Confidence and Risk Estimation Plugins for Multi-Modal Authentication on Mobile Devices using CORMORANT

Daniel Hintze
FHDW Paderborn
Fürstenallee 3 - 5
33102 Paderborn, Germany
daniel.hintze@fhdw.de

Muhammad Muaaz
UAS Upper Austria
Softwarepark 11
4232 Hagenberg, Austria
muhammad.muaaz@fh-hagenberg.at

Rainhard D. Findling
UAS Upper Austria
Softwarepark 11
4232 Hagenberg, Austria
rainhard.findlinge@fh-hagenberg.at

Sebastian Scholz
FHDW Paderborn
Fürstenallee 3 - 5
33102 Paderborn, Germany
sebastian.scholz@fhdw.de

Eckhard Koch
FHDW Paderborn
Fürstenallee 3 - 5
33102 Paderborn, Germany
eckhard.koch@fhdw.de

René Mayrhofer
JKU Linz
Altenbergerstraße 69
4040 Linz, Austria
rene.mayrhofer@jku.at

ABSTRACT

Mobile devices, ubiquitous in modern lifestyle, embody and provide convenient access to our digital lives. Being small and mobile, they are easily lost or stole, therefore require strong authentication to mitigate the risk of unauthorized access. Common knowledge-based mechanism like PIN or pattern, however, fail to scale with the high frequency but short duration of device interactions and ever increasing number of mobile devices carried simultaneously. To overcome these limitations, we present *CORMORANT*, an extensible framework for risk-aware multi-modal biometric authentication across multiple mobile devices that offers increased security and requires less user interaction.

Categories and Subject Descriptors

D.4.6 [**Security and Protection**]: Authentication

Keywords

multi-modal authentication, risk assessment, biometrics

1. INTRODUCTION

Smartphones, tablets, smartwatches and other mobile devices have long become an indispensable part of everyday life, allowing easy access to valuable assets, information and services. Since those small and mobile devices have a high propensity to become lost or stolen, strong user authentication is crucial to protect against the risk of unauthorized access. Therefore, knowledge-based mechanisms like PIN, pattern, and password are commonly applied today. Besides well-studied shortcomings like people being bad at choosing and remembering adequate secrets [29] or vulnerability to shoulder surfing and smudge attacks [2], these authentication techniques require a significant amount of scarce user attention in proportion to the usually short usage sessions [13]. An effect that is even further amplified by the inability of current approaches to scale with the ever growing number of devices used simultaneously. As a result, a survey found that 57,1% of smartphone users do not use authentication, many of which stated that they considered it to be too inconvenient [10]. As a promising approach to overcome the aforementioned drawbacks, continuous unobtrusive user identity verification using different biometrics has been proposed. Commonly used traits include gait, voice, finger pressure, user interface interaction, mouse movement, and keystroke dynamics. While these features can be captured unobtrusively, they are not necessarily available at all times. *Multi-modal* fusion schemes to combine multiple biometrics have hence been developed, usually classified as either feature level fusion, matching score level fusion, and decision level fusion strategies [24].

User authentication on mobile devices is generally applied to defend against the risk of unauthorized access to data and services through an adversary with physical access to the device. This risk, however, is arguably dynamic and highly depends on spatial and temporal context. Considering risk in order to apply as much security as needed but as little as possible potentially facilitates less obtrusive, adequately tailored and thus user-friendly security mechanisms [15].

In this paper, we present the design of *CORMORANT*, an extensible, risk-aware multi-modal authentication framework for continuous user identity verification [12].

2. RELATED WORK

Multi-modal biometric systems, i.e., systems incorporating biometric information from multiple sources, have been well studied for several decades to overcome some drawbacks of unimodal biometrics or to defend against spoofing attacks [23]. Recently, the concept of multi-modal biometrics has been successfully applied to the domain of mobile devices. The authors of [16] utilized face, teeth and voice authentica-

MoMM 2015, December 11 - 13, 2015, Brussels, Belgium

ISBN 978-1-4503-3493-8/15/12...$15.00

DOI: http://dx.doi.org/10.1145/2837126.2843845

tion on mobile devices. Face and voice biometrics are also combined for mobile user identification in [27]. In [5], the authors propose an authentication framework using keystroke dynamics and speaker verification on mobile devices.

