mechanism or taxonomy to classify such devices. For this reason, and in order to scope the type of devices that the AFWD is capable of supporting, a taxonomy or a classification method that capable of
differentiating between these devices is needed. Therefore, based on our knowledge, we devised a taxonomy that can be used to classify devices into one of four groups: wearable, mobile, portable, or other. One of the purposes of this method is to make the difference between wearable and other devices clear. The taxonomy is shown in Figure 3.1.

![Figure 3.1: Device classification method](image)

In order for this method to work efficiently, we listed some assumptions, rules, and other information that need to be taken into consideration when using this taxonomy. The four classes we are considering here are: portable devices (e.g.,
laptop), mobile devices (e.g., smartphones), wearable devices (e.g., smartwatch and smartglasses), and other (those that fit into none of the previous categories).

The assumptions are as follows:

1. The device that we want to classify should have an operating system that can install and run applications, perform at least simple computation tasks, and connect with other devices or the Internet. This assumption may exclude certain types of WD such fitness tracker or smart shoes that are not able to run applications. These devices may be seen as transmitter of information to other devices such as smartphones and not a multifunctional device.

2. The intention or the primary use that the device was built for. For example, it is not common for a laptop to be used in motion or to be worn on the body. The considered devices should be general purpose devices that can be used as computing devices, for Internet communication, storing data, performing tasks, word processing, and has applications. One example of what we are excluding is hand-held gaming devices, such as the Nintendo Switch or Nintendo 3DS [87] whose primary use is gaming even though they have other secondary tasks like texting. Other examples are music players and eBook readers, which are all built for specific purposes. It is worth mentioning that devices, such as tablet computers (e.g., iPads) are different because they were not designed just for playing games or just reading books, but for general purpose.

3. The device can be transferred from place to place easily; the three categories portable, mobile, and wearable will fall under this class most of the time. One example of a device that cannot be considered here is big servers.
4. One way of classifying in this method is to see whether or not a device fits in a typical pocket such as those worn on a daily basis, including jeans pockets, t-shirt pockets, or shirt pockets.

5. Examples of the “Other” class include Bluetooth headsets, remote controls, Apple TVs (or similar devices), servers, GPSs, smart vehicles, and cameras.

Figure 3.2: Using the taxonomy to classify the three different types of devices. From left to right: wearable device, mobile device, and portable device

Figure 3.2 shows a working example of how to use taxonomy to classify a wearable device, mobile device, and portable device respectively. A wearable device is represented by Apple Watch 3 [88], mobile device is represented by Samsung Galaxy S8 [89], and portable device is represented by Macbook Pro 13-inch3 [90].
3.1.2 AFWD and Security

The goal of security is to protect valuable assets. However, security does not mean that we only want to protect private information. In fact, protecting the information is only one aspect of security. A secure system should take into consideration the three security principles: Confidentiality, Integrity, and Availability (CIA) [17].

- **Confidentiality**: is protecting the assets, such as data, communications, and services, from those who are unauthorized to access them.

- **Integrity**: is to ensure that the data is accurate and the assets are only modified by authorized people.

- **Availability**: is to make assets, for those who are authorized, available whenever needed.

![Figure 3.3: This is where AFWD is located within the CIA triad](image-url)

Figure 3.3: This is where AFWD is located within the CIA triad

The CIA triad is a model that can be implemented in any security system to provide an evaluation framework for the security system under discussion. The
AFWD takes into consideration the three security principles: confidentiality, integrity, and availability. The decision whether the current wearer is the owner or not, which includes the SLT calculation is being produced continuously and transparently. Once the AFWD knows whether who is wearing the device is its owner or not, then we can make decisions on whether or not to allow access to data, which is how confidentiality is protected. In addition, by continuously authenticating the identity of the wearer, AFWD also supports the integrity of sensitive data by making sure that it is altered or modified only by an authorized user (owner). With the AFWD, however, the resources are always available whenever needed and no legitimate user is prevented from accessing these resources. Figure 3.3 shows where the AFWD is located in the CIA security model. We use the CIA model as a basis of showing how security principles are is used when designing the AFWD.

3.1.3 AFWD Assumptions

WDs have different types or classes: hand-worn devices such as smartwatches, head-mounted devices such as smartglasses, foot-worn devices such as smart fitness shoes, or body-dressed devices such as smart dress [6]. A specific type may have features that are not available in others. Moreover, a specific device of specific type may also have different crucial features that do not exist in devices of the same type. For example, we could have a smartwatch (from the hand-worn class) that has a numeric keyboard and another with no keyboard at all. This diversity makes it challenging to have clear device assumptions that the AFWD can follow. This also makes it hard to choose a single authentication mechanism that works for all WDs or even for a specific type.

In order to effectively implement an AFWD-based method, the following as-
sumptions were made:

1. **The biometric has to be valid:** the AFWD expects that the chosen biometric has the following seven factors: universality, uniqueness, permanence, measurability, performance, acceptability, and circumvention [28, pp. 29][51]. The absence of a factor may affect the authentication process. For example, if the chosen biometric is not unique enough to distinguish between individuals, it will lead to falsely allowing unauthorized people access to sensitive data. However, with all these requirements, the AFWD assumes that not all the candidate biometrics can solely be used to accurately identify the wearer. For example, voice is not a very unique trait for each individual [28, 91, pp. 33]. However, a non-distinctive biometric can be an effective factor in a multimodal biometric authentication approach, which AFWD can adopt [91, 92]. More discussion about the candidate biometrics and their validity are presented in Chapter 2.

2. **WD is a single user device:** the AFWD assumes that the authentication method will be implemented in a single-user WD in order to support the AFWD verification process (identifying that the wearer of the device is the owner). The single-user concept is also part of the nature of wearing a WD because wearing a WD is like wearing a personal item such as a regular (non-smart) watch or pair of glasses, which are rarely shared with others.
The AFWD consists of two stages: the Enrollment Stage (ES) and the Continuous Authentication Stage (CAS) and these two stages interact with a designated database that resides in the WD to perform the verification. The ES is only used in a limited cases such as enrolling a new wearer or storing new credentials, such as new biometric samples to the AFWD database.

The AFWD distinguishes between two types of wearers: a first-time wearer (FTW) or a returning wearer (RW). The FTW is a person who wants to enroll as a device owner such as those who just started using the WD, they must first enroll into the system. This enrollment is an important process since the AFWD needs to have credentials to compare to during the authentication process. In contrast, the RW is a wearer who has already enrolled into the system and has credentials already stored in the AFWD database. In short, once users wear their WD, one of two processes can happen. They can either enroll into the system (i.e., FTW) or they access using their credentials (i.e., RW). These two options can be seen as sign in process, as represented by the RW, and sign up process as represented by the FTW, which are widely used in the authentication field. More discussion about these two options is presented in Section 3.2.2 and Section 3.2.3.

The general overview of the AFWD is shown in Figure 3.4. The dashed lines show the path that an FTW would follow while the solid lines show the path a RW would follow. In other word, the dashed lines shows the enrollment path and the solid lines show the authentication path. Figure 3.4 also shows numbered steps that each path follows. The shapes with dashed
lines represent the enrollment path, which starts with acquiring the raw data and then storing it in the AFWD database. After this stage finished, the flow moves to the CAS. The shapes with solid lines represent the authentication path, which starts with acquiring the raw data, comparing it to the AFWD database content, and finally calculating the SLT to make the authentication decision.

3.2.1 The AFWD Designated Database

The database is used to store the credentials that the user and the system agreed upon during enrollment. The credentials can be any information about the device owner that represents any of the three authentication factors as described in Chapter 2. For example, it can contain the biometric template or the token that is used in the authentication process. This database resides in the WD so the credentials are kept on the WD to protect its owner privacy.
The amount of data stored in the AFWD database should not exceed the available memory in the WD so it does not cause memory error.

Both ES and CAS interact with the database by inserting new credentials or comparing to the one already stored in the database. If a change or update to the credentials is requested, a set of basic common database operations, such as insert, delete, select, or update, can be run against this database to serve these purposes. For example, the insert operation is used to store new credentials into the database while the select operation is used for comparison.

3.2.2 The Enrollment Stage (ES)

The framework stream initially starts with the ES which is an enrollment process that happens when a FTW starts using a WD for the first time. In this stage, a wearer is explicitly asked to provide some information such as a PIN, password, or biometric samples to be used as credentials. Depending on the WD features, the enrollment could happen on either the WD itself or on a companion device. For example, some WDs do not have the input means such as a keyboard and in this case, the enrollment may take place on a companion smartphone in which the credentials are stored only in this smartphone to protect the wearer privacy. When a wearer finished enrolling, the AFWD transfers to the CAS. ES happens only once and when it is successfully completed, the wearer then becomes a RW, which means the credentials that are needed for authentication in future are already stored in the AFWD database.
The ES consists of three subprocesses: the knowledge-based subprocess (KSP), the token-based subprocess (TSP), and the biometric-based subprocess (BSP). In the KSP, the wearer provides knowledge-based credentials, such as a password or PIN, which is then stored in the AFWD database. In the TSP, the wearer provides an object that is considered trusted such as a smartphone and then the AFWD stores its data in the AFWD database. In the TSP, the kind of data that is stored to identify the trusted devices is what uniquely identifies a device such as the media access control address (MAC) or the international mobile equipment identity (IMEI). In the BSP, the enrollment process that usually happens in any biometric system, also discussed in Section 2.5.3, is applied here. The wearer provides biometric samples to the AFWD that then extracts the distinctive features from them and represents them digitally to create the biometric template, which is then stored in the AFWD database. The structure of the ES and its subprocesses are shown in Figure 3.5.
### 3.2.3 The Continuous Authentication Stage (CAS)

The CAS has three main components: the AFWD input, the AFWD process, and the AFWD output, as shown in Figure 3.6. The result of the CAS is the SLT which represents the level of security in the WD and is periodically reported to its operating system to be used by other WD processes that may require authentication information. The SLT is calculated continuously and transparently during the CAS to decide whether the current wearer of the WD is the owner or not. The CAS is broken into several parts, as discussed in the following sections:

![Figure 3.6: The three components of the continuous authentication stage. TB: token-based data, BB: biometric-based data, and KB: knowledge-based data. The rounded corner rectangles are continuous processes while the sharp corner rectangles are not. The solid arrows represent the initial authentication flow, the dotted arrows represent the new enrollment authentication flow, and the dashed arrows represent the backup authentication flow](image-url)

64
3.2.3.1 The AFWD Input Component

The input component is comprised of information about the current WD wearer. It represents all three main factors in the authentication mechanism; knowledge-based, token-based, and biometric-based. An example of the knowledge-based is a graphical password or a sequence of gestures (in WDs with a touchpad). An example of the token-based mechanism is the WD itself or a tethered smartphone. An example of the biometric-based input is the digital representation of the wearer’s voice, which is available to the AFWD as raw data from the available sensors in the WD such as the microphone, camera, or gyroscope (without extra equipment).

3.2.3.2 The AFWD Process Component

The input data processing in the AFWD depends on the type of input it receives. The process component consists of three paths, each of which represents one of the authentication factors: the knowledge-based path, the token-based path, and the biometric-based path. Each path interacts with the AFWD database that stores all credentials (PIN, biometric templates, or trusted device identifiers) that the system and the wearer agreed upon during the ES. Choosing which path(s) to follow is based on the input type. It is not necessary for all paths to be used. For example, assume that all we have is voice and gait data. In this case, the biometric-based path is the only path to be used. In Figure 3.6, the rounded cornered rectangles represent the continuous process while the sharp cornered rectangle represents the explicit process. The process component paths are as follows:
• **The knowledge-based path:** It is for the “something you know” class. The inputs into this path include any information that depends on the user’s knowledge. This path compares the piece of information provided by the wearer (e.g., PIN) to the one that exists in the AFWD database. The result of the comparison is either match or non-match. Although the knowledge-based path is part of the CAS, it represents an explicit authentication. This path is not one of the regular paths, such as the token-based and biometric-based paths which represent a continuous and transparent authentication. Rather, it is a path where wearers are explicitly asked to provide the knowledge-based credentials to the system in order to authenticate them. The authentication process in the knowledge-based path happens as requested. There are several cases where this explicit authentication may be requested, as follows:

(a) Initial Authentication: If a RW who is already enrolled into the system takes off the WD and then puts it back again, an explicit authentication is needed. This is important because the AFWD cannot guarantee that the owner is the one who puts it back on again. This kind of authentication should not affect the form factor of WD nor the user acceptability to use the AFWD-based authentication method as it happens infrequently. This is because in reality, people wear the WD for a very long periods of time, such as a whole day. In fact, some people even sleep while wearing a wristband to watch their sleeping habits. In this case, they will put the device on, authenticate themselves, and forget about it. This is an average of one explicit authentication per day. The solid arrows in the process
component in Figure 3.6 represent the initial authentication flow in the AFWD.

(b) Backup Authentication: is used to raise the SLT level to perform a high-ranked sensitive task when there is not enough information available to base the authentication. For example, if the SLT was low, the wearer cannot perform a high-ranked sensitive task such as banking. Therefore, in order to perform this task, the wearer needs to wait until sufficient data is collected to raise the SLT, or can raise it instantly by using this backup authentication. The dashed arrow in Figure 3.6 represents the backup authentication flow in the AFWD.

(c) New Enrollment Authentication: explicitly authenticates the wearer before adding any new credentials such as biometric sample or token to the AFWD database. By doing so, the AFWD makes sure that this addition is done by a legitimate wearer. The dotted arrows in Figure 3.6 represent the new enrollment authentication flow in the AFWD.

- **The token-based path:** processes the “something you have” data. There are multiple options that can be used as a token, such as the WD itself or a tethered smartphone. For example, we could check if the tethered smartphone is nearby by checking its Bluetooth connectivity to the WD. Another example could be similar to the Smart Lock feature that Google introduced in Android 5.0 Lollipop [93]. With Google Smart Lock, users can add trusted objects like places, devices, or voices, and whenever the device detects that these objects are nearby, the device
will remain unlocked [93].

This path takes the input token data (e.g., connectivity status) and compares to the stored AFWD database content. Similar to the knowledge-based path, the token-based path comparison result is also either yes or no; the smartphone is nearby or its Bluetooth is on or off. Infrequently, wearers may need to add a new token or delete others. For example, if a wearer sells a smartphone that was a token, the smartphone is no longer a trusted object and therefore should be deleted from the AFWD database. Also, the wearer may want to add their new smartphone as a new token in the AFWD database. To add new tokens to the AFWD database as credentials or to delete any that exist, an explicit authentication is required to make sure that this addition/deletion is done by a legitimate wearer. This required authentication uses the knowledge-based path. If the result of this authentication is a match, then the AFWD goes to the enrollment stage (ES), specifically the TSP, to add the new token to the AFWD database.

• **The biometric-based path:** This path deals with biometric traits such as gait, or voice. This path performs the common biometric system operations, which are the *feature extraction phase* and *comparison phase*[28, pp. 7][29, pp. 58]. See Figure 3.7 for the flow and the representation of this path.

