Delusio – Plausible Deniability For Face Recognition

FELIX DIETZ, University of the Bundeswehr Munich, Germany
LUKAS MECKE, University of the Bundeswehr Munich, Germany
DANIEL RIESNER, LMU Munich, Germany
FLORIAN ALT, University of the Bundeswehr Munich, Germany

Fig. 1. Delusio enables users to unlock their phone into two different modes based on their facial expressions during the authentication process. This enables plausible deniability when they do not have control over their own device and are forced to unlock it for example at a border checkpoint.

We developed an Android phone unlock mechanism utilizing facial recognition and specific mimics to access a specially secured portion of the device, designed for plausible deniability. The widespread adoption of biometric authentication methods, such as fingerprint and facial recognition, has revolutionized mobile device security, offering enhanced protection against shoulder-surfing attacks and improving user convenience compared to traditional passwords. However, a downside is the potential for third-party coercion to unlock the device. While text-based authentication allows users to reveal a hidden system by entering a special password, this is challenging with face authentication. We evaluated our approach in a role-playing user study involving 50 participants, with one participant acting as the attacker and the other as the suspect. Suspects successfully accessed the secured area, mostly without detection. They further expressed interest in this feature on their personal phones. We also discuss open challenges and opportunities in implementing such authentication mechanisms.

Authors' Contact Information: Felix Dietz, felix.dietz@unibw.de, University of the Bundeswehr Munich, Munich, Germany; Lukas Mecke, lukas.mecke@unibw.de, University of the Bundeswehr Munich, Munich, Germany; Daniel Riesner, daniel.riesner@campus.lmu.de, LMU Munich, Munich, Germany; Florian Alt, florian.alt@unibw.de, University of the Bundeswehr Munich, Munich, Germany.

ACM 2573-0142/2024/9-ART249
https://doi.org/10.1145/3676494

CCS Concepts: • **Security and privacy** → **Usability in security and privacy**; **Privacy protections**; **Biometrics**.

Additional Key Words and Phrases: facial authentication; biometrics; plausible deniability

ACM Reference Format:
Felix Dietz, Lukas Mecke, Daniel Riesner, and Florian Alt. 2024. Delusio – Plausible Deniability For Face Recognition. *Proc. ACM Hum.-Comput. Interact.* 8, MHCI, Article 249 (September 2024), 13 pages. https://doi.org/10.1145/3676494

1 INTRODUCTION

Authentication has become omnipresent for users to protect their personal devices and data. While passwords are still widespread to achieve this task, biometric authentication has seen widespread adoption over the past years, in particular on mobile devices where they are commonly available in the form of fingerprint [18] or face recognition [30]. Compared to other authentication approaches, biometrics can be fast and seamless and cannot be stolen, lost, or forgotten [17].

However, biometrics can also lead to a loss of user control over authentication as biometric features are visible and permanent, contrary to, e.g., a secret like a password. As a consequence, there are reports of devices being unlocked by third parties while their owner was asleep[1]. When awake, users can be coerced (possibly with physical force) to unlock their devices leading to user concerns for their physical safety [24] when using biometrics.

In such a scenario, users forfeit their capacity to convincingly disclaim ownership or awareness of the content stored on their devices since they were accessed via their biometric features. This stands in contrast to alternative authentication methods where users can input an incorrect password to reveal a concealed aspect of a computer system that could be meticulously arranged to include sufficient personal information or work-related documents; giving the appearance of authenticity to potential attackers. Ideally, it should not be possible for an attacker to detect whether a device contains hidden systems - even when gaining physical access to the unlocked storage medium [2].

In the past century, this so-called plausible deniability has especially been used in the Cold War to enable covert actions while disclaiming state sponsorship. Methods have varied widely, from money laundry over secret communication to training foreign military to fight in proxy wars [22]. While back then, it might have been enough to burn a written note, e.g., before crossing a checkpoint, today, we need solutions suitable for our digitized world. With the proliferation of personal computing, especially smartphones, privacy and security risks grow accordingly.

A prominent example where plausible deniability can be highly important is domestic violence. In such cases, victims often seek help outside – either through special apps [27] or by messaging friends, social workers, or therapists. However, abusing partners could forcefully gain access to their victim's devices to check if they talked to outsiders about the ongoing abuse. If discovered, this could internally spiral violence in the relationship and even deter victims from seeking help.

Another instance occurs at airport checkpoints during travel between nations embroiled in conflict. Certain frequent travelers, like journalists, may possess dual passports to evade interrogation by border officials regarding their visits to adversary countries. In such scenarios, a smartphone feature that conceals specific information could prove beneficial.

