FAKULTÄT
FÜR !NFORMATIK
Faculty of Informatics

Usable Authentication Ceremonies in Secure Instant Messaging

DIPLOMARBEIT

zur Erlangung des akademischen Grades

Diplom-Ingenieur

im Rahmen des Studiums

Technische Informatik

eingereicht von

Matthias Fassl, BSc
Matrikelnummer 00925194

an der Fakultät für Informatik

der Technischen Universität Wien

Betreuung: Privatdoz. Dipl.-Ing. Mag.rer.soc.oec. Dr.techn. Edgar Weippl
Mitwirkung: Univ.Lektorin Dr.techn. Katharina Krombholz-Reindl

Wien, 23. April 2018

Matthias Fassl	Edgar Weippl

FAKULTÄT FÜR !NFORMATIK
Faculty of Informatics

Usable Authentication Ceremonies in Secure Instant Messaging

DIPLOMA THESIS

submitted in partial fulfillment of the requirements for the degree of

Diplom-Ingenieur

in

Computer Engineering

by

Matthias Fassl, BSc
Registration Number 00925194

to the Faculty of Informatics

at the TU Wien

Advisor: Privatdoz. Dipl.-Ing. Mag.rer.soc.oec. Dr.techn. Edgar Weippl
Assistance: Univ.Lektorin Dr.techn. Katharina Krombholz-Reindl

Vienna, 23rd April, 2018

Matthias Fassl

Edgar Weippl

Erklärung zur Verfassung der Arbeit

Matthias Fassl, BSc
Zentagasse 1/20, 1050 Wien

Hiermit erkläre ich, dass ich diese Arbeit selbständig verfasst habe, dass ich die verwendeten Quellen und Hilfsmittel vollständig angegeben habe und dass ich die Stellen der Arbeit – einschließlich Tabellen, Karten und Abbildungen –, die anderen Werken oder dem Internet im Wortlaut oder dem Sinn nach entnommen sind, auf jeden Fall unter Angabe der Quelle als Entlehnung kenntlich gemacht habe.

Wien, 23. April 2018

Matthias Fassl

Acknowledgements

I would like to thank my advisor Katharina Krombholz-Reindl whose guidance and helpful tips were invaluable for this thesis. She was always available and reassured me when I was unsure and had open questions. Her words of encouragement after every progress report provided more support than I and probably she anticipated.

I would also like to thank all the participants for spending their valuable time in participatory design workshops. Without them this thesis would not have been possible.

I am grateful that my friend and former colleague Mathias Tausig asked me point blank if I did not want to write thesis in Security or Privacy, otherwise I would still be unhappily reading papers on Computer Architecture topics.

I would also like to acknowledge the extensive and quick proofreading feedback that I received from Robert Annessi whenever I asked for it. His experience with scientific writing proved to be helpful and I learned a lot from his feedback.

I want to express my profound gratitude to my parents Regina, Wolfgang, Franz, and Andrea who supported me financially and emotionally through all my years of study without pressuring me to study faster. They were the reason that I had the privilege of never worrying about housing or food, even if the money was sometimes tight.

Finally, I want to thank my girlfriend Kristina for listening to me, reassuring me, and encouraging me throughout my thesis and all my studies.

Kurzfassung

WhatsApp, Signal und die meisten anderen modernen sicheren Messenger haben eine geheime - und für die meisten Nutzer*innen - gut versteckte Möglichkeit Unterhaltungen gegen die gefürchteten Monster-in-the-Middle (MitM) Angriffe abzusichern. Diese so genannte Authentifizierungszeremonie setzt üblicherweise vorraus, dass sich die an der Unterhaltung Beteiligten persönlich treffen und die verwendeten Schlüssel entweder manuell oder automatisch abgleichen. Mehrere Studien bestätigen, dass sich die Sicherheitsvorstellungen und -erwartungen von Entwickler*innen und Benutzer*innen unterscheiden. Folgerichtig verstehen Benutzer*innen, ohne vorherige Erläuterung, oft nicht den Zweck und die Notwendigkeit dieser Authentifizierungszeremonien.

Daher ist ein neuer, explizit auf Benutzer*innen fokussierter Designprozess notwendig um den Ablauf von Zeremonien besser mit deren Vorstellungen und Erwartungen abzugleichen. Mein Ansatz war es Participatory Design Konzepte zu verwenden um die Erwartungen und Wünsche der Benutzer*innen zu verstehen, diese auf Authentifizierungszeremonien anzuwenden, und zu überprüfen wie solche Zeremonien die erforderliche Sicherheit bieten können.

Die Participatory Design Gruppen ergaben (1) Erfahrungsberichte die beschreiben welche Messenger-Eigenschaften den Teilnehmer*innen wichtig sind und welche eher zu Ablehnung führen, (2) Design-Konzepte für Authentifizierungszeremonien und deren Sicherheitsevaluierung, und (3) Design-Implikationen für zukünftige Authentifizierungszeremonien. Um den dargestellten Designprozess zu verdeutlichen werden basierend auf den Vorschlägen der Teilnehmer*innen beispielhafte drei verschiedene Authentifizierungszeremonien vorgestellt.

Abstract

WhatsApp, Signal as well as most other modern secure messaging clients have a secret and – for most users – thoroughly hidden way to protect conversations, especially from the dreaded Monster-in-the-Middle (MitM) attack. This so-called authentication ceremony often requires conversation partners to meet in person to manually or automatically compare their encryption keys. Numerous studies showed that values and mental models differ between users and security engineers. Consequently, users usually do not understand the purpose and necessity of authentication ceremonies without prior explanation.

Therefore, a novel, explicitly user-oriented design process is called for to connect the ceremony design with the users' mental models. I applied the concepts of participatory design in order to (1) understand how users expect resp. want an authentication ceremony to work, and (2) evaluate how ceremonies thus adjusted could be securely implemented.

The participatory design workshops resulted in (1) experience reports, that describe what users are looking for in secure messengers, and which aspects are obstacles to adoption, (2) conceptual designs for authentication ceremonies with corresponding security evaluations, and (3) implications for design of future authentication ceremonies. To illustrate the user-centered design approach, I provided three example ceremonies based on the users' suggestions and their security evaluation.

Contents

CHAPTER 1

Introduction

Sent emails usually traverse several different servers before arriving at the recipient's device. To keep the contents of emails secret from the servers between sender and recipient, it is necessary to use end-to-end encryption (E2EE). The first software that enabled users to use E2EE for their emails was PGP which Phil Zimmermann released in 1991. In 1995 a consortium around RSA Data Security, Inc. defined the first standard for S/MIME, which provides E2EE based on authority-issued certificates. Email clients supported this standard from the beginning because the consortium included the companies developing those clients. Academic researchers developed the Off-the-Record (OTR) plugin in 2004 for popular Instant Messaging (IM) applications because neither PGP nor S/MIME provide perfect forward secrecy or repudiation, two security properties that were important to them. After smartphones became popular in 2009, Whisper Systems developed the first mobile secure instant messaging solution *TextSecure*. At first the application only encrypted SMS messages, but later on an internet based messaging protocol was added. In 2012 other Secure Instant Messaging (SIM) solutions such as *Telegram* and *Threema* became popular as well.

In June 2013 the surveillance capabilities of the Five Eyes countries and especially the NSA and the GCHQ, became public knowledge because the whistleblower Edward Snowden in cooperation with The Guardian kept it in the news for several months" [Gle13; The13]. When Facebook announced their acquisition of *WhatsApp* in February of 2014 for $19 billion US dollars [GG14], the media and the users were still aware of the ongoing massive spy programs. This facilitated a discussion in the media about Facebook's motivation for this acquisition and what the effects on their privacy users could expect [Wor14]. During that discussion media outlets published articles about possible *WhatsApp* alternatives [Gib14]. The Google searches for secure messengers such as *Threema* and *Telegram* spiked as shown in Figure 1.1 after Facebook announced its acquisition on February 19th, 2014.

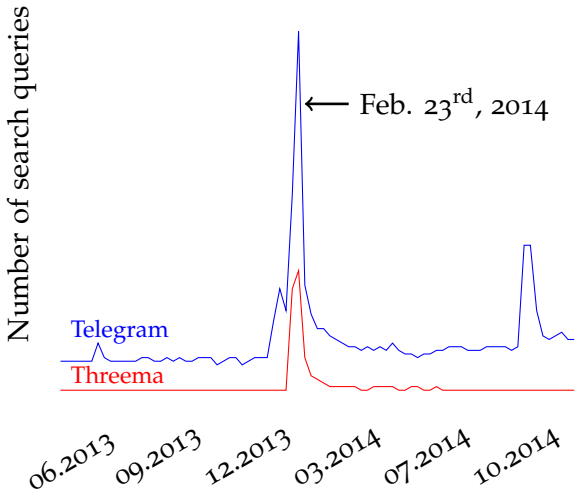


Figure 1.1: Number of search queries[1] concerning *Threema* and *Telegram* after Facebook announced its acquisition of *WhatsApp* on February 19th, 2014

However, these events did not necessarily lead to an increased concern for privacy or security – adoption of secure messaging solutions is usually driven by peer pressure and not by privacy or security concerns [DDO+16]. If users want to be part of a social group they have to follow its communication preferences.

As a response to the criticism and fears concerning acquisition of *WhatsApp*, Facebook introduced E2EE in 2016 based on the Signal protocol. Bey default *WhatsApp* hides all warning messages concerning encryption, which according to Schröder *et al.* [SHWR16] is understandable since they may cause more problems than they solve. However, this makes Monster-in-the-Middle (MitM) attacks without any resulting warning messages possible. This fact triggered a discussion amongst journalists of The Guardian and security engineers if that should be considered a backdoor [Gan17] or if it is a feature that increases usability [Mar17].

1.1 Security

Understanding the discussion about the potential "backdoor" in *WhatsApp* requires some background knowledge on the security features of most modern SIM applications. Most low-risk western users have rather vague threat models and in general focus on protecting their privacy by not revealing too much metadata and protecting themselves from mass-surveillance by the government or application providers [EHM17]. High-risk users have different and more specific threat models, they worried about active targeted attacks and device seizures.

[1] Data source: Google Trends (`www.google.com/trends`).

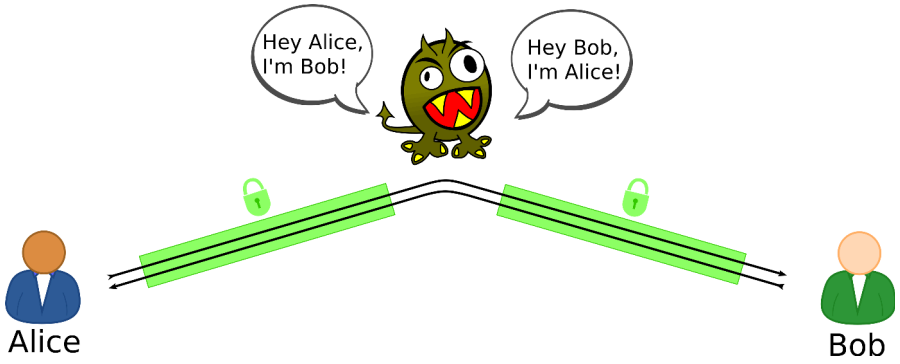


Figure 1.2: Monster-in-the-Middle (MitM) Attack

Modern messengers usually use opportunistic encryption for all communication by default. This does not require any user intervention and protects from passive attackers reading messages between conversation partners. However, opportunistic encryption does not protect from active attackers that either control central messaging services or have the power to arbitrarily manipulate the traffic between the conversation partners [SHWR16].

Under those assumptions a so-called Monster-in-the-Middle (MitM) attack is possible: An attacker can steal the identity of users in a conversation by convincing all participants that it is in fact the intended recipient of all messages. After this impersonation the attacker establishes E2EE connections with both conversation partners Alice and Bob and forwards all messages from Alice to Bob and vice versa as shown in Figure 1.2. Since all messages are forwarded Alice and Bob can communicate as before but the attacker has complete control over Alice's and Bob's communication and can read or manipulate all communication or even insert new messages. However, since the monster does not know the private encryption keys of either Alice or Bob, it has to tell both of them to use its own key so that it can read their communication. If Alice and Bob have communicated before changing the encryption keys will usually triggers a warning in messaging applications.

Since monsters have to use different encryption keys during a MitM attack it is possible for Alice and Bob to detect such attacks it by authenticating their encryption keys. Often, this authentication procedure consists of a face-to-face meeting between the conversation partners during which they read the fingerprints of their encryption keys to each other. If the fingerprints match the encryption key belongs to the conversation partner and not an attacker. Since security protocols that require the users' participation are called ceremonies, such procedure are also called authentication ceremonies.

Currently, the authentication ceremony of all widespread messengers depends on the verification of key fingerprints, sometimes with the help of graphical representations or a QR Code that are supposed to simplify comparisons. *Signal*'s user interface shown in Figure 1.3 offers a digit-only representation that has to be compared manually by

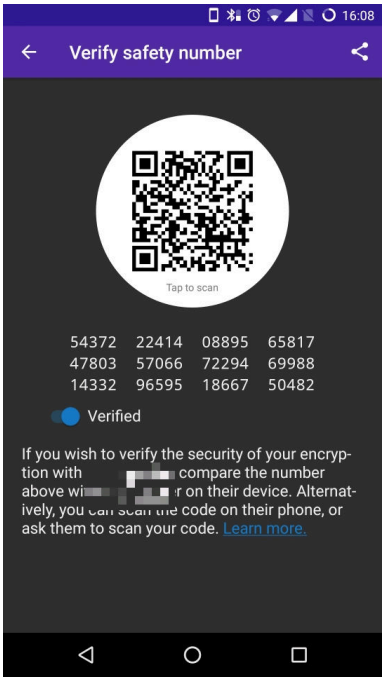


Figure 1.3: Signal's user interface for key verification.

users and a QR Code that enables the application to check the fingerprint automatically after scanning it. Comparing keys during face-to-face meetings guarantees that nobody can manipulate the information in transit, but since meetings are often not possible or practical other verification methods are necessary as well. Remote verification methods can either use in-band comparisons or out-of-band comparisons, the latter makes manipulation less likely. However, most instant messengers do not provide any explicit support for remote ceremony.

Since opportunistic and authenticated encryption protects against different kinds of attacks, it is important that users understand the differences and are able to authenticate each other in order to detect and mitigate MitM attacks. Users that are unaware of necessary security actions might assume that they are as secure as possible. This is called an illusion of security and is outright dangerous for high-risk users.

1.2 Usability

If users manage to detect and mitigate MitM attacks with the provided authentication ceremonies has been the subject of several recent user studies. Schröder *et al.* [SHWR16] found that users often ignored information about changed keys and instead blindly followed the flow of the user interface in order to continue the conversation. The *Signal* messenger did not provide information about potential risks, gave no indication of the verification status and provided unclear instructions or even none at all. These flaws

were part of the reason that participants of their study (all of which were computer science students) often assumed that the authentication ceremony was completed successfully, even if it actually was not.

Vaziripour *et al.* [VWN+17] found that many participants relied on ad-hoc methods, such as using shared knowledge to authenticate the conversation. In many cases they did not know about the existence of the keys or were unsure how they should be handled. The authors asked participants about their threat models and discovered that they are not aware of MitM attacks and consequently do not understand why authentication ceremonies are necessary. Even after participants were told about the existence of an authentication ceremony it took them on average more than 10 minutes to find it in the user interface and complete it. Another common problem was that users only partially verified the encryption keys, either because the keys were seen as too long to compare or because only one conversation partner actually verified them.

A study by Tan *et al.* [TBB+17] about the usability of different key representation formats found no significant correlation between computer and security expertise and the ability to successfully verify keys. This leads to the conclusion that even experienced users could have problems using secure instant messengers in a secure fashion. The representation of key fingerprints has a significant effect on the users' ability to correctly compare them since the attack detection-rate varied between 46% and 94%. However the authors recommend using automatic methods for verification whenever possible.

1.3 Values of Users

Usability studies of secure messengers show that users have problems using current authentication ceremonies, if they use them at all, and interpreting their results correctly. Dodier-Lazaro *et al.* [DSAB17] argue that these failures stem from the fact that security features do not match the users' security goals. Therefore, it is valuable to study those goals in order to design security solutions that make sense to users, that they are able to use, they want to use, and which provide security. Value-Sensitive Design (VSD), which focus on the users' interests, priorities, and values, should be applied during the design process in order to reduce the users' disengagement from security.

Ermoshina, Halpin, and Musiani [EHM17] interviewed developers and users of secure messaging application in order to compare their intentions. They confirmed their hypothesis that there is a mismatch and that developers prioritize features that are not important to either low-risk or high-risk users.

Abu-Salma *et al.* [ASB+17] interviewed 50 non-experts about their reasons for choosing or abandoning secure messengers. They discovered that usability is not the primary obstacle but rather the lacking utility as a communication tool and bad quality of service. The participants only had a vague understanding of basic security properties such as confidentiality, integrity, and authenticity – relying instead on cues from transmitted content to decide if a medium is secure or not.

In order to provide explanations for the low adoption-rate of E2EE Renaud, Volkamer, and Renkema-Padmos [RVR14] constructed a hypothesis-ladder to adoption, starting with "Awareness of Privacy Violations" and leading to "Is able to use E2EE and is not side-tracked". They were able to confirm four of their hypotheses using a series of semi-structured interviews. The users' incomplete mental models about message transmission and their underestimation of possible consequences were two major factors for the non-adoption of security tools. Participants using E2EE tools stated that they gained their knowledge about security issues either by personal experiences or stories told by others. Therefore, media coverage about possible security issues are helpful for raising awareness.

1.4 Research Questions

Authentication ceremonies are important because they provide users a tangible method for establishing trust. However, users have a wide variety of problems when trying to conduct currently used authentication ceremonies. In general they do not know why it is necessary, how to do it correctly, and how the results should be interpreted. The user interface of current implementation also rarely gives them any relevant guidance.

Maybe even more important are the users' motivation for adopting secure messaging solutions and conducting necessary security actions. Their security goals are different to those of the developers, which can be a reason for them to disengage from provided security features. In order to provide users the security features that they want and need it is necessary to focus on their values and priorities. By doing that, this thesis aims to provide usable and secure authentication ceremonies and answer the following research questions:

- How can users be assured that the communication is actually encrypted and that state-of-the-art encryption procedures are used?
- How can users be taught which attacks they are protected from and under which circumstances that protection is effective?
- How can users be motivated to conduct the authentication ceremony?
- In which way should a user interface provide guidance for an (face-to-face or remote) authentication ceremony?
- How should an authentication ceremony be designed such that users can complete it as quick as possible, without sacrificing security?