(a) Feature Extraction Phase: salient and distinctive features of the biometric input are extracted and represented digitally. The extracted features are used in the comparison phase.

(b) Comparison Phase: the comparison phase in the AFWD is based
Figure 3.7: The phases of the biometric path

on comparing extracted features from the newly entered biometric against its counterpart that is stored in the AFWD database. The result of the comparison process is the matching score, which is the similarity between the entered biometric and the stored one. The high value of this score means a high similarity. The match/non-match score is determined based on a predefined threshold where a similarity score that is above this threshold is a match and everything below it is a non-match. For example, we can choose a 90% level of similarity as a threshold, which gives only a 10% chance of incorrect authentication decision. The comparison result, whether it is a match or non-match, is then sent to the the AFWD output component.

The process here is an a verification process because the comparison is a one-to-one comparison [28, pp. 9][67]. Since it is a one-to-one match, a claim of identity is required. Thus, we are assuming that the person who is wearing the WD is the owner, and the act of
wearing it is considered the claim of identity.

3.2.3.3 The AFWD Output Component

In the AFWD output component, the result of the three paths is used to make the final output: the Security Level Tag (SLT). The SLT is the final output of the AFWD and its goal is to determine how secure the WD is. In other words, the SLT represents to what level the WD is used only by its owner. It determines the security level of the WD based on the data gathered; the levels range from low, which means that the data gathered does not match that of the WD’s owner, to high which means the data gathered closely matches that of the WD’s owner. The low level may also mean that we do not have enough gathered data to make an authentication decision. These levels are used to allow or disallow tasks, such as sending text messages and reading email, based on their sensitivity. The SLT starts at the level low and it changes based on the gathered data analysis results; increase with matches and decreases with non-matches. After wearing the WD and performing the explicit authentication (e.g., entering the PIN), the SLT moves to the level high for a predefined period of time CRT (calculation and reporting timing). During this CRT, the SLT is recalculated based on the most recent gathered data and then reported to the operating system. The CRT does not represent the time when the SLT calculation took place. Rather, it represents how often the SLT is calculated and reported and it can be various time units such as a second, a minute, or an hour. The calculation frequency may be limited by the processing power available on the WD since getting more data means processing more data, which may have an effect on
the battery life of the WD.

The process of calculating the SLT shows that the AFWD bases its final
decision on the multifactor and multimodal authentications. For example,
the AFWD uses more than one independent factor, such as a token (some-
thing you have) and biometric (something you are), to calculate the SLT
which is the final authentication decision. On the other hand, the AFWD
relies mainly on biometrics due the nature of WD, especially its limitations
in terms of few having input means. Because of that, having a high accu-

racy in biometric-based authentication is crucial. This can be achieved by
using multiple biometrics in the authentication process as it increases the
accuracy and reliability compared to a single biometric-based authentication
[28]. Therefore, the AFWD also uses multimodal authentication to increase
the security by having multiple biometrics in the SLT calculation process.

Multimodal biometric authentication requires a fusion process. Fusion is a
method to consolidate multiple biometric types or different ways of biometric
data processing in order to improve the performance of the biometric system
[28]. However, the factors that are used to make the final authentication
decision in the AFWD are not all biometrics. To overcome the issue of the
fusion of different factors, such as tokens with biometrics, some researchers
have suggested that the multimodal biometric can be extended to include
factors other than biometrics to improve security [94, pp. 388-389]. There-
fore, the AFWD fuses the multiple different types of factors following the
same idea.

There are several levels in which the fusion can be applied: feature-level
fusion, score-level fusion, and decision-level fusion [28]. Since we only have binary comparison decisions at this stage, the decision-level fusion is the proper type of fusion to use. In the AFWD, a weighted decision-level fusion is used in which each decision is given a different weight based on how accurately it represents the owner. For example, if we have two biometrics, we would give the one that has a high distinctiveness a higher weight than the one with a lower distinctiveness. Further discussion and a worked example are presented in Section 3.2.4.5.

3.2.4 AFWD Data Structure

3.2.4.1 Score Object (SObj)

The score object is the element that is sent to the AFWD output component. Based on SObj, the AFWD output component calculates the SLT. The SObj is a tuple that can be represented as follows:

\[ SObj = (s, w) \]

where \( s \) is the score representing the comparison results that are sent from the AFWD process paths to the AFWD output component, either match or non-match. The element \( w \) represents the weight that is given to the input object. It ranges from 0 to 1 and is used in SLT’s calculation. The process that acts upon SObj is:

- Add SObj to Output Buffer: this process takes each SObj and adds it to the output buffer.
The UML object representation of the SObj is shown in Figure 3.8.

![UML object representation of the SObj](image)

Figure 3.8: The UML object representation of the SObj

### 3.2.4.2 Input Object

The input object represents the raw data entered into the AFWD. This is the data that is used in the authentication mechanism. Each input object represents one of the three possible authentication factors. Therefore, the AFWD has three types of input object: biometric objects, knowledge objects, and token objects. These are the current possible types of objects based on the raw data that a WD sensor can gather. The following are the input object structures followed by the processes or methods that act upon them:

(a) **Knowledge Object (KObj):** The input KObj is a tuple as follows:

\[
KObj = (ID, t, tp, r)
\]

where \(ID\) is a unique identifier for the object, \(t\) refers to the time the raw data was entered, and \(tp\) refers to the type of knowledge authentication method used such as password or PIN. \(r\) refers to the raw data entered
by the user. The element $r$ is used in the matching process during authentication. The processes that act upon $KObj$ are as follows:

- **Add KObj to the AFWD database**: This process creates the KObj from the data collected from the wearer and added to the AFWD database. This process is infrequently used such as during enrollment or when the wearer decides to change or update the knowledge-based credentials.
- **Delete KObj from the AFWD database**: This process deletes a KObj from the AFWD database.
- **Comparison**: During authentication, this process compares the KObj (e.g., PIN) to the one stored in the AFWD database. This process uses the element $r$ for this comparison and the result is either match or non-match. After the comparison, this process creates a SObj and sends it to the AFWD output component. The content of SObj will be the score (match or non-match) and the KObj weight. An explanation of the process and how it interacts with the AFWD database is shown in Figure 3.10.

The UML object representation of the KObj is shown in Figure 3.9

(b) **Token Object (TObj)**: The input token object is a tuple and can be expressed in a similar way to the knowledge object as follows:

$$\text{TObj} = (ID, t, dt, upd)$$

$ID$ is a unique identifier for the object. $t$ represents the time that the token was added to the AFWD database as a trusted object. $dp$ refers to
The device type that is used as token such as smartphone. *upd* represents the unique physical address of the token such as the International Mobile Equipment Identity (IMEI) which identifies wireless devices [95]. The processes that act upon the TObj are as follows:

- Add TObj to the AFWD database: This process creates a TObj
from the data provided by the wearer and, based on the expression of the TObj above, adds to it the AFWD database. Before adding a new TObj, an explicit authentication is needed to ensure the integrity of the authentication data by making sure nothing is added to the AFWD database except by a legitimate user. This explicit authentication is done by the knowledge-based path, as previously discussed in Section 3.2.3.2. This process is used in limited cases, such as during enrollment and it is not meant to be used frequently as this will affect the WD’s form factor.

- Delete a TObj from the AFWD database: When a token is no longer a trusted object, this process is used to delete the TObj that represents that token. Before deleting a TObj, an explicit authentication is required to ensure the integrity of the authentication data by making sure that nothing is deleted from the AFWD database except by a legitimate user.

- List All Token Objects in the AFWD database: The job of this process is to retrieve a list of all TObjs in the AFWD database in order to check their status.

- Check Status: During the authentication, this process continuously calls the previous process and receives a list of the TObjs and then checks the status of each one to see if it is, for example, connected or nearby. Based on the result, this process creates a SObj and sends it to the AFWD output component. The content of SObj is the score, which is either match (connected) or non-match (not connected), and the TObj weight. This process is used frequently to
participate in making the continuous authentication decision and it requires no wearer interaction and therefore will not affect the WD’s form factor.

The UML object representation of the TObj is shown in Figure 3.11

\[
\text{token:Token} \\
- \text{id} = 2 \\
- \text{timestamp} = 2018-02-02 12:02:09 \\
- \text{uniquePhysicalAddress} = \text{E8:22:D3} \\
- \text{deviceType} = \text{smartphone} \\
+ \text{addToDB()} \\
+ \text{deleteFromDB()} \\
+ \text{listAll()} \\
+ \text{checkStatus()}
\]

Figure 3.11: The UML object representation of the TObj

(c) **Biometric Object (BObj):** A biometric object is a tuple as follows:

\[
\text{BObj} = (ID, t, bt, fv, w)
\]

where \( ID \) is a unique identifier for the object, \( t \) refers to the time the raw data was entered, and \( bt \) represents the type of biometric used, such as fingerprint, face, voice, or gait. \( fv \) represents the feature vector extracted from the raw data. The element \( w \) represents the weight that is given to the biometric based on how distinctive it is in identifying the WD’s owner. The more distinctive the biometric, the higher the weight assigned to it. The element \( w \) can be low if the biometric used has low distinctive features, medium if it has promising distinctive features, and
high if the biometric is very distinctive. The weight of each biometric is assigned by the developer of the AFWD-based authentication method.

The processes that act upon the BObj are as follows:

- **Add the BObj to the AFWD database:** This process sends the raw biometric data to the Feature Extraction process and receives the feature vector back. Then, it creates the BObj based on the expression above, which includes the feature vector, $fv$, and then adds it to the AFWD database. Again, since this process includes updating the database, explicit authentication is required.

- **Delete a BObj from the AFWD database:** this process deletes the BObj from the AFWD database. Before deleting, an explicit authentication is required ensure the integrity of the authentication data by making sure that nothing is deleted from the AFWD database except by a legitimate user.

- **Feature Extraction:** This process takes the raw biometric data and extracts the distinctive features from it, and then returns the feature vector.

- **Matching:** During authentication, this process continuously collects biometric data and then sends it to the Feature Extraction process. Once it receives the $fv$ back from the Feature Extraction process, it creates a BObj, then compares it to the one stored in the AFWD database. Based on the comparison result, which is either match or non-match, this process creates a SObj and sends it to the AFWD output component (see Figure 3.10). The content of SObj will be the score (match or non-match) and the BObj weight. After sending
the SObj to the AFWD output component, the BObj used to create it is deleted as it is no longer needed.

The UML object representation of the BObj is shown in Figure 3.12

<table>
<thead>
<tr>
<th>biometric:Biometric</th>
</tr>
</thead>
<tbody>
<tr>
<td>- id = 3</td>
</tr>
<tr>
<td>- featureVector = (x,y,z)</td>
</tr>
<tr>
<td>- biometricType = gait</td>
</tr>
<tr>
<td>- weight= high</td>
</tr>
<tr>
<td>+ addToDB()</td>
</tr>
<tr>
<td>+ deleFromDB()</td>
</tr>
<tr>
<td>+ extractFeature()</td>
</tr>
<tr>
<td>+ match()</td>
</tr>
</tbody>
</table>

Figure 3.12: The UML object representation of the BObj

It is the developer’s preference as to how often and for how long this data is collected and used to create objects. However, several factors must to be taken into consideration. For example, the nature of WDs in terms of being resource-constrained devices needs to be respected by not collecting data, for example, over a very short time. This may result in a heavy load on the WD’s processor, which may negatively affect the performance of the device. In addition, the amount of data should be small to avoid consuming the WD’s resources and it will take too long to reach a classification decision. In contrast, there should be enough data to make accurate authentication decisions. The algorithm for data collection is shown in Algorithm 1.
Algorithm 1 The algorithm of how data is collected in the AFWD

1: set predefinedPeriod \( \triangleright \) How often data is collected
2: set maxDataSize
3: set minDataSize
4: while timePassed < predefinedPeriod do
5:    collect data from sensor
6:    if maxDataSize > dataSize > minDataSize then
7:        create input object
8:        send object to input buffer
9:     end if
10: end while

3.2.4.3 Output Buffer

The AFWD process component sends all SObjs to the AFWD output component. When the AFWD output component receives these objects, they are placed in the output buffer until the SLT calculation starts. The AFWD supports different types of buffers that can be used to hold the SObjs. The recommended data structure for the output buffer is either a queue or stack. Queue, which is First In First Out (FIFO) data structure, and stack which is Last In First Out (LIFO) data structure, are recommended because the age of SObjs does not make a difference in the SLT calculation. After the SLT is calculated and reported to the operating system, the output buffer is cleared so the next calculation is based on fresh data.

3.2.4.4 Output Object

The SLT is the only output of the AFWD. The SLT is a tuple as follows:

\[
SLT = (t, l)
\]
where $t$ is the time the SLT was calculated and $l$ is the value of the SLT: high, medium, or low security level. Once the SObjs are received in the AFWD output component, they are held in the output buffer until the next time the SLT is calculated.

The processes that act upon the output object are as follows:

- **Calculate SLT**: This process collects all SObj from the output buffer and then calculate the SLT. The calculation of the SLT is explained in detail in the next section.
- **Report SLT**: This process reports the final SLT to the WD operating system.
- **Clear all SObj**: this process clears the output buffer to delete all SObjs.

The UML object representation of the output object is shown in Figure 3.12

![UML diagram of output object](image)

Figure 3.13: The UML object representation of the output object

The UML diagram the represents all objects and the relationship between them is shown in Figure 3.14.
3.2.4.5 SLT Calculation Process

As previously discussed in Section 3.2.3, the SLT is calculated every CRT time based on the available objects in the output buffer. Several factors may affect the SLT value; for example, how many biometrics we have, how distinctive these biometrics are, or how many tokens there are.

To calculate the SLT, the AFWD gives each biometric and token object points that represent how much effect they have on the SLT value. The effect is determined by the weight in the case of the biometric object. As
The token and the biometric objects’ weights that are used to calculate the SLT.

<table>
<thead>
<tr>
<th>Objects</th>
<th>Point (score is match)</th>
<th>Point (score is non-match)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$B_h$</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>$B_m$</td>
<td>0.5</td>
<td>0</td>
</tr>
<tr>
<td>$B_l$</td>
<td>0.25</td>
<td>0</td>
</tr>
<tr>
<td>$T$</td>
<td>0.15</td>
<td>0</td>
</tr>
</tbody>
</table>

Table 3.1: The token and the biometric objects’ weights that are used to calculate the SLT.

mentioned, the biometric object weight can be low, medium, or high. The effect of the token is constant for all kinds of tokens since there is no difference between tokens in representing the WD’s owner.

For example, if the result of the comparison in the process component (see Figure 3.6) is *match*, the biometric object is assigned the points 1, 0.50, and 0.25 to the biometrics that have high, medium, and low distinctiveness respectively. If it is *non-match*, the points given is zero and it is still considered in the SLT calculation. All token objects are given the same points (0.15) since they represent one of two values: yes or no. Similarly, if the comparison result of the token object in the process component is *non-match*, the points given is zero and it is still considered in the SLT calculation.