Contribution Statement. In our work, we present a prototype that offers both the convenience and benefits of facial authentication while also providing a method for plausibly concealing sensitive information by using a pre-defined facial expression during the authentication process. We validate our prototype regarding reliability and conspicuousness in a user study with 50 participants. We publish our Android prototype to enable further validation and development.

[1] https://www.engadget.com/2019-10-21-google-pixel-4-face-unlock-eyes-open.html, last accessed August 13, 2024

2 RELATED WORK

We present selected prior work covering technical aspects of plausible deniability in the information age as well as research on authentication mechanisms and users' concerns and usability aspects.

2.1 Plausible Deniable Encryption (PDE)

Because of the sensitive nature of certain data on personal computers and mobile devices, all leading operating system developers have integrated storage encryption to varying extents. Vendors typically implement one of two methods to protect their users' data. One is file-based encryption, where the file system is not encrypted but each individual file. Here, an attacker can see all existing files on a system but cannot read their content. On the other hand, full-disk encryption operates at a deeper level by encrypting the file system itself. This prevents attackers from discerning the number of files present or the amount of space occupied by data on the medium. Full-disk encryption accomplishes this by initially overwriting the entire disk with random data and subsequently establishing an encrypted volume on top of it. As random data cannot be distinguished from the encrypted volume data due to their similar entropy levels, its size becomes undetectable. Therefore, plausible deniable encryption is often based on full-disk encryption by placing multiple volumes on top of the random data. Based on the entered passkey, the normal or the hidden volume will be mounted and accessible.

On desktop computers, Truecrypt and its' successor Veracrypt are well-known tools that can create hidden volumes inside another encrypted volume, enabling the user to mount either one by simply entering one of two different passwords[2]. Anzuoni and Gagliardoni [2] call their project *Shufflecake* a "spiritual successor" to Veracrypt as it integrates natively as an in-kernel tool into Linux while supporting volumes in multiple hidden layers - contrary to the one layer of secrecy possible with Veracrypt.

While yet not implemented by major mobile operating system providers like Apple and Google, researchers have been exploring and developing methods towards plausible deniable storage for mobile devices [9–11, 16, 25, 33]. Chen et al. [12] developed *MobiWear*, a system that allows plausible deniability encryption on wearable devices. It uses image steganography and utilizes a smartwatch's sensors to input a password instead of a keyboard or touchscreen.

2.2 User Concerns and Usability Aspects of Plausible Deniability

Although possessing a mobile device equipped with plausible deniability encryption is advantageous, its usability is equally crucial, especially in high-pressure scenarios like when a border officer questions a journalist. Having this usability concern in mind, Chang et al. [8] improved their system *MobiPluto* by enabling the user to switch from the public mode into the hidden mode within 10 seconds by entering a password or tapping an NFC card at the device. Gründling [15] created an Android application designed to conceal sensitive apps within Android's work profile. This is achieved by entering an incorrect PIN, causing the designated apps to vanish. Similarly, the framework *MobiMimosa* empowers users to establish security protocols capable of removing sensitive apps and erasing sensitive data [16].

While most research focuses on device-related aspects, the *Wink* project has developed a method enabling plausibly deniable communication via text messaging. *Wink* accomplishes this by discreetly embedding hidden messages within the standard random coins (such as salts or initialization vectors) utilized by widespread end-to-end encrypted (E2EE) protocols. This enables covert communication within popular E2EE messaging applications without requiring extensive supplementary tools with minimal effect on system performance and without altering the message formats [7].

[2] https://www.truecrypt71a.com/documentation/plausible-deniability/hidden-volume/, last accessed August 13, 2024

Although technically achievable, as demonstrated by the *Wink* project, the use of plausible deniable messaging raises questions about its desirability and social acceptability in everyday communication. Allowing messages to maintain plausible deniability, particularly those with legal implications, could impede efforts to uncover the truth in a world plagued by misinformation. Yadav et al. [31] explored this issue of acceptability through a survey involving 664 participants and discovered that plausible deniability lacks social acceptance, with most users expressing a preference for non-repudiation.

2.3 Authentication Mechanisms on Mobile Devices

Smartphone authentication methods have gradually evolved over the past few decades. Initially, phones lacked authentication, then progressed to PIN and password-based systems. The advent of touchscreens introduced a new method where users had to swipe along predefined patterns. While this enhanced convenience compared to lengthy passwords, it also made devices more vulnerable to shoulder surfing [3] and smudge attacks [4]. While researchers developed approaches to graphical password input methods that are more resilient against shoulder surfing attacks, such strategies often increase the user's mental load, thus increasing authentication time and error rates [19, 20, 28].