Since risk is ultimately the cause for any security measures, it has been considered in different aspects of computer security. In [7], the authors proposed a contextual risk-based access control system based on a mathematical scoring technique assigning numerical weights to risk factors to improve confidentiality, integrity, and availability of the resulting access control model. The risk-based authentication system introduced in [26] defines risk differently from our work as the likelihood of an intruder impersonating a genuine user, which is continuously evaluated based on mouse and keystroke dynamics. In [3], the concept of risk is applied to develop a risk-aware role based access control system that allows to enforce risk-related constrain in scenarios where certain combinations of permission are considered too powerful (or risky) and should thus not be assigned to the same role.

Surprisingly few authors have so far addressed the scalability problem arising from the ever growing number of personal devices used simultaneously. In [25], a hardware token based approach is introduced in order to avoid password challenges for different accounts and devices. The authors of [11] propose a centralized authentication system to omit authentication for online services or notebook access. Work closest to our approach is [14], who's authors envision an authentication aura formed by trusted devices surrounding the user, even taking location as one major determinant of risk into account.

3. THE CORMORANT FRAMEWORK

3.1 Motivation and Goals

Given a choice, users implicitly conduct a cost–risk analysis, for instance when choosing a password or deciding whether to use a lockscreen or not. *Cost* from the user's perspective in this situation could be the added cognitive effort and perceived overhead. By increasing the usability and user-friendliness of authentication, perceived *costs* of device protection can be reduced, allowing to achieve an higher security level overall. To contribute to this goal, we develop *CORMORANT*, an extensible open source framework that combines various arbitrary implicit and explicit authentication techniques. Our approach is novel in three aspects: We aim to utilize a number of arbitrary biometrics along conventional knowledge-based and potentially possession-based authentication mechanisms. Unlike existing work, features of the utilized biometrics like captured traits, availability and or accuracy are not a priori known but highly dynamic as new biometrics can be added to the system at runtime from third party sources. Secondly, we make extensive use of continuous context-based risk evaluation, allowing to both fine-tune access control and thus reduce user interruption as well as increase overall security. The third distinctive feature is the inclusion of information not only from one but possibly all trusted device a user possesses. This could enable, for instance, devices to derive sufficient confidence in the user's identity from other trusted devices nearby to omit explicit authentication. If for instance a user is continuously authenticated through gait recognition applied on his smartwatch [20], the user's smartphone could establish his identity without explicit authentication if both devices trust each other and are able to determine that they are within close proximity (e.g., $\leq 1\,m$).

3.2 Architecture

In the design and development of *CORMORANT* we strive to enable two properties not commonly found in similar research projects: The designed system is supposed to be deployable on stock Android devices, usable for average consumers (e.g., not require a deeper understanding of how to train biometric classifiers) and provide benefits, e.g., in terms of enhanced convenience that outweigh inevitable disadvantages like reduced battery life. These aspirations introduce limitations like missing hardware access through public APIs as required by many indoor location techniques. The other major design paradigm we follow in designing the framework is to facilitate collaboration with and contributions by researchers working on novel authentication or risk evaluation techniques. The proposed framework, which will be released under an open source license once it reaches alpha status, is easily extensible through a convenient *plugin* mechanism as outlined in fig. 1. The system can be extended at runtime by installing additional authentication and risk plugins.

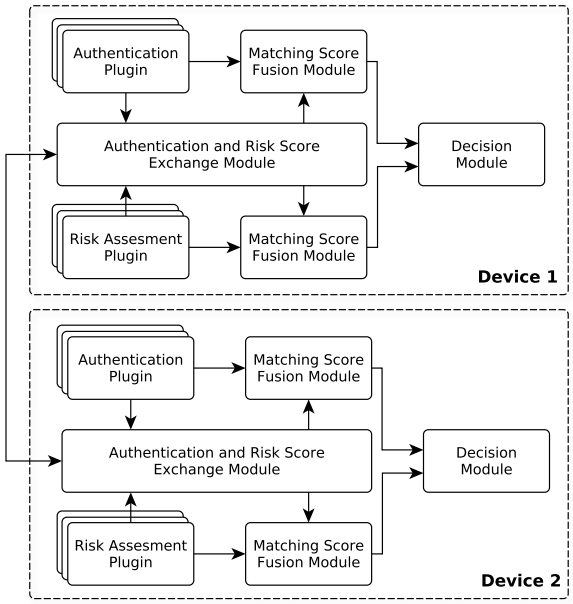

Figure 1: Preliminary architectural overview

While a number of plugins like gait, voice, and face recognition are provided by default, the framework itself is independent and solely relies on the presence of corresponding plugins. It can therefore serve as a platform for researches who, by leveraging the infrastructure provided by *CORMORANT*, can focus efforts on their primary research goals rather than dealing with common tasks like providing fall-back authentication, operating system integration, and collecting usage and performance statistics. Plugins connect to the framework through a lean API based on the OSGi whiteboard pattern [22], posing only minimal implementation requirements. Changes in risk assessment or authentication state are propagated either event-based, periodically pulled or explicitly requested upon necessity, for instance to trigger explicit authentication.