Table 3.1 shows these points in both cases: match and non-match. The $B_h$ means a biometric object that has a *high* distinctiveness, $B_m$ a biometric object that has a *medium* distinctiveness, and $B_l$ a biometric object that has a *low* distinctiveness. The element $T$ represents the token object that is given the lowest weight since it does not represent the owner (i.e., a token can be used by more than one person), but it could help in the authentication process.

The points given in Table 3.1 are not fixed, and the AFWD developer can
adjust these specifications based on their security and usability preferences. In general, these points were given so that in any situation, the final SLT value does not go above 1. The individual choices of the different levels are a reflection of the strength of each level. This means, without loss of generality, the low level should always have a lower weight than the medium level and the medium level should always have a lower weight than the high level.

To explain further, the value 0.15 puts the $T$ at a very low level since it less representative of the owner. Therefore, it will build up more slowly to be a valid factor that accurately identifies the owner. It does not have to be 0.15; it just has to be small enough so that over time, we need at least two tokens to reach the level of $B_l$, four tokens to reach the $B_m$ level, and seven tokens to make them identify a wearer as if we had a one $B_h$. The same applies for all other points. For example, we need at least four $B_l$s to be equivalent to one $B_h$, and two $B_l$s to be equivalent to $B_h$.

The formula to calculate the SLT is as follows:

$$l = \frac{tp}{bw}$$

where $l$ is the level of security, $bw$ is the balanced weight derived from all the objects’ weights that are used in the SLT calculation. We add all objects’ weights together and based on the result, the $bw$ is calculated as follows:

$$bw = \begin{cases} 
1, & \text{if } tw < 1 \\
tw, & \text{if } tw \geq 1 
\end{cases}$$
where \( tw \) is the total of all objects’ weights given during the objects’ creation. \( tp \) is the total points and is calculated using Table 3.1.

To explain how to calculate SLT, we give an example. Assume that we have 3 SObjs coming from the AFWD process component. One is the result of comparing a biometric object that has high distinctiveness (i.e., \( B_h \)) against its counterpart in the AFWD database and the comparison result was \textit{match}. Another is the result of comparing a biometric object that has low distinctiveness (i.e., \( B_l \)) against its counterpart in the AFWD database and the comparison result was \textit{non-match}. Last is a token object \( T \) and the comparison result was \textit{match}. First, we need to calculate the \( tp \) using Table 3.1: \( B_h = 1, B_l = 0, \) and \( T = 0.15 \). Therefore, the \( tp \) is:

\[
\text{tp} = 1 + 0 + 0.15 = 1.15
\]

Next, we calculate \( bw \) using: \( B_h = 1, B_l = 0.25, \) and \( T = 0.15 \). Therefore the \( bw \) is 1.4. Finally, the level of security is:

\[
\text{l} = \frac{1.15}{1.4} = 0.82 = 82\%
\]

Based on this, the developer can set a threshold for each security level. An example of the reported security level based on a threshold can be seen in Table 3.2. Table 3.2 also shows an example of thresholds that can be set for each task based on its sensitivity. The thresholds in this table are not fixed and they were chosen for explanation purposes. For example, the highest threshold, 80%, was chosen to give a 20% chance that the analysis could be
### Table 3.2: An example of what tasks can be performed based on the SLT value.

<table>
<thead>
<tr>
<th>SLT Value (%)</th>
<th>Reported Security Level</th>
<th>Allowed Tasks (sensitivity)</th>
</tr>
</thead>
<tbody>
<tr>
<td>SLT ≥ 80</td>
<td>High</td>
<td>banking (high)</td>
</tr>
<tr>
<td>80 &gt; SLT ≥ 60</td>
<td>Medium</td>
<td>access photo (medium)</td>
</tr>
<tr>
<td>60 &gt; SLT ≥ 40</td>
<td>Low</td>
<td>weather info (low)</td>
</tr>
</tbody>
</table>

The level of security needed for any function or task is the developer’s choice. For example, a developer can ask the operating system to allow the current wearer to access the text message feature only if the SLT was “high.” On the other hand, the developer may decide to allow the wearer to check the time or weather even if the SLT was “low.”

In some cases, the available data that can be used to authenticate the wearer is not enough to raise the SLT to a high enough level, such as when we only have one type of biometric that is not very distinctive. We do not want to completely block the wearer from using the device, but we also do not want to keep it fully available as it would compromise security. The SLT solves this problem by mapping tasks, based on their sensitivity, to the security levels.

It is worth mentioning that AFWD differentiates between wearing and using the device. The WD can be used without being worn, such as the case with smartwatches. However, the data that is needed to authenticate the user should be gathered while the WD is worn. Thus, the output of the AFWD is to decide whether or not the current wearer of the WD is the owner, rather than deciding the identity of whoever using the device. This is why the AFWD assumes that the WD has the ability to know if it is worn or taken off.
3.3 Summary

In this Chapter, we introduced the Authentication Framework for Wearable Devices (AFWD) that provides a framework for building an authentication method for WDs. Based on the AFWD, developers are able to create a transparent and continuous authentication that respects the WDs limitations. The Chapter started by describing the device classification method or taxonomy that is needed to decide whether or not the device under discussion is supported by the AFWD. A detailed discussion of the AFWD design includes the AFWD database, the enrollment stage, the continuous authentication stage, and the data structure. The discussion includes what inputs, processes, and outputs the AFWD has. The SLT, which represents the level of security provided by the AFWD, was also presented.
Chapter 4

AFWD-based Authentication
Method for Hand-worn Devices

4.1 Study Goals

In this study, we aim to determine whether or not a behavioral biometric can be created using the AFWD to identify the WD’s owner. More specifically, we want to see if the walking activity data gathered by a smartwatch’s accelerometer and gyroscope can build a distinctive profile of the WD’s wearer, and consequently be used to identify its owner in a transparent and continuous way. In addition, we want to see if, compared to single biometrics, using multimodal biometrics improves the accuracy of identifying the WD’s owner.

We asked our participants to walk for five minutes while wearing a smartwatch that used its accelerometer and gyroscope sensors gather the x, y, and z coordinates. We chose the walking activity because it satisfies the
requirements of the AFWD regarding its being able to gather data transparently without affecting the WD’s form factor. Also, walking is a behavioral biometric, which respects the WD’s hardware limitations compared to physiological biometrics, which might need additional hardware such as a fingerprint reader. The study was approved by the Florida Institute of Technology Institutional Review Board (IRB number 17-014).

4.2 Equipment

In this study, we used a LG-W100 smartwatch running Android Wear OS version 6.0.1 (shown in Figure 4.1). It was released in 2014 and is a limited smartwatch in terms of its hardware and software. It has a battery with 400 mAh, Qualcomm Snapdragon 400 processor with 1.2GHz CPU, 4GB internal storage and 512 MB RAM, and accelerometer, gyroscope, and compass sensors. There is limited data that it can gather and base behavioral biometrics upon. This left us with limited options: we could either come up with a new biometric or come up with new ways of combining these biometrics. However, in the future, it is expected that these devices will come with more sensors available. In fact, since this smartwatch is an early model (released in 2014), we are starting with the most limited information expected.

Due to the limited sensors, we chose to use the smartwatch’s accelerometer and gyroscope sensors to monitor the device’s motion as the participant walked normally. Both the accelerometer and the gyroscope are available in most of the current wearable and mobile devices.

The accelerometer is a hardware sensor that is used in mobile and wearable
Figure 4.1: The LG-W100 smartwatch that was used in this study
devices for several reasons, such as knowing when a device is in portrait or
landscape mode. The accelerometer has been effectively used in several gait
recognition studies [19, 47, 62, 67]. The accelerometer measures acceleration
based on three directions: x, y, and z that together form the sensor coordinate
system (see Figure 4.2) [63]. The gyroscope measures the orientation of the
smart device by using gravity to determine the orientation of the device.
If the device is on a steady surface, it produces a zero reading in all these
directions (see Figure 4.3).

Figure 4.2: Accelerometers measure changes in velocity along the x, y, and z axes
[64].

We also used two LG Nexus 5 smartphones running Android OS version
6.0.1. The smartphones were used for two purposes. Firstly, we used them to be able to communicate with the smartwatch. In this case, we were able to receive the collected data from the watch in order to analyze it. The second purpose is that the smartphone is also collecting similar data (the accelerometer and gyroscope) that is used in the authentication process. Android Studio 2.1.2 was used to build the needed applications that gather the data. We used the machine learning tool, Waikato Environment for Knowledge Analysis (WEKA) version 3.8.0, for classification purposes [96].

4.3 Method

The experiment included the following steps: data acquisition, data processing, and classification. We followed a similar approach to related work, e.g., [47], [97], and [98] that also used accelerometer and gyroscope data for authentication, although on mobile device instead of WDs. Although WDs and mobile devices are considered different, the idea of studying accelerometer data is the same in terms of how coordinates are collected and analyzed.
4.3.1 Participants

We had 36 participants in our study. Among them, 34 were male and two were female. The participants’ ages ranged between 18 and 54. Sixteen were between the ages of 25 and 34, 11 were between the ages 18 and 24, seven were between 35 and 44, and two were between 44 and 54. The majority had a graduate education level (69%) while the rest were at the undergraduate level. All participants were smartphone users: 80% used an iOS-based smartphone (e.g., iPhone), 14% used an Android-based smartphone, and 6% were Windows phone users. Among the 36 participants, only 5 own a WD, all of which were hand-worn WD.

The 36 participants were asked to walk freely for five minutes while wearing the smartwatch and carrying the two smartphones. One smartphone was carried in the hand that was not used to wear the watch (we call this Phone-Hand throughout this Chapter). The other smartphone was carried in the pocket and the participants were free to use whichever pocket they preferred, right or left (we call this Phone-Pocket throughout this Chapter). The walking activities took place in different locations, including but not limited to school, home, and streets (mostly outside). Using different locations reflects the reality of people walk in various scenarios that may affect walking patterns. The participants were asked to walk normally and to avoid running and jogging.
4.3.2 Data Acquisition

In the Android development field, a wearable device is usually called a “wear” and the smartphone is usually called a “mobile,” “handheld,” or “phone.” We follow the same naming convention in this Chapter. Three Android apps (applications) were built for data acquisition purposes: one Wear app and two mobile apps. All apps are responsible for collecting the x, y, and z coordinates from both the accelerometer and the gyroscope.

In the Android SDK, the SensorEventListener is a Java interface that is used to provide developers with access to accelerometer and gyroscope data. The onSensorChanged method of the SensorEventListener is invoked whenever the built-in sensor detects movement, providing us with the values of the three axes x, y, and z. We thresholded the data since the devices record movement from the vibration of their component parts even when the device is not in motion. The threshold was determined by examining people’s hand movement speed then choosing the average threshold, which is 10 m/s. We do not record x, y, and z values when the hand’s movement is less than 10 m/s.

There were two buttons on the Wear app: the Start button to start collecting the data and also to connect the Wear to the phone, and the Stop button to terminate reading sensor data (see Figure 4.4). In the smartphones’ apps, we also have two buttons to start reading and stop reading.

We calculate the magnitude of each reading as follows:

\[ m = \sqrt{x^2 + y^2 + z^2} \]
Each coordinate reading \((x,y,z)\) and its magnitude \((m)\) are considered one sample, \(S\), as follows:

\[
S = (x, y, z, m)
\]

The classification result performance may be negatively affected if the device orientation is not fixed. Previous research suggests adding the magnitude to the other three coordinates to reduce the amount of negative effect since its sensitivity is low to these orientation changes [99]. Therefore, we added the magnitude to the reading samples.

The Wear app sends the samples of both the accelerometer and the gyroscope to Phone-Hand. At the same time, Phone-Hand collects its own sensors’ coordinates. The apps on the smartphones create a text file for each sensor’s readings; and the phone that is connected to the watch, Phone-Hand, creates two extra text files to store the readings that were sent from the watch. In total, we had six text files that have readings from the accelerometer and gyroscope sensors in each device.

The Wear connects to Phone-Hand to send real-time data via the `MessageApi` Android API, which is a means of communication between two devices. The
connection between devices is established when the Start button on the Wear is tapped. To monitor the devices’ connection so data is not lost, the connection status is shown on the Wear’s screen. The Wear vibrates and the connection status changes to disconnected when connection is lost. In addition, the number of readings is always shown on the Wear screen, which increases whenever a new reading is added (see Figure 4.5). Note that this app is just for the purposes of this study and does not represent a product for users.

![Connected, Reading Counter, Disconnected](image)

Figure 4.5: Screen shots that show the connected status, reading counter, and disconnected status of the connection between Wear and the phone

### 4.3.3 Data Processing

We first cleaned the data to prepare it for classification. During the experiment, one or more of the devices failed to collect data of one or more sensors so we excluded this data. We ended up with 216 datasets of 36 participants representing accelerometer readings and gyroscope readings from each one of the three devices. Other than excluding some readings, the data was not changed in any way.
4.3.3.1 Authentication Scenarios

In this study, we considered several authentication scenarios based on what data could be gathered and what sensors would provide this data. The scenarios reflect using WDs in the real world, such as having a smartphone tethered to the WD. The scenarios are as follows:

- **Single Data Source (SDS):** We use data from only one of the smartwatch sensors: either use the smartwatch’s accelerometer data or its gyroscope data. In this scenario, we want to see if we can rely on only one data source from the smartwatch in the authentication mechanism in terms of classifier error rates. In addition, we want to see what sensor might provide better classification results. We do not consider the single data source approach with the smartphone data because our main focus in this research is on WD authentication.

- **Multiple Data Sources from Single Device (MSSD):** Data gathered from the smartwatch accelerometer and gyroscope are both used in the classification process. In this scenario, we want to test multimodal biometric authentication from the WD’s sensor data only. Multimodal biometric authentication tends to provide lower false accept rate (FAR) and false reject rate (FRR) values compared to single biometric authentication [19] and we want to see if this is also true for WDs.

- **Multiple Data Sources from Multiple Devices (MSMD):** Most WDs depend on a tethered smartphone, for tasks such as setting up the device, updating the software, or installing apps. This means, in most cases, the WD owner needs to carry a smartphone in addition to the WD.
Exploiting this idea, the smartphone’s sensor data might be used to improve the authentication process by adding more data that may be used to identify the device owner. In this scenario, data gathered from the smartwatch accelerometer and gyroscope and one of the smartphones’ accelerometer and gyroscope data are used in the classification process.

4.3.3.2 Analysis Approaches

Data analysis based on accelerometer and gyroscope data can be conducted using different approaches. One way is using raw data acquired from these sensors directly without extracting any features. For example, only the x, y, and z values are used in the classification process. In another way, the raw data is not used in the classification process and we use the features extracted from it instead. An example of these features is calculating basic descriptive statistics such as the mean, standard deviation, and average.