Almost all modern smartphones now incorporate biometric authentication options, such as fingerprint recognition, facial recognition, or a combination. These mechanisms significantly enhance convenience and reduce authentication duration compared to traditional password-based methods. For instance, a user's fingerprint can be processed as soon as they retrieve the phone from their pocket, unlocking it instantly upon the user's screen glance. Over the past few years, researchers have developed biometric authentication mechanisms that operate without requiring direct user input. Instead, these systems employ background authentication algorithms that analyze subtle user micro-movements and inputs to determine whether the current user is legitimate. These mechanisms can be incorporated at various stages, ranging from the virtual keyboard input on the device's touchscreen, where factors such as the location, duration, and intensity of each button press [1] are considered, to analyzing the user's scrolling behavior or how they hold the phone [5, 14]. Such behavior-based authentication methods can also serve for continuous authentication where the system consistently monitors user input using accessible sensors to establish an authentication score. If this score declines below a certain threshold, indicating uncertainty about the legitimacy of the current user, the system prompts the user to input a password or utilizes facial- or fingerprint-based authentication to confirm their identity [23, 26].

Despite advancements in authentication mechanisms, some users still share their credentials with partners, close friends, or family members or inadvertently expose them through proximity, such as during shoulder surfing incidents. Marques et al. [21] examined unauthorized smartphone access through an online survey and discovered that intimate partners accounted for over 60 % of unauthorized users. Additionally, they found that in nearly 70 % of cases, the primary motivation was to exert control over relationships with others.

2.4 Summary

Prior research has delved into plausible deniability encryption for desktop computers and integrated methods for mobile devices like smartphones. User concerns and enhancements in usability have been tackled, as exemplified by *MobiPluto* [9] and *MobiMimosa* [16]. Despite these advancements, there remains an unaddressed gap in implementing plausible deniability for biometric authentication mechanisms such as fingerprint or facial recognition. This highlights a growing research and usability void, particularly as more users transition from PIN-based to biometric authentication.

3 RESEARCH APPROACH

Here, we discuss considerations when designing for plausible deniability using biometrics and introduce the concept for the solution we propose in this paper: *Delusio*. We then provide details on the app's implementation and outline the research questions we aim to answer with this work.

3.1 Considerations when Designing for Plausible Deniability using Biometrics

As outlined above, realizing plausible deniability when using biometrics is challenging. As users *are* the "key" to unlock their devices, they cannot produce a different secret or hand over a different actual key as would be possible for secret- or token-based authentication. This only leaves users with the option to change their appearance (or behavior when interacting with behavioral biometric authentication) or use a side channel captured by the biometric sensor (e.g., finger positioning on a fingerprint sensor). While using other side channels, like pressing a button on the phone during authentication, would theoretically be an option, such approaches might fall short in practice when users do not have control over their devices. Some options for such changes in appearance include translation (e.g., placing the finger to the side of the sensor), rotation (e.g., turning the head), or distortion (e.g., raising an eyebrow). Such triggers can further be augmented with temporal patterns (e.g., tapping), though this may make them harder to remember.

When choosing a trigger, we see three main requirements that should be considered. The interaction needs to be both *easy to perform* and *memorable* so that users can actually use them when under pressure. In addition, the trigger has to be *inconspicuous* so that users can not only perform it but can do so without a third party noticing.

3.2 Concept: *Delusio*

Based on the established design opportunities and requirements, we conceptualized an app called *Delusio* that uses squinting (or blinking) of the eyes as a trigger to enter a hidden mode during authentication using face recognition. We decided on this approach, as users can robustly detect and easily execute squinting. In particular, squinting is sometimes used as a defensive reflex (e.g., to protect the eyes from bright light or when in pain), making it a plausible and natural reaction to situations where a user would be forced to unlock their device. As a side effect, the hidden mode should always be triggered when someone tries to unlock the device in the user's sleep. We decided against any pattern (e.g., blinking multiple times) as it would make the gesture harder to get right, remember, and conceal.

3.3 Implementation of *Delusio*

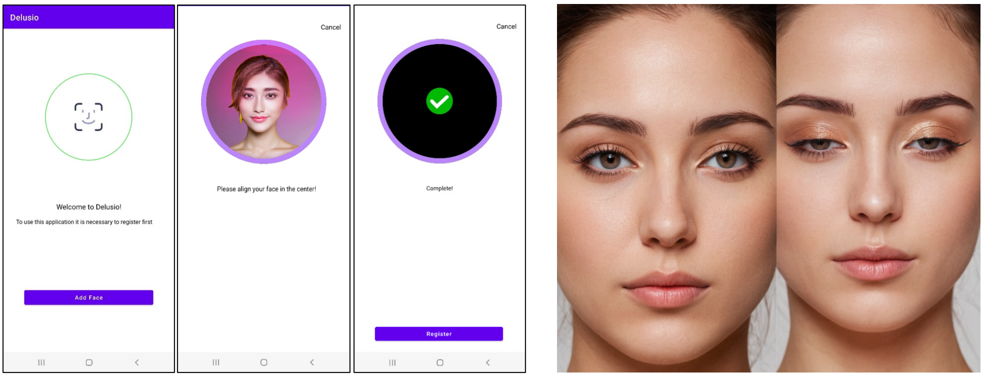
We implemented *Delusio* as an app for Android version 21 and above that replaces the lock screen and triggers a photo gallery with a hidden mode. As such, our implementation is but a prototype and can, in its current form, not be used outside a study setting. We used OpenCV[3] to capture face images and leveraged its implementation of Eigenfaces [29] for face recognition. To detect the squinting gesture we used the Google ML Kit[4] face detection feature, which provides both a map of facial landmarks and a probability for each of the eyes to be open.

