UQAC
Université du Québec
à Chicoutimi

A PRACTICAL APPLICATION OF A TEXT-INDEPENDENT SPEAKER AUTHENTICATION SYSTEM ON MOBILE DEVICES

BY

FLORENTIN THULLIER

B.Sc.A.

THESIS PRESENTED TO THE UNIVERSITÉ DU QUÉBEC À CHICOUTIMI IN PARTIAL FULFILLMENT OF THE REQUIERMENTS FOR THE DEGREE OF MASTER OF SCIENCE IN THE SUBJECT OF COMPUTER SCIENCE

Québec, Canada

ABSTRACT

The growing market of mobile devices forces to question about how to protect users' credentials and data stored on such devices. Authentication mechanisms remain the first layer of security in the use of mobile devices. However, several of such mechanisms that have been already proposed were designed in a machine point of view. As a matter of fact, they are not compatible with behaviors human have while using their mobile devices in the daily life. Consequently, users adopted unsafe habits that may compromise the proper functioning of authentication mechanisms according to the safety aspect.

The first main objective of this research project is to highlight strengths and weaknesses of current authentication systems, from the simpler ones such as PIN (Personal Identification Number) to the more complex biometric systems such as fingerprint. Then, this thesis offers an exhaustive evaluation of existing schemes. For this evaluation, we rely on some existing criteria and we also propose some new ones. Suggested criteria are chiefly centered on the usability of these authentication systems.

Secondly, this thesis presents a practical implementation of a text-independent speaker authentication system for mobile devices. We place a special attention in the choice of algorithms with low-computational costs since we want that the system operates without any network communication. Indeed, the enrollment, as well as the identification process are achieved onto the device itself. To this end, our choice was based on the extraction of Linear Prediction Cepstral Coefficients (LPCCs) (Furui 1981; O'Shaughnessy 1988) to obtain relevant voice features and the Naïve Bayes classifier (Zhang 2004) to predict at which speaker a given utterance corresponds. Furthermore, the authentication decision was enhanced in order to overcome misidentification. In that sense, we introduced the notion of access privileges (*i.e.* public, protected, private) that the user has to attribute to each application installed on his/her mobile device. Then, the safest authority is granted through the result of the speaker identification decision as well as the analysis of the user's location and the presence of a headset.

In order to evaluate the proposed authentication system, eleven participants were involved in the experiment, which was conducted in two different environments (*i.e.* quiet and noisy). Moreover, we also employed public speech corpuses to compare this implementation to existing methods. Results obtained have shown that our system is a relevant, accurate and efficient solution to authenticate users on their mobile devices. Considering acceptability issues which were pointed out by some users, we suggest that the proposed authentication system should be either employed as part of a multilayer authentication, or as a fallback mechanism, to cover most of the user needs and usages.

RÉSUMÉ

La croissance du marché des dispositifs mobiles implique de se questionner au sujet de comment protéger l'identité ainsi que les données personnelles des utilisateurs qui sont stockées sur ces appareils. En ce sens, les mécanismes d'authentification demeurent la première couche de sécurité dans l'utilisation des mobiles. Cependant, il apparaît que la plupart des mécanismes d'authentification qui ont été proposés, ont été conçus suivant un point de vue orienté machine plutôt qu'humain. En effet, ceux-ci ne s'adaptent généralement pas avec l'usage quotidien qu'ont les utilisateurs lorsqu'ils se servent leur téléphone. En conséquence, ils ont adopté des habitudes dangereuses qui peuvent compromettre le bon fonctionnement des systèmes d'authentification. Celles-ci peuvent alors remettre en question la sécurité de leur identité ainsi que la confidentialité de leur contenu numérique.

Le premier objectif principal de ce projet de recherche est de faire ressortir les forces et les faiblesses des méthodes d'authentification qui existent actuellement, des plus simples comme le NIP (Numéro d'Identification Personnel) aux solutions biométriques plus complexes comme l'empreinte digitale. Par la suite, ce mémoire offre une évaluation exhaustive de ces solutions, basée sur des critères existant ainsi que de nouveaux critères que nous suggérons. Ces derniers sont majoritairement centrés sur l'utilisabilité des mécanismes d'authentification qui ont été examinés.

Dans un second temps, ce mémoire présente une implémentation pratique, pour périphériques mobiles, d'un système d'authentification d'orateur indépendant de ce qui est prononcé par l'utilisateur. Pour concevoir un tel système, nous avons porté une attention particulière dans le choix d'algorithmes admettant un faible temps d'exécution afin de se prémunir des communications réseau. En effet, ceci nous permet alors de réaliser le processus d'entraînement ainsi que la reconnaissance, directement sur le mobile. Les choix technologiques se sont arrêtés sur l'extraction de coefficients spectraux (Linear Prediction Cepstral Coefficients) (Furui 1981; O'Shaughnessy 1988) afin d'obtenir des caractéristiques vocales pertinentes, ainsi que sur une classification naïve bayésienne (Zhang 2004) pour prédire à quel utilisateur correspond un énoncé donné. La décision finale, quant à elle, a été améliorée afin de se prémunir des mauvaises identifications. En ce sens, nous avons introduit la notion de droits d'accès spécifiques (*i.e.* publique, protégé ou privé) que l'utilisateur doit attribuer à chacune des applications installées sur son mobile. Ensuite, l'autorisation d'accès la plus adaptée est accordée, grâce au résultat retournée par l'identification de l'orateur, ainsi que par l'analyse de la localisation de l'utilisateur et de l'emploi d'un micro-casque.

Pour réaliser l'évaluation du système que nous proposons ici, onze participants ont été recrutés pour la phase d'expérimentation. Cette dernière a été menée dans deux types d'environnements différents (*i.e.* silencieux et bruyant). De plus, nous avons aussi exploité des corpus de voix publiques afin de comparer notre implémentation à celles qui ont été proposées par le passé. Par conséquent, les résultats que nous avons obtenus ont montré que notre système constitue une solution pertinente, précise et efficace pour authentifier les utilisateurs sur leurs périphériques mobiles. Compte tenu des problèmes d'acceptabilité qui ont été mis en avant par certains testeurs, nous suggérons qu'un tel système puisse être utilisé comme faisant part d'une authentification à plusieurs facteurs, mais aussi comme une solution de repli, en cas d'échec du mécanisme principal, afin de couvrir la majorité des besoins et des usages des utilisateurs.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

ATAP:	ADVANCED TECHNOLOGY AND PROJECTS
ATM:	AUTOMATED TELLER MACHINE
DBFS:	DECIBELS RELATIVE TO FULL SCALE
DTW:	DYNAMIC TIME WARPING
ECG/EKG:	ELECTROCARDIOGRAM
EER:	EQUAL ERROR RATE
FAR:	FALSE ACCEPTANCE RATE
FMR:	FALSE MATCH RATE
FNMR:	FALSE NON-MATCH RATE
FRR:	FALSE REJECTION RATE
GMM:	GAUSSIAN MIXTURE MODEL
HMM:	HIDDEN MARKOV MODEL
IOT:	INTERNET OF THINGS
LPC:	LINEAR PREDICTION COEFFICIENTS
LPCC:	LINEAR PREDICTION CEPSTRAL COEFFICIENTS
MFCC:	MEL-FREQUENCY CEPSTRAL COEFFICIENTS
PCM:	PULSE-CODE MODULATION
PGA:	PICTURE GESTURE AUTHENTICATION
PIN:	PERSONAL IDENTIFICATION NUMBER
ROC:	RELATIVE OPERATING CHARACTERISTIC
SPH:	NIST SPHERE FORMAT

VQ: VECTOR QUANTIZATION

WAV: WAVEFORM AUDIO FILE FORMAT

ACKNOWLEDGEMENTS

First, I would like to express my most sincere gratitude toward my master director Bob-Antoine Jerry Ménélas, assistant professor of Computer Science at Université du Québec à Chicoutimi, as well as my co-director, Bruno Bouchard, for their continuous support, their encouragement, their patience and their valuable advices, which were a key element in the success of this research project.

I also would like to acknowledge every person who participates to the completion of this thesis, to wit, my laboratory collaborators for their support and their help, every reviewer for their conscientious opinion as regards my publications and experimenters for the time they have given in testing my application.

Finally, I would like to express my warmest thanks to my parents for their moral and financial support 5,000 kilometers far from here.

CHAPTER I

INTRODUCTION

1.1 RESEARCH CONTEXT

From the past decade, the market of mobile devices grew up exponentially. The Gartner Institute noticed that smartphone sales were over 400 thousand units in the second semester of 2014 (Laurence Goasduff 2015). These devices take a significant place in people's everyday life. Indeed, people, and more specifically young ones have their mobile devices everywhere and at any time (Wilska 2003). They consider their mobiles as an important part of their life (Goggin 2012). Moreover, according to data from the Nielsen Company, users spent more than 30 hours using applications on such devices in the fourth quarter of 2013 (Nielsen 2014). Besides, it should be noted that a major player of the mobile device industry used to claim that there is an application for everything.

Technical advancements in mobile devices led to significant evolution on the way people communicate and exchange pieces of knowledge. As a result, users do store private data such as pictures, videos, as well as secret information about their personal accounts (*i.e.* emails, social networks, bank accounts) on their devices. These personal contents may be considered as individuals' digital identity. However, they are, most of the time, not adequately wary about the safety of information they save on their mobiles (Falaki *et al.* 2010).

Recently, major information technology companies such as Apple, Google, Microsoft and Yahoo have concentrated efforts in offering concrete novel authentication mechanisms to improve the safety of private contents stored on mobile devices. Indeed, both Apple and Google have suggested a fingerprint solution (*i.e.* Touch ID and Nexus Imprint respectively) to make their digital wallet platforms more secure (*i.e.* Apple Pay and Android Pay respectively). Moreover, Google has also provided, for a few years, several other context-aware and biometric schemes under the Smart Lock feature to secure mobiles' access. In the same way, Microsoft aims at improving the security of every device running the latest major release of their operating system (*i.e.* Windows 10) since they have announced the Windows Hello feature. This feature will allow users to unlock their devices either with their face, iris, or fingerprint. Finally, Yahoo also get involved in biometrics, since they have introduced an under-researched biometric authentication system which allows users to perform the authentication both with their ear and the geometry of their phalanges.

Although latest advances in such a field of research offers more accurate and more reliable authentication solutions, such methods do not take into account several issues induced by users which involve several security threats.

1.2 RESEARCH PROBLEM

Authentication may be defined as the process of an entity that has to become sure of the identity of another one (Lowe 1997). Such a process remains a crucial concern with respect to modern computer systems and more specifically within a mobile

device context. However, proposed authentication mechanisms concede several well-known drawbacks principally led by people's usage and behavior.

Whenever we are facing a machine that has to deal with human users, therefore, we have an interactive system (Benyon *et al.* 2005). As pointed out by the authors, a fundamental challenge with interactive systems is that human beings and machines have different characteristics. Indeed, what may be seen as strengths from a machine's point of view, may also be a weakness for human being. On the one hand, machines can see humans as being vague, disorganized and emotional while they are precise, orderly, unemotional and logical. On the other hand, humans may claim to be attentive, creative, flexible and resourceful while machines are dumb, rigid and unimaginative. Such differences suggest that the key challenge is first to understand the human rather than design an interactive system from the machine's point of view. However, when considering authentication systems that have been proposed over the last three decades, it seems that they have been designed without any concern for the human. As an example, it was reported that half of the population does not lock their phone at all since they estimate that entering a 4-digit-PIN code involves lots of trouble, every time the mobile device has to be unlocked (Ben-Asher *et al.* 2011). Moreover, it is known that users have trouble remembering all passwords they use nowadays (Yan 2004). Consequently, some people prefer to reuse the same password in multiple situations while others choose to write them down. In the same way, since biometric systems do exploit confidential information about the user, their acceptance rate remains slow whereas their devices possess either software, or hardware requirements. It is clear that behaviors reported here may lead a huge impact on the security of mobile devices. Accordingly,

people's authentication usage may generate serious threats for the security that a system initially provides. In fact, an effective mechanism may become a weak one because it is not used as recommended.

Most of proposed authentication mechanisms were also designed in order to be employed with traditional desktop computers, where few of them require costly hardware (*e.g.* retina or iris recognition). Moreover, these methods may involve several series of complex computation, so that, the first mobile devices that appear on the market may not be powerful enough to run these costly processes. Although the continuous growth of mobiles' capabilities with respect to embedded hardware, the need for efficient and reliable authentication solutions on these devices appear to be vital.

Accordingly, the first issue we point out in this research project is that proposed authentication schemes do not take into account most of users' needs and behaviors with respect to his/her involvement in the authentication process. In addition, we identify the importance to provide accurate, effective and reliable solutions which are able to run on weakest mobile devices present in the market, without the need of an extra piece of hardware than the ones already embedded.

1.3 THESIS CONTRIBUTION

The main contribution of this thesis is to design a novel approach in the field of authentication mechanisms for mobile devices, which take into account problems expressed in the previous section (user friendliness and computational cost).

To this end, we first propose a detailed taxonomy of authentication mechanisms used on mobile devices, where both their strengths and their weaknesses are highlighted. Such a review conducted us to suggest an exhaustive evaluation based on the work of Jain *et al.* (2006a), as well as several criteria we introduce, chiefly focused on the usability of these mechanisms for the user. This evaluation allows us to position this research in order to design a system which responds to identified problems.

Moreover, we present a practical authentication system implementation assessed through a rigorous experiment. This experiment was conducted thanks to the participation of several trial users. Hence, thanks to the test results we obtained and several analyses such as computation performances, as well as participants' opinion regarding broadly authentication mechanisms and their opinion concerning our system, we position the proposed system over existing ones.

1.4 THESIS PROJECT ORGANIZATION

This thesis is organized into five main chapters. This first chapter was an introduction to the research project. In the first place, the context of the research was presented. Then, a summary of problems and issues raised by this area of research were highlighted.

The second chapter provides an exhaustive review of current authentication systems used on mobile devices. First, this chapter offers a description of each mechanism classified in three main families: knowledge-based and token-based systems, as well as biometric solutions. We focus on offering practical examples and use case scenarios.

Moreover, both strengths and weaknesses for each scheme are highlighted. Given that, we suggest a detailed evaluation for each authentication family in order to state the future of authentication mechanisms and to guide the design step of the system we introduce.

The third chapter first presents an overview of the state of the art in the field of voice-based authentication mechanisms and more precisely, text-independent speaker authentication systems. Then, the proposed system is detailed. A particular attention in technological choices was paid since we wanted to design a standalone solution for mobile devices to avoid a client/server architecture, which should be able to work properly on weakest mobile devices present in the market.

The fourth chapter introduces the experiment we performed to collect data in order to assess the accuracy and the reliability of our system. Then, several evaluations are proposed for comparison purposes regarding classification performance on few other well-known data sets and computation performance. In addition, the last section of this chapter exposes the opinion collected from participants of the experiment. In that sense, the results obtained allow us to understand the viability of the proposed system as regards to users' needs and behaviors.

Finally, the fifth and last chapter draws a general conclusion for this research. Furthermore, the last part of this thesis ends with a personal assessment of such an experience as a first step into the world of scientific research.

CHAPTER II

AUTHENTICATION MECHANISMS ON MOBILE DEVICES

2.1 INTRODUCTION

Over the past few years, various authentication schemes have been proposed. Hence, we divided them into three broad categories: knowledge-based, token-based and biometrics (Li 2009). Figure 2.1 illustrates the authentication mechanisms that are currently employed with mobile devices.

First of all, knowledge-based authentication schemes focus on what the user knows. Precisely, we differentiate implicit and explicit knowledge-based mechanisms. Explicit ones imply that the user has to retain new data like a 4-digit-PIN code or a password (Lamport 1981). Whereas, implicit knowledge-based mechanisms call upon cognitive functions of the user, to exploit the data they already know (Zviran and Haga 1990).

Secondly, token-based mechanisms need the user to prove they possess a physical token that often involves a two-factor authentication process (Jing *et al.* 2009). As an example, we can mention the smartcard that the user needs to own to authenticate himself or herself on his/her mobile phone.

Finally, biometric mechanisms rely on the uniqueness of users' physiological or behavioral trait to perform the authentication process. Consequently, we subdivide

biometrics into physiological and behavioral sets. Physiological biometrics exploit singularities of the human body like fingerprints (Jain *et al.* 1997) while behavioral ones require users to perform some actions to prove their identity such as gait (Gafurov *et al.* 2006).

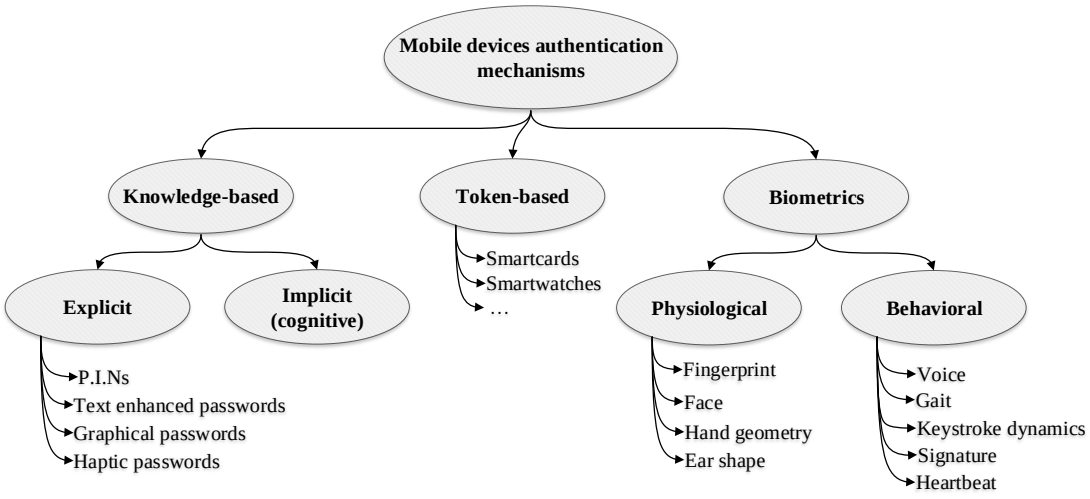

FIGURE 2.1: TAXONOMY OF MOBILE DEVICES' AUTHENTICATION MECHANISMS

The main contribution of this chapter is to emphasize, through a critical analysis, most of the weaknesses of proposed authentication schemes, in real-life situations with a particular focus for implementation on a mobile device. We previously pointed out that users have adopted unsafe habits because of the non-user-centered design of most of the authentication process. Hence, our work aims at analyzing how proposed mechanisms whether suit or not characteristics of human users and why they are not entirely appropriate to their needs.

The remainder of this chapter is organized as follows. Firstly, it will review authentication schemes that have been proposed in a mobile device context, where, for

each of them, both strengths and weaknesses are highlighted. Then, we offer an evaluation guided by criteria that were previously proposed by Jain *et al.* (2006a) in the field of authentication. In addition to, we will increase such an assessment with criteria we suggest that stem from our critical analysis. Finally, this evaluation conducts us to draw a conclusion and state about the aftermath of authentication mechanisms for mobile devices.