In this research, we analyzed our data following both approaches in order to decide, in terms of biometric metrics such as error rates, which one is more representative of the participant’s normal walking pattern. In the following, we show how we organized our data sets to reflect the two approach structure. In Section 4.4, we test each approach’s data with four classifiers and then compare the results.

- **Raw Sensor Data Approach:** We developed Java scripts to organize the data of each scenario and to create the reading sample. The SDS scenario sample of the accelerometer data resulting from the script is as follows (Table 4.1 shows all element descriptions):
<table>
<thead>
<tr>
<th>Item</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>w_acc_x</td>
<td>smartwatch accelerometer x-axis value</td>
</tr>
<tr>
<td>w_acc_y</td>
<td>smartwatch accelerometer y-axis value</td>
</tr>
<tr>
<td>w_acc_z</td>
<td>smartwatch accelerometer z-axis value</td>
</tr>
<tr>
<td>w_acc_m</td>
<td>magnitude of smartwatch accelerometer</td>
</tr>
<tr>
<td>w_gyro_x</td>
<td>smartwatch gyroscope x-axis value</td>
</tr>
<tr>
<td>w_gyro_y</td>
<td>smartwatch gyroscope y-axis value</td>
</tr>
<tr>
<td>w_gyro_z</td>
<td>smartwatch gyroscope z-axis value</td>
</tr>
<tr>
<td>w_gyro_m</td>
<td>magnitude of smartwatch gyroscope</td>
</tr>
<tr>
<td>ph_acc_x</td>
<td>smartphone accelerometer x-axis value</td>
</tr>
<tr>
<td>ph_acc_y</td>
<td>smartphone accelerometer y-axis value</td>
</tr>
<tr>
<td>ph_acc_z</td>
<td>smartphone accelerometer z-axis value</td>
</tr>
<tr>
<td>ph_acc_m</td>
<td>magnitude of a smartphone accelerometer</td>
</tr>
<tr>
<td>ph_gyro_x</td>
<td>smartphone gyroscope x-axis value</td>
</tr>
<tr>
<td>ph_gyro_y</td>
<td>smartphone gyroscope y-axis value</td>
</tr>
<tr>
<td>ph_gyro_z</td>
<td>smartphone gyroscope z-axis value</td>
</tr>
<tr>
<td>ph_gyro_m</td>
<td>magnitude of a smartphone gyroscope</td>
</tr>
<tr>
<td>Label</td>
<td>class: owner or other</td>
</tr>
</tbody>
</table>

Table 4.1: All sample element descriptions. Ph can be either phone-hand or phone-pocket.

\[
S = (w_{acc_x}, w_{acc_y}, w_{acc_z}, w_{acc_m}, \text{Label})
\]

In the MDSSD scenario (accelerometer and gyroscope data from watch only), the sample generated by the Java script is as follows:

\[
S = (w_{acc_x}, w_{acc_y}, w_{acc_z}, w_{acc_m},
    w_{gyro_x}, w_{gyro_y}, w_{gyro_z}, w_{gyro_m}, \text{Label})
\]

An example of a sample generated by the Java script for the MDSSD scenario (accelerometer and gyroscope data from watch and phones) is as follows:

\[
S = (w_{acc_x}, w_{acc_y}, w_{acc_z}, w_{acc_m},
    ph_{acc_x}, ph_{acc_y}, ph_{acc_z}, ph_{acc_m}, \text{Label})
\]
• **Feature Extraction Approach:** for feature extraction, raw data is usually divided into overlapping or non-overlapping segments or windows. Segmentation has been found to be effective in improving the result in related research [100, 101]. Segmentation can be done in several ways such as combining the data that was collected at a certain period of time in one frame or into frames of a fixed size.

In this research, creating windows based on time is not appropriate because we did not have a stream of data or time series data. Instead, we took data readings at set times based on exceeding the 10m/s threshold. For example, we wanted to gather data only when the device (and thus the device owner) was moving, which reduces battery drain and the processing time which is important because we intend to process the data on the device itself. Therefore, we divided our data into non-overlapping (fixed) windows. We tried several window sizes empirically and we found that the best performance is achieved when the frame size is 26 samples. After data segmentation, we extracted the following features (for x, y, z, and the magnitude values in each frame) [102]:

(a) Minimum: the minimum values of the 26 readings of x, y, z, and their magnitudes.

(b) Maximum: the maximum values of the 26 readings of x, y, z, and their magnitudes.

(c) Mean: the mean of the 26 readings of x, y, z, and their magnitudes.

The mean is calculated as follows:

\[ \bar{x} = \frac{\sqrt{t}}{n} \]  

(4.1)
where \( x \) refers to the total of the values in the data set and \( n \) is the number of values in the dataset.

(d) Standard Deviation: shows the spread of the data about the mean. Standard deviation is calculated as follows:

\[
\sigma = \sqrt{\frac{\sum (v - \mu)^2}{n}}
\]  

(4.2)

where \( v \) refers to each value in the data set, \( \mu \) is the mean, and \( n \) is the number of the values in the data set.

(e) Percentile: for a given score, the percentage of scores that are equal to or below it. For example, if a student grade is at 26th percentile, it means the student grade is better than 26% of the grades. The percentile is calculated as follows:

\[
R = \frac{P}{100 (n + 1)}
\]  

(4.3)

where \( P \) refers the percentile we need to find, and \( n \) represents the number of values.

(f) Quadratic Mean: or the root mean square (RMS) which is the square root of the mean square. The mean square is the mean of the squares of numbers. The RMS is calculated as follows:

\[
RMS = \sqrt{\frac{v_1^2 + v_2^2 + ... + v_n^2}{N}}
\]  

(4.4)

where \( v \) represents each value in the data set and \( N \) is the total number of values in the data set.
(g) Variance: represents how the values of the data set are distributed around the mean. It is calculated as follows:

\[
\delta = \frac{\sum (\mu - n_i)^2}{N}
\]

(4.5)

where \(\mu\) is the mean, \(n_i\) is the \(ith\) value, and \(N\) is the total number of values in the data set.

(h) Skewness: is the measure of the data set’s symmetry of distribution. A value of 0 skewness means the data set is perfectly symmetric. It is calculated as follows:

\[
S = \frac{\sum (V_i - M_v)^3}{(N - 1) \sigma^3}
\]

(4.6)

where \(M_v\) is the mean of the \(V\)th score, \(\sigma\) is the standard deviation of the \(v\)th scores, and \(N\) is the sample size.

(i) Kurtosis: a measurement used to determine the peakedness or the flatness of a data distribution. Kurtosis is calculated as follows:

\[
K = \frac{\sum (V - M_v)^4}{(N - 1) \sigma^4}
\]

(4.7)

where \(M_v\) is the mean of the \(V\)th score, \(\sigma\) is the standard deviation of the \(x\)th score, and \(N\) is the sample size.

4.3.3.3 Classification using WEKA

WEKA [96] is an open source tool for machine learning that provides an API for Java that we used for the classification process. We developed WEKA
scripts that read the datasets and performed the classification to produce the performance metrics results. The performance metrics we consider in this study are: false accept rate (FAR), false reject rate (FRR), equal error rate (EER), and Area Under the ROC Curve (AUC) as described in Chapter 2 Section 2.5.3.1. For each metric, we calculated the average of all participants’ results.

Like any other data set, some instances had missing values. We did not want to eliminate all instances that have missing data, so to handle the missing or corrupted values, we used Weka ReplaceMissingValues filter. This filter replaces the missing values with the mean of the numerical distribution.

4.3.4 Classification

In the classification process, we created a model that represents the way a device owner walks. This model is created using a classifier that is trained with part of the data called the “training set.” Then, another part of data (different from the training set), called the “test set,” is compared to this model to identify the owner. A 10-fold cross-validation (one-third for training and two-thirds for testing) was used, which was recommended by researchers [103]. This means that the data set is divided into 10 subsets, and each subset is divided into two subsets for training and testing (9 for training and 1 for testing). This is repeated 10 times, and then the accuracy average is reported. We used two-class classification to reflect how the AFWD-based authentication works. Therefore, each participant in turn was considered an owner, while the rest were other.
We classify using different machine learning methods to see which algorithm performs better. Based on the literature on accelerometer and gyroscope data processing, the related work tends to gravitate toward common classification algorithms, such as Decision Tree (J48), Naive Bayes (NB), k-Nearest Neighbor (k-NN), and Logistic Regression (LR) because of factors such as speed in making the decision, how quickly they train, and the accuracy of the decision [55, 82]. There are other options that might provide good accuracy, such as Support Vector Machine (SVM), but it takes a long time to train which makes it not preferred in the mobile environment as they have limitations in memory, battery, and processing power. More discussion about pattern classification was presented in Chapter 2 Section 2.7.

During classification, we noticed that we had an unbalanced data set. Unbalanced data sets occur when the majority of instances belong to one of the two classes. A classifier may misclassify some data as the larger class in cases where there is uncertainty in classification. This is because we have 36 participants, only one of whom is an owner. This leads to having more negative responses than positive ones, and the classifier is going to predict the most frequent class [104]. This has a significant effect on the error rate values. There were few false positives, and the algorithms are not equipped to classify some samples because they have attributes of both classes; so they are classified as a negative class in our case. Therefore, we balanced the data using WEKA SpreadSubsample filter which resulted in having equal number of instances of the owner’s and other’s data. Then, we classified the data based on the authentication scenarios mentioned previously: SDS, MSSD, and MSMD.
4.4 Results and Discussion

In this section, we present and discuss the results of running the data against the four algorithms. We discuss the results based on two approaches: raw data and feature extraction. In each approach, we only show the best performance in each authentication scenario and what algorithms achieved it. Also, in the discussion below, we present both the best performance when having smartwatch data only and also when combining it with the phone data.

4.4.1 Results Based on the Raw Data Approach

Table 4.2 shows the best achieved performances in each scenario when running the raw data against the four algorithms. The achieved error rates are generally high. Considering the smartwatch data only, the FRR of 25.44% for 7-NN in SDD scenario shows a high probability of rejecting an authorized user. Combining smartwatch data with phone data shows better performances; the lowest FRR was 18.44% in MSMD (with Phone-Hand data) by J48. An authentication system based on this shows that authorized users might be annoyed by the possibility of reauthenticating frequently. For example, in the best case scenario, out of five tries to access, an authorized user will be rejected and have to reauthenticate at least one time.
Table 4.2: Raw data approach: the best achieved result by the four methods in each scenario. Acc: accelerometer and gyro: gyroscope.

<table>
<thead>
<tr>
<th>Scenario</th>
<th>FRR</th>
<th>FAR</th>
<th>EER</th>
<th>AUC</th>
</tr>
</thead>
<tbody>
<tr>
<td>SDD (watch acc)</td>
<td>25.44% (7-NN)</td>
<td>12.24% (NB)</td>
<td>22.07% (7-NN)</td>
<td>81.95% (7-NN)</td>
</tr>
<tr>
<td>SDD (watch gyro)</td>
<td>33.94% (7-NN)</td>
<td>26.15% (J48)</td>
<td>30.98% (7-NN)</td>
<td>72.64% (7-NN)</td>
</tr>
<tr>
<td>MSSD (watch acc &amp; gyro)</td>
<td>25.93% (J48)</td>
<td>13.89% (7-NN)</td>
<td>21.92% (7-NN)</td>
<td>82.90% (7-NN)</td>
</tr>
<tr>
<td>MSMD (watch &amp; Phone-Hand acc)</td>
<td>18.44% (J48)</td>
<td>8.27% (7-NN)</td>
<td>15.06% (7-NN)</td>
<td>88.99% (7-NN)</td>
</tr>
<tr>
<td>MSMD (watch &amp; Phone-Hand gyro)</td>
<td>33.15% (7-NN)</td>
<td>24.64% (7-NN)</td>
<td>30.64% (7-NN)</td>
<td>74.58% (7-NN)</td>
</tr>
<tr>
<td>MSMD (watch &amp; Phone-Pocket acc)</td>
<td>25.76% (7-NN)</td>
<td>14.28% (NB)</td>
<td>21.15% (7-NN)</td>
<td>82.53% (7-NN)</td>
</tr>
<tr>
<td>MSMD (watch &amp; Phone-Pocket gyro)</td>
<td>32.64% (7-NN)</td>
<td>17.88% (NB)</td>
<td>30.27% (7-NN)</td>
<td>73.99% (7-NN)</td>
</tr>
</tbody>
</table>

In contrast, FAR values presented in Table 4.2 show that the MSMD scenario also provides slightly better protection compared to other scenarios. The lowest probability of an attacker gaining an authorized access (i.e., FAR) on smartwatch data only was 12.24% achieved by NB in SSD (accelerometer data). This probability decreased to 8.27% in the MSMD scenario (with Phone-Hand accelerometer data), which was achieved by 7-NN. The value 8.27% still shows some risk as it means, out of 25 tries, an attacker will gain an access two times.

Table 4.2 also shows EER values, which is where FAR and FRR are equal. Considering the smartwatch data only, the lowest ERR was 21.92%. The lowest EER when combining smartwatch data with phone data (i.e., the MSMD scenario) decreased to 15.06% achieved by 7-NN on the smartwatch and the Phone-Hand accelerometer sensor data. The AUC went up from 82% to 88.99% in the MSMD scenario (smartwatch and Phone-Hand accelerometer data), which was achieved by 7-NN.

Table 4.2 also shows that the MSMD scenario provided some improvement
as evidenced by the decrease in FRR by around 5%, FAR by around 4%, and ERR by around 6%, and by the increase of AUC by around 6% compared to the best result of the other scenarios. Also, the results of running the combination of smartwatch and Phone-Hand accelerometer data against the four algorithms was better than that of smartwatch and Phone-Pocket data. This supports the finding that having an authentication system based on multiple biometrics is more accurate than an authentication system based on a single biometric.

4.4.2 Results Based on the Feature Extraction Approach

In the feature extraction approach, the performance of the four algorithms shows a noticeable change compared to that of the raw data approach, as can be seen in Table 4.3. For example, considering only the best results on smartwatch data in the two approaches, the FRR went down from 25.44% to 18.58%, the FAR from 12.24% to 6.46%, the EER from 21.92% to 16%, and the AUC went up from 82.90% to 87.25%. Although the feature extraction approach provided better results compared to the raw data approach on smartwatch data alone, it still provides some high error rates. For example, we still have a possibly annoying authentication scheme shown by an FRR value of 18.58%, which means an authorized user would be rejected and forced to reauthenticate one time out of six tries to access. On the other hand, a somehow acceptable protection was produced against an unauthorized access, evidenced by a FAR value of 6.46%, which means about one unauthorized
access would be allowed out of 16 tries to access.

Comparing the two approaches and considering data from all devices, combining the smartwatch accelerometer data with Phone-Hand accelerometer data has provided the best result over all scenarios. For example, the FRR reduced from 18.44% to 11.07%, the FAR reduced from 8.27% to 5.08%, the EER from 15.06% to 11.72%, and the AUC increased from 88.99% to 91.12% in the raw data approach compared to the feature extraction approach. People may move sometimes slowly, sometimes quickly which makes individual readings have some outliers. In raw data approach, these outliers affected the results. In contrast, in feature extraction approach, summary statistics take the average of the readings and remove the outliers which improved the results.