When interacting with the app, users had to first capture an image of their face as a base for the Eigenface recognition. Afterward, they could trigger authentication and see the public gallery if their eyes were open or gain access to the hidden mode when squinting. This process is illustrated in Figure 2. We publish our prototype for further evaluation and development[5].

[3] OpenCV: https://opencv.org/, last accessed August 13, 2024

[4] Google ML Kit: https://developers.google.com/ml-kit, last accessed August 13, 2024

[5] https://github.com/felix298/delusio-plausibleDeniabilityFaceRecognition



Welcome to Delusio!
To use this application it is necessary to register first

Please align your face in the center!

Complete!

(a) Interaction flow when registering for *Delusio* . (b) Facial expressions to access the hidden mode.

Fig. 2. The Figure demonstrated the interaction with the *Delusio* app. When first opening the app, users needed to (a) register by capturing their face once. Afterward, they could (b) use one of two different facial expressions in order to unlock the device into the two separate modes. By slightly pinching their eyes during the authentication process, users could gain access to a hidden section on their mobile device. Faces have been AI-generated and do not exist.

3.4 Research Questions

With our work, we aim to investigate the general viability of our solution to enable plausible deniability for face recognition. As such, the following research questions guide our work:

RQ1 **Reliability**: Can a mimic-based authentication mechanism be reliably used to unlock a hidden area of a device?

RQ2 **Conspicuousness**: Can users successfully access the hidden area on a device without an observing attacker noticing?

4 USER STUDY

To answer our research questions, we designed a user study where participants would test our system and try to hide content on a phone from an observing attacker. Here, we introduce our study design and procedure as well as the design of our study application and recruitment strategy.

4.1 Study Design

Our study follows a mixed-methods design involving pairs of participants with two independent variables. Participants filled the roles of either *attacker* or *suspect* that would try to hide sensitive information using our system. The study was conducted in two phases with attackers being *unprimed* in the first phase and being *primed* in the second phase to watch out for suspicious behavior from the suspect particularly.

4.2 Procedure

For our study, we invited pairs of participants to our lab, where they would decide how to fill the two roles of attacker and suspect. They then received printed instructions for the study. To make the task more graspable for participants, those instructions contained one of three concrete scenarios revolving around situations where the suspect needed to hide sensitive information (e.g., about their sexual or political orientation) to avoid negative consequences. In all scenarios, the attacker's task was to search for such sensitive information while the users were instructed to use *Delusio* to hide said information.

After reading the instructions, the attacking participant was asked to leave the room shortly, while we introduced the suspect to how the app worked and how to use it. The suspects then had up to five minutes to familiarize themselves with the app and train to enter the hidden mode inconspicuously. In the meantime, we also gave the attacker an introduction about their role and pointed towards the option to either ask the suspect to unlock their phone or hold it up to their face themselves.

If no questions remained, the scenario began with the first phase wherein the attacker tried to find sensitive information on the prepared phone that was unlocked by the suspect. If the attacker was not successful, the study proceeded to phase two, where we gave the attacker the hint to particularly focus on any conspicuous behavior when the suspect unlocked the app before they repeated the scenario. Some of our attackers asked the defender multiple times to unlock their phone in each phase. The defender had only one attempt per unlock request to unlock their phone into the correct mode - just like in a 'real' situation. After both phases of the experiment were complete (between 2 and 5 minutes), we concluded the study with a questionnaire.

4.3 Collected Data

We observed participants' behavior throughout the whole study but used the final questionnaire as our main source of data for this study. The questionnaire was divided into three sections:

1. **Demographics**: In the demographic section, we assessed users' age, gender, and origin. For the latter, Sara Clayton's [13] comments were used as a guide so as not to exclude any ethnic group. We included those questions as previous work suggests face recognition system performance to be sensitive to different user groups. As an example, worse face recognition performance was reported for not-white persons [6, 32].
2. **Security and Privacy Preferences**: In this part, we assessed participants' use of unlocking mechanisms and their need and preferences for hiding certain types on their devices from others.
3. **Role-specific Questions**: In this section, we assessed the scenario perception for the participants and thus divided questions based on the roles. In particular, questions focused on participants' success in using *Delusio* , their trust in the system, and the effectiveness of the app in hiding data from the attacker.