1.5 Methodology

I answered those research questions by using participatory design, a user-centered design approach. Participatory design is used to involve potential users in early stages

of research in order to incorporate their tacit knowledge and their values into the resulting designs. Similar to the study of Weber, Harbach, and Smith [WHS15], I will conduct five design workshops with two or three participants each. Session take about 60 minutes and have three distinct phases: brainstorming of past experiences and problems, introduction to basic security concepts, and design of one or more conceptual designs for authentication ceremonies.

Using Glaser and Strauss' [GS67] Grounded Theory, I will categorize the participants' approaches to authentication by coding the experience reports, conceptual designs, and notes that resulted from the workshop sessions. In addition I will derive implications for the design of authentication ceremonies from those results. Based on the conceptual designs, their security evaluation, and the implications for design I will present one or more example ceremonies that provide security as well as usability.

1.6 Summary of Results

At the beginning of each workshop session the participants listed the secure messengers that they have experience with and discussed the advantages and disadvantages of each. Convenience as a communication tool, large user base, and availability of clients for all platforms were the most common reasons for choosing a messenger. All participants expressed annoyance about the diversity and the lacking interoperability of messaging solutions and some wanted to be able to use messengers without a phone number. They were annoyed by a lacking quality of service, but did not experience usability issues.

The participants proposed numerous ways of authentication in electronic communication and they explained and drew 20 of those concept designs. Those concepts were categorized into six approaches to trust establishment: (1) comparing shared knowledge that only the conversation partners have, (2) showing pictures or videos of conversation partners, (3) asking friends or trusted contacts if they have authenticated the conversation partner, (4) trusting institutions to correctly authenticate people, (5) building up trust in the identity of the conversation partner over long periods of time, and (6) using technological measurements to test if the conversation could currently be under attack. 75% of the suggestions were from the first three categories, which suggests that those are the more intuitive approaches to trust establishment.

From the results of the five workshops I derived implications for design: (1) Successful authentication ceremonies must lead to a higher perceived level of security, (2) use well-known security concepts known from everyday life, (3) friction-less initiation of authentication ceremonies, (4) users work with different levels of trust, (5) the user interface must represent actual level of trust.

Three example authentication ceremonies based on the conceptual designs, their security evaluation, and the implications for design are presented in Chapter 5 to illustrate the design approach.

CHAPTER 2

Related Work

One of the first E-Mail encryption tools that were widely available, PGP, introduced a Graphical User Interface (GUI) in 1997 with version 5.0. Whitten and Tygar [WT99] were not convinced by the promised usability of this graphical version because in their opinion effective security software required a different usability standard. They conducted a cognitive walkthrough of the user interface and a usability test. The cognitive walkthrough discovered flaws and inconsistencies in the user interface – many due to the complexity of the underlying key management system. During the usability test participants were asked to send encrypted mails to a fictive campaign team. The test showed that problems were not constrained to the interface design and that participants did not understand the public key model which led to serious errors. Three of the 12 participants sent secret information unencrypted to the campaign team, seven participants encrypted the message with their own public key instead of the recipients', and only three participants had trust-concerns with the recipients' public keys. Three of the five participants, which progressed far enough to receive mail, had problems decrypting the message. One of the study's conclusion was the need to quickly communicate accurate conceptual models of the security, which requires sufficiently simple security models.

Interestingly, this early study already showed that usable security consists of a multitude of challenges. Some participants did not even progress far enough in the study to discover the complicated user interface - they already failed to understand the underlying conceptual models that no one bothered explaining. This chapter will present previous work related to secure messaging, starting with several usability studies and continuing with studies of the users' values, priorities, and requirements. At the end previous results concerning trust establishment in general and authentication ceremonies specifically will be presented.

2.1 Usability

One of the conclusions of Whitten and Tygar's study was that the usability of email encryption has to improve in order to gain a larger user base. Therefore, researchers invested a considerable amount of effort into conducting usability studies of encrypted communication. In the beginning they mostly studied email but later on also IM and SSL were considered. This section presents several relevant usability studies concerning encryption, mostly focused on email.

In 2005 **Garfinkel and Miller [GM05]** reinterpreted the findings of Whitten and Tygar [WT99] and proposed that the observed usability problems of PGP 5.0 stem from the complex key certification model and the lack of training. They compared PGP with S/MIME and concluded that one usability problem (complex model) was replaced by another one (obtaining X.509 certificates). Garfinkel and Miller suggested the use of Key Continuity Management (KCM), which has a simple trust model (Trust On First Use) and does not require certificates, as a possible solution to the usability problems and implemented a prototype. Their user study, which was constructed to be similar to Whitten and Tygar's original study, concluded that this model works well against attacks using either unsigned messages or new keys but does not work against attacks that use a new identity. KCM seems to be a simple way to provide protection against passive attacks and some active attacks but the fundamental problem of authenticating the used keys and identities remains unsolved.

Sheng *et al.* [SBHK06] tried to replicate Whitten and Tygar's study with a pilot study using PGP 9, which includes major changes such as semi-automatic key creation and distribution, opportunistic encryption, and automatic mail decryption. Half of the participants still had problems verifying the validity of the used keys and did not understand the reasoning behind it. None of the participants was able to sign a message because the interface did not provide any clues how to do that. Even though PGP 9 made life easier for users by automatically encrypting and decrypting messages, the key certification process has not improved at all.

15 years after Whitten and Tygar's assessment of PGP, **Ruoti *et al.* [RAZS15]** conducted a user study with a modern PGP implementation Mailvelope to check for improvements. Only one of the 10 pairs of participants managed to complete the assigned tasks. Both of the successful participants had prior knowledge about public key cryptography. The resulting System Usability Scale (SUS) score fell below the 15^{th} percentile, which is labeled "Not acceptable". Several common user mistakes were discovered and the authors suggested the following improvements: (1) integrated tutorials to explain how Mailvelope is used, (2) an approachable explanation of public key cryptography, (3) automatic email invites for recipients, and (4) better text to accompany the PGP block.

Ruoti *et al.* [RKB+13] conducted a user study of their secure webmail system Pwm, which is tightly integrated with existing webmail services and offers transparent encryption. They discovered that the participants did not have a clear idea of how the system worked and that some even unintentionally sent out plaint text emails.

They hypothesized that a manual encryption would avoid those kinds of users errors and created a mockup solution to test the hypothesis. The automatic and the manual encryption solution resulted in very similar SUS scores. The participants made less errors with their mockup, reported a higher confidence in its security, and accepted the additional effort that it required. Therefore, the authors argue that exposing encryption details produces a system that users trust and helps them avoid mistakes.

Atwater *et al.* [ABH+15] were not convinced by the results of Ruoti *et al.* [RKB+13] that user interaction with cipher text fosters trust and tried to replicate them in their study. They compared a transparent standalone, transparent integrated, and an opaque integrated tool. One of the findings was that about a third of the users trusted standalone tools more than browser solutions because they assumed those would communicate less with external entities. Despite this, all users preferred integrated (browser-based) tools because of their convenience. Another finding was that trust in a security tool depends either on its special security properties or its popularity and reputation. This leads to the conclusion that a better communication of security properties and reviews from reputable authorities are necessary to build trust.

Ruoti *et al.* [RAH+16] studied the usability of several different secure email solutions with pairs of novice users. They found that users prefer integrated solutions over depot-based (standalone) solutions and that tutorials are very important to reduce the amount of mistakes and the task time. They also provide some evidence that the users' trust in a messaging solution is reduced if details regarding the security features are hidden from view.

Even though more than 15 years of usability research on secure email communication led to significant improvements, only few people use it regularly. Focus has shifted to other problems, since email communication has become less popular and there is a an effort to make HTTPS standard for web connection.

Ruoti and Seamons [RS16] pointed out that many user studies of authentication schemes use ad-hoc metrics such as task completion time. They suggest the adoption of two standard scenarios, a bank website and a forum website, and the use of the SUS as a standard metric for calculating the relative usability.

Krombholz *et al.* [KMSW17] studied the usability of deploying HTTPS on Apache webservers. They recruited 28 computer science students from relevant courses whose background knowledge they tested. Data was collected using the think-aloud protocol during the lab sessions, entry- and exit questionnaires, the bash and browser histories, and the resulting configuration files. Additionally, the authors conducted seven expert interviews who could confirm that the discovered issues are also found during security audits. Four of the participants managed to deploy an A grade SSL configuration and four other participants did not manage to deploy a valid configuration at all. The most common usability pitfalls according to the participants were: (1) lack of best practice tutorials, (2) misleading terminology, and (3) weak default configurations.

2.2 Value-Sensitive Design (VSD)

As Whitten and Tygar already noted in their paper, one of users' core problems with PGP was that they did not understand asymmetric encryption and the used trust model. Therefore, it could be argued that the focus should be shifted to finding more appropriate conceptual models, instead of trying to improve the user interface. Despite of years improving the usability of secure communication tools, their user base is still rather low, begging the question why people do not use them and which changes are necessary to motivate them. This section presents papers that cover the mismatch of intentions of users and developers, the reasons for users to disengage from security features, how values and behaviours differ between experts and non-experts, and what can be learned from positive examples.

Dodier-Lazaro *et al.* [DSAB17] argue in their position paper that security goals are often set by security engineers and do not match the interests and priorities of the users. Experts then fail to notice that their imagined users' values differ sometimes drastically from the actual values, which results in disengagement of users. As a consequence they try to fix the users so that they are "able" to use the security features, instead of fixing the technology. However, in order to identify the root of disengagement, researchers have to study the users' reasons for not using security features. According to the authors a VSD approach is necessary that focuses first and foremost on the users interests, priorities, and values before concerning itself with the security implementation necessary to achieve the users' security goals. Since the users' perception of security is entirely independent of actual security features, the designers also need to take care to design visual and interactional cues that communicate a sense of security to the users.

Ermoshina, Halpin, and Musiani [EHM17] interviewed developers, users, and trainers of secure messaging solutions in order to discover how the design intentions diverge from the requirements of low-risk or high-risk users around the world. As the authors assumed they found a developer-user disconnect, with the developers caring about repudiation, decentralisation, standardisation, and licensing - while the users mainly cared about group support and metadata collection. While low-risk users usually had vague threat models consisting of passive attacks, the high-risk users had well-conceived threat models consisting of active attacks and device seizures. Many high-risk users wanted to have some way of key-verification even if it is hard to use, and both low- and high-risk users thought that is important to be able to observe the encryption process. The interviews with trainers uncovered that training sessions in low-risk environments focus on privacy in general, and more on operational security in high-risk settings.

Egelman *et al.* [EJP+14] used structured interviews with 28 participants to discover their reasons for locking or not locking phones and their risk perception. The two most common reasons for locking were keeping out either strangers or relatives. The point was not that participants did not share their phone, but that they wanted to regulate access to it. One-third of the eight participants who did not lock their phone cited a lack of motivation as their main reason. Others wanted it unlocked in case of

emergencies, thought it increased their chance of getting it back in case of loss, or that their data on the phone has no value. The authors conducted an online survey with over 2500 respondents which confirmed those reasons discovered by the interviews. The interview data and the survey results suggest that users choose not to protect their phones because they underestimate the amount of sensitive data stored on them.

McGregor *et al.* [MCHR15] used semi-structured interviews to investigate general and computer security practices of 15 journalists in the US and France. They found out that security researchers should focus on the needs of the sources, since journalists choose methods of communication according to the source's comfort and availability. Additionally, security solutions should not interfere with the journalistic process, if a security tool makes source authentication impossible, journalists will not use it.

Abu-Salma *et al.* [ASB+17] interviewed 50 lay people about their reasons for adopting a secure messenger, their understanding of security and their security-ranking of different modes of communications. The most important factor for adoption seems to be the utility of the communication tool, regarding the user base, offered services, quality of service, and the context of use. Usability is not the primary obstacle of adoption and transmission of sensitive information does not drive adoption of secure messaging tools. The participants think that there are some kinds of information that are sensitive to them but they communicate those face-to-face, with voice calls which is perceived as the next best thing to face-to-face conversation, or with communication modes that are already used for sensitive information like emails and SMS. The participants only have a vague understanding of the basic security properties confidentiality, integrity, and authenticity. Interestingly, they think that confidentiality is unachievable but message integrity is hard to break. There are three reoccurring threat models stated by the participants: intelligence agencies, application service providers, and attackers with technical knowledge. The general notion being that nothing that is communicated is secure from any of those three threats. The study concludes with three suggestions: (1) secure IM solutions that already have utility; (2) in order to have threat models that matter to the users the target population has to be understood - a user-centered design process is necessary; (3) improve quality of service of existing SIM solutions.

Renaud, Volkamer, and Renkema-Padmos [RVR14] wondered why E2EE is still not in widespread use, even though the usability has vastly improved in recent years. They conducted semi-structured interviews with lay people and a survey with computer science students. The authors managed to confirm the four hypotheses: (1) people are privacy aware but not concerned; (2) they are privacy concerned but have misconceptions; (3) they understand the privacy problems but do not see a need to protect themselves; (4) the problems are understood but they do not know how to protect themselves. The two main contributing factors were that participants had an incomplete mental model of how message transmission works and that they were underestimating possible consequences of privacy problems. Guides that explain how to protect ones privacy do not help much in this regard, participants stated that they gained knowledge either by personal experience or stories told by others. Media coverage therefore plays

a significant role in the awareness of risks and consequences. Interestingly, there was no observable difference in the mental models of lay people and computer science students.

Experts and non-experts have different mental models of how the technology they use works. Evaluating the differences of these mental models is worthwhile, since they define the users' interaction with technology. Comparing the different behaviour of those two groups can lead to more appropriate conceptual models.

De Luca *et al.* [DDO+16] used an online survey and a series of structured interviews to find out if expert and lay people have different attitudes towards privacy and security and how those attitudes affect decisions to use a mobile instant messenger. The online survey with over 1500 participants uncovered the two main reasons for choosing a mobile messenger: friends use it and it is free of charge. Another interesting result is that the participants' country correlates with the knowledge about and use of security and privacy solutions - German participants were more than twice as likely to say that they use SIM. The interviews confirmed that the most important reason for using an instant messenger was that either *"everyone is using it"* or *"specific people use it"*. Participants are willing to accept the additional costs of a messenger if they want to stay in contact with even a single person who is important to them. Previously experienced bad usability of security solutions also impact the use of SIM. Five of the 16 participants who did not use SIM, mentioned that it would be difficult to use. Participants also mentioned that IM is time sensitive to them and that they would therefore consider uninstalling a secure solution if it adds additional delay. The importance of the exchanged messages was considered to be rather low. Emails were much more important to the participants and seven of the non-experts even thought it was the more secure medium, which is interesting because the security and privacy of email transmission is usually low compared to IM solutions. Participants did not perceive it as a threat that messages might be leaked to an unknown entity but many were concerned that known people might see their messages. Even though the experts had more technical background knowledge, they exhibited about the same amount of insecure behaviour as the lay people.

Kang *et al.* [KDFK15] conducted semi-structured interviews with lay people and people with computer science background to compare mental models of how the Internet works and the implications for privacy and security. People with technical education have a more articulated mental model and have a better understanding of the possible security risks involved in using Internet services. This knowledge, however, does not translate to secure online behaviour. It was observed that people who identified a higher number of privacy threats also invested more effort in countermeasures. This awareness of privacy threats can be partially explained by education but it is most often shaped by personal negative experience which tends to trigger more secure online behaviour. Those who have not had negative experience are habituated to the convenience of the Internet and are less motivated to protect themselves.

Gallagher, Patil, and Memon [GPM17] conducted semi-structured interviews with

6 experts and 11 non-experts to find out more about their perceptions and usage of the Tor anonymity network. While experts had a fairly accurate mental model of the Tor network, the non-experts treat it as a centralized "black box" that offers a service. Experts understood Tor's threat model but five of the non-experts believed that the network provided more security than it actually does. Non-experts also tended to be more curious about the information and services available on Onion services than about the achievable anonymity.

Positive examples of secure communications are quite rare, and even rarer in larger collaborating groups. It is worthwhile to pinpoint the reasons why secure communication works in some instances, in order to apply those findings to other projects.

McGregor *et al.* [MWA+17] were intrigued by the investigation of the Panama Papers which was kept secret for a year, while protecting the sources, and maintaining control of the documents. They conducted a survey among 118 journalists and semi-structured interviews with the designers of the used collaboration system. The most important reasons for the success were that (1) journalists found the security features both useful and necessary, (2) security practices were normalized, (3) the available secure communication methods minimized workarounds, and (3) social relationships were leveraged.

2.3 Trust Establishment

Unger *et al.* [UDB+15] define trust establishment as *"the process of users verifying that they are actually communicating with the parties they intend"*. This process can be either be an authentication ceremony, which requires an interaction between the communication partners, or an automatic process without necessary interaction. There are currently four different kinds of authentication ceremonies in use: (1) fingerprint verification, (2) Short Authentication String (SAS), (3) secret-based zero-knowledge verification using the Socialist Millionaire Protocol (SMP), and (4) mandatory verification. The automatic processes include: (1) opportunistic mode, (2) Trust On First Use (TOFU), (3) authority-based trust, (4) transparency logs, and (5) blockchains. In general the authentication ceremonies provide the highest level of security but are not usable in many aspects. The automatic processes, especially TOFU and authority-based trust provide excellent usability but do not provide as much security. The authors conclude that a layered approach, such as using TOFU as a baseline and authentication ceremonies as an additional layer of security is a good compromise for messaging solutions.

Fingerprint Verification

Verifying the fingerprints of the used encryption keys is currently the most commonly used approach to user-to-user authentication. Most often this is combined with an opportunistic mode so that the automatically encrypted sessions are trusted by default until proven otherwise. Whatever the reason for the widespread use of this authen-

tication method may be, its resulting security has been subject of several usability studies.

Schröder *et al.* [SHWR16] studied the usability of *Signal*'s authentication ceremony in 2016. They recruited 28 computer science students from an Human-Computer Interaction (HCI) course at the University of Vienna. Participants received a mobile phone with *Signal* preinstalled (Alice) and were tasked with sending sensitive information to Bob (part of the research group). During the conversation a MitM attack was deployed against the conversation and the reactions of the participants were recorded. Four participants immediately accepted the new key and followed the flow of the user interface. Only seven of the 28 participants managed to successfully match keys with Bob (resulting in an error). Exactly one of those seven drew the right conclusion from the mismatch: that a MitM attack is taking place. It was noted that the initial error message does not mention the possible consequences of a key change at all and that the key verification page did not give any instructions on how to perform the key matching. The participants' different verification and attack mitigation strategies can be explained by flawed mental models. The users often have no idea of the fundamentals of E2EE, MitM attacks, and associated risks. The research group gave several recommendations for improving the usability of *Signal*: conversations should have an easily visible security status, error messages need to clearly communicate potential risks, clear instructions on how to perform recommended actions are needed, and the authentication mechanisms should be easily accessible.