2.2 KNOWLEDGE-BASED AUTHENTICATION MECHANISMS

Knowledge-based authentication mechanisms rely upon users' ability to recall secret information. It is possible to distinguish two different kinds of knowledge-based techniques: explicit and implicit schemes. On the first hand, explicit ones need the user to set and learn a piece of knowledge. On the other hand, implicit ones exploit the user memory thanks to either, or both, personal information they already know, or about their everyday life preferences (*e.g.* music they like or food they enjoy).

This section describes in detail both explicit and implicit techniques and exposes that users' capacity to remember a secret remains a common denominator in the weakness of each knowledge-based authentication scheme.

2.2.1 STRENGTH EVALUATION OF KNOWLEDGE-BASED AUTHENTICATION SCHEMES

The strength of a knowledge-based authentication scheme is theoretically measurable through the evaluation of the entropy of the password space. The entropy of a

password represents the measurement of how unpredictable a password is and the password space is the total number of distinct possibilities the authentication system can support. Hence, the evaluation of the entropy of the password space provides a metric of how strong an authentication system is. The size S of the password space for a system having N possible entries is given by the equation (2.1). The length of the input to retain is expressed by k. Finally, the entropy H can be computed by using the equation (2.2), and the result is expressed in bits.

$$S = N^k. \tag{2.1}$$

$$H = log_2(S). \tag{2.2}$$

Real use cases reveal that such an evaluation still not represents an accurate measure of the strength of a knowledge-based authentication mechanism. Indeed, since users have the possibility to choose their own secret input, they often refer to a familiar pattern rather select it randomly. As an example, Yampolskiy (2006) has pointed out that 47.5% of the users chose a family-oriented information as secret input such as a child's name or date of birth. Therefore, a lower subset of the N possibilities is truly used, since the length of passwords are generally less than 8 characters (Yan 2004).

2.2.2 EXPLICIT SCHEMES

Personal Identification Numbers (PINs)

Example of use case scenario: commonly, users have to choose an array of 4-digits that they will need to remember. Then, each time the mobile device has to be unlocked, the system prompts an input field where the user needs to fill these digits in the correct order to be authorized to access the whole content of the device.

Personal Identification Numbers (PINs) are a simple way to restrain access to an entity due to their composition—from 4 to 16 digits. They appear with the growth of ATMs (Automated Teller Machines), and they are mostly used in the banking system. Regarding a mobile device context, PINs currently remain the most dominant authentication method to protect the access of these devices since they are employed by 2/3 of the mobile device users (Clarke and Furnell 2005). PINs can be applied to both the device and the user's Subscriber Identity Module (SIM)—a removable token that contains required cryptographic keys for network access. Both of the two leading mobile device operating systems (*i.e.* Android and iOS) provide this authentication mechanism.

However, PINs present several issues considering memorability or human habits that may compromise the security offers by the system. In that sense, Clarke and Furnell (2005) have assessed that 1/3 of mobile devices users who keep their phone locked through a 4-digits PIN method, consider such protection as an inconvenience in everyday life. As a result, users do need to retain a code that has a familiar signification, such as their date of birth (Yampolskiy 2006). Furthermore, Clarke and Furnell (2005)

also enhance the weakness of this authentication scheme. Indeed, 36% of the respondents have reported using the same PIN-code for multiple services. Thus, it becomes easier for an attacker to determine the correct 4-digits PIN in order to have free access to several other services where the code is employed. The lack of security brought by users can also be underlined through another study that reports that 26% of PIN users shared their own code with someone else (Clarke *et al.* 2002).

While PINs still remain very popular, they may also be considered as the weakest authentication mechanism on the market as they offer a theoretically low entropy (*i.e.* $log_2(10^4)$). Indeed, people adopted several behaviors to cope with the large cognitive load that the system requires. All of this lead PIN authentication to be largely vulnerable to several attacks such as code guessing by social engineering, brute force, or shoulder surfing attacks (Hansman 2003; Orgill *et al.* 2004; Tari *et al.* 2006; Jagatic *et al.* 2007). Social engineering refers to the art of manipulating people to obtain from them few confidential piece of information to get hints about users' PINs code that may be a date of birth. Shoulder surfing happens when a criminal is looking over the shoulder of a user while he/she is performing tasks on his/her mobile device such as unlock it with a PIN code.

Text-enhanced passwords

Example of use case scenario: conversely to PINs, users have to select, at least, an array of 6 characters that are not restricted to only digits. The whole set of characters offered by the keyboard of the mobile device is legitimate. Next, the authentication depends on the same process as the PIN one.

As opposed to PINs, text-enhanced passwords are more complex. Usually, they are composed of several different characters such as lower and upper case letters, digits, and also non-alphanumeric characters. At first, passwords were stored in plain text files without any encryption (Morris and Thompson 1979). Thereby, protect such sensitive information became crucial for numerical system. This mechanism is also provided by both Android and iOS platforms. However, as plainly regards authentication on mobile devices, text-enhanced passwords remain less popular than PINs among mobile device users. Indeed, authenticate users with a complex string of characters is an inherited process that comes from traditional computing and it was not revised at all before its arrival on mobile devices.

The market of mobile applications is vast. As an example, the Apple Store counted up to 600,000 applications, and the total number of downloads surpassed 25 billion in the year 2012 (Cuadrado and Dueñas 2012). Through this immense inventory, numerous applications require the user to login in order to have access to the entire set of features they offer (Sasse *et al.* 2001). Consequently, since text-enhanced passwords also take part in the daily usage of mobile devices as regards authentication; it is now important to illustrate the weakness of this mechanism for users' memory and how such issues lead to affect mobile device security due to behaviors they adopt. Passwords remain theoretically a strong way to secure a system. However, they are usually long and sophisticated. Hence, much more memorizing abilities are required for the user. Indeed, Yan (2004) identifies that without the memorization problem, the maximally secured password would be blended with the maximum number of characters allowed

by the system, randomly arrange. This is possible to do for a machine, but almost impossible to retain for a human. Moreover, text-enhanced passwords do have better entropy than PINs (*i.e.* $log_2(94^k)$) where k is the password length and 94 is the number of printable characters excluding SPACE. Although, because they are everywhere, and because they were designed to a machine point of view, passwords represent a mechanism not as good as claimed. Hence, a study conducted by Riley (2006) shows that more than half password users have conceded using the exact same string of characters for multiple accounts on numerical systems. Moreover, about 15% of them admit that they used to write down their list of passwords in case they forget them, while 1/3 also report using the remember my password function, to produce another password than the one they originally set up. The growth of the number of numerical services we use in our everyday life affects significantly the usage of passwords we have. Another study also highlights the deficiency of this mechanism as they released the "worst passwords" list, which exposed some examples such as "*123456"*, "*password"* or "*qwerty"* that are frequently set up by users (Slain January 19, 2015). Such examples formed perfect cases of vulnerability for the security of mobile devices. Just as PINs, text-enhanced passwords are strongly exposed to brute force, dictionary, social engineering, and shoulder surfing attacks (Hansman 2003; Orgill *et al.* 2004; Tari *et al.* 2006; Jagatic *et al.* 2007).

Graphical passwords

Example of use case scenario: with graphical passwords, users have to recall some pieces of visual information. There are a lot of various implementations, but the

most well-known is probably the one implemented in Android that appears in the earliest version of the mobile operating system. First, the user has to set up a path between dots in a matrix, as shown by the gray stroke in Figure 2.2. To be granted to the full access of the mobile device, the user has to reproduce the path he/she initially set-up. The order of dots where the path is passing by is essential. As illustrated in Figure 2.2, if the user defined a path, from the lower-left-corner dot, to the upper-left-corner dot, the inverse drawing (from the upper corner, to the lower one) will not genuinely authenticate the user.

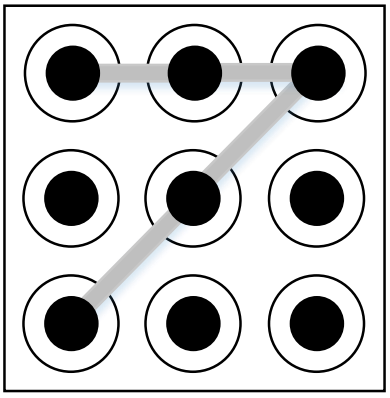

FIGURE 2.2: ANDROID IMPLEMENTATION OF GRAPHICAL PASSWORD

Knowing that humans have better abilities for recognizing and recalling visual information when compared to verbal or textual information (Kirkpatrick 1894), other mechanisms were imagined to use a graphical scheme instead of a sequence of characters as passwords. Since the patent was introduced by Blonder (1996), multiple schemes have been designed. Biddle *et al.* (2012) have grouped these proposed systems into three broad categories: *recall-based systems*; *recognition-based systems*, and *cued-recall systems*.

First of all, in a recall-based system, the first step is to choose a predefined pattern. Then, the user is presented with a selected image or a blank canvas, where the secret sketch has to be reproduced each time he/she wants to authenticate himself or herself. Secondly, in a recognition-based system the user is invited to select a sequence of predefined images among several others. The number of presented images is generally limited to ensure the usability. Finally, in a cue-recall system the user has to recall and target a specific part of a picture. In this way, such systems reduce the memory load that the user needs with recall-based systems. Biddle *et al.* (2012) have pointed out that users are more comfortable to use a graphical password than a digit or a text-based password every day. However, it is known that the Android implementation allows 389,112 possibilities (Uellenbeck *et al.* 2013). Therefore, the password space is not superior to a 6-digit PIN. Moreover, as reported by Uellenbeck *et al.* (2013), users do not exploit the maximum potentiality of the security since some graphical schemes are evident to perform. Another relevant example of a graphical scheme may be the Picture Gesture Authentication (PGA) feature introduced by Microsoft in Windows 8 (Sinofsky 2011). The idea behind this mechanism is to allow users to define some specific gestures (*i.e.* taps, circles and lines) over either pre-defined pictures, or users' personal ones. The whole set of possibilities for such a system largely relies on both the number and the nature of gestures determined. According to Sinofsky (2011), when the user defines two of the most complex gestures (*i.e.* lines), there are 846,183 unique possibilities.

As a matter of fact, as assessed by Biddle *et al.* (2012), it is possible to say that graphical passwords do not offer a higher level of security than PINs or text enhanced

passwords. Indeed, as regards an implementation such as the Android one, Uellenbeck *et al.* (2013) have experienced the ability to find the most common path defined, of numerous graphical schemes. They showed that it was possible to determine the right path statistically by applying a Hidden Markov Model technique on their data set. Concerning gesture-based graphical authentication process, Zhao *et al.* (2015) have demonstrated that the framework they built was able to guess a large portion of the picture passwords set of their study. Moreover, Aviv *et al.* (2010) showed that it was possible to find the graphical scheme through oily residues, or smudges that users leave on the touch-screen surface. They named this vulnerability "smudge attack". Besides, graphical passwords also have the same other vulnerabilities as all knowledge-based authentication mechanisms: social engineering (Orgill *et al.* 2004; Jagatic *et al.* 2007) and shoulder surfing (Tari *et al.* 2006; Lashkari *et al.* 2009). Recently, Gugenheimer *et al.* (2015) have proposed a novel graphical authentication concept that claims to be robust against shoulder surfing. A grid of randomly generated numbers is prompt to the user and the real PIN code is hidden inside. Then, just as the android implementation, the user has to draw the path between each number of the PIN and at the same time, several other paths are painted on the grid. Through this approach, shoulder surfing attacks were reduced down to 10.5% and authors have pointed out that no participant forgot their graphical scheme. However, it is clear that their process involves a high level of memorization to recall the information due to its complexity when compared to a simple 4-digit PIN code.

Haptic passwords

Example of use case scenario: instead of visual information, the user has to recall a sequence of kinesthetic phenomena produced by the mobile devices. The idea is to let the user define his/her own sequence that he/she has to reproduce afterwards to access the mobile device.

With the emergence of modern mobile devices, the desire to exploit haptic in numerical systems, to enhance the user experience, was strong. Mobile devices such as smartphones are composed of a lot of new technologies like touch screens and sensors that provide many more possibilities regarding authentication mechanisms. Consequently, several new knowledge-based authentication schemes have been designed recently. As an example, Bianchi *et al.* (2010) suggested a novel approach through haptic passwords. The initial work of PINs is retained, but the user has to recall a sequence of vibrations scheme instead of a sequence of single digits. As graphic passwords, haptic ones were designed to be more convenient for users than text-enhanced passwords. The implementation proposed by Bianchi *et al.* (2012) attempt to avoid the memorability issues encountered by users and reduce behaviors that conduct to security vulnerabilities. However, this study also highlights that such new authentication mechanisms, still require unreasonable calls from memory. As a result, they are not the answer to fix issues provided by PINs, text passwords, or graphical passwords.

2.2.3 COGNITIVE SCHEMES

Mechanisms we described above are all explicit methodologies for a knowledge-based authentication. However, each person has a unique knowledge.

Thereby, cognitive passwords aim at exploiting personal facts, opinion, and interests as a means of authentication. This process is defined as a challenge-response.

The idea behind these schemes first stems from regular computer security access where users, in addition to a conventional password, have to answer some personal questions to be granted access. However, regarding a mobile device context, such an approach should be more considered. Indeed, as we state in the first chapter of this thesis, users used to store more and more data on their devices. Thus, data such as pictures, music, and information from social media (Zviran and Haga 1990; Lazar *et al.* 2011) may be exploited to build a convenient cognitive process for authentication, revised to be employed with mobile devices.

The experiment led by Bunnell *et al.* (1997) regarding these authentication schemes showed that personal facts were better recalled than others. Nevertheless, people socially close to the user were easily able to guess many answers, that is why, Lazar *et al.* (2011) have proposed a method to personalize cognitive passwords to individual users. Results obtained show that personalization increases the recall of cognitive passwords, but does not help in improving their secrecy.

2.3 TOKEN-BASED AUTHENTICATION MECHANISMS

Example of use case scenario: token-based authentication needs the user to possess a physical piece of hardware which has first to be coupled with the mobile device. Then, the mobile device has to verify credentials of the token to grant access to the user.

Token-based authentication mechanisms require a hardware interaction between the user and his/her device to complete the authentication process. Such mechanisms involve at least a two-factor authentication (multi-factor is used when there are more than two) due to the commitment to attest that both the password is correct, and the user holds the token all along the authentication process. The three major types of tokens are USB token devices, smart cards and passwords that are randomly generated (Council 2005). Such random passwords are generated on a server and usually imply a mobile device in the authentication process, to receive such a secret Aloul *et al.* (2009), but the purpose is not to authenticate a user directly on his/her mobile device.

Thereby, these old implementations—when compared to the existence of mobile devices—are no longer applicable as they were initially designed. Nowadays, we observe the growth of smart objects and connected objects as known as the Internet of Things (IoT) (Miorandi *et al.* 2012). Consequently, modern approaches regarding token-based authentication mechanisms appear to be more convenient with the use of such devices. As examples, smart watches are replacing USB devices while NFC tags will supplant smart cards overtime, since it is possible to bring them everywhere (*i.e.* wallets, clothes). With the fifth version of Android, Google introduced a feature called "Trusted devices". This feature aims at providing an automatic authentication mechanism that uses smart objects the user has to couple to his/her mobile device. As a result, as long as the mobile device detects a connection with the token hardware through Bluetooth, NFC or Wi-Fi, it remains unlocked.

Regarding two-factor authentication mechanisms, Schneier (2005) suggests that "they solve the security problem we had ten years ago, not security problems we

have today". The use of smart objects over the authentication process implies many cases of vulnerability issues. Indeed, whenever someone wears a smart object as a token for the authentication process, the mobile device should not be sighted off. Due to the fact that there is no need to replay the authentication process, the mobile device becomes simply accessible by anyone nearby. Another problematic situation may be observed, where both devices are stolen by the same person.

2.4 BIOMETRICS

For many years, it is known that humans exhibit a various unique set of features. As an example, each human fingerprint describes a unique pattern; in the same way, blood vessels of the retina also have a unique pattern. Biometrics systems exploit these singularities in order to authenticate users. Hence, with a biometric system, there is no need for remembering or recalling any information; instead, the singularity of interest just has to be digitized and compared to the saved one. To this end, several biometric systems necessitate the use of a scanning or recording hardware that is not always adapted to suit with mobile devices. When compared to other approaches, a vast majority of biometric solutions have a greater cost of implementation. Moreover, due to the assessment of data that come directly from the user, biometrics raise some privacy concerns because of the uniqueness of each of us. As a result, even though these systems generally offer a high level of security and a sufficient accuracy, their usage remains limited, particularly with mobile devices.

The review of biometric authentication that we propose below does not heed sophisticated mechanisms such as blood vessels, retina or iris pattern recognitions that

are way too complex in hardware requirements or computational costs, to be applied on mobile devices, presently. Retina-based systems are currently only used in highly classified government and military facilities (Nanavati *et al.* 2002). Moreover, iris pattern recognition requires specific infrared hardware to authenticate users accurately (Wildes 1997; Daugman 2004; Jain *et al.* 2006a). Hence, we focus on offered mechanisms for mobile devices and the ones that may be materialized in coming years.

2.4.1 FINGERPRINT

Example of use case scenario: first, the user has to let her/his device knows the pattern of one or more of his/her fingerprints. To this end, each finger the user wants to use has to be put, several times, on the sensor to digitalize the fingerprint. Then, each time an authentication is required, the user places any previously recorded finger on the sensor to unlock the device.

This technology exploits unique fingerprint patterns that are present in every human's fingers to authenticate users. Ridges that compose the pattern are traditionally classified into *loops*, *arches* and *whorls* motifs. Figure 2.3 illustrates three examples of these patterns.

A) B) C)

FIGURE 2.3: THE THREE RIDGES PATTERN OF THE FINGERPRINT WHERE A) ARCH PATTERN; B) LOOP PATTERN AND C) WHORL PATTERN (SOURCE: FINGERPRINT. 2016, AUGUST 7. IN WIKIPEDIA)

Fingerprint techniques were used for decades using ink to print the pattern onto a piece of paper (Berry and Stoney 2001). However, several sensors were designed to perform the acquisition such as *optical scanning*, *capacitive scanning*, and *ultra-sound scanning (Jain et al. 1997; Kroeker 2002; Jain et al. 2006b)*. Due to the maturity and the flexibility of fingerprint systems it is now the most popular biometric system on mobile devices on the market. Indeed, fingerprint authentication deliver a high accuracy level and may be used in a wide range of environments. Moreover, with the growth of micro-technology and the emergence of mobile devices more and more efficient, fingerprint systems were integrated into these devices as means of authentication. The most well-known example of fingerprints used with mobile devices is the Apple Touch ID technology that comes from the patent of Bond *et al.* (2012). This major player of the mobile device industry builds an extra capacitive sensor in all their latest smartphones that scans sub-epidermal skin layers. Furthermore, they made the acquisition of the fingerprint possible up to 360-degrees of orientation that provide a very high level of accuracy and a small error rate. Moreover, several other phone and tablet

manufacturers also introduced fingerprint sensors built-in their phones. The major difference between each one resides in the location of the sensor to ensure the ease of use for end-users (*i.e.* inside the power button, at the back of the phone).