4.4 Participants

For our study, we recruited 25 pairs of participants (50 total, 21 female, 29 male). We kept the title of the study ("Face recognition and security") vague so as not to prime participants to the topic of plausible deniability. Participants came from a wide spectrum of occupations and levels of education with a median age of 26 years old (min: 18, max: 60). The majority of our participants came from Germany (43) and described their ethnicity as white, European, or a combination of both. Participation took half an hour at most and was compensated with an online gift voucher worth € 5. Our study was IRB-approved under grant no. *EK-MIS-2024-247.*

5 RESULTS

Here we report on the results of our user study. As users filled different roles and the second phase was a direct extension of the first, direct comparisons are of limited value and we thus refrain from reporting statistical test results.

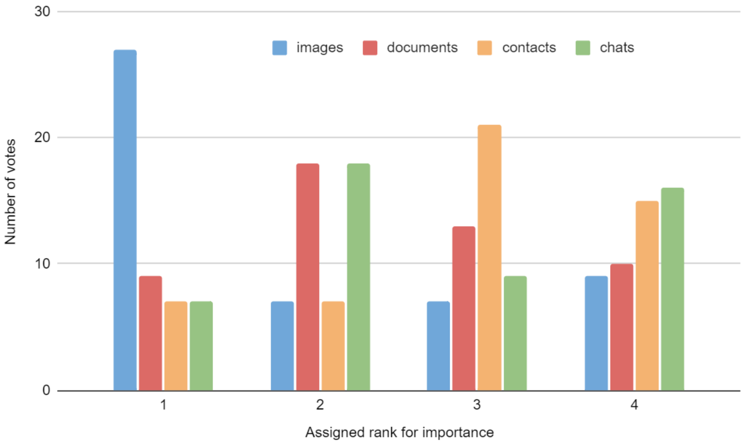


Fig. 3. Ranked results of content types participants wished to being able to hide.

5.1 Security and Privacy Preferences

Participants were equally split between using Android and iOS and 60% of them reported to have previously unlocked their phone with face recognition. A total of 38% of the participants indicated always using face recognition for authentication. Participants completely agreed that the security of their personal data is important to them as well as being interested in using an additional feature to better protect it (Mdn=5 on a 5-point Likert scale for each).

Regarding the question, of whether they had any content on their devices they would want to hide from their friends or family, about half of the participants (24) agreed. Concerning the importance of different content types, participants rated images as the most important (average rank 1.96), followed by documents (2.48), chats (2.68), and contacts (2.88). This is also in line with the most mentioned type of content participants wished to hide being images (see Figure 3).

5.2 Success and Trust when using *Delusio*

All participants in the role of a suspect in our study were able to successfully access the hidden mode on the provided phone by adapting their facial expressions. During the scenario all but one of the suspects were able to successfully hide the sensitive images on the provided phone using *Delusio* . The one participant who failed in doing so reported to have been confused if they had to blink or not to get to the hidden mode. Similarly, all but one suspect reported that they thought the attacker had believed them with one participant feeling as if they had been caught blinking during the unlock process.

Participants strongly agreed (Mdn=5) to feel more secure by using our app. They were overall neutral (Mdn=3) towards being afraid to be caught when accessing the hidden mode and 88% indicated they would use a feature like the one suggested in our approach if it were available for their phones.

5.3 *Delusio* from an Attacker's Perspective

As illustrated above, the users of *Delusio* were successful in accessing the hidden area with the blinking gesture, felt more secure, and believed they were able to deceive the attacker. Here we look into the attackers' perspective to see if this was true.

After the study, 35% of the attackers reported that they felt the suspect was hiding something from them. They explained this by the suspects acting nervous or strange and some also commented, that something with the images seemed to be off. However, only 16% of the attackers found any of the hidden pictures, and 68% believed that the suspect had told them the truth and found them somewhat trustworthy (Mdn=4).

When asked if they had seen through the mechanism for hiding content 16% of the attackers said they did not. The vast majority of 80% indicated having understood the mechanisms only after getting the hint from the experimenter. A single participant saw through the mechanism and also recognized the blinking gesture as the trigger. However, they also reported being familiar with similar approaches for hiding content where a second password was used.