Herzberg and Leibowitz [HL16] quantitatively and qualitatively evaluated *WhatsApp*, *Viber*, *Telegram*, and *Signal* in order to find major barriers to adoption. They found that in general users want to be protected from rogue operators but that required security actions need to be easy for them. Since none of the participants realized that the opportunistic mode (which usually does not require any action by the users) only protects against passive attackers, an illusion of security is created. The participants were unaware of the authentication ceremony that is required to protect themselves against rogue operators. None of the studied messengers' authentication ceremonies were considered usable, 43.6% of the participants described the process as non-intuitive and complicated. All messengers assumed that users have a secure channel for the authentication ceremony, which is a non-trivial assumption that could lead to insecure behaviour. The majority (70% and more) of the participants failed to detect key changes even when explicit warning messages are shown and over 60% stated that they would not (re-)authenticate in regular use. The study concludes with several suggestions: applications need to find simple ways to explain necessary security actions, the ceremonies need to be simpler, gamification could be used to make the ceremony more pleasant, and contacts could be treated with different levels of vigilance.

Vaziripour *et al.* [VWN+17] compared the usability of the authentication ceremonies of three different secure messaging systems (*Viber*, *WhatsApp*, and *Facebook Messenger*). The tasks of the user tests consisted of two phases, in the first one participants were asked to *"make sure that you are really talking to your friend"*. Most of the participants relied on

personal characteristics, such as voice, shared language, visual identification or used shared secrets for the authentication. The success rate for the key verification was only 14%. In the second phase participants received an introduction to secure messaging, the possibility of intercepted communication and the necessary key authentication. The success rate for the key verification rose to 78% and the confidence in privacy increased as well. *Viber* emphasises the concept of "trust" and avoided using technical terms when talking about cryptography. This and its structured user interface led to a high success rate among *Viber* users. The comparison of the used keys seems to be a general problem in all studied messengers: the key can be quite long and many participants did not compare the complete keys (and complained about the length) or in case of a single key the dominant communication partner read the whole key (so that effectively only one party was authenticated). The time it takes to conduct an authentication ceremony was also considered to be too long: it took on average about three minutes to find the option in the user interface and additional seven minutes to complete the key verification.

Key Representations

Since key verification is a widespread method for authentication studies have tried to answer the question how the key material should be presented in order to enable users to detect attacks efficiently. However, this assumes that users are motivated to compare the key material at all.

Dechand *et al.* [DSB+16] conducted an online study with 1047 participants that compared different key-representations used for verification. They measured how much time the participants needed to compare fingerprints and how high the attack detection rate was. The hexadecimal representation fared worst and the authors discourage its use for verification purposes. Large wordlists were the fastest representation and the attack detection rate was highest with generated sentences. However, this study did not measure the users' motivation to verify key fingerprints - even a good representation does not work if users do not verify it.

Tan *et al.* [TBB+17] studied different key representations and tested how well they work in cases were a trade-off between security and usability is necessary. Since attacks are rare, there is an expectation and pressure to have a positive comparison result, which was simulated in the study by introducing a time limit for comparisons. Several textual and visual key representations were tested. Generating sentences from the key led to an attack detection rate of 94%, which was the highest of all the representations. Visualizing keys by generating different unicorn images that have to be compared led to the worst attack detection rate of only 46%. The research group also studied the influence of different modes of comparison (confirmation vs. selection) and the necessity of toggling between different applications (e.g. on mobile phones). The confirm-and-select procedure, where different keys are displayed and users have to select the correct one, apparently habituates users that a correct selection-option always exists. The resulting attack detection rate of only 28% discourages usage of this

comparison mode. In environments where users have to toggle between applications, the textual representations resulted in fewer mistakes since remembering distinct numbers or words is easier than remembering visual features. The study concludes that none of the key representations seem adequate for environments were security is paramount and suggests using automatic key comparison (e.g. by scanning QR Code representations of keys) wherever it is feasible. However, textual representation seem to be an adequate trade-off between security and usability for most common usage scenarios.

Shared Secret

Key verification does not seem to be the most usable mechanism to authenticate users. Therefore, the developers of OTR have chosen a more intuitive approach based on shared secrets.

Alexander and Goldberg [AG07] were unsatisfied with the current state of authentication in messaging. The authentication mechanisms used required an understanding of keys and fingerprints that lies outside the comfort zone of many users. To address this issue, they built an authentication mechanism based on a solution of Socialist Millionaire's problem [BST01] that enables users to verify each others identity based on shared knowledge.

Stedman, Yoshida, and Goldberg [SYG08] conducted a user study with the OTR pidgin plugin in order to test the usability of the shared-secret based authentication ceremony. They recruited four pairs of friends, who they expected to establish a secure shared secret without problems. The most interesting of the participants' four tasks was starting and authentication a private conversation using OTR. The authors discovered several usability flaws with the user interface and the corresponding help pages. The participants understanding of the process and their motivation for it were not part of the study. However, when Participant 2 knowingly authenticated someone in an insecure way, she justified herself by saying that she could not see a reason for anyone to impersonate her chat partner.

Alternatives

It is still an open question how user-to-user authentication can and should work. In general researchers seem to prefer automatic authentication since it minimizes necessary user engagement. The following papers present authentication schemes that have been introduced recently and have not seen widespread use.

Vaziripour *et al.* [VWH+16] discussed that even though there has been progress in the area of authentication, user-to-user authentication is still a problem. They suggest using social authentication for E2EE. Since following users on *Twitter* or accepting a friend request on *Facebook* are authenticated judgements, those trust relationships can be facilitated for verifying key-identity pairs. The authors state that *Keybase* and *SafeSlinger* already provide parts of such a solution. Open questions regarding social authentication

are how managing keys should work, how strangers can be authenticated, and how novices can transition to such a system.

Lerner, Zeng, and Roesner [LZR17] developed *Confidante* an encrypted mail client that avoids common key-management pitfalls by using *Keybase*[1] to discover the receivers' public keys. *Keybase* lets users post cryptographic proofs that associate their public key with their social media accounts. The authors conducted a user study with eight lawyers and seven journalists, both of which are groups that often use email to transmit sensitive data. The participants compared using *Confidante* to a normal email experience. Technical users were concerned about its security, because it seemed "too easy" and two of the participants did not even notice that they sent encrypted emails. However, most participants stated that they would still only encrypt the most sensitive data, suggesting that they are still reasons not to encrypt messages. The authors used the opportunity and asked the lawyers and journalists about their threat models, security needs, and usability preferences.

Farb *et al.* [FLK+13] developed *SafeSlinger* which is used to establish trusted secure communication. *SafeSlinger* users exchange contact details including their public keys when they meet in person, which results in trusted associations of identities and keys. Afterwards, users can exchange secure messages, transfer files, or forward contact information using a mechanism called secure introduction. The authors conducted a user study with 24 participants using a within-subjects design to compare *SafeSlinger* to *Bump*. The participants preferred *SafeSlinger*, since *Bump* does only work between two people at once and nearby users bumping at the same time can result in privacy leaks.

Karlof, Tygar, and Wagner [KTW09] introduce the notion of Condition-safe Ceremonies. Such ceremonies should condition users to only use safe rules, that do no harm in the presence of an adversary, and at least one immunization rule, that causes a potential attack to fail. In order to test their assumptions they build a condition-safe email registration ceremony and conducted a user study with 208 participants. The results suggest that condition-safe ceremonies can be helpful to prevent some kind of social engineering attacks and that the notion of such condition-safe ceremonies could be applicable to other areas.

One of the automatic processes that promises a suitable security level while not needing any user interaction is the CONIKS system introduced by **Melara *et al.* [MBB+15]**. It uses transparency logs to detect MitM attacks in secure messaging solutions. Transparency logs were previously used by certificate authorities to provide irrevocable evidence of issued certificates, which enables auditors to detect rogue operators more quickly. The necessary auditor role and the public name-key bindings made this system unsuitable for messaging applications. Using CONIKS every client audits its own name-key binding and whistleblows to others if it detects a malicious binding. In addition clients have to audit the identity provider for non-equivocation, i.e. verifying that the identity provider hands out the same name-key binding to everyone who

[1] `https://keybase.io`

asks. The system is privacy preserving because those consistency checks do not require knowledge of all name-key bindings. It is not even necessary to disclose the number of users of identity providers. Key changes can be handled in two ways, by default anyone can just upload a new name-key binding and in the strict mode every changed name-key binding has to be signed by the old key. Since every client audits its own name-key bindings even the default mode is secure, since unauthorised changes will be detected. The strict mode provides an extra layer of security but may lead to unusable identities if the key is lost. The performance evaluation showed that the overhead is rather low: clients need to download about 17 kB per day from the CONIKS server and the verification of the key bindings takes a few milliseconds, servers easily support up to 10 million users.

CHAPTER 3

Methodology

Since several studies presented in the last chapter [HL16; SHWR16; VWN+17] suggest that the current design approach for authentication ceremonies does not result in usable solutions, a different design approach is needed. Dodier-Lazaro *et al.* [DSAB17] state that common design approaches are paternalistic because they usually incorporate only the designers mental models and should be replaced by VSD which focuses on the expectations, mental models, and values of prospective users. Participatory design is one of many methods which can be used to implement VSD. Since users are actively involved in the design process their implicit expectations and values are incorporated into the resulting designs. Although this method is well-known in HCI, it has not seen wide-spread use in the field of Usable Security. Weber, Harbach, and Smith [WHS15] used it in 2015 for designing SSL warning messages and described how they achieved interesting results with only a few small workshops, which convinced me to try this approach for designing authentication ceremonies.

This chapter will describe the history of participatory design, the participation requirements, the procedure for the design workshops, and the evaluation of the resulting data. Each workshop is divided into three distinct phases which will be used to understand issues with secure messengers, develop a common language for threats and security properties, and create conceptual designs for authentication ceremonies. The evaluation of participatory design workshops will result in common categories of the participants' conceptual designs, corresponding security evaluations, and more general implications for design of authentication ceremonies.

3.1 Participatory Design

As described by Spinuzzi [Spi05] participatory design was first applied during the 70s and 80s in Scandinavia and is rooted in a Marxist movement. In a time of increasing

automation and mass production it aimed to empower workers to participate in the development of machines that threatened to render their skill set useless. The main concern is bridging the gap between the workers' tacit knowledge, i.e knowledge that is hard to communicate, and the researchers' abstract and analytical knowledge. Researchers applied many methodologies from the preexisting action research to design artifacts together with users instead of designing them on their behalf.

According to the author participatory design studies commonly have three separate stages which are applied in an iterative manner: (1) Exploration; (2) Discovery; (3) and Prototyping. To enable cooperation between researchers and users a shared language is necessary and therefore a central part of participatory design studies. Limitations of participatory design are that it usually results in an evolution and not a revolution, that it focuses more on artifacts instead of workflows, and the tendency of researchers to forget its political commitment to the empowerment of workers.

Even though participatory design has a decade-long history in computer science, it has not been as popular in the field of security. In 2015 Weber, Harbach, and Smith [WHS15] applied participatory design to redesign SSL warnings. SSL warnings are one of the most confusing error messages in existence and most users do not understand them correctly. Akhawe and Felt [AF13] have studied the click-through rates of SSL warnings and confirmed that the user experience of a warning has a significant impact on their behaviour. Weber, Harbach, and Smith used participatory design to identify existing problems from the users' perspective and generate different warning message options. They conducted five focus groups which three participants each, whereby each session consisted of three distinct phases. In the first phase the participants got to know each other and brainstormed experienced problems, in the second phase the group moderator tried developing a shared language for the group by explaining the security issues, and in the third phase the participants were asked to design an alternative warning message.

According to Weber, Harbach, and Smith the major benefit of this method is the generation of many different interface options, since focus groups will always have diverging ideas. In their study they compared the results of two focus groups where all participants studied computer science. The first one which consisted only of women generated a colorful interface that contained a summarized problem description, which was a very different result to the group consisting only of men, which generated an interface with few colors and a detailed problem description. This thesis aims to replicate the methodology used by Weber, Harbach, and Smith as closely as possible. However, alterations are necessary because authentication ceremonies are workflows rather than artifacts like SSL warnings.

3.2 Participants

In order to provide effective security, authentication ceremonies must be usable by a diverse group of people. This means that the study participants should no more

than a layperson's understanding of communication technology and cryptography. However, since I am interested in understanding the problems of existing solutions, it is a requirement for participants to have used a secure messenger before. I assume that people who have experience with secure messaging are interested in the topic of secure communication in general and are therefore easier to recruit for the study. Those considerations resulted in three concrete participation requirements:

1. Potential participants have experience with a messenger that they consider to be secure e.g. Signal, WhatsApp, Viber, Threema, or Wire,
2. they can not explain how asymmetric encryption works which would indicate cryptography knowledge,
3. they can not explain which entities could potentially read a message during transfer which would indicate knowledge of communication technology.

The advertisement for the study focuses on groups of people who have an interest in secure communication but do not have computer science (or related) background. Three different groups are therefore targeted:

- **Political activists groups** are commonly interested avoiding surveillance by rivaling groups and government organizations. They are reachable by contacting a local branch office of a political party, writing to mailing-lists, or word-of-mouth communication.
- **Cryptoparties** are attended by people who already have an interest in secure communication. Attendees are reachable by contacting the event's organisers, writing to the mailing-lists, or announcing the study in person at a regular event.
- **Students from other majors** may have an interest in but not as much knowledge about communication technology and cryptography as computer science students. They are reachable by mailing-lists of student representatives, the local bulletin board, announcement posters in the hallways, and in person announcements during lectures.

All people who were interested but not eligible to participate were asked to recommend other potentially interested participants resulting in a snowball sampling. This recruitment strategy combined with the focus on specific target groups will probably yield participants who are in their mid-twenties and have a higher-tan average education background. This will introduce a bias in the resulting experiences and conceptual designs because older people and people with lower socioeconomic background will most likely be underrepresented.

Since the means to reimburse participants with money were not available, I offered different incentives such as a chance to learn more about secure communication, a

contact person who is available for future questions about secure communication, and free food and (non-alcoholic) drinks during the focus group sessions.

Each participant will receive the following information: (1) detailed explanation of the procedure, including the duration, risks, and benefits; (2) information on how and when results will be shared; (3) guarantee of their anonymity in the study; (4) their right to refuse and withdraw at any time during the study. All participants must sign a consent form to confirm that they have received and understood this information. In addition to the general information the consent form (see figure 3 on page 68) gives permission to use the resulting data for research purposes and includes contact information of the participants which is necessary to share the results. At the end of each workshop sessions participants receive my business card in case questions about secure communication arise in the future.

3.3 Location and Timeframe

The location for the focus group sessions has the following requirements: availability for at least one hour, undisturbed (i.e. not a walk-through room), good lighting, quiet enough for conversation, and at least one table with enough seating for all participants.

We have the option to reserve seminar rooms at TU Wien which fulfill all requirements stated above. Since the focus groups consists of a maximum of four participants, a small seminar room is sufficient. Smaller rooms have the benefit they are unoccupied most of the time and can therefore be reserved on short notice.

All five focus groups are conducted during the course of one week with two alternative dates that are used in case of timing issues. In order to enable the participation of employees the sessions will be scheduled either in the morning from 9:00 to 10:30 or after work hours from 17:30 to 19:00.

3.4 Procedure

Similar to the work by Weber, Harbach, and Smith [WHS15] I aim to have three participants in each focus group. However, the procedure also works well with two or four participants in case of last-minute changes.

Each session is preceded by a **preparation** of about 10 minutes length. This time is used to regulate the room temperature, taking care that the room is well lit, and that sufficient seating and a table is available. Additionally, the documentation of the session needs to be prepared, papers and pens distributed, and the audio recording equipment set up and tested. The moderator prepares the presentation for the common language phase.

After all participants have arrived the **introduction** takes about 10 minutes. This introduction contains an explanation of the procedure, the participants' rights, and

how and when they will receive results of the study. Afterwards, all participants sign the consent form and fill in their contact details if they want to be informed about the study's results.

The three main phases brainstorming, common language, and design are conducted as follows:

The **brainstorming phase** takes about 15 minutes and gives the participants the opportunity to talk about their experienced problems with secure messaging in general and authentication ceremonies specifically. The moderator uses the following questions to start the discussion: (1) *"With which secure messengers have you had experience?"*, (2) *"What kind of positive or negative experiences have you had with secure messengers?"*, and (3) *"Which of those messengers do you not use anymore, and why?"*. Each of the participants has five minutes to elaborate on their experiences and the moderator documents those experiences on paper.

The **common language phase** takes about 10 minutes and aims to teach participants base knowledge and common terms concerning end-to-end encryption, threat models, active attacks, and trust establishment. The moderator presents those basics with the help of a short slide show (see Figure 3.1). Afterwards, the participants have time to ask questions about those concepts. During the presentation and the discussion afterwards, the moderator takes care not to provide details about specific authentication ceremonies, since that information influences the participants' conceptual designs.

During the **design phase** the participants collaboratively design an authentication ceremony which fulfills the requirements described in the common language phase and remedies their experienced problems discovered in the brainstorming phase. Each focus group receives the following scenario as a starting point for the discussion about the prototype design: *"You and a colleague from another branch office want to discuss a surprise party for a mutual friend. Your mutual friend with technical skills has gotten wind of something and wants to find out exactly what you are conspiring but you definitely do not want to ruin the surprise and installed a secure messenger. You do not really know that colleague too well and you may or may not be able to meet him in person. What would you do in order to convince yourself that you are in fact talking to the right colleague and that nobody listens in?"*. The participants have about 10 minutes time to discuss different ideas and design options before drawing and describing the one they think is suited best. In order to document the results, the moderator takes notes and asks the participants to describe their authentication ceremony in detail for the audio recording. This helps interpreting the drawings correctly during the evaluation.

The **cleanup** phase starts after all participants have left. The moderator saves the produced results, i.e. makes a backup copy of the audio recording and scans all paper notes. Afterwards the moderator returns the room to its original state.