In addition, Clarke *et al.* (2002) assessed that 74% of mobile device users positively accept fingerprint biometric as means of authentication. However, this mechanism has weaknesses that engender threats in its usage. It has been discovered that "most devices are unable to enroll some small percentage of users" (Nanavati *et al.* 2002). The accuracy of fingerprint scanning may decrease to null when digits are either too wet, or too dry and also too oily. That is the "moisture effect". Fingerprints also tend to deteriorate over the time because of age, wear or tear (Nanavati *et al.* 2002). Furthermore, fingerprint authentication schemes suffer from confidentiality threats, as well as spoofing attacks that question the security they attempt to provide. Indeed, on the one hand, users may be scanned without their consent. On the other hand, Matsumoto *et al.* (2002) proved that artificial fingers either made of silicon or gelatin, were accepted by 11 fingerprint systems during the enrollment procedure.

2.4.2 FACE RECOGNITION

Example of use case scenario: as well as fingerprint authentication process, the user has to let the device capture his/her facial characteristics. To this end, he has to stand in front of the camera of the device while the system is processing the learning of his/her face. Then, if the face is recognized by the system, the user is allowed to access the device.

Face recognition is the most natural means of biometric authentication because humans also perform this evaluation in their everyday interactions. This authentication scheme may also be plainly integrated into an environment that allows image acquisition such as mobile devices, where a large majority of them have a frontal camera. Moreover, methods to acquire an image of the face are assessed as non-intrusive (Jain *et al.* 2006a). Most of the time, facial recognition systems rely on the analysis of facial features such as the position of eyes, nose and mouth and distances between these features (Dabbah *et al.* 2007). The evolution of both hardware and software led facial recognition to become faster and to provide a better level of accuracy than before. Besides, it should be noted that a facial recognition may be continuously achieved. Indeed, the user may perform the authentication by his/her face, and then, the system may automatically verify that it is always the same face that the device is using when it is not in sleep mode.

From a few years, Google has offered a face-based authentication system. Although it was not genuinely successful, improvements that were made in the new version of the operating system still not accurately identifies the user in numerous cases such as low lighting environments. As a matter of fact, Adini *et al.* (1997) have pointed out this problem as a major drawback in facial recognition, as well as a high complexity in the background. In addition, physical changes, such as hairstyle or beard variations, and wearing hats or glasses (Martínez 2002), may greatly affect the matching rate of face recognition systems. Such systems also have troubles to identify very similar individuals such as twins (Grother *et al.* 2010) and to keep a satisfying level of accuracy when physical changes occur owing to the age (Lanitis *et al.* 2002). Finally, the fact

that a user may be scanned without his/her consent raises serious threats according to confidentiality.

2.4.3 HAND GEOMETRY AND EAR SHAPE

Example of use case scenario: as every biometric system, both hand and ear recognition need to learn from user unique physical features. Depending on the implementation, this process may be performed, several times, through either optical analysis, or through *the capacitive touch-screen of the mobile device. Finally, the user has to repeat once the same process, each time the device has to be unlocked.*

Both hand geometry and ear shape biometric authentication mechanisms are based on the fact that nearly every individual's hands and ears are shaped differently. These body parts also remain practically the same after a certain age.

As regards ear shape as a means of authentication, Alphonse Bertillon's researches helped to develop such biometric systems as he worked particularly on the classification of this body part (Bertillon 1893). Several more recent studies on the subject have shown that the acquisition of the ear was exclusively made with cameras (Yuizono *et al.* 2002; Chen and Bhanu 2005; Choraś 2005). In that sense Descartes Biometric has released, short while ago, the most mature ear-shape-based authentication system of the market: Helix. This software exploits the proximity sensor of the front camera on the mobile device. The user needs to place the device from 6 to 12 inches in front of his/her ear. Then, 30 images per second are recorded, processed and finally compared to the stored template. Moreover, the company offers the possibility

to configure the accuracy threshold at a higher level. However, the record of ear images with the front camera of the mobile device may be disturbing for users in their daily usage. In that sense, the Yahoo research department yet offers an experimental approach that handles capacitive sensors embedded in the screen of mobile devices, to record the topography of the ear. According to Holz *et al.* (2015), this system correctly identified users at 99.8 percent of the time with a false negative rate of 7.8%. This rate is based on a test that involved 12 participants. Such a system appears to deliver a more appropriate ease of use for users since it mimicking the act of calling.

On another side, hand geometry recognition is the ability to compare dimensions of fingers and the localization of joints, shape and size of the palm, and also phalanges disposition. However, hand geometry is not distinctive enough to accurately identify a large set of individuals. Therefore, such systems may not be used in an authentication process, but rather in a verification process. Just as ear recognition, several studies involve a camera through the hand-record process (Ross *et al.* 1999; Sanchez-Reillo *et al.* 2000; Kumar *et al.* 2003). By contrast, researches of Holz *et al.* (2015) also introduce a novel approach for hand geometry recognition solutions. This experimental system allows the authentication of mobile device users in the same way as the ear recognition system. Both phalanges and palm identifications are possible with this system. Results achieved showed a matching rate of 99.5% with a false negative rate of 26.8% and the precision fall down to 86.1% when the false negative rate reached 0.2%.

Ear shape recognition and hand geometry appear to be an encouraging way in order to authenticate users on their mobile devices since they aim to be more usable for

such devices. Examples presented by Holz *et al.* (2015) expose very simple techniques, which are easy to use and which do not require any additional sensor than the ones already built-in the mobile device. Although these systems expose promising results, they also admit several drawbacks. As concern ear shape, when recorded with a camera, hairs, hats or piercings may compromise the identification process. Regarding hand geometry, jewelry and arthritis will involve matching errors in both cases.

2.4.4 VOICE: SPEAKER RECOGNITION

Speaker recognition techniques are classified as a behavioral biometric since they focus on vocal characteristics produced by the speech and not on the speech only. These features depend on the dimension of the vocal tract, mouth, and nasal cavities, but also rely on voice pitch, speaking style, and language (Eriksson and Wretling 1997) as shown in Figure 2.4.

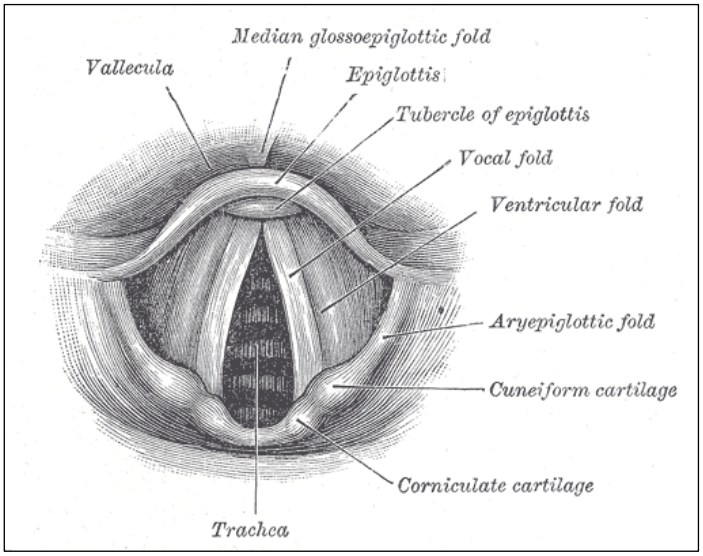

FIGURE 2.4: ANATOMICAL DIAGRAM OF THE VOCAL FOLDS OR CORDS (SOURCE: VOCAL FOLDS. 2016, AUGUST 24. IN WIKIPEDIA)

There are two leading methods to process speaker recognition: *text-dependent* and *text-independent (Doddington et al. 2000; Gold et al. 2011)*. Text-dependent recognition involves the user to pronounce a predefined passphrase. It is considered as a voice password and used both for the enrollment and the verification process. By contrast, text-independent systems are able to identify the user accurately, for any delivered word or locution. However, Boves and Den Os (1998) have identified a third type of speaker recognition technique that is a combination of the two others: the text-prompted method, that randomly select a passphrase the user needs to pronounce each time the system is used. Speaker recognition is an inexpensive solution to authenticate users on their mobile devices since no additional sensor is required. In addition, speaker recognition is another mechanism that may be able to authenticate users continuously. Indeed, a first recognition may be achieved to grant the access to the owner, but such a process may also be performed each time the device either, or both, receive or emit a phone call. As an example of practical implementations, Google also introduces in "Smart-Lock", the "Trusted voice" feature. As all text-dependent-based speaker recognition, the user has to enroll his/her voice by pronouncing "Ok Google" three times. Then, this passphrase must be repeated each time the mobile device needs to be accessed. Whenever the record both matches the voice model and the passphrase, the access to the mobile device is granted.

However, speaker-recognition-based authentication mechanisms admit several major drawbacks. In the first place, since such systems are viewed as behavioral, the current physical, medical or emotional condition of the user may considerably affect the accuracy. Voice is also likely to change over time due to the age. Moreover, speaker

recognition techniques are rarely noise resistant. Then, various loud background noises make such systems almost impossible to use in public places such as bars or public transports. Finally, speaker recognition techniques are singularly exposed to security threats. It is possible for an attacker to record or imitate the voiceprint of the user to perform a fraudulent authentication afterwards (Lau *et al.* 2004).

In this section, we provided an overview of voice-based authentication mechanisms in order to review the entire set of authentication techniques. However, a more specific related work will be presented in the next chapter since this thesis introduces a practical application of a speaker authentication system for mobile devices.

2.4.5 GAIT

Example of use case scenario: the process of authentication through *gait analysis is independent of any action from the user. The mobile device is able to determine whether he/she is walking or not and then perform a recognition to unlock the device automatically in a continuous manner.*

Gait recognition is a technology based on the analysis of the "rhythmic patterns associated with walking stride" (Rani and Arumugam 2010). The observation that each human's walking style is different leads to the development of advanced biometric authentication systems that exploit such behavioral characteristics.

Accordingly, studies have proposed a gait analysis mechanism based on an accelerometer to collect features which create the gait template (Gafurov *et al.* 2006; Derawi *et al.* 2010). It is possible for such a system to be integrated with mobile devices

since they include built-in inertial sensors (*i.e.* accelerometer, gyroscope). As a matter of fact, gait recognition may become a convenient way to authenticate users as they always keep their mobile devices within their pocket or a bag. It is fair to say that it constitutes a human-centered system since the authentication process is wholly imperceptible to the user.

Withal, gait is not as stable over time due to changes in body weight. Such a physical change is not the only way for this human behavior to change. Indeed, brain damages, injury and also inebriety may involve from a short time to a long time or permanent variation in the manner of how individuals walk (Boyd 2004; Sarkar *et al.* 2005).

2.4.6 KEYSTROKES DYNAMICS

Keystroke dynamics are based on the measurement and the assessment of the human's typing rhythm on numerical systems. This process allows the creation of a digital print upon users' interaction with devices that are rich in cognitive qualities (Obaidat 1995). Characteristics of this user behavior are fairly unique to each person and hold a high potential as an authentication mechanism. With the growth of capacitive touch screens, keystrokes patterns are now capable of providing even more unique features for the authentication than only typing rhythm, which includes key press duration, latencies, typing rate, and typing pressure. Such characteristics may be measured up to milliseconds order precision (Senk and Dotzler 2011) and more recent studies have pointed out high accuracy level (Trojahn *et al.* 2013; Deng and Zhong 2015). Thus, it is nearly impossible for an attacker to replicate a defined keystroke pattern

without an enormous amount of effort. The main benefit of keystroke dynamics pattern recognition is that anything except an extra software layer is required. Moreover, keystrokes analysis may be employed as continuous authentication process for free typing instead of one-time authentication such as fingerprint. However, since touch keyboards only appear when users are granted the full access, it is not possible to perform the authentication process beforehand. As a result, keystroke dynamics may only be employed to verify continuously that the current user is truly the owner of the mobile device afterwards. Thus, such a system must not be self-sufficient and requires to be coupled to a non-continuous authentication scheme as regards a mobile device usage.

2.4.7 SIGNATURE

Example of use case scenario: generally, a blank canvas is prompted to the user where he/she has to make his/her personal signature. The analysis is based on the way the signature is produced by exploiting pressure, direction, acceleration; rather than only a comparison of the signature pattern itself. The first step resides in the definition of such a model to compare with, and then, the user has to reproduce the signature each time the mobile device has to be accessed.

Signatures were used for decades in the concrete world while people need to enact documents. The same idea is used over numerical systems for authentication purposes. Signature recognition is considered as a behavioral biometric since it is based on the dynamic of making signature rather than a unique comparison of the signature itself.

Since users have to reproduce their signature on a mobile device touch screen, the identification process may determine dynamics, owing the measurement of the pressure, the direction, the velocity and the acceleration and the length of the strokes. Initially, the hardware used to record individual's signatures was not convenient enough for users (Nalwa 1997). However, with the advent of capacitive touch screens on mobile devices, the use of such an authentication process became user-friendlier. Moreover, a unique extra software layer is required to make it work.

In spite of this, as for every behavioral biometric system, people's physical and emotional condition may considerably affect an authentication mechanism based on signature recognition. Besides, it is important to note that this is the only biometric authentication scheme that is possible to be deliberately changed by the user. Moreover, any replication in the way of proceeding the signature requires lots of effort.

2.4.8 HEARTBEAT

Recently, Sufi *et al.* (2010) have introduced the use of the electrical activity of the heart as a novel biometric authentication mechanism. Since the heartbeat is evaluated as unique for each person (Khan 2008), such a system requires a record of an electrocardiogram (ECG or EKG) as illustrated in Figure 2.5. Hence, as with other biometric authentication schemes we described above, the first step of the process consists in an enrollment phase. Unique features are extracted to build a template, then compared through the identification process.

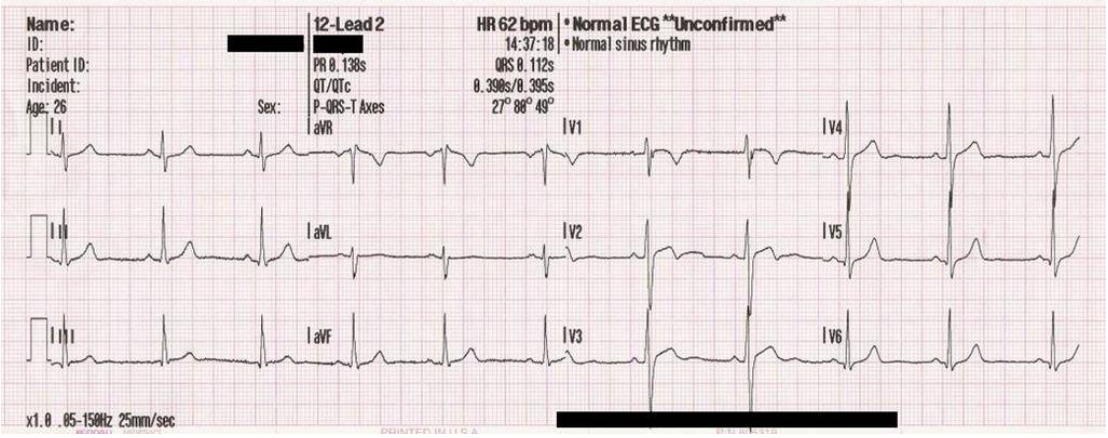

FIGURE 2.5: EXAMPLE OF AN ECG CURVE (SOURCE: ELECTROCARDIOGRAPHY. 2016, JUNE 22. IN WIKIPEDIA)

Exploiting ECG as means of authentication is suitable across a wide range of people. Indeed, heartbeat samples may be collected from any part of the body such as fingers, toes, chest, and wrist. Thus, people who suffer from large injuries may be authenticated, continuously or not, by using such a system. As we saw before, connected objects and in a broader sense, IoT, begin to take a measurable place through the numerical environment (Miorandi *et al.* 2012). In that sense, a Canadian company patented a wristband that is fully compliant with mobile devices through a companion application. This band is able to authenticate a user wearing it by his/her heart signature (Fatemian *et al.* 2010).

Such a system seems to be largely reliable as reproducing heartbeat signature depends upon sophisticated skills and hardware. Therefore, it is nearly impossible for an attacker to spoof such authentication systems. Since it may be considered as a behavioral biometric, several factors may seriously affect its accuracy such as daily activities and nutrition facts, stress level, and weariness.

2.4.9 BIOMETRICS PERFORMANCE

As we have seen in above sections, biometric systems are not devoid of potential errors over the authentication process. These systems can make two types of errors: False Acceptance Rate (FAR), and False Rejection Rate (FRR). Figure 2.6 graphically illustrates these types of errors in detail. On the one hand, FAR or also False Match Rate (FMR) is the probability that the system incorrectly declares a successful match between the input pattern, and the ones stored in the database. FAR is obtained by the equation (2.3) where F_a is the number of false acceptances, and V_i is the number of imposter verifications.

$$FAR = \frac{F_a}{V_i}. \tag{2.3}$$

On the other hand, FRR or also False Non-Match Rate (FNMR) is the probability that the system declares a failure while the input pattern matches with stored templates. FRR is obtained by the equation (2.4) where F_r is the number of false rejections, and V_g is the number of genuine verifications.

$$FRR = \frac{F_r}{V_g}. \tag{2.4}$$

Generally, the matching algorithm performs a decision using some parameters as a threshold. Graphically expressed both FAR, and FRR, by opposition to the given threshold, represent the Relative Operating Characteristic (ROC). This plot allows finding the Equal Error Rate (EER) as shown in Figure 2.6. EER is the rate at which both accept and reject errors are equal. This rate is commonly used to evaluate biometrics. Indeed, the lower the EER is, the more accurate the system is considered to be. The report by Mansfield and Wayman (2002) describes in detail the performance evaluation for biometric systems.

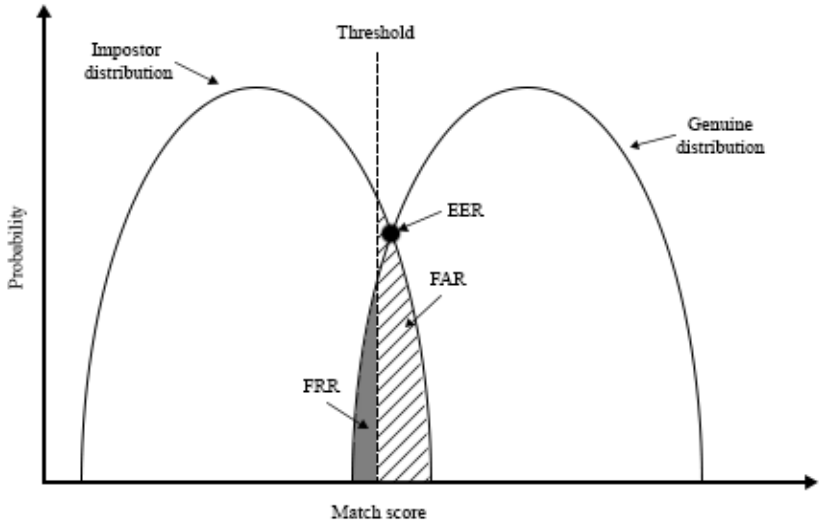

FIGURE 2.6: BIOMETRIC SYSTEM ERROR RATES, WHERE CURVES SHOW FALSE ACCEPTANCE RATE (FAR) AND FALSE REJECTION RATE (FRR) FOR A GIVEN THRESHOLD

Table 2.1 exposes the performance of various biometric authentication mechanisms which stem from several studies (Phillips *et al.* 2000; Jain *et al.* 2004; Jain *et al.* 2005), according to metrics we described above.