With the feature extraction approach, using multiple data sources (multimodal biometrics), specifically from the smartwatch and Phone-Hand accelerometer, reduced the probability of rejecting an authorized user to 11.07% as compared to that of smartwatch data alone, which was 18.58%. The probability of an unauthorized access slightly reduced from 6.46% in the SDD scenario to 5.08% in the MSMD scenario. The best EER went down to 11.72% compared to 16% on smartwatch data alone. The AUC achieved in the MSMD scenario was 91.12%, which was 4% higher than the best AUC in the SDD scenario.
Table 4.3: Feature extraction approach: the best achieved result of the four methods in each scenario. Acc: accelerometer and gyro: gyroscope.

<table>
<thead>
<tr>
<th>Scenario</th>
<th>FRR</th>
<th>FAR</th>
<th>EER</th>
<th>AUC</th>
</tr>
</thead>
<tbody>
<tr>
<td>SDD (watch acc)</td>
<td>19.19% (J48)</td>
<td>6.46% (7-NN)</td>
<td>16.53% (7-NN)</td>
<td>87.25% (7-NN)</td>
</tr>
<tr>
<td>SDD (watch gyro)</td>
<td>32.72% (J48)</td>
<td>10.47% (7-NN)</td>
<td>19.56% (NB)</td>
<td>84.44% (7-NN)</td>
</tr>
<tr>
<td>MSSD (watch acc &amp; gyro)</td>
<td>18.58% (J48)</td>
<td>6.53% (7-NN)</td>
<td>16% (NB)</td>
<td>87.06% (7-NN)</td>
</tr>
<tr>
<td>MSMD (watch &amp; Phone-Hand acc)</td>
<td>11.07% (J48)</td>
<td>5.08% (7-NN)</td>
<td>11.72% (7-NN)</td>
<td>91.12% (7-NN)</td>
</tr>
<tr>
<td>MSMD (watch &amp; Phone-Hand gyro)</td>
<td>23.92% (LR)</td>
<td>13.75% (7-NN)</td>
<td>20.76% (7-NN)</td>
<td>82.33% (7-NN)</td>
</tr>
<tr>
<td>MSMD (watch &amp; Phone-Pocket acc)</td>
<td>17.01% (J48)</td>
<td>7.14% (7-NN)</td>
<td>15.81% (7-NN)</td>
<td>87.97% (7-NN)</td>
</tr>
<tr>
<td>MSMD (watch &amp; Phone-Pocket gyro)</td>
<td>25.41% (J48)</td>
<td>12.78% (7-NN)</td>
<td>21.35% (LR)</td>
<td>81.98% (7-NN)</td>
</tr>
</tbody>
</table>

4.4.3 Discussions and Observations

In summary, the overall results look promising and effective in identifying users but more work is required to have an authentication system that is not annoying. The results show that individuals have a distinctive walking pattern, which provides enough unique information demonstrating the feasibility of using accelerometer and gyroscope data for user identification in WDs. These reported error rates and AUCs show that it is plausible to collect data from WD sensors and provide support for using it as one form of transparent and continuous identification and authentication in WDs.

This is a proof of concept. The AFWD aims to provide a secure and usable authentication method for WDs that also respects their limitations. The results show support of the AFWD security by minimizing the risk of unauthorized access to a WD, which is what the AFWD aims to achieve. They also support usability as it is defined by the seven factors by showing the ability to authenticate WD users in a transparent and continuous way, which is
also an AFWD goal. Also, the used biometrics were chosen because they respect the form factor and the limitations of WDs given that they do not require extra equipment and do not interrupt the user experience, which is also an AFWD objective.

Given the best results in both, we can say that an authentication system based on the feature extraction approach is better than the raw data approach from several aspects as evidenced by the reported metric values. For example, authentication based on the feature extraction approach analysis is less annoying given that the lowest probability of rejecting an authorized user in this approach is 7% less than in the raw data approach. In other words, an authorized user will have to reauthenticate one time out of ten tries based on the feature extraction approach result compared to one out of five tries in the raw data approach. Similarly, the probability of an attacker gaining an access reduced to as low as 5.08%, which was achieved by 7-NN in the MSMD (smartwatch and Phone-Hand accelerometer data) while the lowest achieved in the raw data approach was 8.27% meaning we have about 3% lower risk of unauthorized access.

FAR has a direct effect on security since it measures the probability of an unauthorized user gaining an access to the device. A system based on the reported FAR values will not be very secure but is very promising given its lowest value of 5%. In general, running smartwatch and Phone-Hand accelerometer data with the 7-NN provided the lowest probability of unauthorized access, which also shows that using multiple sources of biometrics is preferred in the authentication compared to a single source of biometrics. This is, however, specific for an authentication based on walking activity.
FRR is more related to usability and therefore a high value might affect users’ willingness to use an authentication method. An authentication system performance similar to the reported FRR is unlikely to be user friendly. The FRR is not as low as the FAR because we have data of one participant labeled as owner and a mix of the other 36 participants’ data labeled as negative (i.e., diversity in the negative class).

Comparing FAR to FRR, having a quite high FRR can be tolerated since the user who is wrongly rejected can simply try to authenticate again. However, a high FAR is dangerous in terms of security since an imposter will gain unauthorized access to assets. Therefore, we see that the big difference between the achieved FAR and FRR is normal since FAR is lower than FRR.

Compared to the European standard discussed in Section 2.5.3.1 (recommended FAR is less than 0.001% and recommended FRR less than 1%) [34], the results that are reported in this study are not that close but are expected. For example, this is a prototype of a brand new system that has not had the long life that biometric authentication has already had. Also, the European standard is about physiological biometrics and because we are using a behavioral biometric, we do not expect to meet these standards. This is because research has shown that behavioral biometrics are less distinctive compared to physiological biometrics [105]. Also, people’s behavior can change over time and may be affected by their health. Although the results do not reach the European standard, this work sets a benchmark for future work.

The best achieved EER in our study was 11.72% in the feature extraction approach (and Phone-Hand accelerometer data), which is still high. However,
whatever affects FAR and FRR is going to eventually affect the EER since it relies on them. Based on the related work, there are not many studies that have been conducted on sensors that are embedded in smartwatches. Most of the previous work was conducted on standalone sensor devices that were attached to the body. An example of research similar to ours was done by Yang et al. [78] who proposed the MotionAuth authentication method discussed previously. MotionAuth achieved 2.6% EER, which is lower than any EER reported in this research because MotionAuth was based on predefined human gestures (arm up, arm down, and forearm rotation). However, using predefined human gestures does not really represent the real world while in our study, participants were asked to walk freely which is more realistic. Another study that uses the smartwatch sensors was done by Al-Naffakh et al. [106] and achieved Euclidean distance scores of 5.5. Since the metrics used in this study are different, we cannot compare it to our results.

The AUC gives the best biometric system performance evaluation techniques since it gives both classes the same weight no matter how different they are [107, 108]. In our study, the best AUC achieved was 91.12% when running a combination of smartwatch and Phone-Hand accelerometer data against the 7-NN, showing good evidence of WD sensors’ distinctiveness. Therefore, using the AUC when calculating SLT in the AFWD-based authentication method seems to be the best choice.

Results in Tables 4.2 and 4.3 also show that performances seem to improve when having a combination of data from different sensors. This provides very strong evidence that using multiple biometrics in WD authentication will result in more secure authentication methods. However, it is important to
carefully choose the right biometrics and right combinations as some of them may not add much improvement. For example, as can be seen in the third row of Table 4.2, combining smartwatch accelerometer and gyroscope sensor data (i.e., the MSSD) came with unnoticeable improvement. We believe that the reason this combination did not add much is that the gyroscope data is not distinctive enough. This is true in this particular experiment, but it does not mean that any gyroscope data in a WD is not distinctive enough to identify the device owner. It depends on how the experiment is conducted and what type of activity participants are asked to do. For example, in our case, if the tasks given to our participants included flipping their hands, the gyroscope data would, most likely, have been more valuable for biometric classification. The participants’ hand movement did not produce very unique gyroscope readings in our study. Similar conclusion was reported by other research [101, 106].

4.4.4 Study Limitations

Despite the contributions of this study in the field of WDs authentication, there are some limitations which are listed below:

(a) The best results in the experiment did not come from the smartwatch alone but the companion smartphone data improved it. Relying on smartwatch data only is preferred since it eliminates the need to have extra equipment and have reliable authentication methods. However, this is not achievable based on the reported result and it is fully expected given that WDs are still new technology. In future, WDs would be more
powerful by having a larger memory size, a more powerful processor, and more sensitive sensors.

(b) The experiment included only one type of user activity which is hand movement while walking. Considering more activities in the authentication such as running, jogging, climbing stairs, or writing would improve the performance.

(c) This study was biased by the fact that 69% of our participants are at a graduate education level. Also, the majority of our participants (34 out of 36) were male. Generalizing the results might be difficult because the general population, which is WD users, is not necessarily at this level of education or dominated by male users.

4.5 Summary

In this chapter, we presented the results of the study we conducted on 36 participants in which we showed that behavioral biometrics can be used in the AFWD to identify the WD’s owner. We also showed that using multimodal biometrics improves the accuracy of identifying the WD’s owner. The results are encouraging to address the first goal of the AFWD, which is providing a base for an efficient and trustworthy authentication method for WDs as proved by the relatively promising FAR value of 5% and the AUC value of 90.12%. Also, the ability to collect the accelerometer and gyroscope sensor data in a transparent and continuous way supports the finding that the AFWD respects the WD form factor and limitations.
Chapter 5

The AFWD Assessment

5.1 Study Goals

The AFWD aims to fill the gap in WD authentication by providing a base create authentication methods to to protect the wearer’s data. Therefore, it is important to evaluate the security and usability of the proposed framework before going further and implementing it. In this chapter, we evaluate the AFWD from a subjective and objective point of view. A subjective assessment can be done in which an individual’s opinion is the main factor that drives the results. In the subjective assessment, we value our potential user’s opinion of the AFWD as to whether or not they are willing to use it. In the objective assessment, we evaluate the AFWD based on a pre-structured metric that is not affected by individuals’ opinion. The objective assessment was based on a framework that was created by Bonneau et al. [73] for comparative evaluation of web authentication schemes. The study was approved by
the Florida Institute of Technology Institutional Review Board (IRB number 17-165).

5.2 Subjective Assessment: User Perceptions of the AFWD

5.2.1 Overview

There are several factors that might affect a user’s willingness to use the AFWD. The most important thing is whether or not wearers have confidence that the AFWD provides trustworthy security. However, security is sometimes maximized at the expense of usability. For example, a very strong security system might be annoying for users by asking them frequently for their credentials, which requires more effort. Therefore, it is important that users are willing to accept and use the AFWD. In short, it is important to balance security and usability because the more security we add, the more effort we require from users. This may annoy them and therefore they may not use the security measures which put their sensitive data at risk.

This study assesses the AFWD from a user’s perspective. The main goal is to determine whether or not the wearers believe that the AFWD-based method provides trustworthy security. Additionally, this study also explores whether or not users accept and consider using the AFWD, if it is available on their WDs.
5.2.2 Equipment and Methodology

In this study, we used an LG-W100 smartwatch running Android Wear OS version 6.0.1. Also, we used one LG Nexus 5 smartphone running Android OS version 6.0.1. The smartwatch and smartphone are connected via Bluetooth at all times so they can communicate and interchange data as needed. We designed this study by following the Wizard of Oz approach, which is a way of simulating or building a prototype of an application that has not been implemented yet [109].

5.2.2.1 AFWD Prototype

We created an Android app called AFWDStudy as a prototype of the AFWD-based authentication method that imitates some of its important functionalities that show its effectiveness and serve the purpose of this study. This app does not collect any data from participants, do any calculations, or do any analysis or pattern classification, although participants were asked to assume so.

The AFWDStudy app (the prototype) was installed on the smartwatch; its main interface is shown in Figure 5.1. The main interface contains a list view that represents the tasks to be accomplished by participants: email, banking, and alarm. This is the first screen that participants see and the red color represents a low security level. Tapping on the email item takes the participants to the email inbox and tapping on each email shows its content. The email content interface also has a button to delete the email. All of these features are shown in Figure 5.2. Similarly, tapping on the alarm item
Figure 5.1: The main interface of the AFWD prototype that shows what apps are available.

takes the participants to a list of alarms and tapping on each alarm takes the participant to an interface in which they can delete that alarm as shown in Figure 5.4. Tapping on the banking item shows the participant their bank app in which they can view their account and pay their bill, as shown in Figure 5.3.

Figure 5.2: The interface of the email app in the AFWD prototype, which lists the inbox, the email body, and email deletion.

At all times, the AFWDStudy app background color is one of the following: red, yellow, or green. Each color represents the current security level or the SLT, which represents how confident the system is that the WD is worn by its owner. The red color means low security level, yellow means medium level, and green means high level. These levels simulated based on the needs
Figure 5.3: The interface of the bank app in the AFWD prototype, which shows the current balance and bill payment.

Figure 5.4: The interface of the alarm app in the AFWD prototype, which shows the current setting of the alarm.

of the study and not from real data from user. Rather, they are controlled by another app we created and installed on the smartphone. As shown in Figure 5.5, this smartphone’s app has only three buttons: one to make the security level low (i.e., make the background color red), one to make the security level medium (i.e., make the background color yellow), and one to make the security level high (i.e., make the background color green). This interface was used by the experimenter to remotely control the smartwatch’s state.
5.2.2.2 Tasks and Security Levels

The tasks given to participants are as follows:

(a) **Alarm Task** (requires low level and above):
   i. View the alarm list
   ii. Delete an alarm from the list

(b) **Email Task** (requires medium level and above):
   i. Read an email
   ii. Delete an email

(c) **Banking Task** (requires high level):
   i. View the bank account
   ii. Pay a current bill

The email, banking, and alarm apps were chosen so that each one represents
one security level in our opinion: an alarm requires a low security level, an email requires a medium level, and banking requires a high level.

If deleting an email, paying a current bill, or deleting an alarm were successful, the participant is notified by pop-up message as shown in Figures 5.2, 5.3, and 5.4 respectively.

Participants are able to perform a given task only if they are able to access the app in which this task is performed. Accessing each app (email, banking, and alarm) to perform a task is restricted by the security level assigned to that app. Participants are not able to perform a task if the current security level in the device is lower than the security level assigned to it. We assigned the security level to tasks as follows: accessing the alarm app requires a low security level or above, accessing the email app requires a medium security level or above, and accessing the banking app requires a high security level. These levels are only the author’s opinion and they are chosen based on how sensitive they are from the author’s viewpoint. However, people in the real world might consider some of these tasks more or less sensitive and therefore require a lower or higher security level that is different from our ranking. Figure 5.6 shows the security levels available in this prototype and what tasks can be performed in each level.