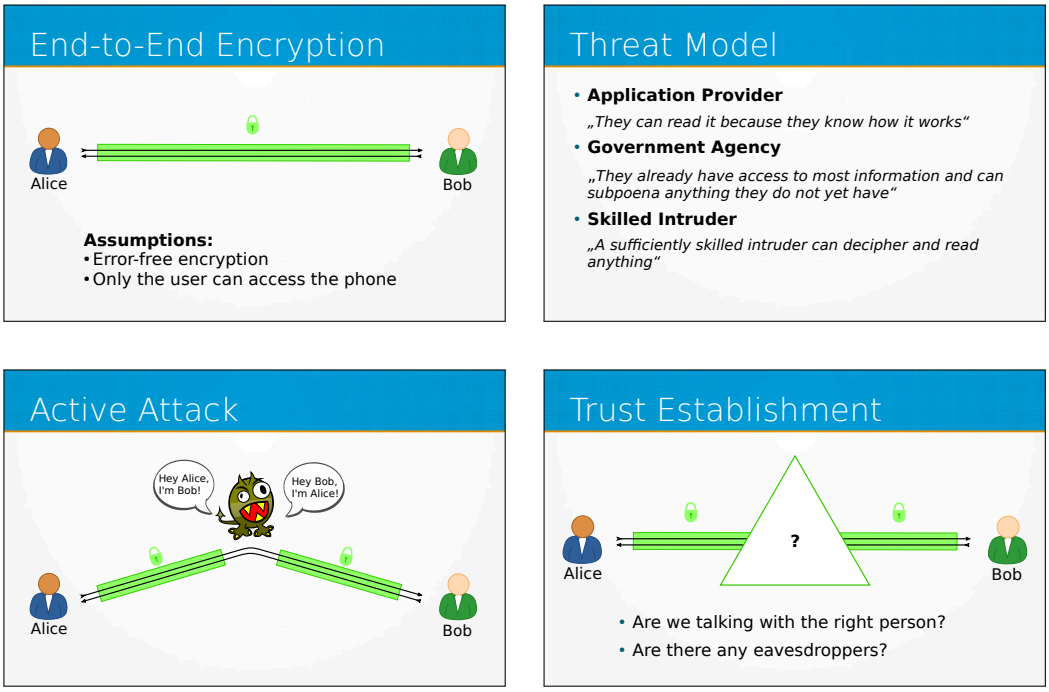


Figure 3.1: Presentation slides for the common language phase.

3.5 Pilot Study

I am conducting several pilot studies in order to collect feedback on the proposed procedure. The conditions are similar to the main study but less formal and with less focus on the results. The participation requirements do not have to apply. The used location for the pilot study needs to be quiet enough to hold a conversation and not have too many distractions but does not need to be prepared in any way. Since testing the procedure is more important than the results the available time in each phase is reduced to 5 minutes.

An additional **feedback phase** is added after the design phase during which the moderator asks participants for feedback about the used procedure. The moderator uses the following questions to elicit responses: *"Did the initial procedure explanation prepare you for the procedure?"*, *"Were the concepts described in the slide show helpful for designing your prototype ceremony?"*, *"What parts of the procedure did you find unexpected?"*, *"What concepts came up during the session that you found confusing or difficult to understand?"*. The moderator will document the resulting feedback points on paper. Should a pilot study lead to significant procedural changes, it needs to be repeated with the adapted procedure and different participants.

3.6 Data Evaluation

The original paper by Weber, Harbach, and Smith does not emphasize the analysis of the warning message prototypes because they can be used in practice without it. However, they compared the prototypes of the focus groups and discussed the implications of the differences.

I analyze the problems discovered during the brainstorming phase for multiple instances of experienced problems. A problem that is experienced by more than 2 participants (i.e. *triangulation*) indicates a widespread problem. I will compare those widespread problems with the ones known from the literature [SHWR16; HL16; VWN+17] to if they were previously unknown, confirm existing results, or if some major problems have not come up in this study.

I use Grounded Theory which was introduced by Glaser and Strauss [GS67] to group similar conceptual designs and assign them to a specific approach to trust establishment. Every conceptual design will be subject to a security evaluation that will suggest a technique known from the literature [UDB+15] to provide the expected security. I use emergent coding to extract implications for the design of authentication ceremonies based on the participants' experience reports and their conceptual designs.

3.7 Ethical Considerations

Our university does not have an ethics board but it has a set of guidelines that we followed in our research. A fundamental requirement of these guidelines is to preserve the participants' privacy and to limit the collection of personal data as far as possible. All participants were informed about the purpose and procedure of the study without any deception before the participatory design workshops began. The consent form explained how the collected data would be used, guaranteed the participants' anonymity, and stated that the participants could leave at any time without explanation. As a reimbursement for their time snacks and non-alcoholic beverages were provided during the workshops and I offered to provide counsel in the future if questions about secure communication should arise. The following personal information was collected for the analysis of the participatory design workshops: age, education, gender, knowledge about communication technology, and knowledge about cryptography. Additionally, I collected the email addresses of the participants in order to organize the workshops and communicate the results of the study.

CHAPTER 4

Workshop Results & Evaluation

In preparation for the participatory design workshops I conducted a pilot study consisting of three sessions. It aimed to confirm that the workshop procedure is feasible in the available time, understood by the participants, and leads to results. At the same time I organized a seminar room for the workshops and advertised the study on social media using snowball sampling when people did not fulfill the participation requirements. Similar to the study by Weber, Harbach, and Smith [WHS15] I wanted to conduct five workshops with three participants each. After one week of advertising I found two participants for each workshop session which was the minimum amount I set for myself.

As described in Chapter 3 each workshop session consists of three phases: (1) the brainstorming phase, (2) the common language phase, and (3) the design phase. During the first phase participants discuss which messengers they use, which problems they encountered, and why they liked using the them. In the design phase participants suggest conceptual designs for authentication ceremonies, which I will present according to their approach to trust establishment. Based on these results I provide security evaluations for the conceptual designs and implications for the design of authentication ceremonies in general.

4.1 Pilot Study

The pilot study comprised several single person sessions, which were used to check the applicability of the procedure, the participants' understanding of the procedure, necessary equipment, recording, and timing.

In the first session I confirmed that the procedure leads to In the beginning of the design stage the participant reiterated that he was familiar with. However, an un-

structured interview about trust in everyday life led to different ideas concerning trust establishment. The feedback phase uncovered that the participant was also unsure when to draw an idea and what material he could use for that. Therefore, sufficiently available drawing materials and encouraging participants to draw is necessary.

In the second session the results from the previous unstructured interview were used to have a structured discussion about trust during the design phase. However, since the type of trust was not sufficiently specified the participant was rather confused about the task and did not suggest any new conceptual design for authentication ceremonies. In order to avoid future misunderstandings I decided to present specific scenarios regarding trust in the identity of others, supported by artifacts.

During the third session I focused on everyday trust in identity. Even though the participant had significant cryptography knowledge he suggested conceptual designs for authentication ceremonies that would invoke some level of trust (but not perfect trust) in him.

The necessary changes to the procedure which the pilot study revealed were: (1) trust in the identity of other people must be discussed explicitly with the help of scenarios and artifacts, and (2) the participants need to be encouraged often to draw their conceptual ideas and they need to have sufficient drawing material to do that.

4.2 Participants

To confirm that the target population for the study has been reached, all participants filled out a short questionnaire shown in Figure 4 on page 69 stating personal details and their knowledge about cryptography and IP networks. The results in Table 4.1 show that with few exceptions most participants had no or not much knowledge about those topics.

I created a website that described the purpose of the study and the participation requirements of minimal knowledge of cryptography and networking. This website was advertised by email, Facebook, and Twitter. Snowball sampling was used to quickly find potentially interested and qualified participants. Interested people who did not fulfill the participant requirements were very helpful in finding potential participants.

All ten participants had either completed a third-level education or were currently attending a third-level education institution. Four participants had some kind of training related to computer science. Seven participants were women and three were men. The participants' aged ranged from 22 to 35, with an average age of 27. The participants were asked to give an indication on how much they knew about cryptography and IP networks on a scale from one ("very little") to six ("very much"). The median knowledge of cryptography on that scale was 2, the lower and upper quartile were at 1 and 2.5, respectively. The median knowledge about IP networks was 1, the lower and upper quartile were at 1 and 3.5, respectively.

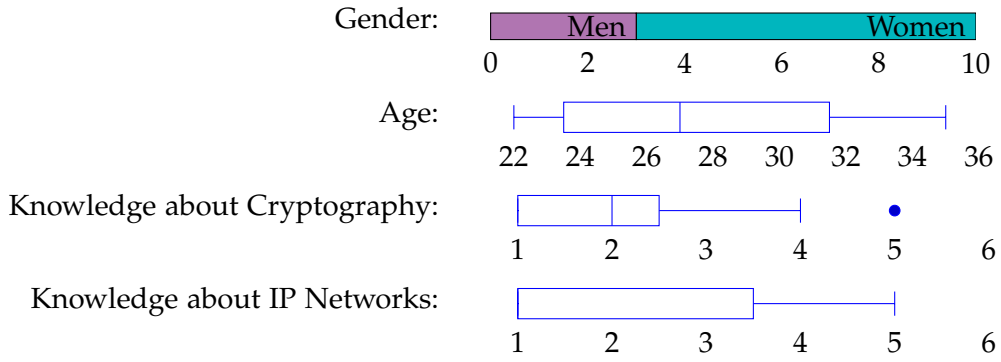


Table 4.1: Results of the Questionnaire

4.3 Experience Reports

Every workshop session started with a discussion about which secure messengers the participants used, what kind of issues they encountered, and which aspects they liked.

Some participants even talked about more general observations with secure communication. They thought it is important how many people use a messenger, since peer pressure leads to other more people using this messenger, which confirms the findings of De Luca *et al.* [DDO+16]. Others reported that they install messengers even for only one friend who uses it and or as soon as they hear about one. One participant explained that it did not matter that someone got their data (immediately going back on that in regards to health care data), but that the real problem was systematic surveillance with the help of communication data. However, as Egelman *et al.* [EJP+14] pointed out, users seem to routinely underestimate the real risks of others getting access to their personal data on their smartphones.

Participants were largely unaware of the existence of authentication ceremonies in their messengers and in cases were participants knew about them they were sometimes more confusing than helpful: *"I don't really understand how it [WhatsApp encryption] works, because it says encryption is used, but when you access the contact data, you can encrypt it again with some kind of code, so i don't get that. [...] and you have to be in the same place to do that, that's very bothersome."*

Usage of Secure Instant Messaging Applications

The most widely used messaging application where in descending order *WhatsApp*, *Telegram*, *Signal*, and *Facebook Messenger*. However, many were not aware that end to end encryption had to be manually enabled in *Telegram* and *Facebook Messenger*. Regardless of the used encryption, the participants did not trust *WhatsApp* and *Facebook Messenger* because they belonged to Facebook which is known for their privacy breaches.

Two participants reported regularly using *Skype* and SMS, which they know are insecure

but used for other reasons such as convenience, interoperability, or peer influence. The following messengers were only named by single participants who had a varying degree of confidence in their security: *Slack*, *Confide*, *Jabber*, *Dashing*, *WeChat*, and *Threema*.

Appraisal

Convenience was the most common appraisal that participants had to offer for messengers. Therefore, the wide availability of *WhatsApp*, i.e. many users use it, is important because it makes it more convenient and cheaper than SMS. Participants praised *Telegram* because messaging clients exist for all platforms, thereby achieving an availability similar to *WhatsApp*. The possibility to use *Telegram* without a cellphone number and its message persistence when changing devices were additional benefits reported by participants.

Telegram was the only messenger that received praise for its *features*, several participants independently named *Telegram*'s stickers as an important feature that they miss in other messaging applications. The aspect of creating and sharing different sticker packages seems to be the main attraction of them. The possibility of voice messages with unlimited length was also named as an important feature by one participant.

Participants rarely praised the messengers' security, but two of them liked the option for ephemeral messages in *Telegram* and *Signal*. One participant each mentioned the secret chats in *Telegram*, and that *Signal* is open source software.

Last but not least the design of messengers were praised, *WhatsApp* seems to have a "pleasant design" and *Signal* is "puristic" which both seem to imply a good choice of colors and typeface.

Disappointments

Participants reported being primarily annoyed by a *lack of convenience*, especially the diversity of inoperable messaging applications. In cases such as *Signal*, which can not be used without a phone number participants stated they would prefer using it without a phone number. Other inconveniences included lack of profile pictures in group chats, not being able to delete photos in chats, and that *Signal*'s replaces the SMS application.

Signal's lack of *quality of service* was a reoccurring theme, participants often reported deleted, missing, or duplicate messages and criticized the small video size they could send. *Telegram* participants reported missing or massively delayed notifications which severely impacted their use and one participant missed the option send files via mobile devices with *Facebook Messenger*.

Interestingly many participants *mistrusted Facebook Messenger* and *WhatsApp*, since they are both owned by Facebook. It was criticized that those apps need to many permissions, upload the list of contacts, and that their privacy policy is too permissive. Those same points were not criticized with other messaging applications which could

be explained by the amount of peer pressure to use *WhatsApp* or *Facebook Messenger* and the media coverage about Facebook's privacy issues. One participant also criticized that *Signal* received money from the US government via the Open Technology Fund which according to the participant hinted at an insecurity despite its open source code.

Usability was a minor source of disappointment, one participant mentioned that adding someone to a group-chat in *Signal* is too complicated. Several other participants mentioned that the interfaces of *WeChat*, *WhatsApp*, and *Facebook Messenger* is ugly or hard to use without specifying details.

4.4 Conceptual Designs

During the design phase of the workshop sessions I presented the scenario mentioned in Chapter 3 and asked how the participants would establish trust in the identity of their conversation partners. The participants proposed numerous ways of authentication in electronic communication and they explained and drew 20 of those concept designs. We categorized the concepts into six approaches to establishing trust: (1) comparing shared knowledge that only the conversation partners have, (2) showing pictures or videos of conversation partners, (3) asking friends or trusted contacts if they have authenticated the conversation partner, (4) trusting institutions to correctly authenticate people, (5) building up trust in the identity of the conversation partner over long periods of time, and (6) using technological measurements to test if the conversation could currently be under attack. 75% of the suggestions were from the first three categories, which suggests that those are the more intuitive approaches to trust establishment. The other three categories of trust establishment were not as popular and only had one or two suggestions each.

Participants assumed a high base level of trust in the identity of the communication partners and the security of their communication. This seems to be a prevalent strategy as long as this base trust does not result in bad consequences. According to them their attitudes would only change after a security incident.

Security features that need increase the effort for communication partners would lead to rejection by the users, since the effort for extra security seems unwarranted for most conversations. This attitude confirms the Egelman *et al.*'s result [EJP+14] that users underestimate the value of the information on their phones. However, participants reported that a manual way of verification provides them an increased sense of security which they value in sensitive situations that matter to them.

Shared Knowledge based Trust

Nine out of ten participants suggested a method based on shared knowledge immediately after I confronted them with a scenario in which they needed to authenticate the identity of their communication partner. The three most common concepts were: (1) agreeing on code-words which are used to obfuscate the conversation, (2) exchanging a

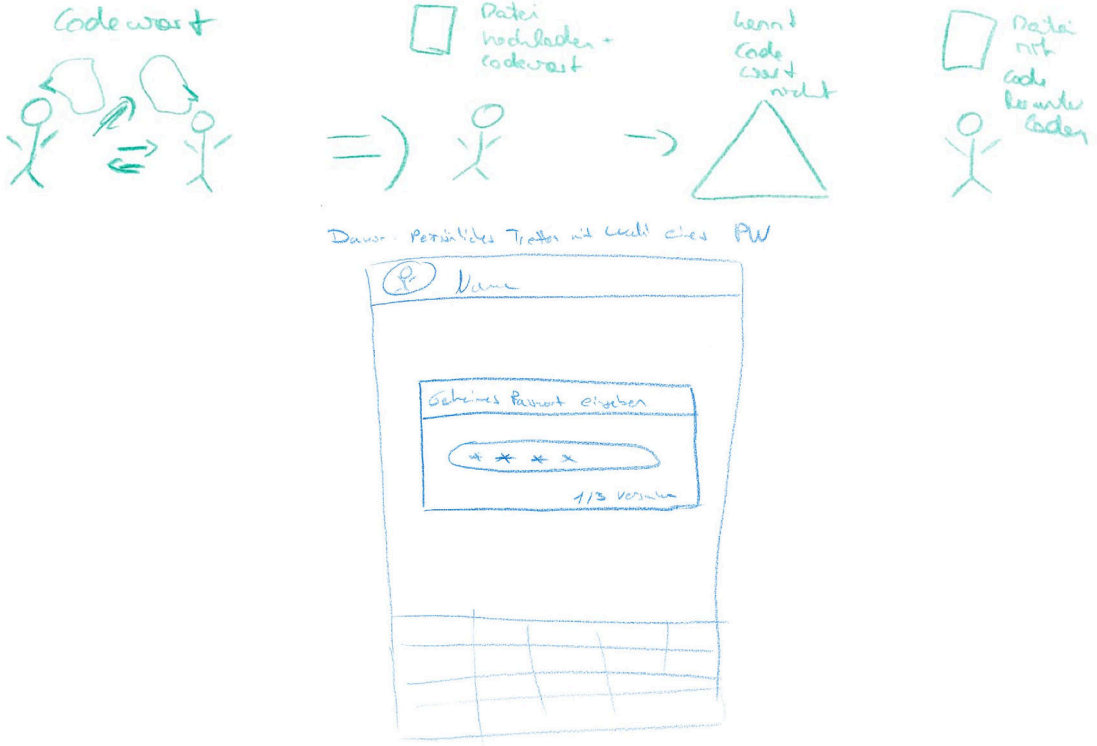


Figure 4.1: Encrypt a document or the entire conversation with an agreed upon password.

password used for accessing the conversation, and (3) asking personal questions that only the other one could answer. They reported high confidence in the results of this method, since they assume that only their conversation partner knows the agreed upon code-words or can answer the personal questions.

Communicating secret plans with code-words is a fairly common TV trope (*Spy Speak* [TV 18]), simple to understand, and straight forward to use (albeit rather annoying) in practice. The participants described the process as following: (1) meet in person to agree on a set of code-words; (2) use those for the electronic communication; (3) end conversation if communicating partners do not use correct codes or none at all. This method provides authentication (if each person has a different set of codes) and a weak form of confidentiality.

A similar version of this procedure shown in Figure 4.1 uses a password instead of a set of code-words. The participants described it as agreeing on a common password, either by meeting in person or using a communication medium which is assumed to be secure (such as mail letters). Then encrypting the entire conversation and only people with knowledge of this password can participate. This method provides authentication and confidentiality.

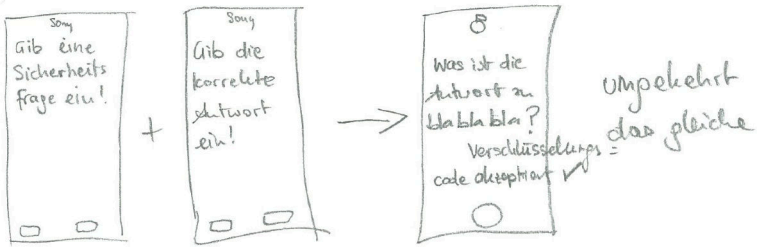


Figure 4.2: Ask personal questions that only the communication partner can answer.