4. RISK PLUGINS

Establishing security always introduces cost of some sort, e.g., in terms of money, performance, or usability. Achieving a higher level of security naturally increases those associated costs: A long password is stronger than a short one, but takes more time to enter. Hence, a trade-off between cost and security exists when evaluating the appropriated level of security, which depends on the *risk* to defend against. Risk is commonly considered to be the product of the probability of an adverse event occurring as well as the resulting impact. These two dimensions, however, are by no means static but highly dependent on the current situation or context. Considering that authentication on mobile devices is applied to defend against the risk of unauthorized access, the following examples of contextual properties influencing this risk can be unobtrusively captured by modern mobile devices:

Probability of Unauthorized Access.

- **Macro location:** The risk of being mugged or robbed varies drastically on country level. It is, according to the United Nations Office on Drugs and Crime, highest in Southern Africa and lowest in South Asia [1].
- **Micro location:** Crime rates also vary on city or even district level. The South Bronx, NY, USA, for instance, is known for having a notably higher crime rate than, e.g., Brooklyn[1].
- **Contextual location:** The probability of unauthorized access also depends on the contextual nature of a location. Smartphone loss, for instance can occur in public transport but not at home.
- **Time of day:** In the US, robberies, for instance, are statistically roughly three times more likely at night (5:00 PM to 4:59 AM) than during daytime [8].

Impact of Unauthorized Access.

- **Sensitivity of data:** The adverse impact of unauthorized access is arguably less if this device contains no or only publicly available data compared to very private or compromising data.
- **Accessible services:** The number (and significance) of accessible services affects the potential harm. An active business VPN, for instance, might endanger a remote company network.
- **Value of transactions:** The associated damage of an unauthorized mobile banking transaction literally depends on the pecuniary value of the transaction.

While these examples arguably have an objective influence on the risk of unauthorized access, one has to keep in mind that the perceived or even actual risk vary across users. While some might consider their home sufficiently safe to omit authentication, others might not appreciate their device being accessed by a wary spouse or children [10].

5. AUTHENTICATION PLUGINS

5.1 Gait Recognition

Gait as a biometric trait offers an unobtrusive way of recognizing individuals from their walking styles. Various studies have explored the feasibility of deploying gait authentication on off-the-shelf mobile devices [6, 18, 20, 21].

[1]http://maps.nyc.gov/crime/

We have developed a gait plug-in that implicitly processes accelerometer data and delivers authentication results to the *CORMORANT* framework. Figure 2 outlines the steps involved in the enrollment and the verification phases, details of which can be found in [20]. To enroll, a gait template is created by walking $\approx$ 300 meters at normal pace, carrying the mobile phone in the trousers' front pocket. Once the gait template is generated, the plugin automatically switches to verification phase, which is similar to the enrollment phase. Capturing acceleration data is fairly power-intensive. We therefore utilize a low-powered ever-on step detector sensor to avoid recording accelerometer data when the user is not actually walking. Once the user starts walking, the step detector triggers the plugin which in turn registers to the accelerometer sensor to start recording acceleration values. Accelerometer data are record for a period of fifteen seconds, after which the application checks if user is still walking by monitoring the timestamps between the steps taken by the user, and if so, continues to record acceleration data. In parallel, previously recorded data are processed and authentication results computed by applying a matching engine that uses Dynamic Time Warping distance to compare live gait cycles with the enrolled template.