6 DISCUSSION

In our investigation, we assessed the effectiveness of our application *Delusio* in terms of reliability and inconspicuousness through a user study. We observed that nearly all participants could successfully conceal sensitive information without being detected by potential attackers. While this outcome is promising for our prototype and study, there is uncertainty regarding how well users could activate the hidden mode in a more authentic setting, such as during a high-stress encounter at a border checkpoint. Additionally, it's worth noting that in our study, suspects had the opportunity to practice the unlocking process just before the actual test. However, in a real-world scenario, the critical situation might arise unexpectedly, potentially with several months passing since the last practice unlock.

Another aspect of uncertainty arises from the strategy of security by obscurity. In this method, rather than cryptographically securing a system, the suspect relies on obfuscation to conceal the system's access point. Here, the defense strategy hinges on maintaining the secrecy of the hidden entry point. In our user study, security by obscurity certainly contributed to the suspects' success. However, it's likely that border patrol officers, who routinely question individuals, would become aware of the *Delusio* approach if its usage were to become more widespread.

In our prototype, we opted for having the public mode activated by default rather than the hidden one. While it could be argued that the opposite approach would offer greater security, it would also entail more effort in daily use, potentially leading users to disable the feature. In a real deployment, such decisions should be made based on the individual risk and acceptable inconvenience levels.

Roughly half of our participants expressed a desire to be able to conceal specific information, particularly images, when sharing their phones with others. This demand for a feature to hide sensitive content extends beyond highly exposed individuals like investigative journalists, as handing over personal phones is a common practice in everyday interactions, such as exchanging contact information or sharing pictures.

7 FUTURE WORK

Although our prototype and user study represents an initial stride towards achieving usable plausible deniability for a biometric authentication method, further investigation is needed to determine the reliability of users' ability to switch between modes in various situations and activities. These may involve different postures, viewing angles, and light conditions, which could influence the reliability of the authentication process.

Another area of research involves assessing users' proficiency in concealing triggers to access the hidden system, particularly once the interaction method is known to adversaries. An approach to enhance resilience could entail combining multiple triggers, each chosen individually by users. This would significantly increase the difficulty for attackers, as they would need to monitor each potential trigger throughout the authentication process.

8 CONCLUSION

In this paper, we presented *Delusio* , an Android application designed to provide plausible deniability for face recognition by leveraging facial expressions during the authentication process. Our prototype underwent evaluation in a role-playing user study, where nearly all participants assuming the role of the suspect were able to effectively conceal sensitive content on their phones. Likewise, the majority of participants acting as attackers did not detect any suspicious behavior on the suspect's end. We discuss our study's outcomes and limitations, and provide insights into further interaction techniques for plausible deniability that can be explored. We hope that our work inspires designers and engineers to integrate plausible deniability features into biometric authentication mechanisms and thus improve the security of exposed users such as journalists.

ACKNOWLEDGMENTS

The presented work received funding from the German Research Foundation (DFG) under project no. 425869382 and Technology Research Center of the Bundeswehr (Voice of Wisdom). dtec.bw is funded by the European Union - NextGenerationEU.

REFERENCES

[1] Abdulaziz Alzubaidi and Jugal Kalita. 2016. Authentication of smartphone users using behavioral biometrics. *IEEE Communications Surveys & Tutorials* 18, 3 (2016), 1998–2026.

[2] Elia Anzuoni and Tommaso Gagliardoni. 2023. Shufflecake: Plausible Deniability for Multiple Hidden Filesystems on Linux. In *Proceedings of the 2023 ACM SIGSAC Conference on Computer and Communications Security*. 3033–3047.

[3] Adam J Aviv, John T Davin, Flynn Wolf, and Ravi Kuber. 2017. Towards baselines for shoulder surfing on mobile authentication. In *Proceedings of the 33rd Annual Computer Security Applications Conference*. 486–498.

[4] Adam J Aviv, Katherine Gibson, Evan Mossop, Matt Blaze, and Jonathan M Smith. 2010. Smudge attacks on smartphone touch screens. In *4th USENIX Workshop on Offensive Technologies (WOOT 10)*.

[5] Attaullah Buriro, Bruno Crispo, Filippo Delfrari, and Konrad Wrona. 2016. Hold and sign: A novel behavioral biometrics for smartphone user authentication. In *2016 IEEE security and privacy workshops (SPW)*. IEEE, 276–285.

[6] Ángel Alexander Cabrera, Will Epperson, Fred Hohman, Minsuk Kahng, Jamie Morgenstern, and Duen Horng Chau. 2019. FairVis: Visual analytics for discovering intersectional bias in machine learning. In *2019 IEEE Conference on Visual Analytics Science and Technology (VAST)*. IEEE, 46–56.

[7] Anrin Chakraborti, Darius Suciu, and Radu Sion. 2023. Wink: deniable secure messaging. In *32nd USENIX Security Symposium (USENIX Security 23)*. 1271–1288.