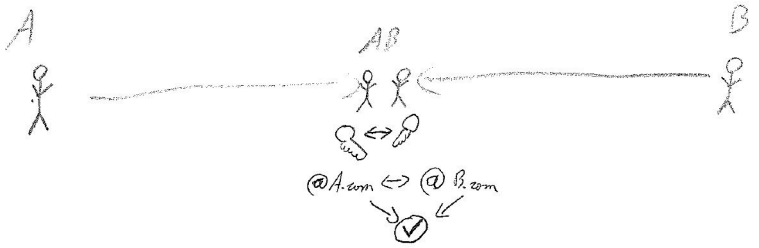


Figure 4.3: Meet in person and check nickname-key binding.

Asking personal questions was the third common suggestions using the shared knowledge approach to establish trust in others identity. The participants assume that they are talking to the right person if the other person can answer one or more questions about themselves correctly. One participant drew an interface shown in Figure 4.2) which would support this procedure. This procedure has several disadvantages: (1) the communication partners have to know each other quite well; (2) the communication partners have to think of questions that only the conversation partner can answer correctly; (3) in many cases this only provides a weak form of authentication since many common questions can be answered with (semi-)public information.

Two of the participants also mentioned that they would verify the identity of the communication partner by meeting in person and comparing the used encryption keys as shown in Figure 4.3. According to them this idea stems from prior knowledge about authentication ceremonies in existing messaging application, such as *Signal* and *WhatsApp*. One of the participants mentioned that comparing the keys is bothersome when not done automatically by QR Code and that he would like to compare something simpler, like a username.

Picture based Trust

Six participants mentioned the possibility of authenticating conversation partners by showing each other pictures of one another and three of them provided a conceptual design for this method. Participants who suggested this were quite confident that they were talking to the right person afterwards, since they commonly know the face of their regular communication partners. However, most of them realized that someone could

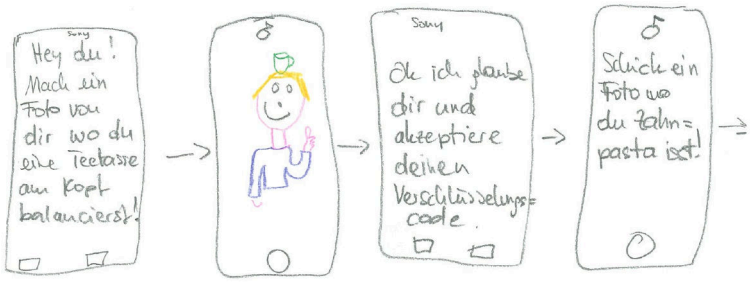


Figure 4.4: Send a series of pictures showing the conversation partners fulfilling each others requests in order to prove that their pictures are recent.

send them old pictures or that the pictures could be spoofed by attackers. The resulting trust would increase if senders could prove that the picture is recent or if real-time communication, i.e. a video-chat, is used.

Therefore, one participant described the procedure shown in Figure 4.4 as a series of picture requests of the conversation partner. In this case the first requested picture was that the chat partner should put a tea cup on their head, the second request was that the other should eat toothpaste in the picture. This approach is known as a "proof of life" in hostage situations, known from pop-culture movies where hostages have to hold up a recent newspaper or a well-known TV station runs in the background in order to confirm the age of the picture or video. This approach has two main advantages: (1) remote authentication is possible, and (2) it invokes a high level of trust in the identity of chat partners. Possible disadvantages on the other hand are: (1) chat partners have to have seen each others faces before, otherwise they will not recognize each other; (2) chat partners have to come up with action requests that do not require extra resources; (3) in case of a MitM attack all images and videos could be manipulated by the monster.

This procedure can be used for verification that both parties use the same correct encryption keys since information about them can be embedded in the pictures. If the kind of action requests are based on the used key material, then confirming that the correct action is visible in the pictures also confirms that the correct key material is used for encryption. Conversation partners could also be requested to say SAS phrases in addition to gestures or movements if a video stream is used for the procedure. Possible MitM attackers can manipulate transmissions, so real-time communication is preferred for this procedure since it makes high quality image manipulation more difficult.

Social Trust

All participants mentioned that in everyday life they receive information about identities and trust from their social contacts, but most were unsure if and how that process could be translated to electronic communication in a meaningful way. For instance, in everyday life, they might ask their trusted friends explicitly if they know a specific person and if that person can be trusted, alternatively they might only trust people who

were introduced to them by trusted friends. The three participants who provided a conceptual design for social authentication wanted the messaging client to automatically establish which of their contacts is trusted by one or more friends. However, opinions were divided on how this form of trust should be represented in the user interface, some wanted to have color-coded categories, others suggested to show the number of friends that verified a contact, and some wanted to see names of the friends that verified a contact. The participants reported that the resulting trust from social authentication would be medium to low, suggesting that social authentication can only be a part of a more extensive authentication concept and that the effect depends on which friend actually verified a contact.

One of the strategies was the explicit transfer of trust from already trusted friends. So establishing trust with only a few of their contacts would suffice to create a large trust network. Transferring trust works either by requesting it manually i.e. *"Did you establish trust with Hannah?"* or automatically, so that the interface shows either the names of the friends or just the number of friends that established trust with the contact in question. Broadcasting a loss of trust to the list of contacts works similarly *"I checked Hannah's identity and it was compromised - do not trust her"*. The resulting trust network can be visualised in different ways, one participants drew an interface shown in Figure 4.5 where different levels of trust are colored similar to a traffic light: green for trusted contacts, yellow for second order trust (i.e. a trusted contact confirmed the identity), and red for unknown trust. PGP based mail encryption uses a similar method to establish trust, however the problems of the encryption model and the usability are so overwhelming that the trust model can not be reliably evaluated on its own. Additionally, mail communication makes it harder to build trust networks, since the list of mail contacts is usually quite large and impersonal compared to a personal phone-book.

According to the participants, implicit trust transfer in everyday life works by observing interactions and behaviour of other people. Although this leads to usable results in everyday life it is very hard to transfer this mechanism to electronic communication. Therefore, it is not surprising that none of the participants drew a conceptual design for this.

An automatic approach to establishing trust in social networks would be beneficial of two distinct reasons: (1) Participants mentioned that they find additional verification procedures bothersome and that they would not do them for everyone of their contacts. Therefore, an automatic approach provides a base line of trust that does not require any user action; and (2) Participants stated that they feel uncomfortable voicing their distrust to the person in question, since that could be perceived as rude and unfriendly. Using third parties to establish trust removes this social awkwardness. In order to avoid voicing distrust, authentication ceremonies need to have a unrelated side-effect so that people do not have to utter their wish for authentication directly.

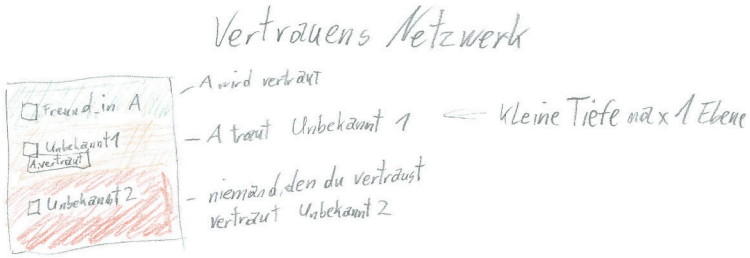


Figure 4.5: Representation of a trust network, where trusted contacts confirm the identity of other contacts.

Institutional Trust

Participants discussed two scenarios: verifying the identity of an unknown bank employee or someone fetching a package. Many stated that they would ask for an ID card or check the name-tag of the employee. This form of trust is based on the issuing organization, if a bank or a government confirms the identity of someone, many would be convinced by that. However, since none of the participants saw how they could translate this form of trust establishment to electronic communication nobody described a process for it.

One participant's first idea to the problem of authentication was a fingerprint check. People now unlock their phones with a fingerprint and even buy groceries or apps with a fingerprint, so it should be good enough to authenticate people in conversations. The process shown in Figure 4.6 locks conversations by default, and they can be accessed by scanning a fingerprint. Participants trust this procedure because they know that everyone has to do confirm their identity before being able to access the conversation. This is a form of institutional trust since the users have to trust the application to verify everyone's fingerprints before they can participate in the conversation. However, since the fingerprint reader in a mobile phone actually verifies the identity of a user to the mobile phone and nobody else, it is not possible to remotely verify fingerprints. Additionally, users might feel uncomfortable when they get the impression that their fingerprint leaves their device.

Similar to quality control in the food industry participants suggested certification marks could be used to communicate to end-users that a known institution has checked the security of applications. This would be quite useful because users rarely have an idea which product is secure and have to rely on external input when deciding which one to use. However, this approach does not lead to a technical solution for establishing trust between communication partners.

Participants mentioned that trust in banking employees was established in major part through the location of their conversations. A person in a bank that says that he is a banker is much more trustworthy than someone ringing your door bell offering you banking services. However, this location based institutional trust does not really

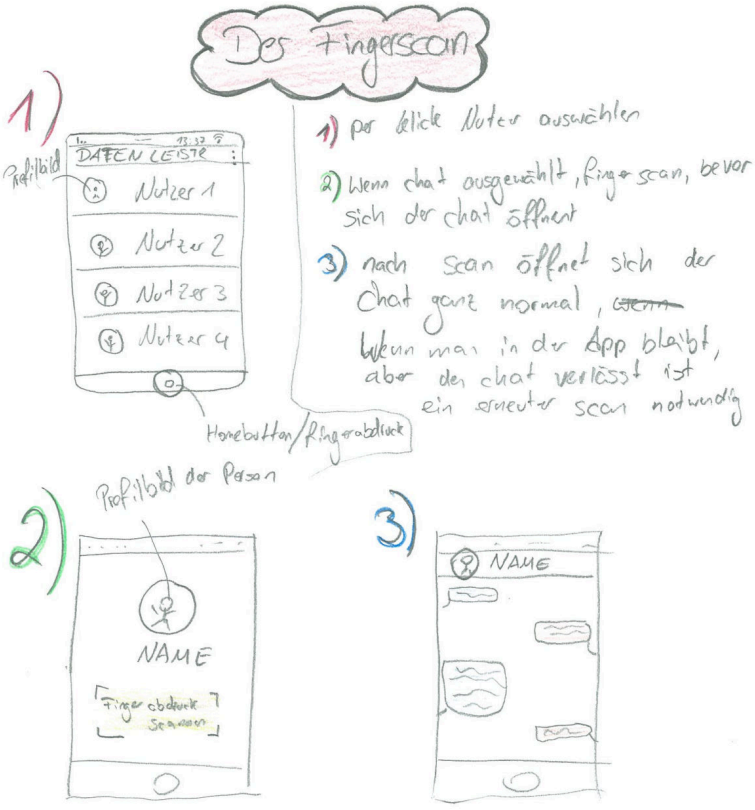


Figure 4.6: All users access their secured conversation with a fingerprint confirming their identity.

lend itself to mobile messaging application, since conversations are independent of the physical location of the participants.

Habituation of Trust

Offline relationships with neighbours, colleagues, and even bank employees indicate that some kind of trust can be built up over time. Almost all participants said that this is not a fool proof way of establishing trust, but that they nonetheless depended on this method in some ways. Participants usually agreed that this method could be useful in electronic communication as well. They said that over time more information can be collected that can be matched to information known from other sources which validates their identity. Measuring this trust involves either counting the number of messages between conversation partners or measuring the time since the last key change. This method could in fact warrant some form of cryptographic trust, since targeted attacks would have to be kept alive for long periods of time in order to habituate trust, which is possible but rather unlikely.

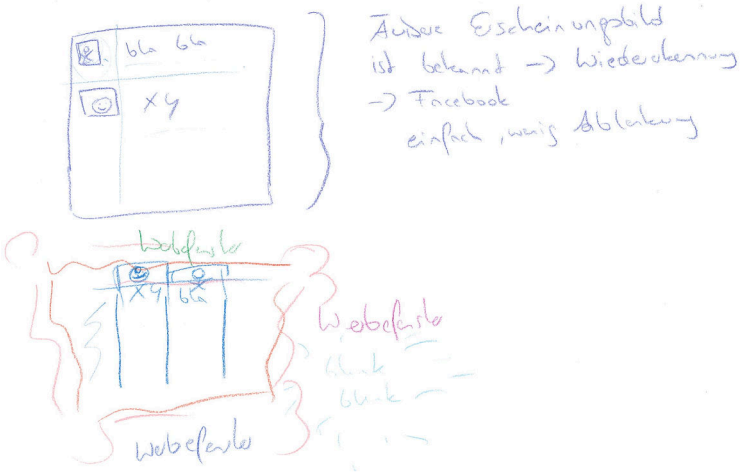


Figure 4.7: Different user interface designs for missing trust (bright colors, ads, and different font-sizes) and established trust (sleek, uniform, and soft colors).

Not only the time or the amount of messages, but also the content and style of the messages were important to participants. They thought it might be feasible to have the messages analyzed automatically. A text analysis algorithm could check if the style and content of the messages is consistent and show warning messages that it might be a different person or a person under duress if the writing style is not consistent. Most participants who mentioned this method also found it creepy, invasive, and paternalistic and would not want to use it even if it were available.

Since it is difficult to draw the passing of time and the process of habituation the workshop participants focused on how different trust levels should be represented in the user interface One idea, shown in Figure 4.7, was that the interface itself should change from looking like a scam website to a serious chat interface. This could be achieved initially using bright colors, showing ads, using non-uniform font-sizes and improving it one by one according to the increased trust status. However, apart from the issue of privacy-invading ads, a messaging application that initially looks like a scam website will probably not be very successful, since a low quality of service is an obstacle to adoption [ASB+17].

Testing based Trust

As mentioned at the beginning of this section, even participants who are reluctant to conduct authentication ceremonies with every of their contacts still want to be able to verify the communication security in cases of sensitive topics. Testing based trust establishment reflects this belief and offers different ways to test the communication channel for eavesdroppers.

In case of a known attacker two participants independently suggested insulting the

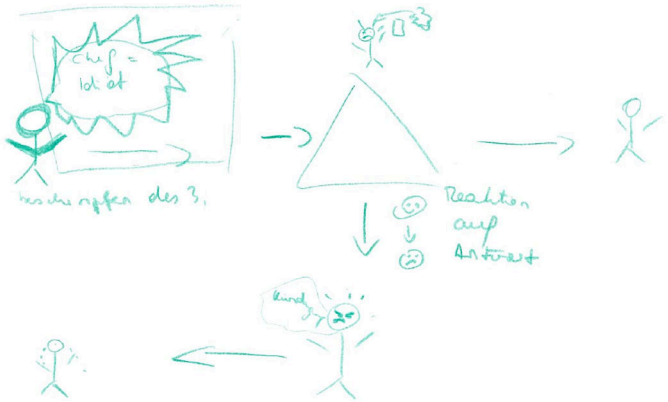


Figure 4.8: Insulting possibly known MitM in order to get some reaction.

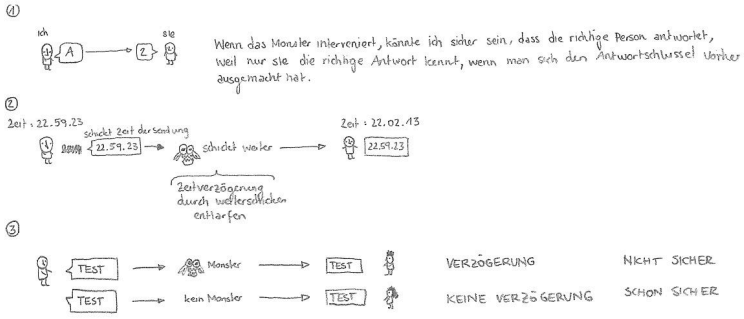


Figure 4.9: Comparing the latency of two different conversations to see if one of them contains a MitM relaying all messages.

MitM as shown in Figure 4.8 hoping to get a reaction from the attacker. Both participants were unsure how much trust this procedure warranted, but they probably feel revenged in case the known MitM is actually eavesdropping.

Other approaches to testing based trust were mostly technology based, one included meeting up and comparing the received messages in order to reveal if any manipulations took place, and another checked the quality of transmission as shown in Figure 4.9. The participant in question thought that if one conversation has a higher transmission latency than another conversation, there would be a chance that a MitM is relaying all messages. The latter approach could be implemented rather easily, but it probably would not be a good way to detect a MitM since latency varies a good deal in mobile environments.

4.5 Implications for Design

In order to give security engineers support in designing authentication ceremonies, I not only present the participants' conceptual designs but also derive some fundamental implications for design. These were derived by generalizing design hints from the notes and the conceptual designs into theories.

1. **Successful authentication ceremonies must lead to a higher perceived level of security.**
   Invoking a sense of trust is only possible if users have some (not necessarily technical) understanding of the effects of an authentication ceremony. Additionally, the messaging client needs to represent this gained trust in the user interface so that (1) users will be continuously reminded of the gained trust and (2) that users receive clear feedback for their actions. Users have no reason to conduct an authentication ceremony if the implications are not communicated to them well enough, e.g., via visible changes in the user interface.

2. **Use well-known security concepts known from everyday life.**
   In order to understand the effects of an authentication ceremony and possibly explaining its necessity to a conversation partner, users need to understand the underlying security concept. There are in general two ways to achieve that: either explain the underlying security concept in lay-people's terms to all users or use an established security concept that users know from everyday life. *"The key takeaway from mental models research is that non-experts do not understand abstract security properties. They can only understand why a property matters in the context of a specific threat model that matters to them."* [ASB+17]. Key fingerprint verification is an abstract security concept that most users do not know from everyday life, and Tan *et al.* [TBB+17] found that users who are tasked with fingerprint verification have unacceptable failure rates. Therefore, I recommend to refrain from using manual verification of key fingerprints.

3. **Friction-less initiation of authentication ceremonies.**
   Initiating an authentication ceremony should be quick and easy, not only should the user interface make it easy for users, but the social process of starting an authentication ceremony should be as easy. In cases were the authentication ceremony needs cooperation of the conversation partners, the need for this ceremony and its requirements must be negotiated. Compelling answers are needed if one of the conversation partners asks why an authentication ceremony is necessary, otherwise the process will never be started. Since it could be socially awkward to voice distrust in the identity of conversation partners, a necessary negotiation could result in an additional barrier to the initiation of the ceremony and should therefore be avoided. A (possibly fun) side-effect of the authentication ceremony could reduce the awkwardness and also provide a compelling reason to conduct the ceremony.