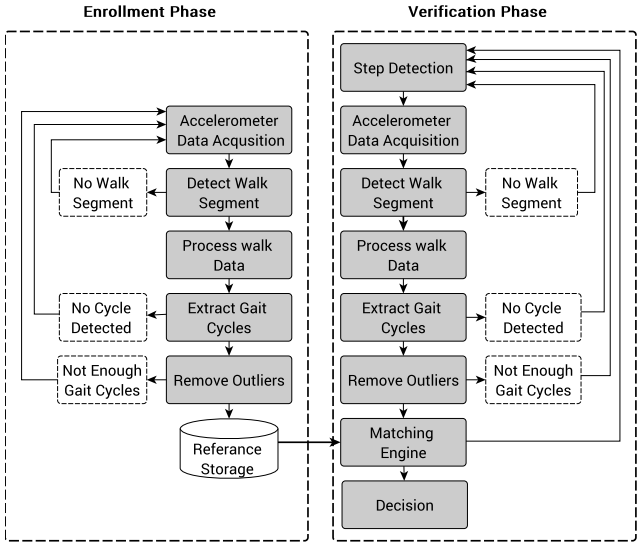

Figure 2: Continuous gait authentication overview

5.2 Speaker Recognition

Speaker recognition is a technique that allows to identify individuals from their voice samples. Speaker recognition systems can be divided in two types: text-dependent and text-independent [4]. In text-dependent speaker recognition systems users use the same utterance for enrollment and verification phase. In text-independent system, however, users are not bound to use the same utterance for enrollment and verification process. We have developed a *text-independent* speaker verification plug-in that processes speaker data recorded by a mobile device's microphone and delivers authentication results to the *CORMORANT* framework. Figure 3 shows various steps of the enrollment and the verification process. In the enrollment step, speakers provide their voice samples. These samples undergo voice activity detection which removes no-voice parts from the

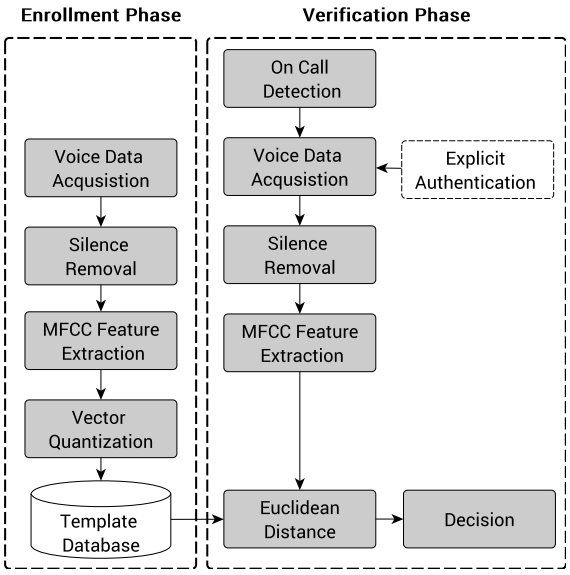

Figure 3: Voice authentication overview

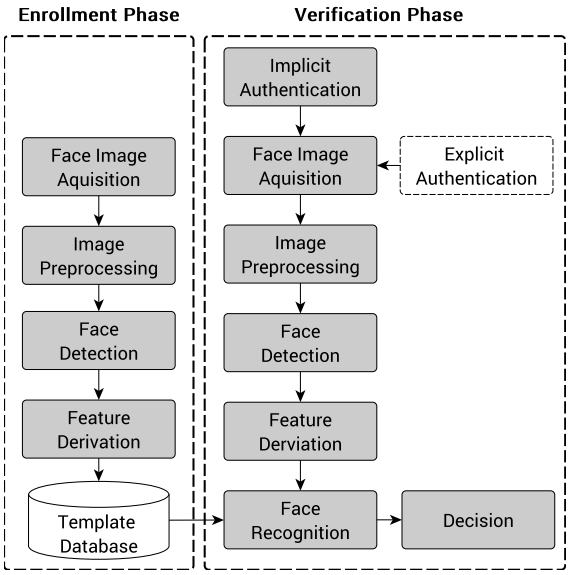

Figure 4: Face authentication overview.

samples. Subsequently, standard Mel Frequency Cepstral Coefficients (MFCC) technique is applied to extract features from the voice samples. Following, Vector Quantization (VQ) is used to generate a speaker-specific codebook by clustering the MFCC features using K-Means clustering algorithm [17]. During verification phase, MFCC features are extracted from the test voice. The minimum average distance of test feature vectors to all code books is computed by eq. (1) where $D_Q$is the average quantization distortion between test feature vector V and speaker model codebook Y. $\mathcal{D}(x, y)$ is the multi-dimension Euclidean distance function.