TABLE 2.1: EVALUATION OF VARIOUS BIOMETRICS PERFORMANCES

	EER	FAR	FRR	CONDITIONS
FACE RECOGNITION	-	1%	10%	Varied light: indoor, outdoor.
FINGERPRINT	2%	0.1%	2%	Rotation and exaggerated skin distortion.
HAND GEOMETRY	1%	0.14%	2%	With rings and improper placement.
IRIS	0.01%	≈ 0%	0.99%	Indoor.
VOICE RECOGNITION	6%	3%	10%	Text dependent and multilingual.
KEYSTROKES	1.8%	7%	0.1%	Data record during 6 months.

The required accuracy of biometric systems depends chiefly on the need of the application. Presently, biometric system designers aim at reducing false acceptance rates. However, the false rejection rates undeniably grow. To cover FRR up, most of current designs offer to bypass the biometric system that fails and suggest the user to perform another authentication scheme instead. The use of such a fallback mechanism represents a multi-factor authentication scheme. Biometric was often introduced as an encouraging way to end with the use of passwords (*i.e.* knowledge-based schemes). However, since both Apple and Google recently released their respective biometric technologies "Touch ID", "Nexus Imprint" and "Smart-Lock", mobile device users still have to set up a more traditional authentication mechanism such as PIN before enabling these features. As a matter of fact, we may affirm that passwords are still not dead contrary to everything said. Nevertheless, as claimed by Kokumai (2015), threats that can be thwarted by biometric authentication that operated together with rescue pass-words, still remain less secure than just a knowledge-based authentication mechanism. Indeed, a two-factor authentication system must be treated as a conjunctive statement in opposition to a disjunctive statement. In other words, both the main system and the

fallback one have to authenticate the user properly, and not just one, as offered by the main biometric solutions available on the market presently.

2.5 EVALUATION CRITERIA AND DISCUSSION

Based on the above critical analysis, it is possible to say that each authentication mechanism we assessed concede some advantages, as well as drawbacks. In order to state about the aftermath of authentication, this section will discuss the whole set of proposed mechanisms.

First, we offer an examination of knowledge-based authentication schemes through both the theoretical and real measure of the password space entropy that each of them provides, as shown in Table 2.2. Theoretical entropy is inevitably greater than the real one since users do choose their own piece of knowledge. The main deficiency of PINs resides in their simplicity, as they remain easy to crack by a force brute attack. Moreover, users do prefer a significant code to them instead a random selection of numbers. Text-enhanced passwords are an excessively complex solution for mobile devices users that force them to choose easily-findable identification codes. Finally, graphical and haptic passwords provide an adequate level of security, but still remain easy to obtain through shoulder surfing attacks or smudge attacks.

TABLE 2.2: EVALUATION OF KNOWLEDGE-BASED AUTHENTICATION MECHANISMS FOR MOBILE DEVICES VIA THE PASSWORD SPACE ENTROPY METRIC, WHERE H: HIGH, M: MEDIUM, L: LOW

	KNOWLEDGE-BASED		
	EXPLICIT		
	PIN	*TEXT-ENHANCED PASSWORDS*	*GRAPHICAL & HAPTIC PASSWORDS*
THEORETICAL PASSWORD-SPACE ENTROPY	L	H	M
REAL PASSWORD-SPACE ENTROPY	L	L	M

Secondly, we suggest an empirical evaluation of biometrics authentication mechanisms based on previous related work of Jain *et al.* (2006a). As assessed in the previous sections, several novel studies aim at improving biometric processes to be more accurate and fast, as regards mobile devices (Nickel *et al.* 2012; Belgacem *et al.* 2015; Deng and Zhong 2015; Fathy *et al.* 2015). As a matter of fact, this study requires some improvement. Hence, changes we introduce mainly focus on the performance criteria. Moreover, we include heartbeat authentication that was not discussed in their work. Our evaluation is based on the same criteria suggested by Jain *et al.* (2006a) which were defined by biometric experts. Such a guideline is described as follows:

1. *Universality*: Biometric solutions rely upon singularities of the human body or behavior, but the ability of such mechanisms to accurately identify the genuine user largely varies between each one. Consequently, Jain *et al.* (2006a) have suggested to quantitate the fact that each person should have the characteristic.

2. *Uniqueness*: Some physical traits of the human body (*e.g.* face) remain largely close in some cases (*e.g.* twins). Therefore, Jain *et al.* (2006a) have proposed to evaluate the probability that two individuals are potentially the same, in terms of characteristics.

3. *Permanence*: Physical or behavioral features of the human used with biometrics may gradually evolve. This criterion figures out the invariance of these characteristics with time (Jain *et al.* 2006a).

4. *Performance*: The inability of a biometric system to identify a user with a 100% accuracy, lead to identifying the related performance offered by each one (Jain *et al.* 2006a).

5. *Collectability*: Most of the time, the biometric authentication process involves additional hardware or a major computing complexity with mobile devices. This criterion refers to the evaluation of how simple a characteristic is quantitatively measurable (Jain *et al.* 2006a).

6. *Acceptability*: Users may be hesitant to use some biometrics because of their association to certain use cases (*i.e.* forensics) or because of the hardware they require. Thus, it is important to state the users' acceptance rate, according to such mechanisms (Jain *et al.* 2006a).

7. *Circumvention*: The last criteria reported by Jain *et al.* (2006a) refers to the easiness of mimicking a singular trait or behavior with biometric

systems. Such an evaluation delivers the strength rate of biometric systems in front of fraudulent attacks (*e.g.* spoofing attack).

The evaluation of biometric systems according to the previously proposed work, as well as improvement we suggest according to the evolution of these mechanisms and the introduction of heartbeat as mean of authentication is provided by the Table 2.3.

TABLE 2.3: EVALUATION OF BIOMETRIC AUTHENTICATION MECHANISMS SUGGESTED BY JAIN *ET AL.* (2006A) REVISED THROUGH A PERSONNAL OPINION AND ADAPTED FOR MOBILE DEVICES USAGES, WHERE H: HIGH, M: MEDIUM, L: LOW. IMPROVEMENTS ARE IDENTIFIED IN *BOLD-ITALIC*

	BIOMETRICS								
	PHYSIOLOGICAL				BEHAVIORAL				
	FINGERPRINT	*FACE RECOGNITION*	*EAR SHAPE*	*HAND GEOMETRY*	*VOICE RECOGNITION*	*GAIT RECOGNITION*	*KEYSTROKE DYNAMICS*	*SIGNATURE*	***HEARTBEAT***
UNIVERSALITY	M	H	M	M	M	M	L	L	***M***
UNIQUENESS	H	L	M	M	L	L	L	L	***M***
PERMANENCE	H	M	H	M	L	L	L	L	***L***
COLLECTABILITY	M	H	M	H	M	H	M	H	***M***
PERFORMANCE	***H***	***M***	***H***	***H***	***M***	***M***	***M***	L	***M***
ACCEPTABILITY	M	H	H	M	H	H	M	H	***H***
CIRCUMVENTION	H	L	M	M	L	M	M	L	***H***

Finally, we estimate that several criteria were missing to perform a proper evaluation of authentication mechanisms. Consequently, we introduce four new criteria

based on the above critical analysis of these schemes we proposed. As a matter of fact, we define the following guideline:

1. *Confidentiality*: Several authentication scheme designs such as fingerprint involve either, or both, record or store data that relate to the secrecy of the user. To ensure that the information the user provide remain in his/her privacy regards, is a crucial interest in the case of security threats.

2. *Intrusive*: Several authentication mechanisms reviewed in this chapter involve providing information or features relating to the user. This criterion focuses on the ethical concern of the data that use to be collected during the authentication process. This criterion mainly focuses on cognitive schemes, as well as biometric ones. Indeed, the authentication process requires to handle data that directly concern users, and that may be performed without their broad agreement.

3. *Ease of use*: The major drawback of authentication mechanisms is predominantly due to the way that Humans have to interact with these systems. In that sense, we suggest evaluating how the user is involved in the interaction process and how the system focuses on what people want to do rather than possibilities offered by the technology. In other words, with this criterion, we suggest evaluating how an inexperienced user is able to use an authentication system without a lot of difficulty in a short period of time.

4. *Usage frequency*: Since most of more used authentication schemes also remain the weakest ones, we offer to identify to the popularity of the mechanisms when applied to secure a mobile device.

The evaluation of the whole set of authentication mechanisms as regards criteria we introduce in this section is provided by the Table 2.4. These values were mainly determined through a personal opinion.

TABLE 2.4: PERSONNAL ASSESSMENT OF THE WHOLE AUTHENTICATION MECHANISMS FOR MOBILE DEVICES THROUGH THE CRITERIA WE SUGGEST, WHERE H: HIGH, M: MEDIUM, L: LOW

			CONFIDENTIALITY	INTRUSIVE	EASE OF USE	USAGE FREQUENCY
KNOWLEDGE-BASED	EXPLICIT	*PINS*	L	N/A	M	H
		TEXT-ENHANCED PASSWORDS	L	N/A	L	L
		GRAPHICAL & HAPTIC PASSWORDS	M	N/A	M	H
	IMPLICIT	*COGNITIVE-BASED PASSWORDS*	H	M	M	L
TOKEN-BASED			H	N/A	M	M
BIOMETRICS	PHYSIOLOGICAL	*FINGERPRINT*	M	M	H	M
		FACE RECOGNITION	M	H	M	M
		EAR SHAPE	M	M	M	L
		HAND GEOMETRY	M	M	M	L
	BEHAVIORAL	*VOICE RECOGNITION*	M	L	M	M
		GAIT RECOGNITION	L	L	H	L
		KEYSTROKE DYNAMICS	H	L	H	M
		SIGNATURE	M	L	M	M
		HEARTBEAT	L	H	H	L

Based on such a personal assessment, it is possible to observe that most of authentication schemes raise some inconveniences for users since most of them are considered as not easy to use. Besides, proposed mechanisms that do not involve users in

the process of authentication, namely transparent for them such as gait recognition or keystroke dynamics, were identified as convenient. As a matter of facts, it is clear that the future trend for authentication on mobile devices will turn into systems that focus on the users first. As an example, it is known that people do spend a considerable amount of time in a few key locations such as home or work as assessed by Hayashi and Hong (2011). In that sense, ubiquitous mechanisms that will be able to learn about users' habits and that will not require any passwords or tokens are close. Indeed, studies in that field of research expose promising results (Micallef *et al.* 2015), but efforts should probably pay more attention to such a key idea in coming years.

Nowadays, authentication mechanisms remain an important field of interest for researchers and leaders of mobile device industry. According to major players on the market, it should be noted that both Apple and Microsoft took the biometric band respectively, with the "Touch ID" technology and the launch of Windows 10 (Belfiore 2015). Despite, Google seems to want to see further since their Advanced Technology and Projects (ATAP) division is currently working on an experimental multi-modal biometrics system based on behavioral analysis: the "Project Abacus". The system will identify a genuine user through a "trust score" calculated through a real-time analyze of users' voice recognition, keystroke dynamics and touch gestures, facial recognition, and location. The firm presented research results at its annual conference of 2015 and claimed that the entropy of such a system is now ten times higher than the most valuable fingerprint system of the market. While the project is still in development, Google aims

at integrating such a system into one of the next version of the Android mobile operating system. Hence, this research project simply confirms the trend in the evolution of authentication for a near future.

2.7 CONCLUSION

The present review of mobile device authentication mechanisms leads us to affirm that each of the schemes we have reviewed concedes several strength and weakness aspects. Since knowledge-based mechanisms were designed to a machine point of view, they first involve an enormous amount of memory efforts from users. Behaviors they consequently adopt to overcome a system they are not comfortable with, yield several threats and weaknesses as regards the security of their mobile devices. Token-based authentication schemes are not devoid of weaknesses as well. However, with the advent of the Internet of Things (IoT), such mechanisms appear to be more convenient than knowledge-based systems and will certainly keep growing. Presently, biometric methods become more and more popular and easy to reach for everyone. Some remain just too much intrusive for users or lead them to believe that providing personal and unique features describing them is an important threat according to their privacy. Nevertheless, biometric techniques are, overall, very accurate, but also accept certain dysfunctions. However, gathering a number of biometric mechanisms together, allow the entropy of the entire multi-factor system to extend and consequently, increase its accuracy. The major drawback of biometrics resides in the fallback mechanism designed to cope with false rejections and let a genuine user proceeds to his/her authentication

properly. Nevertheless, most of current biometric solutions for mobile devices available on the market such as “Touch ID” or “Smart-Lock” have adopted this method.

As we stated that the human factor is the fundamental drawback for authentication mechanisms since they involve lots of interaction with users; we consider that the aftermath of such systems should singularly take care of this criterion. Such mechanisms already evolve to become no password systems. Although they currently provide a better convenience, they do not provide a better security. Based on such an evaluation, we assume that the optimal solution should be able to recognize a genuine user without the need for any interaction of his/her part. To be really accurate and secure, such a mechanism should be based on users’ habits of the everyday life (what does he/she do all along his/her day, in which order?) that involve collecting the most of possible relevant patterns through the mobile device. However, it is important to consider that such a solution implies a large set of information, more than just for one user. The processing of all of this knowledge, considering each mobile device users, will undeniably increase the cost in hardware requirements. This observation now questions us about ecological issues related to the growth of the number of brand new data centers everywhere in the world.

CHAPTER III

A TEXT-INDEPENDANT SPEAKER AUTHENTICATION SYSTEM FOR MOBILE DEVICES

3.1 INTRODUCTION

Speaker authentication systems may be designed according to two leading methods: text-dependent and text-independent (Doddington *et al.* 2000; Gold *et al.* 2011). A text-dependent authentication requires the user to pronounce a predefined passphrase that is considered as a voice password. It is used both for the enrollment and the identification process. For instance, Google recently introduced the trusted voice feature on Android, where users have to enroll their voice by pronouncing "Ok Google" three times. Then, this passphrase must be repeated each time the mobile device needs to be accessed. By contrast, text-independent schemes are able to identify the user accurately, for any delivered word or locution that it is impossible to recover afterwards.

Since we have stated, in the previous section, that the previously proposed authentication mechanisms involve lots of interaction with users, this research leans toward a user-centered design. Since we have mentioned the ability to authenticate users in a continuous manner with voice-based authentication, our choice was based on a speaker authentication system. Moreover, existing speaker authentication techniques offered on mobile devices, always require network communications. Indeed, matching

templates are usually stored in the cloud. In that sense, they may represent costly authentication solutions for certain users. However, since no additional sensors are needed, they remain inexpensive solutions to authenticate users as regards hardware requirements. Furthermore, such approaches offer a sufficient acceptance rate with end-users and remain less intrusive than fingerprint or retina scan (Clarke *et al.* 2002; Jain *et al.* 2006a). Hence, these mechanisms may play a major role in some real-world applications to secure identity management systems such as e-commerce solutions, attendance systems, mobile banking or forensics.

It is known that several proposed speaker recognition and identification systems achieve accurate results (Reynolds and Rose 1995; Kumar *et al.* 2009; Nair and Salam 2014). Despite the effectiveness of these mechanisms, few of them are presently implemented on mobile devices. Moreover, the considerable amount of users who still do not secure the access to their mobile devices (Ben-Asher *et al.* 2011) reveals a need for novel methods, that should take into account the diversity in user profiles and usages. This research targets these needs. In that sense, this chapter presents a practical use of text-independent speaker authentication system applied to mobile devices. The system that we suggest relies on Linear Prediction Cepstral Coefficients (LPCCs) (Furui 1981; O'Shaughnessy 1988) and the Naïve Bayes algorithm (Zhang 2004) for patterns classification. Since authentication mechanisms usually either grant, or deny the access to the whole content of the phone, we further suggest enhancing such a final decision to overcome false positive and negative identification that may occur. Therefore, we introduce the notion of access privileges, that enable restricting certain access, based on a simple evaluation of the user's location and the presence of a headset. Moreover, we

pay attention to opt for low complexity algorithms since we want to avoid network communications and achieve both the training and the identification, on the mobile device itself.

The rest of this chapter is structured as follows: section 3.2 provides an overview of related work to proposed speaker identification and verification systems. Finally, section 3.3 details the suggested system, specifically designed to entirely operate on mobile devices.

3.2 RELATED WORK

In order to achieve speaker authentication, several techniques have been described for years through disparate features extraction techniques and classification algorithms. This section first exposes suggested text-independent speaker identification and authentication systems to determine their suitability as regards a usage on mobile devices. Finally, we will examine proposed schemes which were explicitly designed to operate on mobile devices.

First of all, Reynolds and Rose (1995) have proposed a text-independent speaker identification which exploits Mel-Frequency Cepstral[1] Coefficients (MFCCs) as features and a Gaussian Mixture Model (GMM) to predict which person is speaking. MFCCs are widely used in speaker recognition as they accurately represent the envelope of the short-time power spectrum of the signal. Although such coefficients appear to be more robust against noisy conditions, their acquisition remains theoretically very

[1] The cepstrum is defined as the inverse DFT of the logarithm of the estimated spectrum of a signal.

expensive (Rabiner and Juang 1993). The main motivation of using a GMM was based on an empirical observation that a large number of unlabeled classes of sample distribution, may be represented as a linear combination of Gaussian basis functions. To evaluate the system, a subset of the KING speech database (Godfrey *et al.* 1994) was used. This database provides utterances from speaker conversations over both signal-to-noise radio channels and narrowband telephone channels. An accuracy of 80.8% was obtained for 49 telephone speech samples of 15 seconds. Besides, authors claimed that this model is computationally inexpensive and easy to implement on real-time platforms. However, the main drawback of such a system lies in the initialization of the training procedure, where parameters such as mean, covariance and prior of each distribution have to fit the data. Indeed, such a process may be achieved through several costly methods like a Hidden Markov Model (HMM), or a binary k-means clustering algorithm.

Secondly, Kumar *et al.* (2009) have suggested another text-independent speaker identification approach which aims at predicting utterances thanks to a backpropagation neural network, where LPCs (Linear Prediction Coefficients) parameters were used as input features. The goal of the backpropagation method is to optimize weights between each neuron layers so that, the neural network can learn how to correctly map arbitrary inputs to outputs. Hence, outcomes provide the resulting decision in determining at which speaker corresponds each given utterance. The evaluation of the system was performed over a collection of 25 speech samples in different languages. An overall accuracy measure of 85.74% was achieved. This led authors to state that such a technique remains appropriate and reliable. However, the theoretical complexity of a

standard back-propagation neural network training phase is $O(nmh^k oi)$, where n are training samples, m refers to features, k are hidden layers, each containing h neurons, o refers to output neurons and i is the number of iterations (LeCun *et al.* 2012). Hence, such a computation time remains overly expensive in a mobile device context.