4. **Users work with different levels of trust.**
   Similar to everyday life, users do not think that all of their conversations are equally important or trust everyone in their contact list equally. We argue that a messaging client should reflect these different levels of trust and importance. Users have a base level of trust for their contacts and expect a base level of security for all of their contacts, but for important people or conversations they prefer an additional layer of security that enables them to manually confirm that the conversation is secure. This leads to the conclusion that automatic authentication should be used wherever feasible but that manual options would provide an additional benefit to users.. The study by Ruoti *et al.* [RKB+13] provided evidence that manual encryption invokes an increased sense of security even if it is less convenient than automatic encryption.

5. **The user interface must represent actual level of trust.**
   Since users work with different levels of trust, and an authentication ceremony leads to a different level of security, this should be reflected in the user interface. This can be achieved by representing different trust levels in different colors, or providing a different look and feel according to the level of trust. Our results suggests that pictures of faces invoke a sense of trust, so those should only be shown if the conversations has been secured appropriately. Additionally, our data suggests that a high quality of service is associated with trustworthiness, this could be used to express trust in a more intuitive way by decreasing quality of service for conversations with unauthenticated contacts. However, as Abu-Salma *et al.* [ASB+17] have shown, low quality of service is an obstacle to adoption – so this method of communicating trust should be used with care.

CHAPTER 5

Example Ceremonies

Contrary to traditional approaches to designing security features, VSD focuses on the users' values. One of this thesis' aims is demonstrating how a user centered design process, such as participatory design, improves the resulting authentication ceremonies. This involves using the conceptual designs from the participatory design workshops as a starting point, designing the procedure, the social interactions, and the user interface before thinking about how to secure the authentication ceremony. To illustrate this approach I present three example ceremonies based on the participants' conceptual designs, their security evaluation, and the implications for design in this chapter.

5.1 Combination Lock

A combination lock is a security mechanism users know from either bicycle locks or lockers in school. It is used to deny access to information or objects to anyone without knowledge of the correct combination. This example ceremony transfers this concept to electronic conversation, everyone who knows the combination can participate in the conversation and all others are locked out.

The advantages of this approach are: (1) users who are familiar with combination locks intuitively understand that knowledge of the combination regulates access; and (2) the authentication ceremony can either be conducted in person or remotely using a secure channel. The disadvantages are that users may get the impression (1) that conversations, which they did not protect with a combination, are insecure - which is not true; and (2) that attackers can break its security like a traditional combination lock.

Conceptual Design

The combination lock approach is based on the suggestion to require a password to access conversations as shown in Figure 4.1 of the previous chapter.

The participant's conceptual design shows the conversation partners meeting in person to exchange a password, which they need to enter upon opening their conversation. This concept uses shared knowledge to establish trust, which participants suggested most often during the workshops. The participant reported high levels of trust in the identities of the conversation partners after this procedure, since all of them need to prove their knowledge before being able to access the messages.

Even though password protection is a well-known security concept, this ceremony uses a combination lock for several reasons: (1) it is almost equally well-known; (2) since combination locks usually only use three or four digit codes, users do not expect long shared secrets; (3) combination locks always show a combination, so suggesting random combinations does not break the metaphor; (4) specific character sets can be prescribed without annoying the users; and (5) it is less error-prone to communicate short sequences of recognizable icons than long passwords consisting of several different character sets.

The result of this operation is a higher level of perceived security since the mental model of a combination lock suggests an increased security. However, the user interface must also give continuous visual feedback about the increased level of security.

Procedure

At least one conversation participant has to understand that the current level of security has to be improved and initiate the ceremony. Therefore, the user interface shown in the first part of Figure 5.1 provides several hints on unauthenticated conversations. Profile pictures are not shown since they may convey a sense of trust, an open red combination lock is visible in the top bar, and a warning message explains possible consequences of keeping the conversation unlocked.

Conversation partners can motivate each other to authenticate by asking *"We should lock the conversation against eavesdroppers."* or *"Show me your symbols!"* if the procedure is already known to everyone involved. However, since every user can start the locking procedure on their own, cooperation is not necessary to initiate the ceremony. Even if the conversation partner does not see the need for authentication, the ceremony needs to be completed once it has been started, otherwise communication is not possible.

Users initiate the locking procedure by tapping the lock icon, which leads to a locking dialog as shown in the second part of Figure 5.1. It suggests a random combination of four symbols and gives the option to remember this combination, since most users will not want to enter it every time they open the conversation. Contrary to traditional digit-based combination locks, a set of recognizable symbols is used which results in a higher amount of possible combinations and makes the communication of the

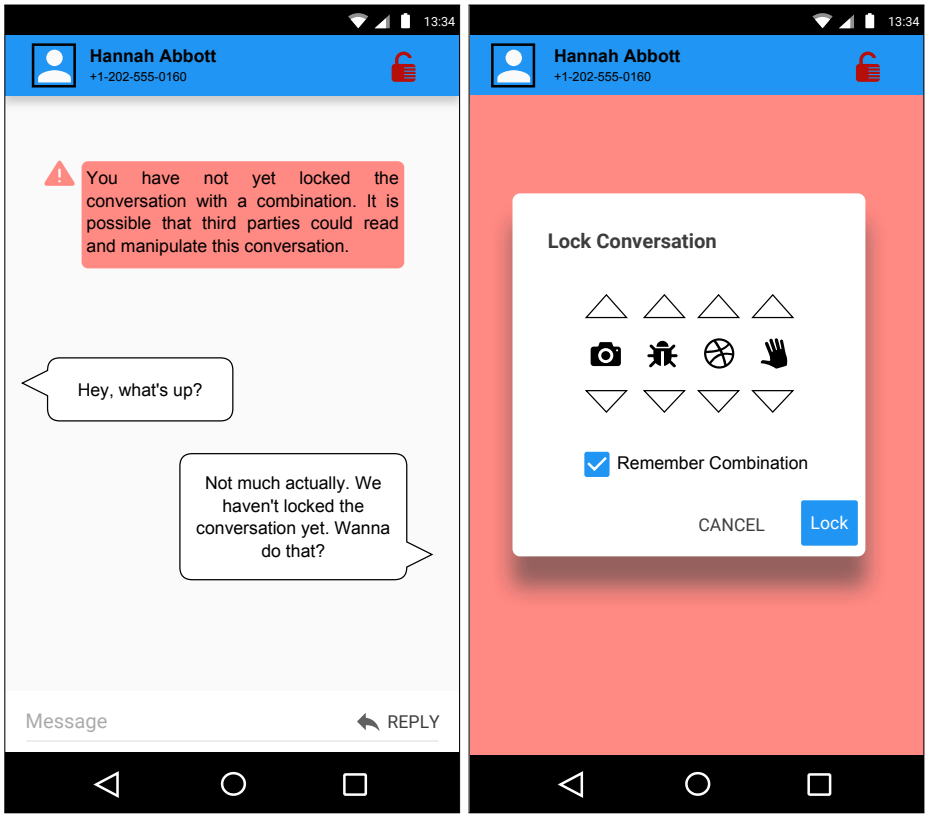


Figure 5.1: Locking an unauthenticated conversation.

combination by voice or text less error-prone, since users have to describe each symbol. After completing this locking procedure, the conversation partner will immediately be locked out of the conversation and prompted for the correct combination.

After a conversation has been locked in this way, the participants lose access to the conversation until they have set the correct combination on their device as shown in Figure 5.2. The conversation partners are forced to share the secret combination using a different channel. In order to provide the highest level of security the authentication ceremony needs to be conducted using an in-person meeting, during which the partner who locked the conversation shows the other one which combination they used. In cases of remote authentication either the locked-out partner will ask for the correct combination or the locking partner will proactively send the correct combination. A different channel for this communication has to be chosen, since the locked conversation is not usable until both partners know the correct combination. The procedure design makes it hard for users to transmit their combination using the unauthenticated conversation because they do not know the suggested combination before opening the locking dialog and the locked conversation can not be used for this purpose.

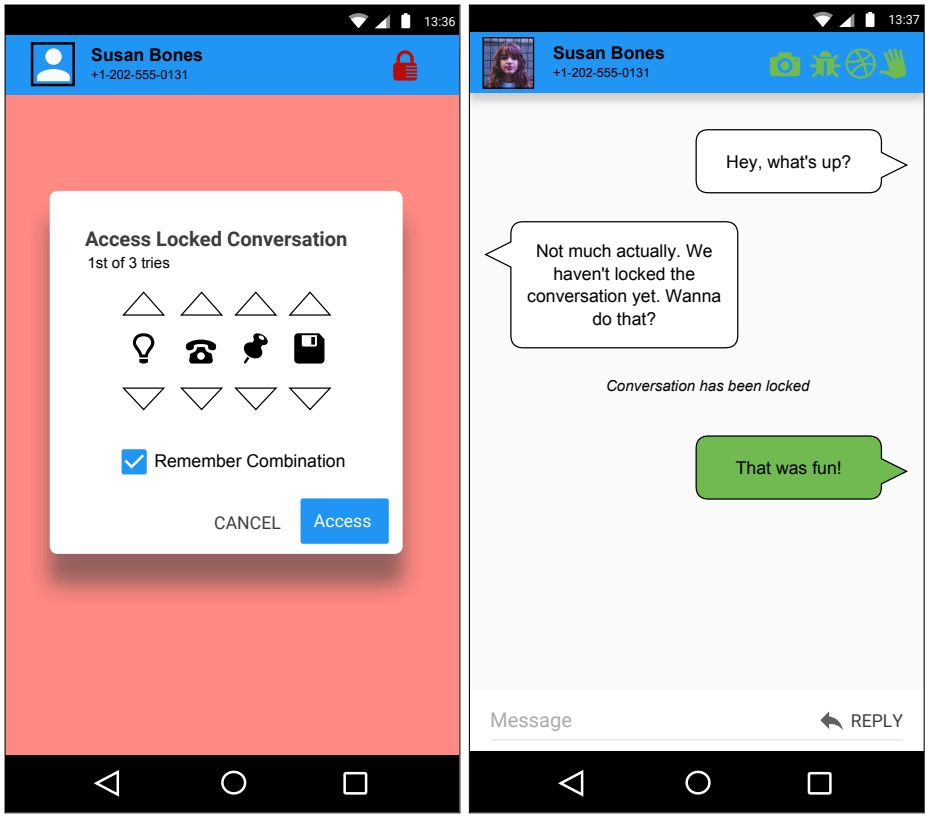


Figure 5.2: Accessing a locked conversation.

Consequences

A successful authentication ceremony will convince the users that only people with knowledge of the combination can access the conversation and the user interface will express this new trust status. In the second part of Figure 5.2 a conversation with a verified trust status is shown: the profile picture is now visible, the combination lock is now closed and green, the warning message has vanished, and all messages sent in a locked conversation have a green background.

If conversation partners try to access the conversation three times unsuccessfully, the conversation will stay locked in order to stop users from communicating over potentially insecure channel. Since the authentication could simply have failed because of miscommunication a fallback mechanism is provided to unlock such conversations again. The fallback mechanism requires the users to meet in person and verify the used encryption keys automatically.

Security

The security of this ceremony depends on a short common secret that has been exchanged over a secure channel. After locking the conversation users only have three tries to enter the correct combination before access is denied permanently. Those security properties are provided by using the SMP which Alexander and Goldberg [AG07] have previously suggested for OTR messaging.

It builds on the Socialist Millionaire's Problem, which concerns itself with millionaires wanting to check if their wealth is equal without actually revealing how rich they are. Boudot, Schoenmakers, and Traoré [BST01] solved this problem in 2001 using the Decision Diffie-Hellman assumption and Alexander and Goldberg [AG07] adapted their solution for use in OTR. Instead of the conversation partners wealth, they compare a hash of the session ID, the key fingerprints, and a shared secret. MitM attacks on this protocol fail because the attacker will either (1) block SMP messages, which will make the protocol fail, (2) forward the SMP messages, which will make the protocol fail because the session ID and the key fingerprints will not be equal, or (3) spoof the SMP messages, which will make the protocol fail because the shared secret is unknown to the attacker.

The two main security advantages of this approach are that (1) it does not depend on the entropy of the shared secret, so its entropy can be very low, and (2) restrictions on the amount of tries are possible which stops brute-force attacks. The shared secret should be chosen using a secure channel, but even shared secrets that were exchanged over insecure channels can provide some level of security.

5.2 Mimic-ID

Identity checks are a common occurrence in the offline world, if anyone has doubts about someone's identity they can ask for a government issued ID, and match the given name and picture on the card to the person in front of them. This works because government issued ID cards are well-known and trusted. Since users are already familiar with those kind of ID checks, it can be used in electronic communication to establish trust in the identity of conversation partners.

The advantages of this approach are that: (1) users are usually already familiar with the process; (2) an electronic ID check can be combined with a key verification, thereby matching perceived and actual level of security; and (3) the process does not require knowledge about encryption, so all references to it can be eliminated from the user interface. The disadvantages are that: (1) contrary to the usually informal conversations, checking ID cards is a formal process; (2) users may feel anxious to show distrust in the identity of their conversation partners and consequently refrain from initiating the authentication ceremony; and (3) checking the identity of conversation partners may reduce the social acceptability of anonymity.

Conceptual Design

During the workshops two scenarios concerning trust in identities from the offline world were discussed (as described in Section 3.4). In the first one the identity of a package recipient was unclear and in the second one the identity of a new bank advisor was in question. In both cases several participants mentioned that they would establish trust in the identity of them by checking some form of identification, like a bank-issued name tag or a government-issued ID card.

However, the workshop participants did not transfer those suggestions to conceptual designs because they were unsure how that process could work using electronic communication. Other forms of institutional trust (see also Section 4.4) were also discussed and one resulting concept design used fingerprints for authentication as shown in Figure 4.6.

Since the process of ID card checking is well-known and understood by the users, it seems advantageous to translate this process to electronic communication. This adapted process includes asking for an ID card and verifying that information on the ID card. It is assumed that both offline and electronic versions invoke a similar amount of trust in the identity of the conversation partners, since the verification procedure is similar for both.

Procedure

Conversations with unauthenticated contacts as shown in the first part of Figure 5.3 will provide hints about its trust status. A warning message is shown that explains that possibility of identity theft and eavesdroppers and how users can mitigate those attacks. Unverified personal information about the conversation partner is either not shown at all (in case of the profile picture and the status message) or in an unobtrusive way (such as the name taken from the phone's contact list). Additionally, the ID card icon on the top signifies a necessary action involving an ID card and the color scheme avoids green colors, so as not to suggest security.

Security concerned users might initiate the authentication ceremony by asking the conversation partner to meet in person after reading the warning message and concluding that the ID cards have to be verified to protect against eavesdroppers. However, since in many cases users might not care about eavesdroppers and showing distrust in the others identity could be socially inappropriate, this solution provides other reasons for conducting the ceremony as well.

Conversations with unauthenticated contacts do not show a profile picture or the current user-set status message at all, and the name of the contact is only visible in small type underneath the telephone number. Users might want to have those non-essential features and ask the conversation partner for them, instead of awkwardly asking for identity verification. Therefore, conversation partners could motivate each

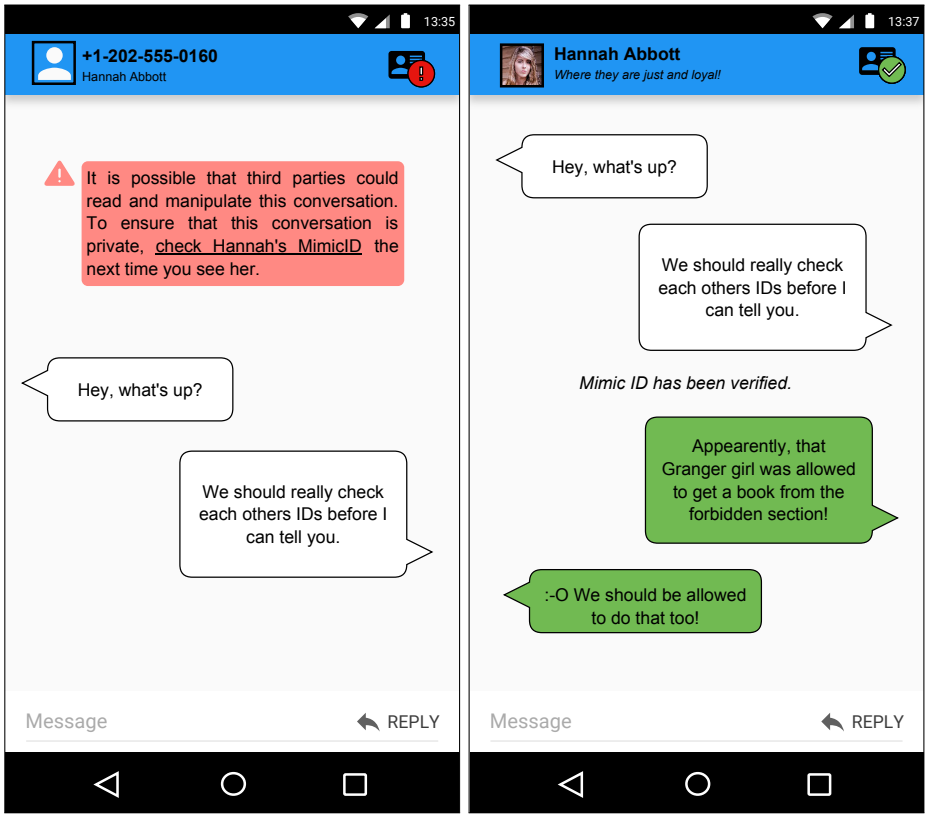


Figure 5.3: User Interface of a conversation before and after verifying the Mimic-ID

other to conduct the authentication ceremony by asking: *"You're one of the few in my contact list without profile picture, can I quickly check your ID?"*.

In order to conduct this authentication ceremony the conversation partners have to meet in person. Frequent conversation partners will often meet in their day-to-day lives anyway, but the assumption that in-person meetings are possible does not hold for everyone. An alternative authentication strategy is necessary for those cases.

When meeting each other conversation partners initiate the ceremony by opening the Mimic-ID / Verification interface (as seen in Figure 5.4) in their messaging app. This is accessible by tapping the ID card icon in the conversation, but other shortcuts, for example in the corresponding warning message exist as well. The Mimic-ID contains all the profile-information that is missing from unauthenticated conversations, so a profile picture, a current status message, the name, and the telephone number. Each conversation partner matches the profile information to the person in front of them and afterwards scans the QR Code in order to mark the contact as authenticated and transfer the profile information to their own device.