$$D_Q(V, Y) = \frac{1}{M} \sum_{i=1}^{M} \min_{1 \leq j \leq n} \mathcal{D}(\mathrm{v}_i, \mathrm{y}_j) \quad (1)$$

5.3 Face Recognition

Face authentication deals with verifying or identifying individuals based on their facial features, which are usually derived from face images. Therefore, face authentication systems employ a number of different steps to perform face authentication – including face image recording, image preprocessing, face detection and face recognition. The authentication information is usually derived from face recognition results then. With our face authentication plugin [9] (see fig. 4) we utilize similar mechanisms. At first, the mobile device camera is used to record face images. These images are preprocessed (grayscaling, brightness and contrast adjustments), then Viola and Jones face detection [19, 28] is applied to detect and segment faces present in the images. For facial face feature derivation we currently employ principal component analysis (PCA) and either K-nearest-neighbor (KNN) or support vector machines (SVM) classification as face recognition. The plugin's authentication confidence used in our framework is derived from these classification results then. In the enrollment phase users take multiple pictures of themselves in different illumination conditions, which serve as training data. In the verification phase, face authentication could be used both in an explicit and implicit manner in our framework. With explicit usage, users are requested to perform face authentication, then explicitly take a picture of their face. The authentication confidence of this face is used by the framework then. In contrast, for implicit usage, faces visible to the mobile device camera are monitored continuously while the device is used (e.g. screen is turned on). In case users leave their device unlocked and unattended and a unauthorized person starts interacting with the device, face authentication will not recognize the face as authorized, which will further lead to a drop in authentication confidence.

Face authentication could be used as well in the context of risk assessment: faces of multiple people being present in images used for face authentication could indicate increased risk of shoulder surfing attacks (e.g. input of PINs, passwords and graphical patterns) or overhearing of pass phrases (e.g. non-text-free voice authentication scenarios).

6. CONCLUSION AND FUTURE WORK

In this work we presented the preliminary design of *CORMORANT*, an extensible framework for risk-aware multimodal authentication on mobile devices. By continuously assessing the risk of unauthorized access while evaluating the user's identity using various biometrics it facilitates both convenient and more user-friendly security while it can also be configured to achieve a higher level of overall security. Open research questions to be address next are how to exchange risk and authentication scores across devices, initial device pairing, group key exchange, device exclusion, and spatial distance estimation. We intent to eventually evaluate the usability, performance, and practicality of the *CORMORANT* framework by conducting extensive in situ user study.

7. ACKNOWLEDGMENTS

We gratefully acknowledge funding and support by the German Federal Ministry of Education and Research under grant number 03FH030IX5, Christian Doppler Gesellschaft, A1 Telekom Austria AG, Drei-Banken-EDV GmbH, LG Nexera Business Solutions AG, NXP Semiconductors Austria GmbH, and Österreichische Staatsdruckerei GmbH.

References

[1] International Statistics on Crime and Justice. In S. Harrendorf, M. Heiskanen, and S. Malby, editors, *European Institute for Crime Prevention and Control*. 2010.

[2] A. J. Aviv, K. Gibson, E. Mossop, M. Blaze, and J. M. Smith. Smudge Attacks on Smartphone Touch Screens. *Proceedings of the 4th USENIX conference on Offensive technologies*, pages 1–10, 2010.

[3] K. Z. Bijon, R. Krishnan, and R. Sandhu. A framework for risk-aware role based access control. *2013 IEEE Conference on Communications and Network Security (CNS)*, pages 462–469, 2013.

[4] F. Bimbot, J.-F. Bonastre, C. Fredouille, G. Gravier, I. Magrin-Chagnolleau, S. Meignier, T. Merlin, J. Ortega-García, D. Petrovska-Delacrétaz, and D. A. Reynolds. A tutorial on text-independent speaker verification. *EURASIP J. Appl. Signal Process.*, 2004:430–451, Jan. 2004.

[5] H. Crawford, K. Renaud, and T. Storer. A framework for continuous, transparent mobile device authentication. *Computers and Security*, 39:127–136, 2013.

[6] M. O. Derawi. *Smartphones and Biometrics: Gait and Activity Recognition*. PhD thesis, Gjøvik University College, November 2012.

[7] N. N. Diep, S. Lee, Y.-K. Lee, and H. Lee. Contextual Risk-Based Access Control. *Security and Management*, 2007.