On the other hand, concerning speaker authentication, Nair and Salam (2014) have proposed a text-independent system which exploits both LPCs and LPCCs to compare their strength. The decision was made through the Dynamic Time Warping (DTW) algorithm. DTW allows calculating the distance between two given sequences which provides the optimal match. Authors have experimented their system over the TIMIT speech corpus which provides 630 real speech signals of American English speakers. An overall accuracy of 92.2% was obtained with LPCs features while it rose up to 97.3% with the derivative cepstral coefficients. Thus, combining LPCCs with the DTW algorithm involves an accurate and reliable solution to authenticate users by their voice. Since DTW requires a quadratic time and space complexity (Salvador and Chan 2007), it may not be the most suitable method to achieve speaker authentication, directly on the mobile device. Nevertheless, real speaker authentication scenarios usually imply few distinct samples. In that sense, DTW as decision-making still stays an acceptable choice for such an authentication mechanism on limited-performance devices.

According to literature, it appears that few text-independent speaker authentication solutions for mobile devices were applied in practice. For instance, Vuppala *et al.* (2010) have suggested a recognition model which lies in several speech enhancements to improve the overall performance faced with different noisy conditions. In that

sense, authors aimed to prove the robustness of the method when used in varying background environments. However, their evaluation was performed through noise simulations over speech samples from the TIMIT corpus. Hence, the efficiency of such a method with real life noises in the background, may considerably decrease.

Conversely, Brunet *et al.* (2013) have introduced a practical text-independent speaker authentication system which is entirely usable on mobile devices. The approach suggests extracting MFCCs features from speech samples. Then, a reference model is built thanks to a Vector Quantization (VQ) method. The Euclidean distance between stored centroids and testing samples is calculated and compared to a given threshold in order to accept or reject the attempt. Authors have performed an experiment over their own database, where training and testing samples were collected thanks to a mobile device, as well as the Sphinx database which contains 16 American English speakers' utterances. Since the method was implemented as a stand-alone biometric system, only the equal error rate was computed to evaluate the performance. Hence, they obtained better performances on their database (4.52 of EER at best), than the ones on the public database (5.35 of EER at best). However, achieved results largely rely on initial parameters required for the quantization step (*i.e.* the number of centroids) that must be optimized according to the training data.

With this brief review, it appears that few text-independent authentications that focus on mobile device computation capabilities and generic usages were proposed. Hence, this chapter introduces a novel text-independent speaker authentication system

for mobile devices, where a special attention was paid to low-computational cost algorithms and any parameter to optimize with expensive techniques as regards processing time.

3.3 PROPOSED SYSTEM

The proposed system is a text independent speaker authentication mechanism for mobile devices. This method works as stand-alone and does not require any costly architectures such as client/server. Hence, the entire computation is done end-to-end on the mobile device. As illustrated in Figure 3.1, this mechanism consists of three main processes. The first one involves extracting individual voice features from a raw audio input to build a data set. The following operation lies in training such data with a Naïve Bayes classifier. The last process is the authentication decision. It aims at enhancing the conventional speaker verification mechanism. To achieve this, we suggest granting a specific access privilege to the user of the mobile device through the evaluation of two different states. The first one concerns the current location of the user that is compared to the ones defined beforehand. Secondly, we evaluate the presence of a headset, that is, whether if it is plugged in the mobile device or not.

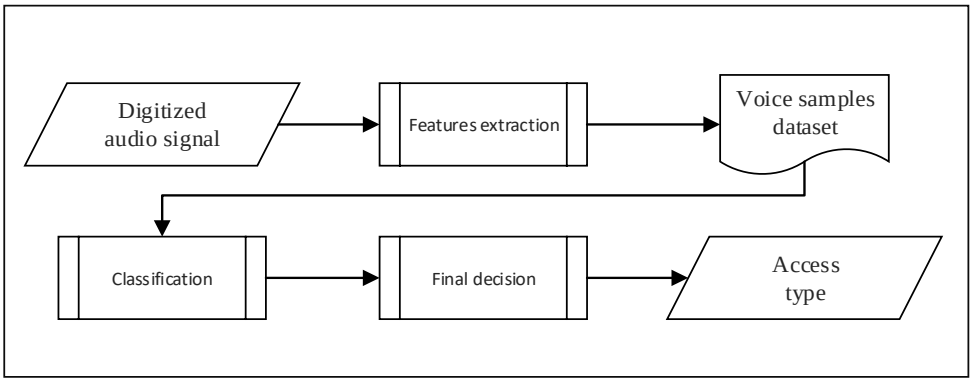

FIGURE 3.1: FLOWCHART OF OUR PROPOSED SPEAKER AUTHENTICATION SYSTEM

3.3.1 INPUT

Audio files are recorded using a 16-bit signed integer PCM (Pulse-Code Modulation) encoding format in bi-channels. The sampling rate of such audio files is up to 44.1 kHz.

3.3.2 PREPROCESSING

Silence Removal

Given an audio record as input, the first step that we produce is a preprocessing phase that aims at removing every silence area only to keep speech segments. First, we have defined a threshold close to zero (*i.e.* 0.0001). Then, our main focus in such a preprocessing step is to identify which section of the input signal is close to this threshold in order to remove it. To achieve this, we apply the autocorrelation function $r_x(t, k)$ suggested by Sadjadi and Hansen (2013) onto a windowed audio segment $s_w(n)$ of the entire input signal $s(n)$ as follows,

$$r_x(t,k) = \frac{\sum_{n=0}^{N_w-1} s_w(n)w(n)s_w(n+k)w(n+k)}{\sum_{n=0}^{N_w-1} w(n)w(n+k)}, \quad (3.1)$$

where t and k are frame and autocorrelation lag indices, respectively, and $w(n)$ is a Hamming window given by,

$$w(n) = \begin{cases} 0.54 - 0.46\ \cos\left(\frac{2\pi n}{N_w - 1}\right), & 0 \le n \le N_w - 1 \\ 0, & otherwise. \end{cases}, \quad (3.2)$$

where its length (N_w) is based on the frequency of the signal. For each processed segment $S_w(n)$, if the mean value of the computed coefficients, resulting from the autocorrelation function, gets close to the defined threshold then, it is identified as a silence area and we finally remove it. Figure 3.2. graphically illustrates this process.

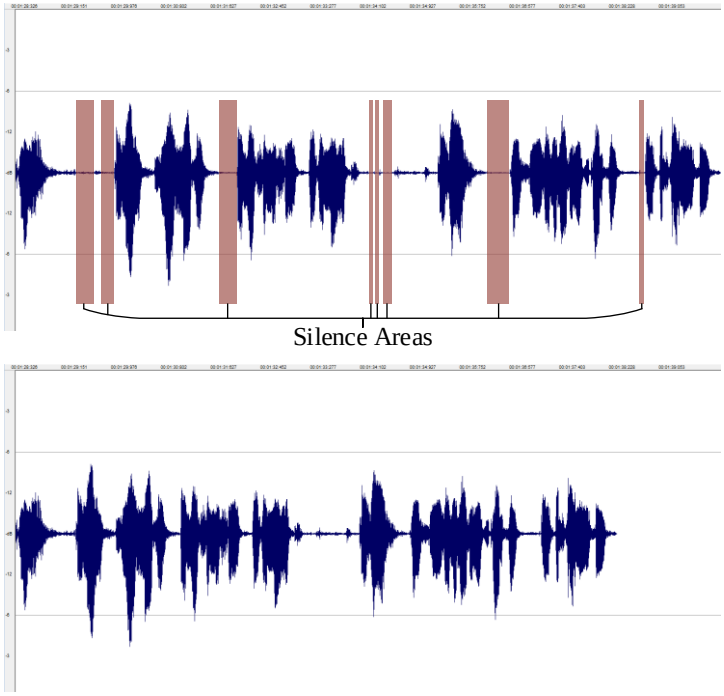

FIGURE 3.2: THE FIRST SIGNAL IS THE RAW INPUT WHERE SILENCE AREAS ARE HIGHLIGHTED. THE SECOND IS THE OUTPUT OF THE SAME SIGNAL AFTER THE SILENCE REMOVAL PROCESS

Audio Normalization

Succeeding the silence removal phase, a peak normalization is performed. The goal is to change the gain of the input to the highest peak of the signal, uniformly. Traditionally, this process is used to ensure that the highest peak remains at 0 dBFS (deciBels relative to Full Scale)—the loudest level allowed in a digital system. Since the entire signal is adjusted, it is indistinguishable and does not affect the original information. Moreover, the process of normalization ensures that the audio will not clip in any manner. Figure 3.3 shows the graphical result of such a process.

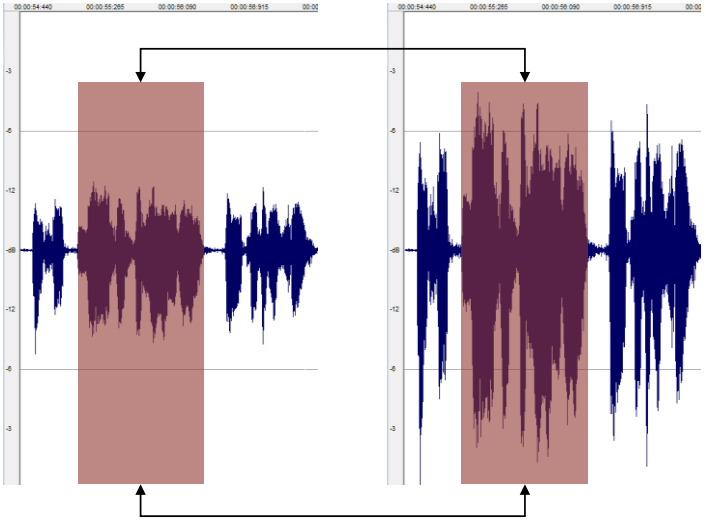

FIGURE 3.3: THE LEFT SIGNAL IS THE INPUT SIGNAL AND THE RIGHT ONE IS THE SAME SIGNAL WITH PEAK NORMALIZATION, WHERE THE SAME SEQUENCE IS HIGHLIGHTED ON BOTH SIGNALS

3.3.3 FEATURE EXTRACTION

Since the voice is considered as a signal containing a lot of information about the speaker—the process of extracting several discriminative features from the speech remains a critical part of both speaker identification and authentication systems. In that sense, we decide to favor the use of the Linear Prediction Cepstral Coefficients (LPCCs). Such coefficients are directly derived from the Linear Prediction analysis that aims at estimating the relevant features or characteristics from a speech signal (Benesty *et al.* 2007). We justify such a choice by its ability to provide extremely accurate estimates of the speech parameters, and by its relative speed of computation (Rabiner and Juang 1993). This last point was a crucial criterion since mobile devices

presently remain less powerful than traditional desktop computers. Figure 3.4 graphically summarizes the steps of the features extraction from a preprocessed signal to the resulting data set containing voice features.

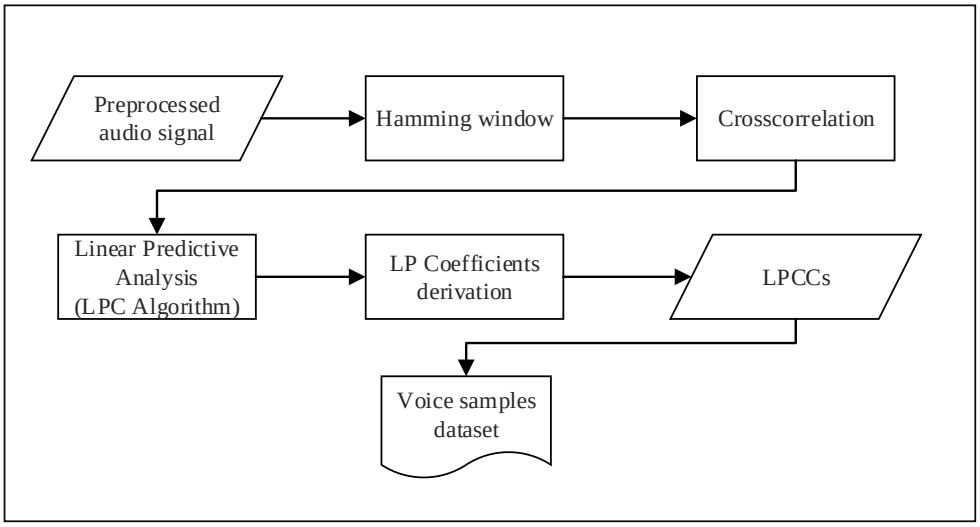

FIGURE 3.4: FLOWCHART OF FEATURES EXTRACTION PROCESS

To compute the LP analysis, we have implemented the Linear Predictive Coding algorithm. It was designed to exploit the redundancy present in the speech signal by assuming that each sample may be approximated by a linear sum of the past speech samples (p). Hence, the predicted sample $S_p(n)$ may be represented as,

$$S_p(n) = \sum_{k=1}^{p} a_k s(n-k), \tag{3.3}$$

where $a(k)$ are the Linear Prediction Coefficients (LPCs), $s(n-k)$ are past outputs and p is the prediction order. In our case, the speech signal is multiplied by an over-lapped Hamming window of 25ms to get a windowed speech segment $S_w(n)$ as,

$$s_w(n) = w(n)s(n), \tag{3.4}$$

where $w(n)$ is the windowing sequence given in equation (3.2). The error between the actual sample and the predicted one $e(n)$ may be expressed as,

$$e(n) = s_w(n) - \sum_{k=1}^{p} a_k s_w(n-k). \tag{3.5}$$

The main objective of the LP analysis is to compute the LP Coefficients that minimize this prediction error. To this end, our system exploits the autocorrelation method that is usually preferred since it is computationally more efficient and more stable than the covariance one (Al-Hassani and Kadhim 2012). Thus, the total prediction error E is given as,

$$E = \sum_{n=-\infty}^{\infty} e^2(n) = \sum_{n=-\infty}^{\infty} \left(s_w(n) - \sum_{k=1}^{p} a_k s_w(n-k) \right)^2, \tag{3.6}$$

values of $a(k)$ that minimize this total prediction error may be computed by finding,

$$\frac{\delta E}{\delta a_k} = 0, \qquad 1 \leq k \leq p, \tag{3.7}$$

thus, each a_k gives p equations with p unknown variables. The equation (3.8) offers the solution to find LP Coefficients,

$$\sum_{n=-\infty}^{\infty} s_w(n-i)s_w(n) = \sum_{k=1}^{p} a(k) \sum_{n=-\infty}^{\infty} s_w(n-i)s_w(n-k), \quad 1 \leq i \leq p. \tag{3.8}$$

Consequently, it is possible to express the linear equation (3.8) in terms of the autocorrelation function $R(i)$ as follows,

$$R(i) = \sum_{n=i}^{N_w} s_w(n)s_w(n-i), \qquad 0 \leq i \leq p, \tag{3.9}$$

where N_w is the length of the window. Then, by substituting values from equation (3.9) in the equation (3.8) with the autocorrelation function $R(i) = R(-i)$ we obtain the following equation,

$$\sum_{k=1}^{p} R(|i-k|)a_k = R(i), \qquad 1 \le i \le p. \tag{3.10}$$

The set of linear equations is expressed by the relation $Ra = r$ and may be represented in a matrix form as,

$$\underset{R}{\begin{bmatrix} R(0) & R(1) & \cdots & R(p-1) \\ R(1) & R(0) & \cdots & R(p-2) \\ \vdots & \vdots & \ddots & \vdots \\ R(p-1) & R(p-2) & \cdots & R(0) \end{bmatrix}} \underset{a}{\begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_p \end{bmatrix}} = \underset{r}{\begin{bmatrix} R(1) \\ R(2) \\ \vdots \\ R(p) \end{bmatrix}}, \tag{3.11}$$

where a is the vector of LP coefficients and r is the autocorrelation. The resulting matrix is a Toeplitz matrix where all elements along a given diagonal are equal. For example, with $p = 3$, we get,

$$\underset{R}{\begin{bmatrix} R(0) & R(1) & R(2) \\ R(1) & R(0) & R(1) \\ R(2) & R(1) & R(0) \end{bmatrix}} \underset{a}{\begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix}} = \underset{r}{\begin{bmatrix} R(1) \\ R(2) \\ R(3) \end{bmatrix}}. \tag{3.12}$$

Towards the computation of the LP Coefficient a_k, it is possible to derive cepstral coefficients c_n directly through the following relationship,

$$c_n = \sum_{k=1}^{n-1} a_k c_{n-k} + a_n, \qquad 1 < n \leq p\,. \tag{3.13}$$

where p refers to the prediction order.

It is known that speaker recognition requires more cepstral coefficients than speech recognition which employs around 15 of them. Although it was pointed out that increasing the number of such coefficients does not affect the recognition (Kinnunen 2003), we suggest using 20 LPCCs to preserve a relatively good computation speed.

3.3.4 CLASSIFICATION

Several classification algorithms were employed for speaker recognition (*i.e.* GMM, ANN). However, it is known that the Naïve Bayes classifier is fast, very effective and easy to implement. As a supervised and statistical learning method for classification—it simply computes the conditional probabilities of the different classes given the value of attributes. Finally, it selects the class with the highest conditional probability. Accordingly, Table 3.1 exposes the theoretical time and space complexity evaluations of the Naïve Bayes classifier (John and Langley 1995).

TABLE 3.1: NAÏVE BAYES TIME AND SPACE COMPLEXITIES, GIVEN K FEATURES FOR BOTH TRAINING AND TESTING OPERATIONS

OPERATIONS	TIME	SPACE
TRAINING ON n SAMPLES	$O(nk)$	$O(k)$
TESTING ON m SAMPLES	$O(mk)$	$\Theta(1)$

Once the feature extraction process is completed, a set of samples denoted $s_1, s_2, \dots, s_i$ with their associated class labels $c_{s_1}, c_{s_2}, \dots, c_{s_i}$, where $c_{s_i} \in \Omega = \{c_1, c_2, \dots, c_j\}$ is obtained. Each sample has k features (*i.e.* LPCCs) represented by floating numbers (with $k = 20$), that are denoted as $a_1, a_2, \dots, a_n$. The objective of the Naïve Bayes classifier is to exploit these samples to build a model (*i.e.* the training phase) that will be reused to predict the label of the class c_p for any future sample (*i.e.* the identification phase). Figure 3.5 shows a simplified block diagram of this process.