If a participant tries to access an app while the current security level is lower than required, a pop up message appears telling the participant that higher security level is required. In Figure 5.7, the alert message is shown when an attempt to access the email and the banking apps was attempted while the security level was low as represented by the red color.
Figure 5.6: The security levels available in the AFWD prototype and what apps are accessible in each level

<table>
<thead>
<tr>
<th>Low Security Level</th>
<th>Medium Security Level</th>
<th>High Security Level</th>
</tr>
</thead>
<tbody>
<tr>
<td>Apps accessible:</td>
<td></td>
<td>Apps accessible:</td>
</tr>
<tr>
<td>• Alarm</td>
<td></td>
<td>• Email</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Alarm</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Banking</td>
</tr>
</tbody>
</table>

Figure 5.7: The alerts presented to wearer when trying to perform a task that requires a security level higher than the current security level

5.2.2.3 Participants

The targeted age in this study was 18 years old and above and the number of participants was 27; only one of them was female. They were recruited using a convenience sampling strategy that includes, but is not limited to, mailing lists, social media, word of mouth, and personal invitation. The percentage of people who were between the ages of 25 and 34 was 71%, 19% between 18 and 24, and 9% between 35 and 44. Most of the participants’ (66%) current education level is graduate while the level of the rest is undergraduate. All participants were smartphone users, and 90% of them used an iPhone while
the rest used Android-based devices. The number of participants who own a WD was 33.3%; all of them are the hand-worn type such as smartwatch or fitness tracker. The experiment took place in common public spaces and at our university meeting rooms and labs. All participants completed the entire experiment and no one withdrew or refused to sign the consent form.

5.2.2.4 Procedure

Participants were first welcomed to the study and then the investigator briefly explained the study process. Each participant was then given the informed consent to read and sign. If the participants agreed to sign the informed consent and wanted to participate, a unique ID was given to each of them and associated with the informed consent. Participants cannot participate in the study if they do not agree to sign the informed consent. The participant was then asked to wear the smartwatch while the smartphone was with the investigator. The investigator started by briefly explaining the AFWD’s main features. The explanation included, but was not limited to, why we need AFWD and what problems it is trying to solve, what data might be gathered and used in the authentication, the idea of continuous and transparent authentication, the SLT (Figure 5.6 is shown), and the backup authentication. After that, a brief tour was given on the AFWDS Study app to familiarize the participants with it. All of the experiment sessions were observed by the investigator to make sure it is done properly and to take some observational notes that might help during the analysis.

The participants were asked to assume that they had already enrolled into the system and that they were the owner of the smartwatch they wore. Also, they
were asked to assume that an AFWD-based authentication mechanism was already implemented in the device and that the device was able to collect and analyze data about its current wearer in order to authenticate them based on it. Initially, the security level was set to low (red background) and as the time of the experiment passed, the security level gradually was switched by the experimenter to medium (yellow background) and then high (green background). While the security level was low, the participants were asked to access the banking app or email. This access then would be rejected since both require a higher security level than low. In this way, we purposely made them fail to access in order to match the model we created so that they would have the idea that sometimes they would get blocked because of a lower security level than that required by the given task. After that, the participants were asked to perform the tasks in the order listed in Section 5.2.2.2. After finishing all tasks, a survey about the user’s experience with the tasks and the AFWD in general was given to the participants. The order of questions was as follows: email task questions, banking task questions, alarm task questions, and general AFWD questions.

The first part of the survey was demographic questions, from which we presented the information about our participants that is given in Section 5.2.2.3. The second part was to assess the security in the AFWD from the user’s perspective. As discussed in Section 3.1.2, a secure system should take into consideration the three security principles: confidentiality, integrity, and availability (CIA). The questions were grouped according to these principles to get a per task assessment of the AFWD. We broke down our questions into: how confidentiality is affected, how integrity is affected, and how avail-
ability is affected. We were interested in whether or not the AFWD is able to manage the CIA principles from the user’s perspective. Specific 6-point Likert Scale questions were formed about the tasks that participants were asked to perform, with answers ranging from “strongly disagree” to “strongly agree.” The third part of this survey is a general assessment of the AFWD framework. This assessment also takes into consideration the CIA triad and other aspects, such as comparing the proposed AFWD to the currently implemented authentication methods.

5.2.3 Results and Discussion

5.2.3.1 Per Task Assessment

All participants were successfully able to perform all given tasks that are listed in Section 5.2.2.2. In the email task, participants were required to perform two things: read an email from the list and then delete it. In the questionnaire about this task, participants were asked to first rank the email app accessibility in terms of what level of security that they felt it should require regardless of what rank we gave it. We ranked email as a medium security level task and only nine participants (36% of participants) agreed with our ranking. The eighteen remaining participants felt that email should be ranked as a high security level task. We see these numbers as understandable because people think of email differently depending on what they use it for.

In the banking task, participants were required to perform two things: view their balance and pay the current bill. Most of the participants felt that this
task should require a high security level, which was expected. However, two participants had a different opinion and ranked this task either as medium or low. This was not expected for a sensitive action like this which involves a financial transaction. One reason could be that a participant was randomly filling out the questionnaire. For example, by examining the responses, we noticed that one participant responded with “agree” for all questions.

In the alarm task, participants were required to perform two tasks: view the alarm list and delete an alarm from the list. In the questionnaire about this task, participants were asked first to rank the alarm app accessibility in terms of what level of security that they felt it should require regardless of what rank we gave it. The majority (80%) of the participants felt that the alarm should require a low security level. It was unexpected to have 20% of the people ranking alarm higher than low because setting and deleting and alarm is not that sensitive and can sometimes be accessible on a locked screen, such as in some smartphones. Those who ranked alarm as medium or high justified their choices by saying that a low security level might cause privacy issues as unauthorized people could view their schedule or even delete an alarm, causing them to miss an important meeting.

Overall, our participants’ ranking of the given tasks varies. What is sensitive for some might not be sensitive for others and vice versa. This suggests that users should be given the choice about which security levels each app should require. This can be easily applied in the AFWD and one suggestion is to give users the ability to rank tasks during the enrollment stage.

The 6-point Likert scale’s three main questions and the responses for each
alarm task, email task, and banking task are listed in Figures 5.8, 5.9, and ?? respectively. The results show that the confidentiality is managed and not affected in the AFWD in the medium and high security levels. This is supported by the fact that all participants agreed that they feel that they are the only owner who is able to view emails and bank balances, which is what confidentiality is about. This positive result might be encouraged by our deliberate access rejection (as discussed in Section 5.2.2.4) that participants experienced when the security level was low in the beginning of the experiment, and then the ability to access when the security level went up to the required level (medium for email and high for banking).

Figure 5.8: The participants’ opinions toward per task AFWD assessment in terms of the CIA principles (Alarm Task): [C] is confidentiality, [I] is integrity, and [A] is availability.
In the alarm task, 64% of the participants agreed that the owner is the only person who is able to view the alarm list, which suggests that the AFWD also manages confidentiality in such a task. Although confidentiality means only those who are authorized can view the alarm list, it should be accessible by all people since it only requires a low security level. Therefore, it was not expected that the majority of participants would agree that only the owner is able to view the alarm list. Some participants may have created and inaccurate mental model that affected their responses. Also, they may think of security as a binary value in which it either allows or disallows (protecting all or nothing).
From the user’s perspective, the integrity of the data in tasks that require medium and high level security is not affected and data modification can only be performed by an authorized wearer. This is supported by our participants’ opinions in which they felt that only the owner is able to delete an email or pay the current bill. In the alarm task, the effect on integrity is evaluated by the ability to delete an alarm from the alarm list. Since the alarm task requires a low security level, it is not expected that deleting an alarm is restricted to those who are authorized. However, 64% of participants felt the opposite. This might also be related to an inaccurate mental model while filling out the questionnaire.
<table>
<thead>
<tr>
<th>General AFWD Assessment</th>
<th>Strongly Agree</th>
<th>Slightly Agree</th>
<th>Slightly Disagree</th>
<th>Disagree</th>
<th>Strongly Disagree</th>
<th>N</th>
<th>Mean</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>I feel I’m the only person who is able to view the alarm list (C)</td>
<td>7</td>
<td>7</td>
<td>4</td>
<td>2</td>
<td>5</td>
<td>27</td>
<td>4.11</td>
<td>1.70</td>
</tr>
<tr>
<td>I feel that I’m the only person who will be able to delete an alarm (I)</td>
<td>7</td>
<td>7</td>
<td>3</td>
<td>2</td>
<td>5</td>
<td>3</td>
<td>4</td>
<td>1.797</td>
</tr>
<tr>
<td>I feel that I will be able to view my alarm list whenever I want to (A)</td>
<td>18</td>
<td>5</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>5.41</td>
<td>1.047</td>
</tr>
<tr>
<td>I feel that I’m the only person who will be able to view my email (C)</td>
<td>11</td>
<td>6</td>
<td>10</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>5.04</td>
<td>0.898</td>
</tr>
<tr>
<td>I feel that I’m the only person who will be able to delete an email (I)</td>
<td>13</td>
<td>7</td>
<td>7</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>5.22</td>
<td>0.847</td>
</tr>
<tr>
<td>I feel that I will be able to view my email whenever I want to (A)</td>
<td>9</td>
<td>6</td>
<td>9</td>
<td>1</td>
<td>2</td>
<td>0</td>
<td>4.7</td>
<td>1.203</td>
</tr>
<tr>
<td>I feel that I’m the only person who will be able to view my bank account (C)</td>
<td>15</td>
<td>7</td>
<td>5</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>5.37</td>
<td>0.792</td>
</tr>
<tr>
<td>I feel that I’m the only person who is able to withdraw money from my bank account (I)</td>
<td>17</td>
<td>5</td>
<td>5</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>5.44</td>
<td>0.801</td>
</tr>
<tr>
<td>I feel that I’m able to view my bank account whenever I want to (A)</td>
<td>13</td>
<td>6</td>
<td>6</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>5.04</td>
<td>1.192</td>
</tr>
<tr>
<td>I feel that using an authentication method that is based on the AFWD will prevent unauthorized access to my sensitive data (C)</td>
<td>19</td>
<td>2</td>
<td>5</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>5.41</td>
<td>1.04</td>
</tr>
<tr>
<td>I feel that using an authentication method that is based on the AFWD will prevent any unauthorized alteration of my sensitive data such as editing or tampering (I)</td>
<td>16</td>
<td>4</td>
<td>7</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>5.33</td>
<td>0.877</td>
</tr>
<tr>
<td>I feel that using an authentication method that is based on the AFWD will allow me to access my data whenever I need to (A)</td>
<td>11</td>
<td>9</td>
<td>2</td>
<td>1</td>
<td>3</td>
<td>1</td>
<td>4.78</td>
<td>1.5</td>
</tr>
<tr>
<td>Having multi factor authentication such as that provided by the AFWD would encourage me to use it compared to a single factor method such as password or fingerprint</td>
<td>19</td>
<td>5</td>
<td>2</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>5.52</td>
<td>0.93</td>
</tr>
<tr>
<td>The idea of the transparent authentication implemented in the AFWD will encourage me to use AFWD-based authentication when available</td>
<td>17</td>
<td>6</td>
<td>3</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>5.41</td>
<td>0.97</td>
</tr>
</tbody>
</table>

Figure 5.11: The number of people who chose each option with mean and standard deviation for both per task assessment and general AFWD assessment. (C) means this question assesses confidentiality, (I) means this question assesses integrity, (A) means this question assesses availability, N is the total number of people responded to that question, and SD is the standard deviation. The scale used for the words is as follows: Strongly Agree = 6, Agree = 5, Slightly Agree = 4, Slightly Disagree = 3, Disagree = 2, and Strongly Disagree = 1.

Our participants felt that the availability of their data in all tasks is not affected in the security measures applied by the AFWD. A percentage of 93% of participants felt that they have the ability to access their devices. In email and banking tasks, 7% of the participants felt that some access attempts would be rejected at some point and their data would not be available whenever needed. Those participants believed that behavioral biometrics might be affected by emotions or medical issues such as having a fast heartbeat or sore throat. Similarly, in the alarm task, participants also felt that availabil-
ity is not affected by the AFWD security measures. This was expected since the alarm task requires only a low security level in which all people are able to perform the task successfully.

Much more detail is presented in Figure 5.11 which contains multiple table that show the three task assessment results and the general AFWD assessment results. It shows what options the participants chose for each question. The mean values, which were between 4 and 5, in the tables show that participants were agreeing that AFWD supports the CIA triad based on per task assessment. Same thing applies on the general AFWD assessment in which participants show positive opinion toward the AFWD in terms of maintaining the security principles (CIA triad). The standard deviation (SD) values were also reported in these tables which show that individual responses to a question are close to the mean, on average.

5.2.3.2 General AFWD Assessment

In the previous section, we reported a per task assessment of the AFWD in which we covered the three security levels specifically. Per task assessment makes participants think about the system more concretely. The tasks we chose are not the only tasks that can be done nor are they the only tasks that are affected by the AFWD. Therefore, we cannot assess the system based only on them. In this section, we present the results of users’ assessment of the AFWD in general in which we want to make a broader statement based on participants’ inputs about whether or not they feel the whole system is trustworthy and secure. This assessment evaluates the perceived level of security produced by the AFWD in terms of several aspects. First, the assessment is
based on the three security principles in which we see how users feel about the confidentiality, the integrity, and the availability of their data while using a WD with an AFWD-based authentication method implemented. Second, we present an assessment based on the comparison between their current authentication method and the AFWD-based authentication method. Third, we test whether or not having multi-factor and transparent authentication as provided by the AFWD would encourage people to use it when it is available.

(a) Assessment Based on The CIA Triad: In general, from the user’s perspective, confidentiality is not affected in the AFWD security measures considering all of the features and functionalities as presented to each participant in the beginning of the experiment. The majority of the participants (96%) strongly felt that the AFWD-based authentication manages the confidentiality by preventing an unauthorized accesses to the owner data. Similarly, all participants felt that an AFWD-based authentication method preserves the integrity of the owner’s data by preventing any data alteration by unauthorized people. The majority of the participants (81%) felt that availability is not affected in the AFWD-based authentication method.

(b) Assessment Compared to Current Authentication Method: In this part, participants were asked to compare the proposed AFWD-based authentication method to their current authentication method. By current authentication method, we mean what people use to protect their data in their smart devices whether they are smartphones or WDs. All of our participants said they activate an authentication method,
which includes a password, PIN, fingerprint, or face recognition. The responses are shown in Figure 5.12 and it can be seen that most of the participants felt that the AFWD-based authentication method is more secure or at least similar to their current authentication method.