[8] Bing Chang, Yao Cheng, Bo Chen, Fengwei Zhang, Wen-Tao Zhu, Yingjiu Li, and Zhan Wang. 2018. User-friendly deniable storage for mobile devices. *computers & security* 72 (2018), 163–174.

[9] Bing Chang, Zhan Wang, Bo Chen, and Fengwei Zhang. 2015. Mobipluto: File system friendly deniable storage for mobile devices. In *Proceedings of the 31st annual computer security applications conference*. 381–390.

[10] Bing Chang, Fengwei Zhang, Bo Chen, Yingjiu Li, Wen-Tao Zhu, Yangguang Tian, Zhan Wang, and Albert Ching. 2018. Mobiceal: Towards secure and practical plausibly deniable encryption on mobile devices. In *2018 48th Annual IEEE/IFIP International Conference on Dependable Systems and Networks (DSN)*. IEEE, 454–465.

[11] Bo Chen. 2020. Towards Designing A Secure Plausibly Deniable System for Mobile Devices against Multi-snapshot Adversaries–A Preliminary Design. *arXiv preprint arXiv:2002.02379* (2020).

[12] Niusen Chen, Bo Chen, and Weisong Shi. 2021. MobiWear: a plausibly deniable encryption system for wearable mobile devices. In *EAI International Conference on Applied Cryptography in Computer and Communications*. Springer, 138–154.

[13] Sara Clayton. 2020. Rethinking the design of the race and ethnicity question on surveys. Retrieved January 13, 2023 from https://uxdesign.cc/rethinking-the-design-of-the-race-and-ethnicity-question-on-surveys-6f9066c69392

[14] Mario Frank, Ralf Biedert, Eugene Ma, Ivan Martinovic, and Dawn Song. 2012. Touchalytics: On the applicability of touchscreen input as a behavioral biometric for continuous authentication. *IEEE transactions on information forensics and security* 8, 1 (2012), 136–148.

[15] Bernhard Gründling. 2020. *App-based (Im) plausible Deniability for Android.* Ph. D. Dissertation. Master Thesis. Johannes Kepler University Linz.

[16] Shuangxi Hong, Chuanchang Liu, Bingfei Ren, Yuze Huang, and Junliang Chen. 2017. Personal privacy protection framework based on hidden technology for smartphones. *IEEE Access* 5 (2017), 6515–6526.

[17] Anil Jain, Ruud Bolle, and Sharath Pankanti. 1996. *Introduction to biometrics.* Springer.

[18] A.K. Jain, S. Prabhakar, and Lin Hong. 1999. A multichannel approach to fingerprint classification. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 21, 4 (1999), 348–359. https://doi.org/10.1109/34.761265

[19] Teoh joo Fong, Azween Abdullah, NZ Jhanjhi, and Mahadevan Supramaniam. 2019. The coin passcode: A shoulder-surfing proof graphical password authentication model for mobile devices. *International Journal of Advanced Computer Science and Applications* 10, 1 (2019).

[20] Sung-Hwan Kim, Jong-Woo Kim, Seon-Yeong Kim, and Hwan-Gue Cho. 2011. A new shoulder-surfing resistant password for mobile environments. In *Proceedings of the 5th International Conference on Ubiquitous Information Management and Communication*. 1–8.

[21] Diogo Marques, Tiago Guerreiro, Luis Carriço, Ivan Beschastnikh, and Konstantin Beznosov. 2019. Vulnerability & blame: Making sense of unauthorized access to smartphones. In *Proceedings of the 2019 chi conference on human factors in computing systems*. 1–13.

[22] Michael Poznansky. 2022. Revisiting plausible deniability. *Journal of Strategic Studies* 45, 4 (2022), 511–533.

[23] Praveen Kumar Rayani and Suvamoy Changder. 2023. Continuous user authentication on smartphone via behavioral biometrics: a survey. *Multimedia Tools and Applications* 82, 2 (2023), 1633–1667.

[24] Chris Riley, Kathy Buckner, Graham Johnson, and David Benyon. 2009. Culture & biometrics: regional differences in the perception of biometric authentication technologies. *AI & society* 24 (2009), 295–306.

[25] Adam Skillen and Mohammad Mannan. 2013. Mobiflage: Deniable storage encryptionfor mobile devices. *IEEE Transactions on Dependable and Secure Computing* 11, 3 (2013), 224–237.

[26] Ioannis C Stylios, Olga Thanou, Iosif Androulidakis, and Elena Zaitseva. 2016. A review of continuous authentication using behavioral biometrics. In *Proceedings of the SouthEast European Design Automation, Computer Engineering, Computer Networks and Social Media Conference*. 72–79.