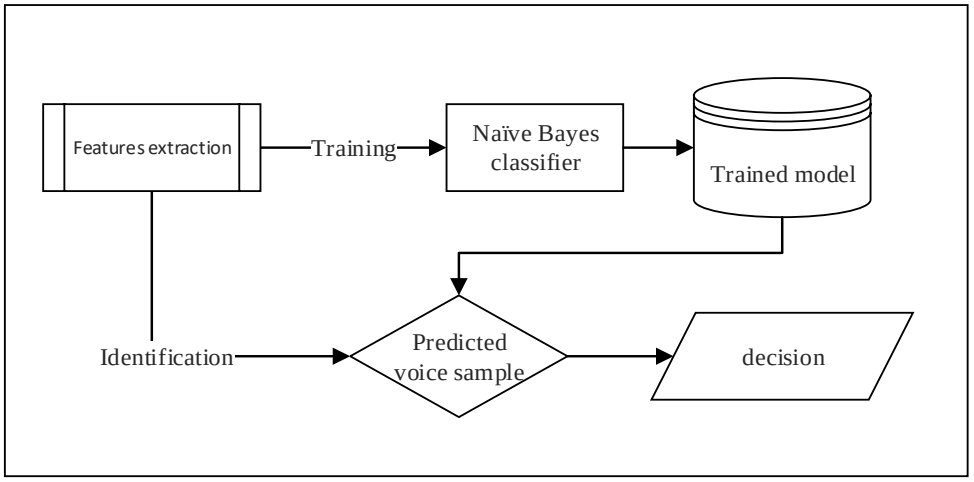

FIGURE 3.5: FLOWCHART OF THE CLASSIFICATION PROCESS

The algorithm, strongly relies on the Bayes theorem and imposes two assumptions. Firstly, all features $a_1, \dots, a_n$ should be independent for a given class c. This is

the class-conditional independence. Secondly, all features $a_1, \dots, a_n$ should be directly dependent on their assigned class c. Given that, it is possible to describe the classifier as,

$$P(c|a_1, a_2, \dots, a_n) = \frac{P(c)\prod_{i=1}^{n} P(a_i|c)}{P(a_1, a_2, \dots, a_n)}. \tag{3.14}$$

Since $p(a_1, a_2, \dots, a_n)$ is common for a certain sample, it may be ignored in the classification process. As a result, we can derive equation (3.14) to predict the class c of a given sample during the identification phase as follows,

$$c = \underset{c\in\Omega}{\operatorname{argmax}}\, P(c)\prod_{i=1}^{n} P(a_i|c). \tag{3.15}$$

However, as we obtain the LP coefficients through an autocorrelation method, resulting LPCCs remain strongly dependent and consequently, violate the independence assumption of the Naïve Bayes classifier. Nevertheless, Zhang (2004) has demonstrated that such a condition is not necessary to satisfy in reality. Indeed, no matter how strong dependencies among attributes are, Naïve Bayes can still be optimal if they are distributed evenly in class, or if they cancel each other out. Moreover, we have ob-

served that the distribution of our features, for all classes, when compared to their frequency, follows a normal distribution. Hence, it is possible to assume a valuable classification rate with Naïve Bayes according to the supposed quality of the LPCCs.

Here, we propose a simple concrete example of the Naïve Bayes classification algorithm. We describe the context as follows. You are talking online with someone called Alex you do not know and you want to determine if Alex is a male or a female, without asking him/her directly. First, we assume that there are two classes, $c_1 = male$ and $c_2 = female$ and that we have a training model with names and sex given by Table 3.2.

TABLE 3.2: TRANING MODEL OF THE CONCRET EXAMPLE OF THE NAÏVE BAYES CLASSIFIER

NAME	SEX	NAME	SEX
Alex	Male	Matthew	Male
Emily	Female	Sarah	Female
Alex	Female	Alyssa	Female
Alex	Female	Ethan	Male

Then, we can use this model to apply the Bayes rule given in equation (3.14) as follows,

$$P(male|Alex) = \frac{P(Alex|male) \times P(male)}{P(Alex)} = \frac{{}^{1}/_{3} \times {}^{3}/_{8}}{{}^{3}/_{8}} = \frac{0.125}{{}^{3}/_{8}}$$
$$P(female|Alex) = \frac{P(Alex|female) \times P(female)}{P(Alex)} = \frac{{}^{2}/_{5} \times {}^{5}/_{8}}{{}^{3}/_{8}} = \frac{0.250}{{}^{3}/_{8}} \quad (3.16)$$

Since the probability of being named Alex is the same for all the two classes, it is actually irrelevant in this example. Hence, seeing as $P(female|Alex) = 0.250 > P(male|Alex) = 0.125$, Alex that we are chatting with, is a woman according to our Naïve Bayes classification.

3.3.5 DECISION-MAKING

The result provided by the classification task may return two kinds of errors. On the one hand, a false-negative outcome refers to a failure in a genuine authentication, while a false-positive result concerns an impostor attempt mistakenly identified. According to authentication on mobile devices, false-negative does not compromise the security of the private content. However, it may be disturbing for the user since either the process has to be repeated, or a fallback mechanism has to be used. In contrast, false-positive exhibits a serious vulnerability for the security of such devices since the objective is to avoid fraudulent accesses. Besides, speaker authentication systems are not devoid of other drawbacks that may also lead to security threats. Indeed, they remain vulnerable to voice mimicry or mock authentication through legitimate voice records.

Hence, we suggest improving the authentication process by introducing the notion of access privileges in order to ward against misidentification. Firstly, we assume that users have assigned a right for each application installed on their mobile device beforehand. Therefore, we define three privileges as follows. The public privilege allows the access to only non-critical content and applications. The protected privilege

restricts the access to the most critical pieces of data (*i.e.* bank account). Finally, the private privilege gives the access to the entire content of the mobile device.

The process of determining the safest authority to grant that we suggest begins by verifying the result produced by the identification process. If the voice does not match with a genuine one then, the system allows the user to have a public access. In case of false-negative, the user has to repeat the entire process otherwise, an impostor identification is avoided. If a match does exist, a protected access is granted and the current location is fetched. In that case, the system verifies that the position is inside a given radius between 200 and 500 meters of one trusted location—where trusted locations refer to a predefined set of places connected with the user (*i.e.* home, work). This verification allows us to be quite more robust against fraudulent authentication attempts. However, a risk still exists, especially when we are facing users living together such as a family or roommates who obviously share at least, one same location. Hence, to reduce chances for a user to be unwillingly authenticated on his own device, we offer to proceed another verification. Indeed, we suggest that the private access level must only be allowed when the authentication process is achieved while using a headset and all previous verification are satisfied. Therefore, we both verify that the headset is plugged into the output of the device and that it provides an extra microphone to bypass the built-in one. Thus, we judge that it represents an additional level of security when there are shared trusted location. In that sense, we assume that false acceptance rates must considerably decrease. Figure 3.6 graphically summarize this process.

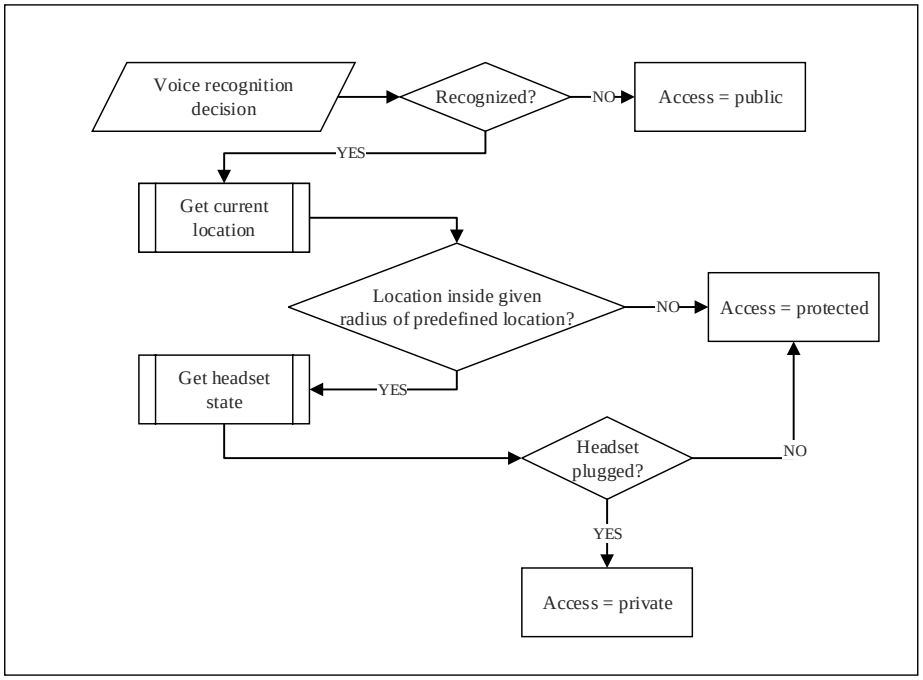

FIGURE 3.6: FLOWCHART OF THE DECISION-MAKING PROCESS

3.4 CONCLUSION

In this chapter, we have proposed a text-independent speaker authentication system for mobile devices. This implementation operates as stand-alone which does not require any network communications. Indeed, both training and identification phases are achieved onto the device itself.

The first objective in designing such a system was about to take the user into account during the authentication process. In that sense, the identification step was enhanced thanks to a decision-making that substantially relies on users' locations and the presence of a headset. This improvement in the final decision led us to introduce the notion of access privilege that each user has to set up, beforehand, for each application.

To this end, users may still perform, in risky situations, several tasks on their devices which do not involve more sensitive contents. Consequently, it remains less cumbersome than several binary biometric authentication schemes such as a fingerprint system, and users may be authenticated in a continuous manner.

Moreover, since an efficient system according to mobile devices computational capabilities was desired, a particular attention was paid in the choice of techniques for voice features extraction, as well as classification. Hence, a total of 20 LPCCs was decided to be extracted from the voice input signal. Then the statistical Naïve Bayes classifier was selected to perform the classification process because of fitting input features, as well as its linear time complexity.

CHAPTER IV

EXPERIMENTS AND RESULTS

4.1 INTRODUCTION

Firstly, in order to assess the accuracy of the proposed text-independent speaker authentication system, we suggest an experimental protocol. Hence, to achieve such an evaluation, every participant used the system to authenticate himself or herself on a provided mobile device, with a headset plugged. Two distinct environmental conditions were exploited in this experiment. The first one was a quiet environment, where the training process and a first authentication attempt were completed. Finally, another authentication attempt was performed within a noisy environment in order to evaluate the robustness of such a mechanism. Consequently, this experiment let us suggest a public data set of 11 speakers voice features (*i.e.* UQAC-Students) as a comparison purpose for other researches in the field of speaker authentication. The section 4.2 describes in detail the experimental protocol we adopted.

Secondly, section 4.3 exposes and discusses results we obtained further to the experiment. Moreover, a comparison with existing data sets was performed in order to provide a more complete evaluation of our system. Furthermore, we provide computation performance, to state about its efficiency and its reliability in running on weakest mobile devices present on the market. Finally, the last point of this section exposes results from the participants' survey we supervise. Such an opinion collection let us

better understand users' habits and needs in terms of authentication mechanisms, as well as their point of view concerning the usability of our proposed system.

Finally, section 4.4 closing this chapter by drawing a conclusion which sums up the interpretation of obtained results for each evaluation. Then, the section also exposes possible future work concentrations concerning a life-size experiment.

4.2 EXPERIMENTAL PROTOCOL

4.2.1 PARTICIPANTS

We recruited 11 university students as participants, 7 males, and 4 females from 19 to 36 years. All participants were speaking French but some distinct accent, such as French Canadian, were observed. Moreover, they were either iOS, or Android users and owned at least one recent mobile device (*i.e.* smartphone or tablet). Furthermore, 9 of the participants used an unlocking mechanism for their smartphone (PIN: 4, pattern: 2, fingerprint: 3) and fingerprint users either had a PIN code, or a pattern as fallback mechanism.

4.2.2 DATA COLLECTION

The proposed text-independent system was implemented as an Android application which requires at least the 4.0.1 version of the mobile operating system. All participants performed the experiment on the same smartphone (*i.e.* LG Nexus 5 running Android 6.0.1 with a Snapdragon 800 Quad-core at 2.3 GHz CPU and 2 GB of

RAM) with the same headset (*i.e.* Bose SoundTrue I) in same conditions (*i.e.* room and public place).

Since it was desired to have real-environment recording conditions, a quiet room was selected to achieve the training, as well as the quiet identification session. Conversely, the noisy session was performed in the cafeteria of the University. The sound level of each distinct place was measured thanks to a sound level meter embedded in the application. The mean value evaluated reached 16.5 dB in the quiet environment while 95 dB was observed in the noisy one.

4.2.3 PROCEDURE

In the beginning, participants were introduced to the experimental procedure and the current position was added to the trusted location list.

Then, training participant voices was the first phase of the experiment. To complete such an operation, a text was randomly selected in a database and displayed on the screen of the device. Participants were instructed to wear the headset and to familiarize with the content. Once they were ready, participants were advised to start the recording by themselves and next, to begin reading the text aloud. The record was automatically stopped after one minute by the application and participants were warned through both a vibration and a text-to-speech synthesis system. At that point, participants were asked to wait until the end of the computation. In the meantime, the main recorded file was split into 12 seconds chunks, being five instances per class in total. Each set of features from each instance were written in the data set which was used to

create the training model of the Naïve Bayes classifier, as described previously. Finally, participants were advised of the completion of the process thanks to a pop-up message.

At the end of the training process, the authentication process starts. This procedure was performed twice. In the first place, participants were asked to wear the headset and to pronounce the locution of their choice in the quiet environment. In the second place, they were requested to execute the same task in the noisy environment. Insofar as there was no restriction on the locution which had to be said—participants were able to use either two different expressions, or the same one for the two authentication sessions. Since every authentication attempts were performed in the same place, our decision-making has always stated that users stood in a trusted location. Therefore, we have mocked a location which was not considered as a trusted one afterwards, in order to verify the reliability of our technique. Figure 4.1 summarizes the proceedings of the experiment we conducted using a sequence diagram.

Finally, in the last step of the experiment, participants were sounded out about their habits concerning authentication on their own device, as well as their opinion as regards the proposed system.

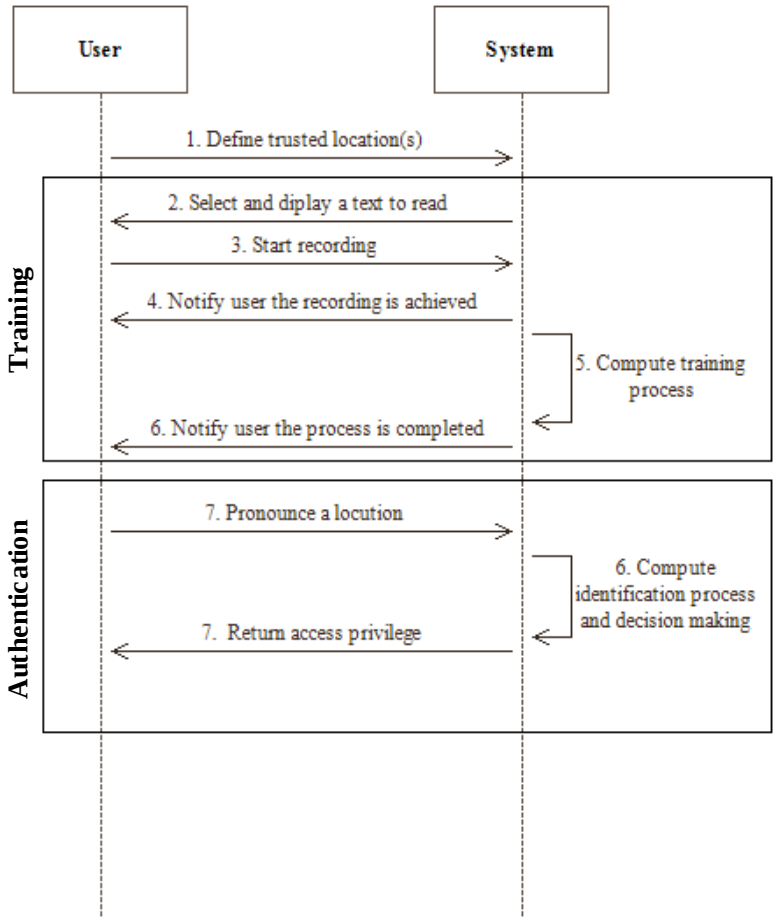

FIGURE 4.1: SEQUENCE DIAGRAM OF THE EXPERIMENT

4.3 RESULTS AND DISCUSSION

4.3.1 SPEECH CORPUSES

In this research, we have evaluated the performance of our system by exploiting two additional speech corpuses for comparison purpose with the data set we suggest.

The first one is the Ted-LIUM corpus which has been proposed by Rousseau *et al.* (2014). It includes a total of 1495 audio files extracted from TED talks, where all

speeches are English-based with multiple distinct accents. These records are mono-channel, and they are encoded in 16-bit signed integer PCM at a 16 kHz sampling rate. Although the corpus was published using the NIST Sphere format (SPH), we required to convert the whole files in Waveform Audio File Format (WAV). Furthermore, we took care of removing the first 20 seconds of each file, as they correspond to the talk opening sequence.

The second speech corpus we have exploited in this research, is the Ahumada_25 corpus, which has been suggested by Ortega-Garcia *et al.* (2000). This is a subset of the full database which provides several recording sessions in different conditions, as well as different hardware for 25 Spanish speakers. However, we have only exploited recording sessions where a microphone has been used to collect free speech talks, being 6 distinct training sets in total. Such provided files are also mono-channel and encoded in 16-bit signed integer PCM at a 16 kHz sampling rate.

4.3.2 CLASSIFICATION PERFORMANCE METRICS

Since classification let us predict at which registered speaker corresponds a given utterance; it is important to evaluate the performance of our system thanks to representative metrics. To this end, the accuracy is probably the most dominant measure in the literature, because of its simplicity. This measure provides the ratio between the correct number of predictions and the total number of cases given as,

$$accuracy = \frac{TP + TN}{TP + TN + FP + FN}. \tag{4.1}$$

where TP and TN refer to true positive and true negative predictions respectively, and the total additionally include false positive (FP) and false negative (FN) predictions.

Despite its popularity, accuracy alone does typically not provide enough information to evaluate the robustness of prediction outcomes. Indeed, accuracy does not compensate for results that may be expected by luck. Indeed, a high accuracy does not necessarily reflect an indicator of a high classification performance. This is the accuracy paradox. For instance, in a predictive classification setting, predictive models with a given level of accuracy may have greater predictive power than models with higher accuracy. In that sense, as suggested by Ben-David (2007), we decided to provide the Cohen's kappa evaluation metric as well. This measure takes into account such a paradox and remains a more relevant metric in multiclass classification evaluations such as our system. The kappa measure $\in [0,1]$ is given by,

$$k = \frac{P_o - P_e}{1 - P_e}. \tag{4.2}$$

where P_0 and P_e are observed and the expected probabilities respectively. As an example, we suppose that we are classifying data related to a group of 14 people applying for a PhD degree admission. Each application is evaluated by two professors (A and

B) and they either say “Yes” or “No”. Their opinion was collected and exposed in a matrix form as follows,

		B	
		YES	NO
A	YES	a	b
	NO	c	d

		B	
		YES	NO
A	YES	8	1
	NO	1	4

In this case, $P_o = \frac{(a+d)}{a+b+c+d} = \frac{8+4}{14} = 0.857$. To obtain the expected probability P_e, we first calculate both probabilities of A and B denoted P_A and P_B respectively, where $P_A = \frac{(a+b)\times(a+c)}{a+b+c+d} = \frac{81}{14}$ and $P_B = \frac{(c+d)\times(b+d)}{a+b+c+d} = \frac{25}{14}$. Thus, the overall expected probability is given by, $P_e = \frac{P_A+P_B}{a+b+c+d} = 0.541$. Finally, when applying the formula for the Cohen’s Kappa we get, $k = \frac{P_o+P_e}{1-P_e} = \frac{0.857-0.541}{1-0.541} = 0.688$.