![Figure 5.12: How participants compare the AFWD to their current authentication methods.](image)

Specifically, 66.7% (18 of the participants) felt that an authentication method based on AFWD is more secure than their current authentication method while 14.8% (four participants) felt it is similar. About 18% (five participants) felt that the proposed framework will be less secure than their current authentication method. This might be affected by the lack of knowledge or trust in behavioral biometrics, which play the main role in the AFWD, because they are not as familiar to users as other traditional authentication methods such as a password, PIN, or fingerprint. Those who saw the AFWD as less secure than their current authentication method still think it is secure or protected when they were asked how they would describe it.
(c) **Multi-factor Authentication and Transparent Authentication**

**Effect:** The system has some features that we believe support its strength and effectiveness, such as multi-factor authentication and transparent authentication. It was explained to the participants that multi-factor authentication would support security and transparent authentication would support usability and would respect the limitations of WDs. We asked participants about their feeling about whether or not having these features would encourage them to use the AFWD-based authentication if it was made. All participants but one believed these two features would encourage them to use the AFWD-based authentication method when available. This positively supports the acceptability of the AFWD when available since the prospective users felt that it provides some aspects of two important features: security and usability.

(d) **How Do You Describe the AFWD?** In this part, we listed fourteen words to describe the AFWD. The list includes both positive and negative adjectives. We did not give any description to any of the adjectives in the list; the interpretation was left to the participants’ understanding. The list includes adjectives that are related to some important aspects in the authentication such as security and usability. Specifically, the positive ones include: secure, protected, efficient, reliable, easy to use, learnable, vulnerable, trustworthy, effective, and necessary. The negative ones include: complex, confusing, optional, and time consuming. We considered “optional” because selecting it shows the opinion that there is no need for the AFWD. We asked our participants to choose all words from the list that they thought describe the AFWD from their
Figure 5.13: The adjectives chosen by our participants to describe the AFWD point of view. We used the concept of a Word Cloud, which is a way to visualize the responses by giving the most frequent adjectives bigger size than others. The participants were allowed to choose as many words as they wished. Figure 5.13 shows the responses to this question and it can be seen that the most frequent adjectives were the positive ones. The responses show that our participants felt that the AFWD is both secure and usable. This is supported by the fact that the three most important words chosen by our participants were protected and secure which are related to security and ease of use, which is related to usability.
5.3 The AFWD Objective Assessment

5.3.1 Overview

Although the outcome of the AFWD assessment from the user’s perspective is valuable, we do not want to base all of our assessment only on something that is subjective. This is because subjective assessment might be affected by issues we cannot control, such as the demand effect which is when people misinterpret questions or try to please the researcher. Therefore, an objective assessment is used to mitigate such concerns. In this section, we assess the AFWD from an objective point of view.

Bonneau et al. [73] developed a framework to analyze the authentication methods that aim to replace a traditional password. It evaluates authentication schemes in terms of usability, deployability, and security; the authentication scheme either “offers the benefit”, “almost offers the benefit,” or “does not offer the benefit” [73]. Although Bonneau et al.’s framework was directed at evaluating the authentication methods used on traditional devices such as a PC and laptop, the framework can be adopted to work on other types of devices since the authentication idea is still the same. Researchers have used Bonneau et. al.’s framework to assess their proposed authentication methods on devices other than PCs. For example, Chan et al. [61] noted that Bonneau et. al.’s framework can in fact be used to assess their proposed Glass OTP authentication scheme based Bonneau et al.’s framework, as discussed in Section 2.6.1.3.

In this section, we evaluate the possible AFWD-based authentication meth-
ods against Bonneau et. al’s framework, as shown in Tables 5.1, 5.2, and 5.3 for usability, deployability, and security respectively. Some benefits are not applicable in our scheme and if this is the case, we mark it not applicable (NA). In this assessment, we do not consider the AFWD explicit authentication a fundamental element that the framework relies on and therefore we give it a small weight in the assessment. This is because the AFWD mainly depends on biometrics and on tokens. Also, the AFWD’s explicit authentication happens infrequently, as per the AFWD specification. However, in limited situations and whenever necessary, especially in security assessment, we might need to discuss it because it imposes some risk.

5.3.2 Usability Assessment

Our usability assessment based on Bonneau et. al’s framework shows that the AFWD-based authentication methods are usable since they offer almost all benefits. This also matches our participants’ choices in the subjective assessment study in which the majority described the AFWD as “easy to learn,” and “learnable” which are two of the usability aspects.

(a) Memorywise-Effortless: The AFWD has a backup authentication mechanism that might be a knowledge-based method, which might require memorization. However, this process is not a fundamental feature in the AFWD since it depends mainly on behavioral biometrics in the whole session. This suggests that the AFWD needs nothing memorized during its core function, the continuous and transparent authentication. However, we conservatively chose “almost offers the benefits.”
Table 5.1: The usability assessment of the AFWD-based authentication method. √ means it offers the benefit, ∼ means it almost offers the benefit, χ means it does not offer the benefit, and NA is not applicable.

<table>
<thead>
<tr>
<th>Usability Benefits</th>
<th>Assessment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Memorywise-Effortless</td>
<td>∼</td>
</tr>
<tr>
<td>Scalable-for-Users</td>
<td>√</td>
</tr>
<tr>
<td>Nothing-to-Carry</td>
<td>√</td>
</tr>
<tr>
<td>Physically-Effortless</td>
<td>√</td>
</tr>
<tr>
<td>Easy-to-Learn</td>
<td>√</td>
</tr>
<tr>
<td>Efficient-to-Use</td>
<td>√</td>
</tr>
<tr>
<td>Infrequent-Errors</td>
<td>∼</td>
</tr>
<tr>
<td>Easy-Recovery-from-Loss</td>
<td>√</td>
</tr>
</tbody>
</table>

(b) **Scalable-for-Users**: The AFWD assumes that the WD is for a single user and therefore if the wearer uses the AFWD on other WDs, it does not increase the load on them.

(c) **Nothing-to-Carry**: Although a token is considered one of the factors that can be used in the authentication decision in the AFWD, it is only used upon availability. For example, if we have an AFWD-based authentication that uses a smartphone as a token, the authentication can happen even though the smartphone is not available nearby the wearer. Therefore, we say that the AFWD-based authentication method requires nothing to carry beyond the WD itself.

(d) **Physically-Effortless**: Users are not required to make an effort because all needed data is gathered transparently.

(e) **Easy-to-Learn**: The AFWD is easy to use, users are not required to have specific skills in order to use it. Enrollment might require some learning but it should be straightforward since users are asked to provide behavioral biometric which is part of their daily activities such as voice
and walking.

(f) **Efficient-to-Use:** The authentication happens transparently in the AFWD and users do not spend time on each authentication. The enrollment may take some time that is longer than the authentication process but this is considered reasonable according to Bonneau et al.’s framework.

(g) **Infrequent-Errors:** Users are not required to perform specific tasks in order to authenticate themselves in which an input error could occur. Rather, the framework exploits what WD sensors can gather about users as they go about their regular tasks and uses data from this as the authentication mechanism’s input. In general, being behavioral biometrics oriented, the AFWD-based authentication method may make some mistakes. This also supported by the results we reported in Chapter 4 in which we had a high FRR. Therefore, we chose to say that the AFWD “almost offers the benefit.”

(h) **Easy-Recovery-from-Loss:** One of the AFWD features is that it has a backup authentication that is used when the WD does not have enough information to authenticate current wearer before performing certain tasks. Also, this feature can be used to replace old biometric templates that are no longer valid for authentication. Therefore, we stated that the AFWD-based authentication offers this benefit.
Table 5.2: The deployability assessment of the AFWD-based authentication method. √ means it offers the benefit, ~ means it almost offers the benefit, χ means it does not offer the benefit, and NA is not applicable.

<table>
<thead>
<tr>
<th>Deployability</th>
<th>Assessment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Accessible</td>
<td>~</td>
</tr>
<tr>
<td>Negligible-Cost-per-User</td>
<td>√</td>
</tr>
<tr>
<td>Server-Compatible</td>
<td>NA</td>
</tr>
<tr>
<td>Browser- Compatible</td>
<td>NA</td>
</tr>
<tr>
<td>Mature</td>
<td>χ</td>
</tr>
<tr>
<td>Non-Proprietary</td>
<td>√</td>
</tr>
</tbody>
</table>

5.3.3 Deployability Assessment

As can be seen in Table 5.2, our deployability assessment shows that the AFWD is not yet mature. It also shows that there are some accessibility issues especially with biometrics, which the AFWD mostly depends on, since not all users can provide the required traits. In general, given that we ranked three benefits one of each kind, “offer”, “does not offer”, and “almost offers” the benefit, we can say that the AFWD is not 100% deployable but to some extent. This is expected since the AFWD is in its infancy.

(a) **Accessible:** We say that the AFWD-based authentication method is almost accessible where a user is not prevented from using it because it depends on what kind of method is used. For example, the authentication method based on heartbeat is accessible by all people. However, methods that depend on gait are not accessible by disabled people and those that depend on voice are not accessible by mute people.

(b) **Negligible-Cost-per-User:** There is little cost in order to deploy an authentication method based on the AFWD. In fact, the AFWD aims
to provide a strong authentication base that asks for no extra equipment to authenticate users. Using a token is part of the AFWD environment, but having it is not required. It is only used if it is available. Therefore, the AFWD-based authentication method offers this benefit.

(c) **Server-Compatible:** We chose “NA” because the authentication in the AFWD does not need an interaction with any server and it is only a client-side process. The data is stored locally in the WD and the process is done on the WD.

(d) **Browser-Compatible:** We chose “NA” because there is no web browser needed in the AFWD-based authentication methods.

(e) **Mature:** the AFWD is still in the development process and has not been implemented on any scale. There has been research on WD authentication in general but it has not come to the point where we can safely call it mature. Therefore, we say that the AFWD-based authentication method does not offer this benefit.

(f) **Non-Proprietary:** The AFWD was developed to be hardware and software independent and therefore, we stated that the AFWD-based authentication method offers this benefit.

### 5.3.4 Security Assessment

Our security assessment shows that the AFWD offers almost all of the benefits, as can be seen in Table 5.3 and as described in the list below. Our assessment shows that the AFWD either offers or almost offers ten security benefits out of twelve from the Bonneau et al.’s framework. This indicates
Table 5.3: The security assessment of the AFWD-based authentication method. √ = offers the benefit, ~ = almost offers the benefit, χ does not offer the benefit, and NA not applicable.

<table>
<thead>
<tr>
<th>Security Benefits</th>
<th>Assessment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Resilient-to-Physical-Observation</td>
<td>~</td>
</tr>
<tr>
<td>Resilient-to-Targeted-Impersonation</td>
<td>~</td>
</tr>
<tr>
<td>Resilient-to-Throttled-Guessing</td>
<td>√</td>
</tr>
<tr>
<td>Resilient-to-Unthrottled-Guessing</td>
<td>√</td>
</tr>
<tr>
<td>Resilient-to-Internal-Observation</td>
<td>χ</td>
</tr>
<tr>
<td>Resilient-to-Leaks-from-Other-Verifiers</td>
<td>√</td>
</tr>
<tr>
<td>Resilient-to-Phishing</td>
<td>√</td>
</tr>
<tr>
<td>Resilient-to-Theft</td>
<td>√</td>
</tr>
<tr>
<td>No-Trusted-Third-Party</td>
<td>√</td>
</tr>
<tr>
<td>Requiring-Explicit-Consent</td>
<td>√</td>
</tr>
<tr>
<td>Unlinkable</td>
<td>√</td>
</tr>
</tbody>
</table>

that the AFWD provides a trustworthy security, which is in accordance with how the potential WD users feel toward the AFWD security as presented in the subjective assessment. Although the explicit authentication is not a fundamental part of the AFWD authentication process, it affected our choices when ranking some benefits. The only benefit that the AFWD-based authentication method struggles to provides is the resilience to internal observation. It is mostly about biometric security and how it can be stolen and used to impersonate legitimate users.

(a) **Resilient-to-Physical-Observation**: is an example of a shoulder surfing attack, which is where an attacker observes the victim to get the secret key like a password or PIN. The authentication in the AFWD happens transparently and the user has no explicit interaction that can be observed by an attacker. The only situation that can be considered risky would be during explicit authentication, which can be observed
by an attacker. Attackers would then be able to use the credentials to gain access to high level tasks or possibly enroll their biometrics instead of those of the legitimate wearer. This, however, would require having a physical possession of the WD and any companion device if the explicit authentication is not being implemented on the WD. Having these conditions is less likely to happen as it requires much effort from the attacker and it will catch the owner’s attention in a short time.

(b) **Resilient-to-Targeted-Impersonation:** It depends on the type of biometric being used in the AFWD-based authentication. For example, it is hard for an acquaintance to impersonate users when using an authentication method that depends on heartbeats or body temperature. However, it is possible when using an authentication method based on gait or voice, as an acquaintance may be able to imitate how people walk or talk. This is still limited by the attacker’s knowledge of the owner and what type of biometric is currently used, so we conservatively chose “almost offer the benefits.” Another kind of attack that might be launched is a replay attack. As an example of this, assume an AFWD-based authentication method uses voice; an attacker may record or possess a recording of the victim’s voice and then play it back. In this case, the AFWD-based method would think it is the legitimate user’s voice when it is not.

(c) **Resilient-to-Throttled-Guessing:** Although we say that the AFWD offers this benefit, the explicit authentication might provide risk. This risk depends on what kind of mechanism is used by the AFWD explicit authentication. Also, it depends more on the security measures on the
device in which the explicit authentication is performed. For example, if the explicit authentication is applied in a companion smartphone, the risk varies whether it is password, PIN, or fingerprint. Having a long password is probably stronger than having a four-digit PIN in terms of guessing. Bonneau et al.’s framework suggests that this benefit might be granted if an attacker is limited to around 10 guesses per account per day. Many smartphones lock the user out after a small number of failed attempts to enter. For example, on the iPhone 7, the user is locked out and has to reset the device after ten failed attempts to enter the correct password or PIN [110]. This supports our choice to grant this benefit.

(d) **Resilient-to-Unthrottled-Guessing:** The risk comes from the explicit authentication that is subject to a guessing attack. However, having a limited number of guesses per account per day is widely available in the smart devices which drove our choice.

(e) **Resilient-to-Internal-Observation:** There are no user inputs in the AFWD core authentication process, which lowers the chance to impersonate a user by intercepting their inputs. However, a biometric template attack could happen in which the attacker steals the biometric template and then replays it against the stored one in the database. The AFWD depends on multimodal biometrics, which makes it harder since this attack would require an attacker to steal all current biometric templates used in the authentication to increase the SLT level and therefore gain the access to the sensitive resources. Intercepting the input requires that the system is affected by a malware as this benefit assumes. Therefore, it depends on whether or not the device resources,
such as the operating system, were built with security in mind.

(f) **Resilient-to-Leaks-from-Other-Verifiers:** The authentication happens locally in the WD, which suggests there are no other verifiers such as a remote server. The data collected are stored, analyzed, and classified on the WD.

(g) **Resilient-to-Phishing:** The authentication happens locally so the probability of simulating a valid verifier is missing and therefore no credentials can be stolen.