[27] Mehreen Sumra, Sohail Asghar, Khalid S Khan, Juan M Fernández-Luna, Juan F Huete, and Aurora Bueno-Cavanillas. 2023. Smartphone apps for domestic violence prevention: a systematic review. *International journal of environmental research and public health* 20, 7 (2023), 5246.

[28] Hung-Min Sun, Shiuan-Tung Chen, Jyh-Haw Yeh, and Chia-Yun Cheng. 2016. A shoulder surfing resistant graphical authentication system. *IEEE Transactions on Dependable and Secure Computing* 15, 2 (2016), 180–193.

[29] Matthew Turk and Alex Pentland. 1991. Eigenfaces for recognition. *Journal of cognitive neuroscience* 3, 1 (1991), 71–86.

[30] Esteban Vazquez-Fernandez and Daniel Gonzalez-Jimenez. 2016. Face recognition for authentication on mobile devices. *Image and Vision Computing* 55 (2016), 31–33. https://doi.org/10.1016/j.imavis.2016.03.018

[31] Tarun Kumar Yadav, Devashish Gosain, and Kent Seamons. 2023. Cryptographic deniability: a multi-perspective study of user perceptions and expectations. In *32nd USENIX Security Symposium (USENIX Security 23)*. 3637–3654.

[32] Adrienne Yapo and Joseph Weiss. 2018. Ethical implications of bias in machine learning. In *Proceedings of the 51st Hawaii International Conference on System Sciences*.

[33] Xingjie Yu, Bo Chen, Zhan Wang, Bing Chang, Wen Tao Zhu, and Jiwu Jing. 2014. Mobihydra: Pragmatic and multi-level plausibly deniable encryption storage for mobile devices. In *Information Security: 17th International Conference, ISC 2014, Hong Kong, China, October 12-14, 2014. Proceedings 17*. Springer, 555–567.

A DEMOGRAPHICS QUESTIONNAIRE

(1) Age
(2) Sex
(3) Country of Origin
(4) Which category describes you?
  - American Indian or Alaska Native
  - Asian
  - Black or African American
  - Hispanic, Latino, or Spanish Origin
  - Middle Eastern or North African
  - White
  - Multi-ethnic
  - European
  - Prefer not to disclose
  - Other: _

(5) Highest degree

B QUESTIONNAIRE ON SECURITY AND PRIVACY PREFERENCES

(1) What operating system do you use for your smartphone?
- Android
- iOS
- Other

(2) Have you ever unlocked your smartphone by facial recognition?
- Yes
- No

(3) Do you regularly unlock your smartphone using facial recognition? (1: "I'm not using face recognition" to 5:"always")

(4) If *yes*, approximately how many times per day do you unlock your smartphone with your face?

(5) What is your favorite way to unlock your smartphone?
- PIN
- Pattern
- Fingerprint
- Face recognition
- No unlock protection

(6) The security of my data is important to me. (1: 'completely disagree" to 'completely agree")

(7) I would use an additional feature to protect my data. (1: 'completely disagree" to 'completely agree")

(8) Is there data on your smartphone that you want to hide from friends and family?
- Yes
- No

(9) If *yes*, which of your data would you like to hide?

(10) What data on your smartphone is most important to you (Rank the following options from 1 to 4: Pictures, Documents, Contacts, Chat messages)

C ROLE-SPECIFIC QUESTIONNAIRES

We finished the study with role-specific questionnaires for both the roles of *attackers* and *suspects.*

C.1 Attacker Questionnaire

(1) Did you spot any photos during the case study?
- Yes
- No

(2) If *yes*, what kind of data?

(3) Did you feel like something was hidden from you?
- Yes
- No

(4) If *yes*, why?

(5) Do you believe your opponent told you the truth?
- Yes
- No

(6) If *no*, why?

(7) Would you consider the other person trustworthy? (1: "not trustworthy" to 5: "very trustworthy")

(8) If *no*, why?

(9) Have you figured out how to access the secret mode?
  - Yes
  - No
  - By indication of the study director

C.2 Suspect Questionnaire

(1) I felt safe using the app (1: ‘completely disagree” to ‘completely agree”)
(2) Were you able to successfully hide your data from your counterpart?
  - Yes
  - No

(3) If *no*, what betrayed you?
(4) Did you feel like the person you were talking to believed you?
  - Yes
  - No

(5) If *no*, why?
(6) I was able to get into the hidden mode of the app. (1: ‘completely disagree” to ‘completely agree”)
(7) During the experiment, I was worried about being exposed. (1: ‘completely disagree” to ‘completely agree”)
(8) Would you use the feature if your smartphone offered it to you?
  - Yes
  - No

Received February 2024; revised May 2024; accepted June 2024