4.3.3 RESULTS OBTAINED

The performance of our proposed system was evaluated according to several analyses. First of all, results of the experiment we described previously are shown in Table 4.1. In this evaluation, we have exploited testing instances we obtained over our experiment for both quiet and noisy environments. Thanks to such achieved results it is possible to observe that our system yields an accurate identification of voices in real-environmental conditions with our own instances.

TABLE 4.1: RESULTS OF THE EXPERIMENT BASED ON THE REALIZED DATA SET: UQAC-STUDENTS

	UQAC-STUDENTS	
	QUIET ENVIRONMENT	NOISY ENVIRONMENT
KAPPA	0.90	0.80
ACCURACY	91%	81.82%
NUMBER OF CLASSES	11	
NUMBER OF INSTANCES USED FOR TRAINING	5	
NUMBER OF INSTANCES USED FOR TESTING	1	1

In order to compare such results, we have constructed related data sets thanks to the Ted-LIUM and the Ahumada_25 corpuses. However, since the Ahumada_25 corpus contains a total of 6 distinct recording sessions of 25 classes, we judged that it was a necessity to unify the Ted-LIUM corpus accordingly. In that sense, we have created 6 different training sets by selecting 25 samples randomly over the 1494 files. Moreover, we have also ensured that a sample was not chosen more than once for a given batch. For each corpus, samples of all sessions were split into 7 instances of 12 seconds. In order to be more consistent with our experimental procedure, the first 5 instances were used in the training phase; while the last two were exploited for the identification. Obtained results over these two data sets are exposed in Figure 4.2. Although achieved results on Ted-LIUM data set were valuable (0.94 of kappa at best), the ones we got on Ahumada_25 data set were not as much accurate (0.75 of kappa at best), but remain promising. Such a change between these two data sets may be related to several factors such as the quality of microphones or the language of the speaker that

may affect the quality of the input features even if it is a text-independent system (Kinnunen and Li 2010).

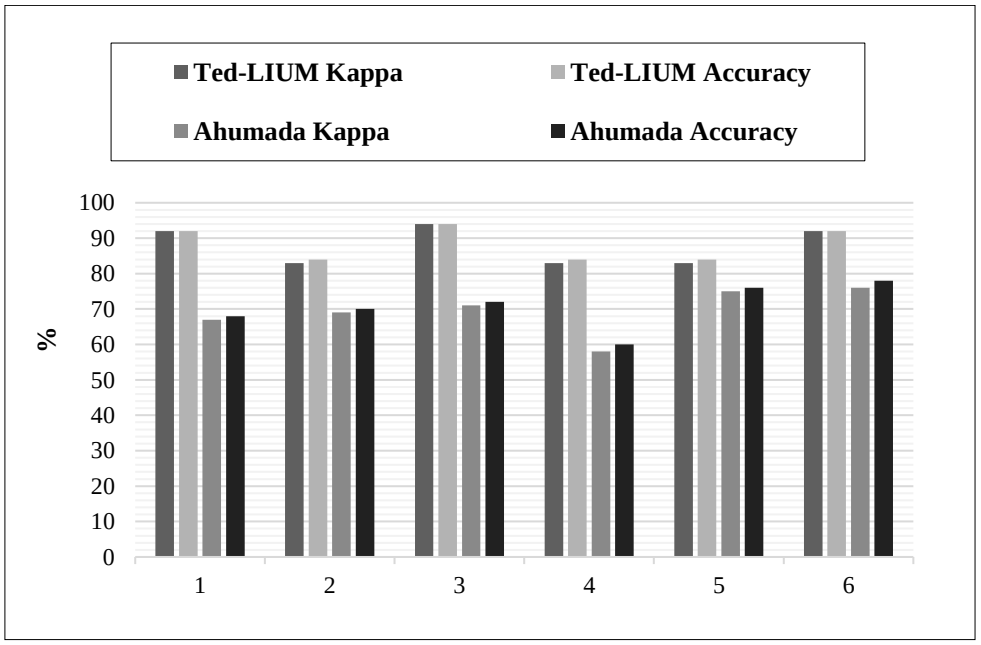

FIGURE 4.2: ACCURACY AND KAPPA MEASURES ACHIEVED BY OUR SYSTEM OVER THE 6 DATA SETS OF BOTH TED-LIUM AND AHUMADA CORPUSES

However, as these evaluations involve a relatively low number of distinct classes, we point out the analysis of the evolution of the kappa measure when increasing the number of classes. The Ted-LIUM corpus let us perform such an appraisal since it is the largest corpus we used in this research. Hence, we did not change the number of instances that we have exploited in the previous evaluation, 5 instances per class for the training and 2 for the identification phase. We chose to compute the kappa by increasing exponentially the number of classes to the total of 1495. Figure 4.3 shows that the more there are classes, the more the kappa measure tends to decrease. Indeed, our

system obtains a kappa of 0.51 where the entire set of classes was used in the identification process. Such a result was expected since we are not facing a binary classification problem.

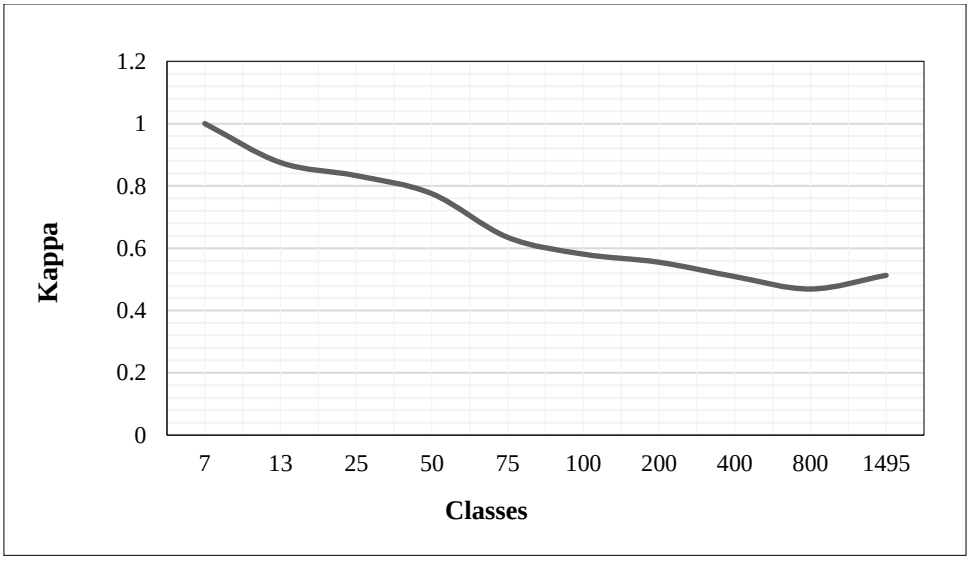

FIGURE 4.3: EVOLUTION OF THE KAPPA MEASURE OVER THE TED-LIUM CORPUS WHEN ICREASING EXPONENTIALLY THE NUMBER OF CLASSES

4.3.4 COMPUTATION PERFORMANCE CONSIDERATIONS

The primary concern in designing such a system was about building an efficient authentication mechanism for mobile devices.

To this end, we have chosen fitting techniques with attention to time complexity and memory consumption. Figure 4.4 exposes a profiling of both CPU and memory utilization of the mobile device, over every phase of the experiment we conduct. As a result, we reach the highest peak through the training phase, being 45% CPU and 70

MB usage. Indeed, since the features extraction relies on a sophisticated series of operations, we parallelized the algorithm by distributing the calculation over the different cores provided by the processor to reduce the time complexity. Since we wanted to exploit all the available hardware resources, the CPU consumption undoubtedly grows but its usage remains largely satisfying.

Thanks to these measurements, it is possible for us to assess that our system remains broadly satisfactory regarding both time and memory consumption when running on a Nexus 5 Android mobile device. With regards to less efficient mobile devices, our system should be subject to further testing to be assessed as effective. Although memory consumption should not be a real problem, we expect a slight increase in the computation time with less powerful devices.

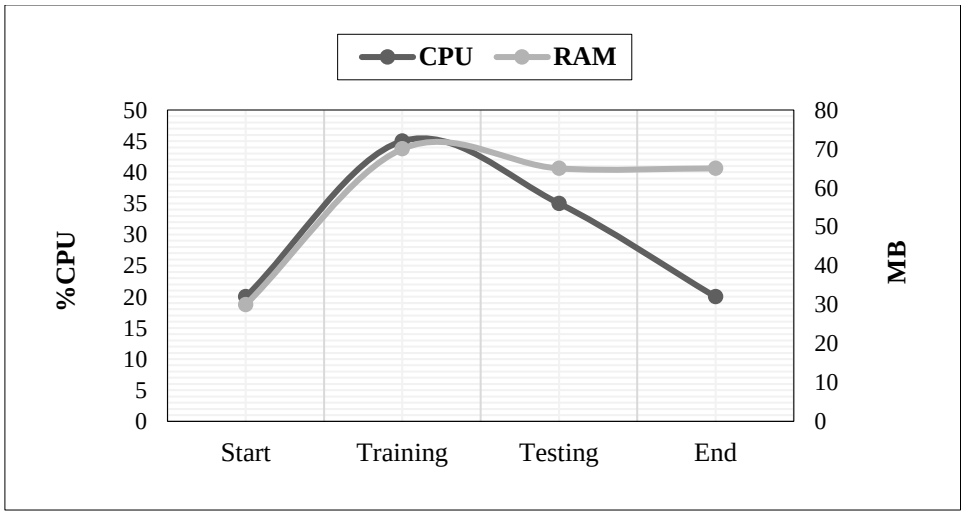

FIGURE 4.4: CPU AND RAM CONSUMPTION, RESPECTIVELY EXPRESSED IN %CPU AND MB, OVER EVERY STEP OF THE EXPERIMENT

4.3.5 PARTICIPANTS OPINION CONSIDERATIONS

Here we report the participants' opinions concerning the proposed system. Hence, it aims at better understanding users' needs and habits as regards authentication in order to replace present mechanisms offered on mobile devices.

This survey shows that 2 of users who enable a knowledge-based authentication mechanism (*i.e.* PIN or pattern), have reported that it is overly repetitive and lead them to make mistakes several times a day. Besides, every fingerprint user has mentioned a bothering malfunction with finger moisture. As a result, 3 participants over the 9 who lock their device, as well as 1 over the two participants who do not employ such security, would use this system as a replacement of their present authentication scheme because of its simplicity. Moreover, 8 respondents have mentioned they could place their confidence in the described system. However, remaining participants have declared that talking to their mobile device could be annoying in public areas and consequently, they have claimed that they do not trust any voice-based authentication scheme.

This opinion collection now allows us to state that our system could become a relevant authentication mechanism for several users. However, an opinion collection on a larger population of users should better reflect such observation. In addition, since it is text-independent, such a system is able to perform the authentication in a continuous manner, without any involvement from the user. In that sense, anxieties, as regards the discomfort in talking to a device in public places which were reported in the survey, may be reduced to void. Therefore, we esteem that such a technique may be a more

significant option as part of a multilayer authentication. Moreover, it should also be better employed as a more reliable fallback solution in order to eradicate PIN codes.

4.4 CONCLUSION

Results we have obtained over the different analysis we performed, suggest that our proposed system embodies an accurate and reliable authentication in both quiet and noisy environments, with a plugged headset (*i.e.* 0.90 and 0.80 of kappa in quiet and noisy environments, respectively) for the 11 French people involved in the experimentation. Besides, computation performance outcomes demonstrate that this mechanism is also an efficient way to authenticate users since we have observed encouraging results in its ability to run on weakest mobile devices. However, further experimentations with less powerful devices are required to prove such an observation.

We found that 7 users were still not ready to switch from their present authentication mechanism. Moreover, 27% of the participants have reported that they could not place their confidence in such a system, as it may be disturbing when used in public places. However, since it is text-independent, legitimate users may be implicitly authenticated as they start speaking, insofar as the mobile device is neither in their pocket, nor their bag (*i.e.* during a conversation). In that sense, we suggest that this technique should be either used in a multilayer authentication system, or as a fallback mechanism, namely when the first one fails, to cover most of users' needs and usages.

Future works will focus on offering the application on the *Google Play Store* to better assess the accuracy and the robustness of the proposed authentication system.

However, the current implementation will be adapted in order to let us track user authentication attempt outcomes and locations. In this way, such a large-scale evaluation will provide more reliable results in front of real life condition usages and the location-based decision will be better exploited and significant, as it was in the experiment we have conducted in this research.

CHAPTER V

CONCLUSION

Over the past few years, several authentication mechanisms were proposed. These systems may be divided into three broad categories: knowledge-based, token-based and biometric authentications. However, the taxonomy presented in chapter 2 led us to maintain that existing authentication systems concede several advantages and drawbacks. Moreover, we pointed out that most of these schemes were designed to a machine point of view. Consequently, users have adopted wrong behaviors which lead to important security threats and weaknesses.

Given that, the first effort of this research project was to introduce several evaluation criteria based on strengths and weaknesses which were underlined in our review. These criteria allowed us to evaluate each authentication systems to understand their pros and their cons as regards users' considerations and usability. Then, by way of conclusion, we proposed an appraisal lean toward the aftermath of authentication on mobile devices that head for new perspectives concerning *no password* and *ubiquitous* authentication mechanisms.

As stated in chapter 3, speaker authentication systems remain an inexpensive solution to authenticate users. Indeed, no additional sensors are required and these mechanisms may play a major role in some real-world applications to secure identity

management systems such as e-commerce solutions, attendance systems, mobile banking or forensics. However, we showed that few speaker authentication systems are presently implemented on mobile devices.

Accordingly, the second concern of this research was to suggest a practical implementation of a text-independent speaker authentication mechanism for mobile devices. Our main concentrations in designing such a system were about both its efficiency in terms of computational requirements, as well as its accuracy in authenticating users. To this end, the chapter 3 introduced the proposed authentication mechanism. It is based on the extraction of LPCCs to obtain relevant voice features, as well as the Naïve Bayes classifier to predict at which speaker a given utterance corresponds. Moreover, the authentication decision was enhanced through the affectation of a given access privilege (*i.e.* public, protected or private), for each authentication attempt. Such an authority is granted through the analysis of the user's location and the presence of a headset.

The chapter 4, allowed us to state that the proposed system entirely responds to research problems formulated in chapter 1. Indeed, we first evaluated its accuracy through an experiment we conducted over eleven participants' voice samples, in different sort of environments (*i.e.* quiet and noisy). Although the results we obtained on our data set are accurate, we also experienced the system with other existing data sets (*i.e.* Ted-LIUM and Ahumada_25). Over again, results remained similarly accurate on the Ted-LIUM dataset and stayed relevant when increasing the number of classes to predict. However, we observed a minor drop of accuracy with Ahumada_25 samples which may be mainly due to their initial recording quality.

Finally, as personal assessment, I would say that this research project was a rewarding experience as a first step into the world of research. I was able to successfully conduct this project because of its interesting characteristics and all the knowledge I could acquire in working on it. I particularly collect strong knowledge in the field of machine learning and signal processing. Moreover, this experience allowed me to develop other important new skills such as a rigorous research methodology, solid writing skills in English, as well as communication skills. This project also let me produce two scientific publications, where the first one is accepted (Thullier *et al.* 2016) and the second one is, for now on, submitted to a scientific journal about *Information Security*. Finally, such a positive experience pushes me to pursuing doctoral studies, since I want to challenge myself even further.

REFERENCES

Adini Y, Moses Y and Ullman S. 1997. Face recognition: The problem of compensating for changes in illumination direction. Pattern Analysis and Machine Intelligence, IEEE Transactions on, 19: 721-732.

Al-Hassani MD and Kadhim AA. Design A Text-Prompt Speaker Recognition System Using LPC-Derived Features. In: The 13th International Arab Conference on Information Technology ACIT, 2012. p. 10-13.

Aloul F, Zahidi S and El-Hajj W. Two factor authentication using mobile phones. In: AICCSA, 2009. p. 641-644.

Aviv AJ, Gibson K, Mossop E, Blaze M and Smith JM. 2010. Smudge Attacks on Smartphone Touch Screens. WOOT, 10: 1-7.

Belfiore J. 2015. Making Windows 10 More Personal and More Secure with Windows Hello. Accessed 2016 June 16th, http://bit.ly/2cpnNsP

Belgacem N, Fournier R, Nait-Ali A and Bereksi-Reguig F. 2015. A novel biometric authentication approach using ECG and EMG signals. Journal of medical engineering & technology, 39: 226-238.

Ben-Asher N, Kirschnick N, Sieger H, Meyer J, Ben-Oved A and Möller S. On the need for different security methods on mobile phones. In: Proceedings of the 13th International Conference on Human Computer Interaction with Mobile Devices and Services, 2011. ACM, p. 465-473.

Ben-David A. 2007. A lot of randomness is hiding in accuracy. Engineering Applications of Artificial Intelligence, 20: 875-885.

Benesty J, Sondhi MM and Huang Y. 2007. Springer handbook of speech processing. Springer Science & Business Media.

Benyon D, Turner P and Turner S. 2005. Designing interactive systems: People, activities, contexts, technologies. Pearson Education.

Berry J and Stoney DA. 2001. The history and development of fingerprinting. Advances in fingerprint Technology, 2: 13-52.

Bertillon A. 1893. Identification anthropométrique: instructions signalétiques. Impr. administrative.

Bianchi A, Oakley I and Kwon DS. The secure haptic keypad: a tactile password system. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, 2010. ACM, p. 1089-1092.

Bianchi A, Oakley I and Kwon DS. 2012. Counting clicks and beeps: Exploring numerosity based haptic and audio PIN entry. Interacting with computers, 24: 409-422.

Biddle R, Chiasson S and Van Oorschot PC. 2012. Graphical passwords: Learning from the first twelve years. ACM Computing Surveys (CSUR), 44: 19.

Blonder GE. 1996. Graphical password. U.S. Patent No. 5,559,961. Washington, DC: U.S. Patent and Trademark Office.

Bond RH, Kramer A and Gozzini G. 2012. Molded fingerprint sensor structure with indicia regions. U.S. Patent No. D652,332.

Boves L and Den Os E. Speaker recognition in telecom applications. In: Interactive Voice Technology for Telecommunications Applications, 1998 IVTTA'98 Proceedings 1998 IEEE 4th Workshop, 1998. IEEE, p. 203-208.