Figure 5.4: The Mimic-ID contains profile information which will be transferred between the conversation partners by scanning the QR Code

Consequences

After a successful authentication ceremony the user interface of the corresponding conversations changes as shown in the second part of Figure 5.3. The previously invisible or unobtrusive profile information is now shown prominently. The ID card icon on the top bar now shows a green check mark, the warning message has disappeared, and all messages that are sent after a successful authentication ceremony have a green background.

Since key verification with a QR Code is not error-prone a failed authentication ceremony indicates a high probability of an attack and has drastic consequences. The color scheme of the conversation changes to red and yellow, a warning message appears after each message, and the quality of service is decreased. Only texting is allowed, all text messages are limited in length, and messages can only be sent every three minutes.

If authenticated conversation partners change phone the identity key will change. Since this is also happens during MitM attacks, the authentication status of the conversation has to be reset, thereby removing the previously visible profile information. Additionally, a warning message describes what happened and advises users to take action: *"Your contact renewed their Mimic-ID, to ensure you are secure from eavesdroppers check it the next time you see them"*

Security

The security of this example ceremony is based on key verification, which works by comparing the encryption keys of the conversation partners over a secure channel. In this case they meet in person and the compare keys automatically as suggested strongly

by Tan *et al.* [TBB+17] in their study about fingerprint representations and modes of comparison.

Several messengers use an authentication ceremony based on (automatic) key verification, amongst others *Signal*, *WhatsApp*, *Facebook Messenger*, and *Telegram*. Usability studies about those authentication ceremonies [VWN+17; SHWR16; HL16] showed low success rates between 14% and 25%. However, those rates improved as soon as the purpose of authentication ceremonies and how they are conducted was explained to the participants. The security properties are therefore equal to those for traditional key verification. The main security benefit of this approach is the user-friendly metaphor that potentially leads to higher verification rates.

5.3 Verification Dance

Many workshop participants mentioned that they would trust the identity of their communication partner when they see a recent photo or video of them. Therefore, pictures or a video chat could be used to establish trust. Information about the used key can be transferred using pictures or videos by asking the users to make key-dependent gestures or say key-dependent phrases. Since a picture of a person carries some trust, the user interface does not show a profile picture before the authentication ceremony has been conducted. As an added benefit profile pictures are chosen from the gestures.

The advantages of this approach are: (1) users have a high trust in the identity of the chat partner; (2) the procedure can be conducted remotely; and (3) could result in interesting and fun profile pictures. However, using pictures as a form of authentication also has several disadvantages: (1) communication partners have to know each others faces in order to verify their identity; (2) users who do not like to be photographed will likely reject the procedure; (3) since the video chat transmitted over an insecure channel the stream could be manipulated; and (4) if chat partners happen to stand next to each other a picture verification might seem weird.

Conceptual Design

Similar to all suggestions described in Section 4.4 about picture-based trust this example ceremony derives its trust from the conversation partners being able to see and recognize each other. This works with pictures and even better with audio-visual streams because recognizing the conversation partner invokes trust in the identity.

The ceremony is based on the conceptual design of a workshop participant shown in Figure 4.4 which requires conversation partners to send each other pictures of themselves. In order to prove that a picture is recent they have to fulfill tasks set by the requesting conversation partner.

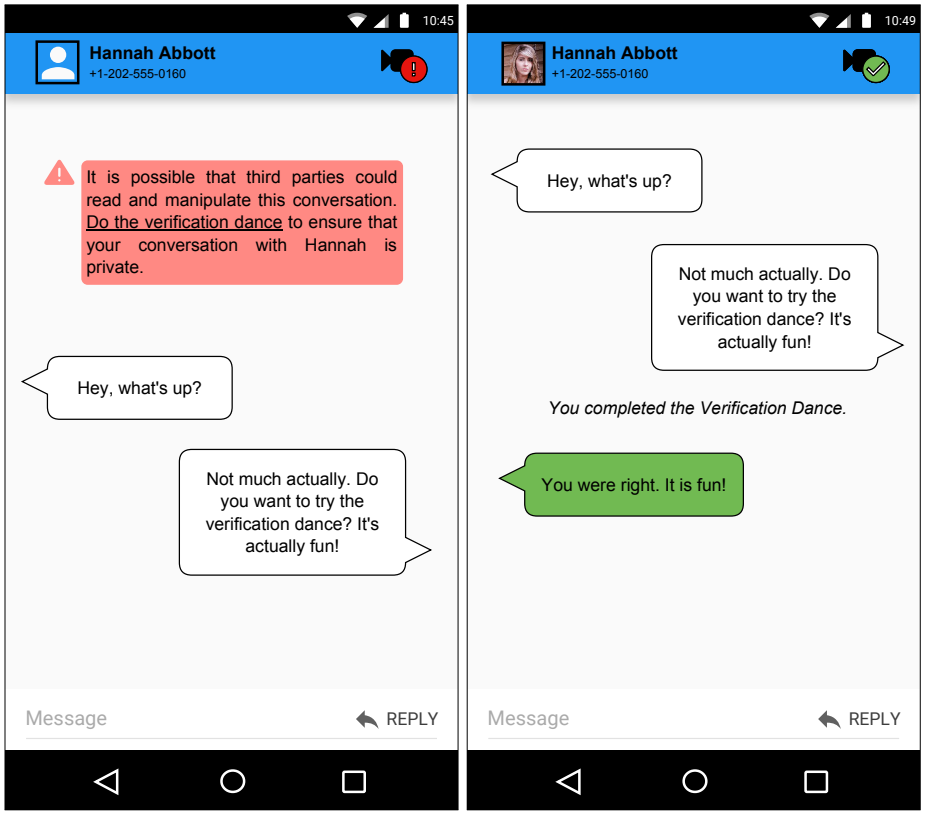


Figure 5.5: User Interface of the conversation before and after the Verification Dance

Procedure

As seen in the first part of Figure 5.5 conversations with unauthenticated contacts do not show a profile picture so as not to suggest trust. A video-camera icon with red exclamation point is shown, hinting at some necessary action and a warning message is visible that educates the users about the potential dangers that can be mitigated by doing the verification dance.

Conversation partners who already know of the risk of unauthenticated contacts or who have read the corresponding warning message might want to conduct the authentication ceremony without further motivation. If the extra security is not of concern, users could also want to have a profile picture of their contacts, which is only available after conducting the ceremony. As to limit the amount of necessary negotiation the users can either request an identity proof or proactively send an identity proof to the conversation partner. The conversation partners could also motivate each other by writing *"I don't have a profile picture of you yet, send me an identity proof!"*, *"Let's protect our conversation from eavesdroppers, I'll start by sending you an identity proof."*, or just *"Come on, let's do the verification dance!"*.

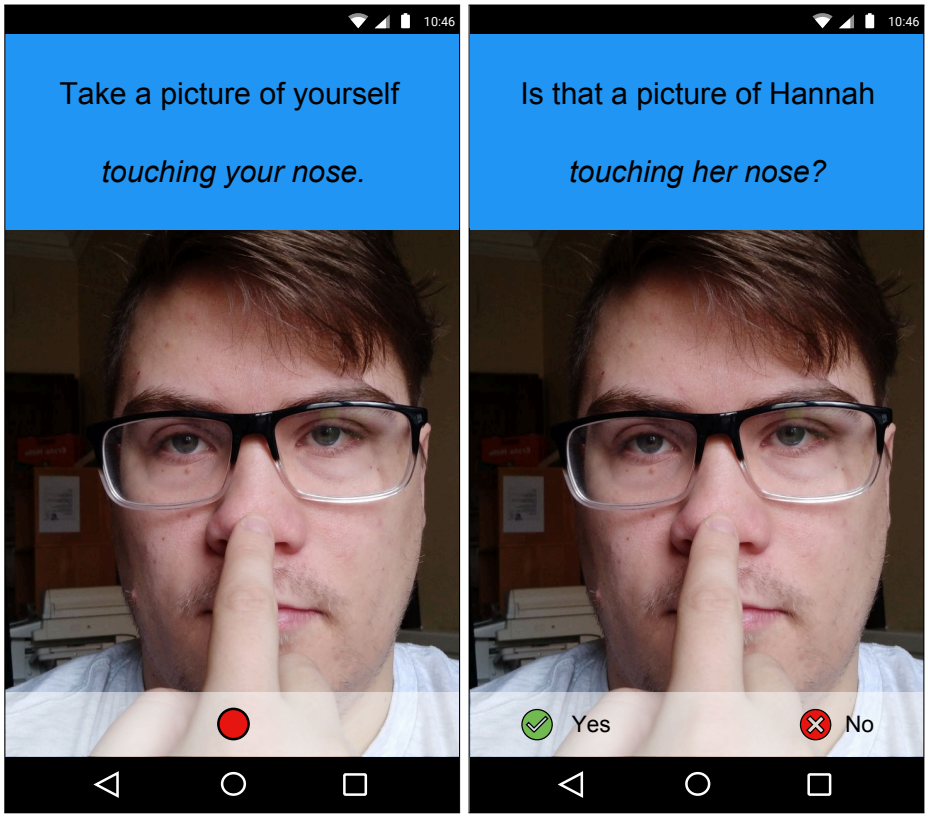


Figure 5.6: Verification Dance

Tapping the video-camera icon opens a menu offering a choice of either requesting an identity proof or sending one. Sending an identity proof leads to the interface shown in the first part of Figure 5.6 which shows a description of the required gesture and a record button. Several different gestures have to be recorded until the identity proof is complete. The interface of the receiving conversation partner is shown in the second part of Figure 5.6 which displays the description of the required action and the recorded gesture. If all of the recorded gestures match the required ones the identity proof is accepted. Both conversation partners have to verify each other in order to complete the authentication ceremony.

Consequences

After the conversation partners have successfully completed all required actions the authentication ceremony ends and the user interface will reflect the successful result as shown in the second part of Figure 5.5. A picture showing the conversation partner executing one of the gestures is used as a profile picture and the icon that initially hinted at a required action changes to a green check-mark. Additionally, the warning message vanishes and the all future messages have a green background.

Conversation partners have three tries to complete a required action, after which the authentication ceremony is aborted with the assumption of an ongoing MitM attack. Since either of the conversation partners could have misunderstood the required action, a negative result need not be an attack. The user interface reflects that fact by limiting the quality of service and requiring the conversation partners to meet in person and compare the used encryption keys. If that fallback mechanism also fails the users will be blocked from using the conversation under attack.

If a conversation partners changes keys after having been authenticated, the user interface reverts back to the unauthenticated state and a warning message informs the users of a possibly ongoing attack and that the authentication ceremony needs to be repeated.

Security

The picture-based approach shown in Figure 5.6 invokes trust in the identity of the conversation partners by showing a recognizable face. The required tasks serve as a reassurance that the pictures of the faces are recent and that they are harder for attackers to spoof. This example ceremony suggests gestures and tasks for the conversation partners based on the used key material. Those are used to transmit information about the used key in an unobtrusive way and add actual security to the procedure. If the resulting pictures match the expected gesture or task then the key material is equal and the conversation is secure from eavesdropping third parties.

According to Dechand *et al.* [DSB+16] a 112 bit fingerprint needs to be compared to withstand a brute-force attack comparable to a classical 2^{128} brute-force attacker. In case of 50 available gestures or tasks users would need to match 20 pictures in order to authenticate a conversation partner. This amount can be reduced by increasing the number of available gestures, but since this relation is exponential 2350 different available gestures are necessary to reduce this amount to 10. A compromise could be 12 necessary pictures, which would require 650 available gestures. This amount could be further minimized by reducing the attack resiliency to a 2^{80} brute-force attacker.

Since a high amount of necessary picture matches decreases the adherence rates the set of available gestures and tasks must be as large as possible. Using a video-chat could increase the usability of this method by requiring users to say several SAS phrases. The secure instant messaging application *Wickr* uses a similar authentication ceremony which they have patented [SHC17] in February of 2017. They require users to record an *identity verification video* where they have to say nine different SAS phrases. All other users have the option to retrieve the resulting video and verify it. Even though this approach has not yet been tested in a usability study, the results from the participatory design workshops suggest that it works better than other more common authentication ceremonies.

CHAPTER 6

Discussion

Dodier-Lazaro *et al.* [DSAB17] argued that security experts often do not notice that their values are different from those of the users. As a result, security as perceived by the users can be completely unrelated to actual security. In order to narrow this gap, user interfaces and social interaction need to be used to express the actual security level while embracing the users' values and mental models. The authors suggest that researchers should stop trying to "fix" the users so that they are able to correctly use security mechanisms and instead adopt a VSD approach in order to understand the users' values and consequently design for them.

A whole spectrum of methodologies can be used for VSD, as long as the research focuses on the users' values, mental models, and expectations. Research teams can therefore choose a method that they have most experience in. We chose the method of participatory design in order to explore how authentication should be represented in the user interface and which interactions invoke a sense of security. We built on the experience of Weber, Harbach, and Smith [WHS15] who used a similar approach to explore different design options for SSL warning messages. We expanded this approach by not only exploring different user interfaces, but also different kinds of interactions. We confirmed a result by Weber, Harbach, and Smith [WHS15] that simply structured participatory design workshops with a small number of participants are sufficient to provide a rich set of conceptual designs that can be used to understand users' values and serve as a basis for future implementations.

Each participatory design workshop started with a discussion about the participants' experiences with secure messengers. Convenience was the most important reason for participants to praise messengers, *Telegram* and *WhatsApp* were favored because so many people use them and clients exist for all platforms. Features were equally important for users and *Telegram* was praised by almost all the participants for its variety of sticker packages. All participants expressed annoyance about the diversity and inoperability of messaging solutions. *Signal* but also to a lesser degree *Telegram* were criticized for

their lacking quality of service. *WhatsApp* as well as *Facebook Messenger* were mistrusted because of Facebook's bad reputation concerning privacy. Interestingly, usability did not seem to be a major concern for participants of our workshops.

Based on a comprehensive analysis of the data gathered during the workshops I presented the participants' conceptual designs and evaluated how they could be adapted to provide the security that users expect. Participants most often suggested conceptual designs based on shared knowledge or recognizing the face of the person they are communicating with, which confirms an observation made by Vaziripour *et al.* [VWN+17] that users tend to rely on personal characteristics such as a person's voice, face, or shared knowledge for authentication.

Additionally, I distilled a set of implications for design from the data: (1) successful authentication ceremonies must leave users with a higher perceived level of trust, (2) authentication ceremonies should use well-known security concepts, (3) initiating authentication ceremonies must be friction-less, (4) messaging solutions should support users working with different levels of trust, and (5) the different levels of security and trust must be represented as such in the user interface.

From the users' conceptual designs, the implications for design, and the preliminary security evaluations several example ceremonies were derived that include the mental model behind it, a user interface design, necessary social interactions, and an appropriate security mechanism. Those example ceremonies were based on common types of trust discovered during the workshops: (1) a combination lock based on shared knowledge, (2) an identification card based on institutional trust, and (3) a verification dance using picture-based trust.

The related literature, the results of the participatory design workshops, the analysis of the resulting data, and the implications for design led to insights about the research questions proposed in the first chapter:

- *How can users be assured that the communication is actually encrypted and that state-of-the-art encryption procedures are used?*
  Users can not tell which encryption procedures are state-of-the-art and should not need to, instead they derive their knowledge about security from friends and media [RVR14]. The implications for design suggest that users assume a base level of security if a messaging solution is advertised as secure. In cases were manual actions are necessary, users need to understand why those are necessary, and see an effect of their actions. This effect can be visual, related to features, or the quality of service.

- *How can users be taught which attacks they are protected from and under which circumstances that protection is effective?*
  Several studies found that the success rate for authentication ceremonies work improves significantly after participants received an explanation of the purpose

and design of the ceremonies [VWN+17; HL16; SHWR16] and that users felt more confident in the security of their communication after receiving such explanations.

The implications for design suggest that users cannot be expected to be interested in the explanations and that a better way would be to use simple well-known security concepts that invoke a sense of security matching the provided security. This would provide users a simplified working model of the security mechanism that is usable without further explanation.

- *How can users be motivated to conduct the authentication ceremony?*
  In order to motivate users to conduct the authentication ceremony they need to have some (not necessarily correct) understanding of its purpose and its effects. The primary focus need not be the added security, since for many users this is not that important, but can also be an additional feature set, or fun side-effect of the ceremony. Additionally, the initiation of the authentication ceremony needs to be friction-less and the quick.

- *In which way should a user interface provide guidance for an (face-to-face or remote) authentication ceremony?*
  In the best case the underlying security concept of the authentication ceremony should be well-known and simple so that guidance is not necessary. If a rather uncommon security-concept is chosen the user interface needs to be simple and strict in order to provide helpful guidance to users.

- *How should an authentication ceremony be designed such that users can complete it as quick as possible, without sacrificing security?*
  The time that users are willing to invest in extra security is depended on how important they perceive their communication to be. Communicating sensitive topics or with close people will warrant time investment to stay secure for many users. Since users will not want to invest any time if they do not value their communication, the time efficiency of an authentication ceremony does not matter too much. If the communication is valuable to them, time is not the issue. Therefore, every messaging solution should provide an adequate level of security even without any time investment by the users.

The takeaway from those answers is that (1) users need a better way to establish which messaging applications are in fact secure, this could be achieved through media coverage or well-known certification marks, (2) users expect a high level of security by default, which requires trust to be established in an automatic way whenever possible, and (3) users are willing to invest time and effort for increased security when the situation warrants it, but only if the procedure makes sense to them and they understand its benefit.

CHAPTER 7

Future Work

The usability of the ceremonies described in the last chapter is the most concrete concern resulting from this thesis. In order to evaluate if they lead to a greater success rate than the authentication ceremonies all of them could be implemented using an open source messenger, such as *Signal*, as a basis. A usability study with a between-group design and structured interview could be used to compare the different implementations. The interview would be a necessary part in order to collect information about the resulting trust levels, the incentives to use it, and the perceived usability of the different solutions.

The implications for design which were presented at the end of Chapter 4 need to be evaluated and researched in detail. One of the main concerns was that either users need an explanation of the purpose of authentication ceremonies or that a well-known security concept without necessary explanation is used. A serious of semi-structured interviews could be conducted in order to discover which well-known security concepts fit the expectations of the users and match the security properties of an authentication ceremony. The results would give an immediate feedback of how authentication ceremonies need to look like in order to minimize explicit explanations.

To date it is unclear why users choose to initiate or not to initiate authentication ceremonies and what design features make it easier or harder for users. The implications for design suggest that initiation should be friction-less both in the user interface as well as in the necessary social interaction. More research is necessary to discover the users' reasons for not conducting authentication ceremonies and how potential hurdles could be removed altogether.