[8] M. Felson and E. Poulsen. Simple indicators of crime by time of day. *International Journal of Forecasting*, 19:595–601, 2003.

[9] R. D. Findling. Pan shot face unlock: Towards unlocking personal mobile devices using stereo vision and biometric face information from multiple perspectives. Master's thesis, University of Applied Sciences Upper Austria, Hagenberg, Austria, Sept. 2013.

[10] M. Harbach, E. V. Zezschwitz, A. Fichtner, A. De Luca, and M. Smith. It's a Hard Lock Life: A Field Study of Smartphone (Un) Locking Behavior and Risk Perception. *Symposium on Usable Privacy and Security (SOUPS)*, pages 213–230, 2014.

[11] E. Hayashi and J. I. Hong. Knock x Knock : The Design and Evaluation of a Unified Authentication Management System. 2015.

[12] D. Hintze, R. D. Findling, M. Muaaz, E. Koch, and R. Mayrhofer. CORMORANT: Towards Continuous Risk-Aware Multi-Modal Cross-Device Authentication. *Proc. UbiComp 2015: Adjunct Publication*, 2015.

[13] D. Hintze, R. D. Findling, S. Scholz, and R. Mayrhofer. Mobile Device Usage Characteristics: The Effect of Context and Form Factor on Locked and Unlocked Usage. In *Proc. MoMM 2014*, 2014.

[14] C. G. Hocking, S. M. Furnell, N. L. Clarke, and P. L. Reynolds. Authentication Aura - A distributed approach to user authentication. *Journal of Information Assurance and Security*, 6(2):149—-156, 2011.

[15] A. Hurkala and J. Hurkala. Architecture of Context-Risk-Aware Authentication System for Web Environments. *ICIEIS'2014*, pages 219–228, 2014.

[16] D. J. Kim, K. W. Chung, and K. S. Hong. Person Authentication using Face, Teeth and Voice Modalities for Mobile Device Security. *IEEE Transactions on Consumer Electronics*, 56(4):2678–2685, 2010.

[17] T. Kinnunen and H. Li. An overview of text-independent speaker recognition: from features to supervectors. *Speech Communication*, 52(1), 2010.

[18] J. R. Kwapisz, G. M. Weiss, and S. A. Moore. Cell phone-based biometric identification. In *Biometrics: Theory Applications and Systems (BTAS), 2010 Fourth IEEE International Conference on*, pages 1–7. IEEE, 2010.

[19] R. Lienhart and J. Maydt. An extended set of haar-like features for rapid object detection. In *IEEE International Conference on Image Processing 2002*, pages 900–903, 2002.

[20] M. Muaaz and R. Mayrhofer. Orientation Independent Cell Phone Based Gait Authentication. *Proc. MoMM 2014*, pages 161–164, 2014.

[21] C. Nickel. *Accelerometer-based Biometric Gait Recognition for Authentication on Smartphones*. PhD thesis, TU Darmstadt, June 2012.

[22] OSGi Alliance. Listeners Considered Harmful: The "Whiteboard" Pattern. 2004.

[23] A. Ross and A. K. Jain. Multimodal Biometrics: an Overview. *Signal Processing*, (September):1221–1224, 2004.

[24] P. S. Sanjekar and J. B. Patil. An Overview of Multimodal Biometrics. *Signal & Image Processing (SIPIJ)*, 4(1):57–64, 2013.

[25] F. Stajano. Pico: No more passwords! *Lecture Notes in Computer Science*, 7114 LNCS:49–81, 2011.

[26] I. Traore, I. Woungang, M. S. Obaidat, Y. Nakkabi, and I. Lai. Combining Mouse and Keystroke Dynamics Biometrics for Risk-Based Authentication in Web Environments. *2012 Fourth International Conference on Digital Home*, pages 138–145, 2012.

[27] P. Tresadern, T. F. Cootes, N. Poh, P. Matejka, A. Hadid, C. Lévy, C. McCool, and S. Marcel. Mobile Biometrics: Combined Face and Voice Verification for a Mobile Platform. *IEEE Pervasive Computing*, 12(01):79–87, 2013.

[28] P. Viola and M. Jones. Rapid object detection using a boosted cascade of simple features. *Proceedings of these 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 1:511–518, 2001.

[29] J. Yan, A. Blackwells, R. Anderson, and A. Grant. Password Memorability and Security: Empirical Results. *IEEE Security & Privacy*, 2(5):25–31, 2004.