Boyd JE. 2004. Synchronization of oscillations for machine perception of gaits. Computer Vision and Image Understanding, 96: 35-59.

Brunet K, Taam K, Cherrier E, Faye N and Rosenberger C. Speaker Recognition for Mobile User Authentication: An Android Solution. In: 8ème Conférence sur la Sécurité des Architectures Réseaux et Systèmes d'Information (SAR SSI), 2013. p. 10.

Bunnell J, Podd J, Henderson R, Napier R and Kennedy-Moffat J. 1997. Cognitive, associative and conventional passwords: Recall and guessing rates. Computers & Security, 16: 629-641.

Chen H and Bhanu B. Contour matching for 3D ear recognition. In: Application of Computer Vision, 2005 WACV/MOTIONS'05 Volume 1 Seventh IEEE Workshops on, 2005. IEEE, p. 123-128.

Choraś M. 2005. Ear biometrics based on geometrical feature extraction. Progress in Computer Vision and Image Analysis: 321.

Clarke NL and Furnell SM. 2005. Authentication of users on mobile telephones–A survey of attitudes and practices. Computers & Security, 24: 519-527.

Clarke NL, Furnell SM, Rodwell PM and Reynolds PL. 2002. Acceptance of subscriber authentication methods for mobile telephony devices. Computers & Security, 21: 220-228.

Council FFIE. 2005. Authentication in an internet banking environment. Financial Institution Letter, FIL-103-2005 Washington, DC: Federal Deposit Insurance Corp(FDIC) Retrieved March, 18: 2005.

Cuadrado F and Dueñas JC. 2012. Mobile application stores: success factors, existing approaches, and future developments. Communications Magazine, IEEE, 50: 160-167.

Dabbah M, Woo W and Dlay S. Secure authentication for face recognition. In: Computational Intelligence in Image and Signal Processing, 2007 CIISP 2007 IEEE Symposium on, 2007. IEEE, p. 121-126.

Daugman J. 2004. How iris recognition works. Circuits and Systems for Video Technology, IEEE Transactions on, 14: 21-30.

Deng Y and Zhong Y. 2015. Keystroke Dynamics Advances for Mobile Devices Using Deep Neural Network. In: Recent Advances in User Authentication Using Keystroke Dynamics Biometrics. p. 59-70.

Derawi MO, Nickel C, Bours P and Busch C. Unobtrusive user-authentication on mobile phones using biometric gait recognition. In: Intelligent Information Hiding and Multimedia Signal Processing (IIH-MSP), 2010 Sixth International Conference on, 2010. IEEE, p. 306-311.

Doddington GR, Przybocki MA, Martin AF and Reynolds DA. 2000. The NIST speaker recognition evaluation–overview, methodology, systems, results, perspective. Speech Communication, 31: 225-254.

Eriksson A and Wretling P. How Flexible is the Human Voice? A Case Study of Mimicry. In: Proceedings of the European Conference on Speech Technology (Eurospeech ’97), Rhodes, Greece, 1997. p. 1043-1046.

Falaki H, Mahajan R, Kandula S, Lymberopoulos D, Govindan R and Estrin D. Diversity in smartphone usage. In: Proceedings of the 8th international conference on Mobile systems, applications, and services, 2010. ACM, p. 179-194.

Fatemian SZ, Agrafioti F and Hatzinakos D. HeartID: Cardiac biometric recognition. In: Biometrics: Theory Applications and Systems (BTAS), 2010 Fourth IEEE International Conference on, 2010. IEEE, p. 1-5.

Fathy ME, Patel VM and Chellappa R. Face-based active authentication on mobile devices. In: Acoustics, Speech and Signal Processing (ICASSP), 2015 IEEE International Conference on, 2015. IEEE, p. 1687-1691.

Furui S. 1981. Cepstral analysis technique for automatic speaker verification. IEEE Transactions on Acoustics, Speech, and Signal Processing, 29: 254-272.

Gafurov D, Helkala K and Søndrol T. 2006. Biometric gait authentication using accelerometer sensor. Journal of computers, 1: 51-59.

Godfrey J, Graff D and Martin A. Public databases for speaker recognition and verification. In: ESCA Workshop on Automatic Speaker Recognition, Identification and Verification, 1994.

Goggin G. 2012. Cell phone culture: Mobile technology in everyday life. London: Routledge.

Gold B, Morgan N and Ellis D. 2011. Speech and audio signal processing: processing and perception of speech and music. John Wiley & Sons.

Grother PJ, Quinn GW and Phillips PJ. 2010. Report on the evaluation of 2D still-image face recognition algorithms. NIST interagency report, 7709: 106.

Gugenheimer J, De Luca A, Hess H, Karg S, Wolf D and Rukzio E. ColorSnakes: Using Colored Decoys to Secure Authentication in Sensitive Contexts. In: Proceedings of the 17th International Conference on Human-Computer Interaction with Mobile Devices and Services, 2015. ACM, p. 274-283.

Hansman SL. 2003. A taxonomy of network and computer attack methodologies. Master's Thesis, University of Canterbury, New Zealand.

Hayashi E and Hong J. A diary study of password usage in daily life. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, 2011. ACM, p. 2627-2630.

Holz C, Buthpitiya S and Knaust M. Bodyprint: Biometric User Identification on Mobile Devices Using the Capacitive Touchscreen to Scan Body Parts. In: Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems, 2015. ACM, p. 3011-3014.

Jagatic TN, Johnson NA, Jakobsson M and Menczer F. 2007. Social phishing. Communications of the ACM, 50: 94-100.

Jain A, Hong L and Bolle R. 1997. On-line fingerprint verification. Pattern Analysis and Machine Intelligence, IEEE Transactions on, 19: 302-314.

Jain A, Bolle R and Pankanti S. 2006a. Biometrics: personal identification in networked society. Springer Science & Business Media.

Jain A, Maltoni D, Maio D and Wayman J. 2005. Biometric Systems Technology, Design and Performance Evaluation. London: Spring Verlag.

Jain AK, Ross A and Pankanti S. 2006b. Biometrics: a tool for information security. Information Forensics and Security, IEEE Transactions on, 1: 125-143.

Jain AK, Pankanti S, Prabhakar S, Hong L and Ross A. Biometrics: a grand challenge. In: Pattern Recognition, 2004 ICPR 2004 Proceedings of the 17th International Conference on, 2004. IEEE, p. 935-942.

Jing X, Wen-Tao Z and Deng-Guo F. 2009. An improved smart card based password authentication scheme with provable security. Computer Standards & Interfaces, 31: 723-728.

John GH and Langley P. Estimating continuous distributions in Bayesian classifiers. In: Proceedings of the Eleventh conference on Uncertainty in artificial intelligence, 1995. Morgan Kaufmann Publishers Inc., p. 338-345.

Khan MG. 2008. Rapid ECG interpretation. Humana Press, New Jersey.

Kinnunen T. 2003. Spectral Features for Automatic Text-Independent Speaker Recognition. Licentiate's Thesis. University of Joensuu, Department of computer science, Finland.

Kinnunen T and Li H. 2010. An overview of text-independent speaker recognition: From features to supervectors. Speech Communication, 52: 12-40.

Kirkpatrick EA. 1894. An experimental study of memory. Psychological Review, 1: 602.

Kokumai H. 2015. Password-dependent password-killer. Accessed 2015 June 20th, http://bit.ly/2cpIULB

Kroeker KL. 2002. Graphics and security: Exploring visual biometrics. Computer Graphics and Applications, IEEE, 22: 16-21.

Kumar A, Wong DC, Shen HC and Jain AK. Personal verification using palmprint and hand geometry biometric. In: Audio-and Video-Based Biometric Person Authentication, 2003. Springer, p. 668-678.

Kumar R, Ranjan R, Singh SK, Kala R, Shukla A and Tiwari R. 2009. Multilingual Speaker Recognition Using Neural Network. Proc of the Frontiers of Res on Speech and Music, FRSM: 1-8.

Lamport L. 1981. Password authentication with insecure communication. Communications of the ACM, 24: 770-772.

Lanitis A, Taylor CJ and Cootes TF. 2002. Toward automatic simulation of aging effects on face images. Pattern Analysis and Machine Intelligence, IEEE Transactions on, 24: 442-455.

Lashkari AH, Farmand S, Zakaria D, Bin O and Saleh D. 2009. Shoulder surfing attack in graphical password authentication. International Journal of Computer Science and Information Security, 6: 145-154.

Lau YW, Wagner M and Tran D. Vulnerability of speaker verification to voice mimicking. In: Intelligent Multimedia, Video and Speech Processing, 2004 Proceedings of 2004 International Symposium on, 2004. IEEE, p. 145-148.

Laurence Goasduff JR. 2015. Market Share: Devices, All Countries, 4Q14 Update. Gartner.

Lazar L, Tikolsky O, Glezer C and Zviran M. 2011. Personalized cognitive passwords: an exploratory assessment. Information Management & Computer Security, 19: 25-41.

LeCun YA, Bottou L, Orr GB and Müller K-R. 2012. Efficient backprop. In: Neural networks: Tricks of the trade. Springer, p. 9-48.

Li SZ. 2009. Encyclopedia of Biometrics: I-Z. Springer Science & Business Media.

Lowe G. A hierarchy of authentication specifications. In: Computer Security Foundations Workshop, 1997 Proceedings, 10th, 1997. IEEE, p. 31-43.

Mansfield AJ and Wayman JL. 2002. Best practices in testing and reporting performance of biometric devices. Centre for Mathematics and Scientific Computing, National Physical Laboratory Teddington, Middlesex, UK.

Martínez AM. 2002. Recognizing imprecisely localized, partially occluded, and expression variant faces from a single sample per class. Pattern Analysis and Machine Intelligence, IEEE Transactions on, 24: 748-763.

Matsumoto T, Matsumoto H, Yamada K and Hoshino S. Impact of artificial gummy fingers on fingerprint systems. In: Electronic Imaging 2002, 2002. International Society for Optics and Photonics, p. 275-289.

Micallef N, Just M, Baillie L, Halvey M and Kayacik HG. Why aren't Users Using Protection? Investigating the Usability of Smartphone Locking. In: Proceedings of the 17th International Conference on Human-Computer Interaction with Mobile Devices and Services, 2015. ACM, p. 284-294.

Miorandi D, Sicari S, De Pellegrini F and Chlamtac I. 2012. Internet of things: Vision, applications and research challenges. Ad Hoc Networks, 10: 1497-1516.

Morris R and Thompson K. 1979. Password security: A case history. Communications of the ACM, 22: 594-597.

Nair R and Salam N. A reliable speaker verification system based on LPCC and DTW. In: Computational Intelligence and Computing Research (ICCIC), 2014 IEEE International Conference on, 2014. IEEE, p. 1-4.

Nalwa VS. 1997. Automatic on-line signature verification. Proceedings of the IEEE, 85: 215-239.

Nanavati S, Thieme M and Nanavati R. 2002. Biometrics, Identity Verification in a Networked World. Wiley Computer Publishing.

Nickel C, Wirtl T and Busch C. Authentication of smartphone users based on the way they walk using k-NN algorithm. In: Intelligent Information Hiding and Multimedia Signal Processing (IIH-MSP), 2012 Eighth International Conference on, 2012. IEEE, p. 16-20.

Nielsen. 2014. Smartphones: so many apps, so much time. Accessed 2015 November 24th http://bit.ly/TFXobf

O'Shaughnessy D. 1988. Linear predictive coding. IEEE potentials, 7: 29-32.

Obaidat MS. A verification methodology for computer systems users. In: Proceedings of the 1995 ACM symposium on Applied computing, 1995. ACM, p. 258-262.

Orgill GL, Romney GW, Bailey MG and Orgill PM. The urgency for effective user privacy-education to counter social engineering attacks on secure computer systems. In: Proceedings of the 5th conference on Information technology education, 2004. ACM, p. 177-181.

Ortega-Garcia J, Gonzalez-Rodriguez J and Marrero-Aguiar V. 2000. AHUMADA: A large speech corpus in Spanish for speaker characterization and identification. Speech Communication, 31: 255-264.

Phillips PJ, Martin A, Wilson CL and Przybocki M. 2000. An introduction evaluating biometric systems. Computer, 33: 56-63.

Rabiner L and Juang B-H. 1993. Fundamentals of speech recognition. Prentice Hall.

Rani MP and Arumugam G. 2010. An efficient gait recognition system for human identification using modified ICA. International journal of computer science and information technology, 2: 55-67.

Reynolds DA and Rose RC. 1995. Robust text-independent speaker identification using Gaussian mixture speaker models. Speech and Audio Processing, IEEE Transactions on, 3: 72-83.

Riley S. 2006. Password security: What users know and what they actually do. Usability News, 8: 2833-2836.

Ross A, Jain A and Pankati S. A prototype hand geometry-based verification system. In: Proceedings of 2nd Conference on Audio and Video Based Biometric Person Authentication, 1999. p. 166-171.

Rousseau A, Deléglise P and Estève Y. Enhancing the TED-LIUM Corpus with Selected Data for Language Modeling and More TED Talks. In: LREC, 2014. p. 3935-3939.

Sadjadi SO and Hansen JH. 2013. Unsupervised speech activity detection using voicing measures and perceptual spectral flux. Signal Processing Letters, IEEE, 20: 197-200.

Salvador S and Chan P. 2007. Toward accurate dynamic time warping in linear time and space. Intelligent Data Analysis, 11: 561-580.

Sanchez-Reillo R, Sanchez-Avila C and Gonzalez-Marcos A. 2000. Biometric identification through hand geometry measurements. Pattern Analysis and Machine Intelligence, IEEE Transactions on, 22: 1168-1171.

Sarkar S, Phillips PJ, Liu Z, Vega IR, Grother P and Bowyer KW. 2005. The humanid gait challenge problem: Data sets, performance, and analysis. Pattern Analysis and Machine Intelligence, IEEE Transactions on, 27: 162-177.

Sasse MA, Brostoff S and Weirich D. 2001. Transforming the 'weakest link'—a human/computer interaction approach to usable and effective security. BT technology journal, 19: 122-131.

Schneier B. 2005. Two-factor authentication: too little, too late. Commun ACM, 48: 136.

Senk C and Dotzler F. Biometric Authentication as a service for enterprise identity management deployment: a data protection perspective. In: Availability, Reliability and Security (ARES), 2011 Sixth International Conference on, 2011. IEEE, p. 43-50.

Sinofsky S. 2011. Signing in with a picture password. Accessed 2016 January 13th, http://bit.ly/2cd72fK

Slain M. January 19, 2015. Worst Passwords of 2015. In: SplashData éd.

Sufi F, Khalil I and Hu J. 2010. ECG-based authentication. In: Handbook of Information and Communication Security. Springer, p. 309-331.

Tari F, Ozok A and Holden SH. A comparison of perceived and real shoulder-surfing risks between alphanumeric and graphical passwords. In: Proceedings of the second symposium on Usable privacy and security, 2006. ACM, p. 56-66.

Thullier F, Bouchard B and Ménélas B-AJ. 2016. Exploring Mobile Authentication Mechanisms from PIN to Biometrics Including the Future Trend. In: Protecting Mobile Networks and Devices: Challenges and Solutions. Taylor & Francis.

Trojahn M, Arndt F and Ortmeier F. Authentication with keystroke dynamics on touchscreen keypads-effect of different N-Graph combinations. In: MOBILITY 2013, The Third International Conference on Mobile Services, Resources, and Users, 2013. p. 114-119.

Uellenbeck S, Dürmuth M, Wolf C and Holz T. Quantifying the security of graphical passwords: The case of android unlock patterns. In: Proceedings of the 2013 ACM SIGSAC conference on Computer & communications security, 2013. ACM, p. 161-172.

Vuppala A, Chakrabarti S, Rao K and Dutta L. Robust speaker recognition on mobile devices. In: Proc of IEEE Int Conf on Signal Processing and Communications, 2010.

Wildes RP. 1997. Iris recognition: an emerging biometric technology. Proceedings of the IEEE, 85: 1348-1363.

Wilska T-A. 2003. Mobile phone use as part of young people's consumption styles. Journal of consumer policy, 26: 441-463.

Yampolskiy RV. Analyzing user password selection behavior for reduction of password space. In: Carnahan Conferences Security Technology, Proceedings 2006 40th Annual IEEE International, 2006. IEEE, p. 109-115.

Yan J. 2004. Password memorability and security: Empirical results. IEEE Security & privacy: 25-31.

Yuizono T, Wang Y, Satoh K and Nakayama S. Study on individual recognition for ear images by using genetic local search. In: WCCI, 2002. IEEE, p. 237-242.

Zhang H. The optimality of naive bayes. In: Proceedings of the Seventeenth Florida Artificial Intelligence Research Society Conference, Miami Beach, 2004. AAAI Press, p. 562–567.

Zhao Z, Ahn G-J and Hu H. 2015. Picture Gesture Authentication: Empirical Analysis, Automated Attacks, and Scheme Evaluation. ACM Transactions on Information and System Security (TISSEC), 17: 14.

Zviran M and Haga WJ. User authentication by cognitive passwords: an empirical assessment. In: Information Technology, 1990'Next Decade in Information Technology', Proceedings of the 5th Jerusalem Conference on (Cat No 90TH0326-9), 1990. IEEE, p. 137-144.

APPENDIX 1

ETHICS CERTIFICATE

UQAC

Comité d'éthique de la recherche
Université du Québec à Chicoutimi

APPROBATION ETHIQUE

Dans le cadre de l'*Énoncé de politique des trois conseils : éthique de la recherche avec des êtres humains 2* (2014) et conformément au mandat qui lui a été confié par la résolution CAD-7163 du Conseil d'administration de l'Université du Québec à Chicoutimi, approuvant la *Politique d'éthique de la recherche avec des êtres humains* de l'UQAC, le Comité d'éthique de la recherche avec des êtres humains de l'Université du Québec à Chicoutimi, à l'unanimité, délivre la présente approbation éthique puisque le projet de recherche mentionné ci-dessous rencontre les exigences en matière éthique et remplit les conditions d'approbation dudit Comité.

Responsable(s) du projet de recherche :	*Monsieur Florentin Thullier,* *Étudiant, Maîtrise en informatique*
Direction de recherche :	*Monsieur Bob-Antoine Jerry Ménélas* *Professeur, Département d'informatique et de mathématiques*
Codirection de recherche :	*Monsieur Bruno Bouchard,* *Professeur, Département d'informatique et de mathématiques*
Projet de recherche intitulé :	*Authentification d'utilisateurs sur plateformes mobiles par le biais de caractéristiques de la voix*
No référence :	*602.499.01*
Financement :	*N/A*

La présente est valide jusqu'au 31 août 2016.
Rapport de statut attendu pour le 31 juillet 2016 (rapport final).
N.B. le rapport de statut est disponible à partir du lien suivant : *http://recherche.uqac.ca/rapport-de-statut/*

Date d'émission initiale de l'approbation : *25 février 2016*
Date(s) de renouvellement de l'approbation :

Nicole Bouchard,
Professeure et présidente