(h) **Resilient-to-Theft:** The possession of the WD by itself imposes little risk since the attacker would be locked out after a short time when the security level goes down as a result of the AFWD’s low confidence that the current wearer is the owner of the device. If a companion smartphone is used for the AFWD backup authentication, the possession of it does not guarantee the unauthorized access. It depends on whether a security measure is used and whether or not these measures control the access well. This criterion suggests that the modest strength of a PIN would be enough to mark this as “offers benefit.” Furthermore, to gain access to the WD, an attacker would have to have both the WD and the companion smartphone at the same time.

(i) **No-Trusted-Third-Party:** The AFWD does not rely on any third party. Therefore, we say that it offers the benefit meaning that there is no risk imposed since its possible source, the third party, is absent.

(j) **Requiring-Explicit-Consent:** The AFWD requires that the wearer use explicit authentication every time they put on the WD. Users are
asked to provide credentials before using the AFWD which can be skipped and therefore they are not enforced to use the AFWD without their consent.

(k) **Unlinkable:** This benefit states that “colluding verifiers cannot determine, from the authenticator alone, whether the same user is authenticating to both” [73]. However, authentication happens locally and there is no interaction with any verifier that might be affected by any colluding.

### 5.4 Study Limitations

As in any study, we had some limitations that might have affected our results as follows:

(a) In the subjective assessment study, we gave a presentation about the AFWD to participants so they would know what it was before starting the experiment. It is possible that the positive nature of that presentation could have biased them in terms of their opinion after using the system.

(b) The subjective assessment study was a Wizard of Oz study type so it did not work the way it would in real life. Although we tried to cover the most important aspects of our framework, several other features were not tested. The prototype we provided did not cover some potential challenges that might face WD technologies, such as processing performance and battery drain.
(c) The subjective assessment study was biased by the fact that 70% of our participants were between 25 and 34 years old. Also, all of our participants had at least a Bachelor’s degree. Furthermore, 66% of them are currently at a graduate education level. We might have some difficulties in generalizing the result because the general population, which is WD users, is not necessarily at this age and does not have the same education level. In fact, the population of WD owners is likely to be far more diverse than our sample. For example, Figure 5.14 shows a recent survey by Statista Global Consumer Survey, an online statistics portal for marketing purposes, in which the WD users’ ages appear very diverse and the age range 25-34 is only 33% of the population.

![Figure 5.14: The WD users by age.](image)

(d) The results of the subjective assessment study might have been affected by an experimenter demand effect which is where participants become biased by trying to respond to the questionnaire in a way that supports the research hypothesis to please the researcher [111].
In the objective assessment, Bonneau et al.’s framework is used to evaluate authentication methods. However, in our case, we evaluated the whole AFWD framework and took into consideration some possible methods that could be built based upon it. This assessment might miss some of the potential methods so this should be kept in mind when considering using the framework in the future.

5.5 Summary

This chapter provided the results of the AFWD subjective and objective assessment. The AFWD subjective assessment was to measure how secure the framework is from user’s perspective and whether or not they are willing to use it once available. The AFWD objective assessment was based on Bonneau et al.'s framework that assesses authentication schemes based on security, usability, and deployability. Both assessments showed that the AFWD provides usability, deployability, and security, which encourages going further to implement the suggested framework.
Chapter 6

Conclusions and Future Work

6.1 Contributions

6.1.1 Main Contribution: The Authentication Framework for Wearable Devices (AFWD)

WDs are an immature technology that lacks a viable authentication mechanism that is capable of protecting its sensitive data. Solving such a problem is constrained by limitations such as the unique form factor of WDs, which requires the device to always be on, and the lack of input methods that might be used in traditional authentication methods, such as passwords or PINs. This research aims to solve this problem by developing the AFWD, which is a framework that works as a basis to build a transparent and continuous authentication method for a WD that protects its sensitive data and respects its limitations at the same time.
(a) WDs do not yet have a viable authentication method to protect the integrity of this private data. The AFWD addresses this issue by providing the ground to create an effective security mechanism for WDs based on their sensors’ data, which can be used in multi-factor and multi-modal biometric authentication.

(b) WDs have a unique form factor in which they have to be always on and always accessible [5]. Balancing the sensitive data protection and keeping the device always available for authorized users is challenging. Therefore, a WD must have an authentication method that respects this unique form factor. The AFWD respects this unique form factor by authenticating wearers transparently as they go about their regular tasks.

(c) The lack of a keyboard or a screen limits the possibility of implementing traditional authentication methods such as a password or PIN in WDs. The AFWD addresses this issue by authenticating the wearer based on factors, such as behavioral biometrics that do not require input means nor extra equipment to gather them other than the WD’s built-in resources.

(d) An authentication method for WDs should not annoy wearers and should require as little effort as possible from them. The AFDW-based authentication is less annoying to wearers because it depends on factors such as behavioral biometrics that do not require much effort from the wearer compared to other methods.
6.1.2 Other Contributions

Working on this solution has resulted in other contributions to the field of WD authentication as follows:

(a) A study was conducted using a smartwatch and smartphones on 36 participants to show that behavioral biometrics can be used to identify the WD’s owner.

(b) A subjective and objective assessment of the AFWD was done. The subjective assessment involved 27 participants and it was conducted to evaluate the security of the AFWD from the user’s perspective. The objective assessment was conducted based on Bonneau et al.’s. [73] framework, which evaluates the usability, deployability, and security of authentication methods.

(c) A method or a taxonomy has been created to classify devices as wearable, mobile, portable, or other. This method not only helps to identify whether or not the AFWD supports the targeted devices but it also can be used for other classification purposes as there is no such method in the field.

6.2 Research Question and Research Hypothesis

This research was driven by a research question and related hypothesis. The research question is: In terms of error rate, how trustworthy is an AFWD-
based authentication method in identifying whether or not the current WD’s wearer is its owner in a transparent and continuous way? Does this method minimize the wearer’s effort and respect the form factor, the diversity, and the limitations of the WD?

The hypotheses that were drawn from the above research question and how they were addressed are as follows:

**H1:** It is possible to create an authentication method for WDs that is transparent, continuous, trustworthy, and respectful of their unique form factor and limitations.

This hypotheses is *supported*. The AFWD provides the grounds to create such a method. The transparency is satisfied by the ability to collect the authentication data in the background. Examples of this data are behavioral biometrics such as voice and gait. It can be continuous where the AFWD provides the ability to repeatedly check the owner’s identity. It is trustworthy because it uses multi-factor and multi-modal biometric authentications in which the security is improved. This is shown by the results of our hand-worn study presented in Chapter 4. It is also trustworthy and it provides sufficient security based on our subjective and objective assessments as shown by the results of the AFWD assessment study presented in chapter 5. It respects the unique form factor by authenticating users transparently. In addition, it respects limitations of WDs such as the lack of keyboard or screen by using behavioral biometrics.

**H2:** WD sensors such as an accelerometer and gyroscope have low error rates to support using them to determine whether or not the current wearer
of the WD is its owner.

This hypothesis is supported. As previously mentioned, the error rates reported in the hand-worn authentication study is not low enough to make a solid authentication decision. However, we say this hypothesis is supported for the following reasons. First, it was an authentication based on behavioral biometrics, which are known to have low distinctiveness compared to physiological biometrics [105]. Second, in a real world implementation of the AFWD, it is expected that more biometrics would be used than we used in our study, which would provide a lower error rate. The evidence of this claim was shown by the results of combining accelerometer data from both the smartwatch and the hand-phone, which produced the lowest error rate in our study.

H3: Wearers feel that an AFWD-based authenticating method is secure and prefer to use it, if available, over their current authentication method to protect their private data.

This hypotheses is supported. This decision is based on our study result which was to evaluate the AFWD security from the user’s perspective. The majority of the 27 participants believe that the three security principles, confidentiality, integrity, and availability are not affected in the AFWD. The majority of them would also prefer the AFWD-based authentication method, if available, over their current one.

Finally, to answer the research question: based on the fairly low FAR (5%) and high AUC (91%) reported in this research and based on the way we gathered data, we can create an authentication method that provides a trustwor-
thy, transparent, and continuous authentication based on the AFWD. This method minimizes the wearer’s effort and respects the form factor, the diversity, and the limitations of the WD

### 6.3 Future Work

The work presented here is just the beginning of a long journey to provide the best authentication possible to protect sensitive data in WDs. WDs have been improving on a frequent basis as more advanced technologies are being implemented. These include a faster processor, longer battery life, as well as new and improved sensors. This improvement is opening a door of research to enhance the security of WDs.

Some work that can be done in the near future to enhance WDs’ security and we list some aspects below:

(a) A simulation of the AFWD is needed so it can be studied and evaluated in a realistic environment. This practical simulation would help researchers to see the effectiveness of the AFWD in terms of how much security it provides, how much effect it would have on WDs, how much data is sufficient for authentication, how much data the device can handle without complications, and other related issues.

(b) Although the error rates reported in the study of the hand-worn device study in Chapter 4 are promising, it is still not low enough to solely rely on. More research and investigation are needed to improve the error rate. This might include exploring new ways to improve the result
of walking-based authentication, such as using different classification algorithms or extracting more features. Moreover, different behavioral biometrics such as heartbeat and body temperature should also be investigated.
Bibliography


Appendix A

The AFWD Assessment

Questionnaire

A.1 Demographics Questions

(a) ID [...] 

(b) What is your gender? [...] 

(c) What is your age?
   i. 18-24
   ii. 25-34
   iii. 35-44
   iv. 45-55
   v. 56-64
   vi. 65-74
vii. 75 or older

(d) Current Education Level:
   i. High school
   ii. Undergraduate level
   iii. Graduate level
   iv. Other

(e) Are you a smartphone user?
   i. Yes
   ii. No

(f) If yes, what brand do you own?
   i. iPhone
   ii. Android-based devices
   iii. Windows Phone
   iv. Not applicable

(g) Are you a wearable device owner?
   i. Yes
   ii. No

(h) If yes, what type of wearable device?
   i. Hand-worn devices (e.g.: Smartwatch fitness devices)
   ii. Head-mounted (e.g.: Smartglasses)
   iii. Body-worn device (e.g.: pedometer and heart rate monitor)
   iv. Foot-worn device (e.g.: Nike+)
   v. No applicable
A.2 Per-Task Assessment Questions

(a) Did you successfully view the email?
   i. Yes
   ii. No

(b) Assume that we have three levels of security: low, medium, and high.
   In your opinion, what level of security does this task require?
   i. Low
   ii. Medium
   iii. High

(c) Because of the level of protection the AFWD provides, I feel that I am the only person who will be able to view my email on the smartwatch:
   i. Strongly Agree
   ii. Agree
   iii. Slightly Agree
   iv. Slightly Disagree
   v. Disagree
   vi. Strongly Disagree

(d) Did successfully delete the email?
   i. Yes
   ii. No

(e) Because of the level of protection the AFWD provides, I feel that I am the only person who will be able to delete an email:
   i. Strongly Agree
ii. Agree

iii. Slightly Agree

iv. Slightly Disagree

v. Disagree

vi. Strongly Disagree

(f) With of the level of protection the AFWD provides, I feel that I will be able to view my email (e.g. will not be prevented from accessing) it whenever I want to:

i. Strongly Agree

ii. Agree

iii. Slightly Agree

iv. Slightly Disagree

v. Disagree

vi. Strongly Disagree

(g) Did you successfully view your bank task?

i. Yes

ii. No

(h) Assume that we have three levels of security: low, medium, and high. In your opinion, what level of security does this task needs? *

i. Low

ii. Medium

iii. High
(i) Because of the level of protection the AFWD provides, I feel I'm the only person who will be able to view my bank account:

   i. Strongly Agree

   ii. Agree

   iii. Slightly Agree

   iv. Slightly Disagree

   v. Disagree

   vi. Strongly Disagree

(j) Did you successfully pay your bill?

   i. Yes

   ii. No

(k) Because of the level of protection the AFWD provides, I feel I'm the only person who is able to withdraw money from my bank account:

   i. Strongly Agree

   ii. Agree

   iii. Slightly Agree

   iv. Slightly Disagree

   v. Disagree

   vi. Strongly Disagree

(l) With the level of the protection that the AFWD provides, I feel that I'm able to view my bank account and I will not be prevented from accessing it whenever I want to:

   i. Strongly Agree
ii. Agree

iii. Slightly Agree

iv. Slightly Disagree

v. Disagree

vi. Strongly Disagree

(m) Assume that we have three levels of security: low, medium, and high. In your opinion, what level of security does this task require?

i. Low

ii. Medium

iii. High

(n) Because of the level of protection the AFWD provides, I feel I'm the only person who is able to view the alarm list on my smartwatch:

i. Strongly Agree

ii. Agree

iii. Slightly Agree

iv. Slightly Disagree

v. Disagree

vi. Strongly Disagree

(o) Did you successfully delete an alarm from the alarm list?

i. Yes

ii. No

(p) Because of the level of protection the AFWD provides, I feel that I'm the only person who will be able to delete an alarm on my smartwatch:
(q) With the level of the protection that the AFWD provides, I feel that I will be able to view my alarm list and will not be prevented from accessing it whenever I want to:

i. Strongly Agree  
ii. Agree  
iii. Slightly Agree  
iv. Slightly Disagree  
v. Disagree  
vi. Strongly Disagree

A.3 General AFWD Assessment Questions

(a) I feel that using an authentication method that is based on the AFWD will prevent unauthorized access to my sensitive data:

i. Strongly Agree  
ii. Agree  
iii. Slightly Agree  
iv. Slightly Disagree  
v. Disagree  
vi. Strongly Disagree
iv. Slightly Disagree
v. Disagree
vi. Strongly Disagree

(b) I feel that using an authentication method that is based on the AFWD will prevent any unauthorized alteration of my sensitive data such as editing or tampering:

i. Strongly Agree
ii. Agree
iii. Slightly Agree
iv. Slightly Disagree
v. Disagree
vi. Strongly Disagree

(c) I feel that using an authentication method that is based on the AFWD will allow me to access my data whenever I need to:

i. Strongly Agree
ii. Agree
iii. Slightly Agree
iv. Slightly Disagree
v. Disagree
vi. Strongly Disagree

Compared to my current authentication method, I feel that an authentication method based on AFWD will be:

i. More secure
ii. Similar

iii. Less secure

(d) Having multi factor authentication such as that provided by the AFWD would encourage me to use it compared to a single factor method such as password or fingerprint.

i. Strongly Agree

ii. Agree

iii. Slightly Agree

iv. Slightly Disagree

v. Disagree

vi. Strongly Disagree

(e) The idea of the transparent authentication implemented in the AFWD will encourage me to use AFWD-based authentication when available:

i. Strongly Agree

ii. Agree

iii. Slightly Agree

iv. Slightly Disagree

v. Disagree

vi. Strongly Disagree

3. How do you describe the AFWD (please select all that applies):

(a) Secure

(b) Protected
(c) Efficient
(d) Reliable
(e) Easy to use
(f) Complex
(g) Learnable
(h) Confusing
(i) Vulnerable
(j) Trustworthy
(k) Effective
(l) Optional
(m) Time consuming
(n) Necessary