A common point of discussion during the participatory design workshops was that users assign different levels of importance to conversations with different people or about different topics and therefore work with different levels of trust. This is important because users are prepared to invest more time and effort if security matters to them.

Currently it is unclear how to design a set of authentication strategies that provide a base level of security for everyone but also considers the expected levels of security for different people or topics.

It is unclear how those different levels of security should be communicated effectively to the users. This could either be achieved using a visual representation in the user interface or a necessary action by the users. The communication should neither understate nor overstate the actual level security backing up the users' trust. More research into the communication of different levels of security to the users is necessary for designing appropriate authentication ceremonies.

This thesis demonstrated that participatory design is a useful method for VSD, but since this is by far not the only applicable approach, a few other methods to explore the users' values are suggested: (1) semi-structured interviews might be a good method to understand why users choose specific instant messaging applications and what their corresponding threat model is; (2) elicitation diaries could be used to explore the variety of reasons why users choose less secure ways of communicating resp. which are the instances when security mechanisms reduce the usability of a messenger; (3) case studies could be used to understand the usage of messengers in the context of every day family or business live; (4) ethnographies could be used to understand the values and mental models of specific user groups, such as students, business people, or activists.

CHAPTER 8

Conclusion

This thesis demonstrated that a VSD approach is applicable to designing authentication ceremonies. Only five participatory design workshops with two participants each lead to several intuitive conceptual designs worth of further investigation. Based on the results of the workshops I was able to derive three concrete authentication ceremonies that provide security and conserve the user's understanding of the original process. A comprehensive usability study of those resulting authentication ceremonies is necessary in order to compare it to existing approaches.

During the participant's experience reports it became clear that users have a hard time figuring out which messaging applications are secure and what that actually means. Literature suggests that media coverage and personal anecdotes are important for users information about potential risks and their mitigation, but further research is necessary. Participants reported that they do not see any reason to invest the time and effort needed to secure all of their conversations, since they think that only a fraction of them are important enough to warrant the effort. Therefore, automatic methods for trust establishment are necessary to provide the default level of security that users expect. However, participants also reported that they want to manually verify the security of conversations if contacts are important or the discussed topics sensitive. Therefore, Usable and secure authentication ceremonies provide a security benefit that automatic trust establishment can not provide.

Appendix

Workshop Procedure

We split the workshops into three phases as common in Participatory Design studies according to Spinuzzi [Spi05] and similar to the study by Weber, Harbach, and Smith [WHS15] about SSL warning messages.. Of those three phases the design phase was the longest with a length of about 40 to 45 minutes and I tried to limit each workshop to about an hour. In the following sections I describe the scenarios and questions used to elicit responses from the participants.

Discussion of Experiences

- With which secure messengers have you had experience?
- What kind of positive or negative experiences have you had with secure messengers?
- Which of those messengers do you not use anymore, and why?

Creating a Shared Language

I presented a short slide show explaining a few basic concepts needed to discuss authentication in instant messaging. The following questions covered in that slide show:

- What is end-to-end encryption and under which circumstances is it secure?
- What are common threat models of users and for which of them does end-to-end encryption help? *(as mentioned by Vaziripour* et al. *[VWN+17] and Renaud, Volkamer, and Renkema-Padmos [RVR14])*
- What is a Monster-in-the-Middle attack?
- What are the unknown properties in unauthenticated conversations?

I encouraged the participants to ask question during and after the presentation to ensure that everyone understood the basics.

Designing Concept Ceremonies

At the beginning of the design phase I presented the following scenario to the participants: *"You and a colleague from another branch office want to discuss a surprise party for a mutual friend. Your mutual friend with technical skills wants to find out exactly what you are conspiring but you definitely do not want to ruin the surprise and installed a secure messenger. You do not really know that colleague too well and you may or may not be able to meet him in person. What would you do in order to convince yourself that you are in fact talking to the right colleague and that nobody listens in?"*.

In the discussion that ensued afterwards I encouraged the participants to draw their ideas in order to explain them. Additionally, I frequently asked which conversation partner had to act and in which order they had to act according to their concept ideas.

In order talk about intuitive authentication approaches commonly used in the offline world, I presented a picture of an entrance hall of a bank (Figure 2) and a picture of a delivered package in front of a door (Figure 1). With these artifacts in mind we asked the following questions:

- You accepted a package of unknown value for an unknown neighbour and a few days later someone comes by to pick it up. What convinces you that you are handing the package over to the right person?
- You have been assigned a new bank advisor and you are waiting in the entrance hall to talk about a loan. What convinces you that you talking to your bank advisor and are not victim of a fraud?

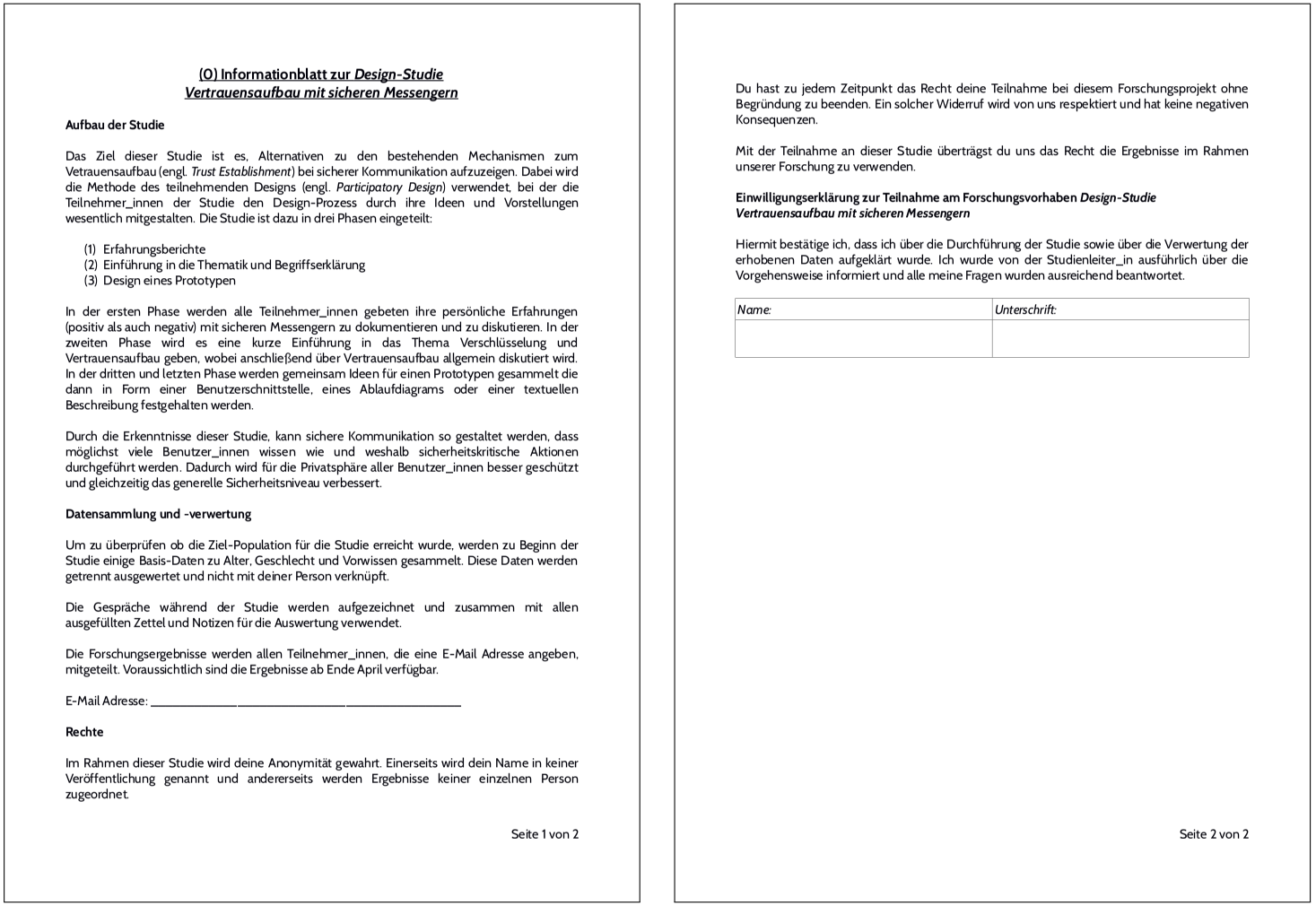
Afterwards, we discussed if any of those approaches could be translated to electronic communication and if so, how that would work.

Workshop Forms

Before each workshop session starts the participants receive an explanation of the workshop procedure and their rights. All the information is also stated in the consent form shown in Figure 3 that each participant signs. All participants also fill out a questionnaire shown in Figure 4 with basic information about themselves.

Figure 1: Artifact showing a package with fragile contents delivered at a door

Figure 2: Artifact showing an empty bank lobby



(0) Informationblatt zur *Design-Studie*
Vertrauensaufbau mit sicheren Messengern

Aufbau der Studie

Das Ziel dieser Studie ist es, Alternativen zu den bestehenden Mechanismen zum Vetrauensaufbau (engl. *Trust Establishment*) bei sicherer Kommunikation aufzuzeigen. Dabei wird die Methode des teilnehmenden Designs (engl. *Participatory Design*) verwendet, bei der die Teilnehmer_innen der Studie den Design-Prozess durch ihre Ideen und Vorstellungen wesentlich mitgestalten. Die Studie ist dazu in drei Phasen eingeteilt:

(1) Erfahrungsberichte
(2) Einführung in die Thematik und Begriffserklärung
(3) Design eines Prototypen

In der ersten Phase werden alle Teilnehmer_innen gebeten ihre persönliche Erfahrungen (positiv als auch negativ) mit sicheren Messengern zu dokumentieren und zu diskutieren. In der zweiten Phase wird es eine kurze Einführung in das Thema Verschlüsselung und Vertrauensaufbau geben, wobei anschließend über Vertrauensaufbau allgemein diskutiert wird. In der dritten und letzten Phase werden gemeinsam Ideen für einen Prototypen gesammelt die dann in Form einer Benutzerschnittstelle, eines Ablaufdiagrams oder einer textuellen Beschreibung festgehalten werden.

Durch die Erkenntnisse dieser Studie, kann sichere Kommunikation so gestaltet werden, dass möglichst viele Benutzer_innen wissen wie und weshalb sicherheitskritische Aktionen durchgeführt werden. Dadurch wird für die Privatsphäre aller Benutzer_innen besser geschützt und gleichzeitig das generelle Sicherheitsniveau verbessert.

Datensammlung und -verwertung

Um zu überprüfen ob die Ziel-Population für die Studie erreicht wurde, werden zu Beginn der Studie einige Basis-Daten zu Alter, Geschlecht und Vorwissen gesammelt. Diese Daten werden getrennt ausgewertet und nicht mit deiner Person verknüpft.

Die Gespräche während der Studie werden aufgezeichnet und zusammen mit allen ausgefüllten Zettel und Notizen für die Auswertung verwendet.

Die Forschungsergebnisse werden allen Teilnehmer_innen, die eine E-Mail Adresse angeben, mitgeteilt. Voraussichtlich sind die Ergebnisse ab Ende April verfügbar.

E-Mail Adresse: _____

Rechte

Im Rahmen dieser Studie wird deine Anonymität gewahrt. Einerseits wird dein Name in keiner Veröffentlichung genannt und andererseits werden Ergebnisse keiner einzelnen Person zugeordnet.

Du hast zu jedem Zeitpunkt das Recht deine Teilnahme bei diesem Forschungsprojekt ohne Begründung zu beenden. Ein solcher Widerruf wird von uns respektiert und hat keine negativen Konsequenzen.

Mit der Teilnahme an dieser Studie überträgst du uns das Recht die Ergebnisse im Rahmen unserer Forschung zu verwenden.

Einwilligungserklärung zur Teilnahme am Forschungsvorhaben *Design-Studie Vertrauensaufbau mit sicheren Messengern*

Hiermit bestätige ich, dass ich über die Durchführung der Studie sowie über die Verwertung der erhobenen Daten aufgeklärt wurde. Ich wurde von der Studienleiter_in ausführlich über die Vorgehensweise informiert und alle meine Fragen wurden ausreichend beantwortet.

Name:	*Unterschrift:*

Figure 3: Consent form for participants of the participatory design focus groups

Basis-Fragen

Um zu überprüfen ob das Zielpublikum dieser Studie erreicht wurde, bitten wir Dich folgende Fragen zu beantworten:

Alter	
Geschlecht	
Wissen über Kryptographie	Sehr wenig ○ ○ ○ ○ ○ ○ Sehr viel
Wissen über IP-Netzwerke	Sehr wenig ○ ○ ○ ○ ○ ○ Sehr viel

Figure 4: Questionnaire for participants of the participatory design focus groups

Codebooks

Experience Reports

Category	No.	Name	Freq.	Location(s)
Prevalance	2	WhatsApp Prevalence	1	5055:7
	7	Tedious Diversity	2	5055:12 10916:8
	26	Large User Base	1	10916:16
	27	Peer pressure as a reason to use messengers	2	10916:19 21784:19
	31	Clients available for many platforms	2	18197:7 21784:9-10
Privacy	8	Personal Data unimportant	1	5055:14
	9	Concern about Mass-Surveillance	1	5055:14-15
	10	Health-Data is important	1	5055:16-17
	12	Secret Chats available	1	5442:5
	13	Ephemeral Messaging	2	5442:5 5442:6
	17	Images not deletable	1	5442:17
	24	Too many required app authorizations	2	10916:12-13 21784:16
	25	Uploading Contact Information	1	10916:14
	28	Wiretapping Device	1	10916:15-19
	29	Minimal or no data storage as a sign of quality	1	18197:2
Trust	6	Loss of Trust due to Sponsoring	1	5055:11
	11	Facebook as a sign of insecurity	1	5442:3
	22	Open Source as a sign of quality	2	10916:6 21784:20
	34	No trust in Telegram	1	21784:4
Asthetics	3	Pleasent Design	1	5055:8
	15	Puristic Design	1	5442:8
	40	Ugly	1	21784:17
Obstacles	5	Message-Loss after re-install	1	5055:10
	7	Tedious Diversity	2	5055:12 10916:8

	16	Complicated Group-Chats	1	5442:10
	19	Missing notifications	1	5442:15
	20	Intermittent message-loss	2	5442:16 10916:10-11
	21	Incomprehensible Authentication Ceremony	1	5442:17-19
	23	Phonenumber necessary	1	10916:9
	32	Forced to use app because of technical reasons	1	18197:10
	33	Technical Limitations of Device	1	18197:15
	37	No pictures next to names	1	21784:12
	38	Difficult to use	2	21784:13 21784:14
Features	1	Telegram Sticker	3	5055:5 5442:7 18197:6
	4	More convenient than SMS	1	5055:7
	14	Sticker-Selection and -Trading	1	5442:7
	18	Only short videos sendable	1	5422: 14
	30	Easy to use voice messages	1	19197:5
	35	Usable without phonenumber	1	21784:7
	36	Messages available on different devices	1	21784:8
	39	File-transfer impossible	1	21784:16

Conceptual Designs

Category	No.	Name	Freq.	Location(s)
Shared Knowledge	1	Personal Questions	3	130:1 321:7 530:29
	7	Exchange of secret keys	3	130:6 424:1 424:6
	25	Exchange code-words	2	226:1-2 530:4
	27	Exchange symmetric encryption keys	3	226:5 530:7 530:10-13
	44	Knowledge of personal information	1	424:15

	46	QR-Code Confirmation	1	530:1 130:26
	13	Continously confirm identity with known information	6	321:13 321:22 424:9 424:23-24 530:14-15
Pictures	22	Profilepicture not important	1	130:32 130:14
	12	Recognition of person	4	226:10 321:17 530:16
	31	Recognition from photo	2	226:13 424:20
	38	Picture-Check	8	226:32 321:9 321:26 321:27 424:21 530:21 530:24 530:27
Social	2	Inconspicuous identity-check	2	130:1 130:25
	3	Check reactions person in question	1	130:3
	11	Avoid seeming suspicious	1	130:12
	16	Lookism	1	130:20 130:22
	17	Check reactions of social environment	5	226:19 321:20 424:14 530:18
	24	Trust network	2	130:34-35 424:28
	28	Trust a third-party	3	226:6 424:19 530:23
	32	Suspicious Behaviour	3	226:15 321:14 424:13
	33	Fits the environment	1	226:16

	41	Friendliness	1	321:16
Institution	15	ID-card	5	130:19 226:8 226:25-26 321:12 321:19 424:18
	19	Name-tag	2	130:24 226:17
	30	Authentication by possession	2	226:12 321:10
	34	Verification Institution	1	226:20
	36	Verification by experts	1	226:27
	39	Fingerprint	1	321:1
	40	Signature	1	321:11
	42	Place of conversation	4	226:14 424:16 530:19 530:22
	45	Crypto-ID-card	1	424:29
Habituation	20	Time and Reoccuring Messages	7	130:27 130:28 226:22 226:30 321:15 321:31 424:12 530:26
	21	Visualize habituation	1	130:31
	37	Writing-Style as expected	2	226:28 321:23
	47	Style of speech	1	530:25
Testing	5	Compare Chat-History	1	130:5
	6	Uncover Message manipulation	2	130:5 321:6

	13	Continously confirm identity with known information	6	130:26 321:13 321:22 424:9 424:23-24 530:14-15
	14	Manual Check by the users	3	130:18 226:11 226:31
	18	Check if identity is plausible	2	130:22 130:25
	26	Measure Transmission Latency	1	226:4
Wishes and Requirements	23	Base-Level of Trust	1	130:33
	14	Manual Check by the users	3	130:18 226:11 226:31
	11	Avoid seeming suspicious	1	130:12
	10	Automatic detection of attacks	1	130:11
	8	Small choice when selecting shared secrets	1	130:7
	2	Inconspicuous identity-check	2	130:1 130:25
	9	Warning after Security-Errors	1	130:9
	35	Careful after incidents	2	226:24 424:10-11
	43	Minimal personal information required	1	321:24-25
	29	Effort only worth it for few	2	226:7 530:28-29
Reactions	4	Change Communication-Medium	1	130:4

Bibliography

[ABH+15] E. Atwater, C. Bocovich, U. Hengartner, E. Lank, and I. Goldberg, "Leading Johnny to Water: Designing for Usability and Trust", *Eleventh Symposium On Usable Privacy and Security (SOUPS 2015)*, pp. 69–88, 2015. [Online]. Available: https://www.usenix.org/conference/soups2015/proceedings/presentation/atwater.

[AF13] D. Akhawe and A. P. Felt, "Alice in warningland: a large-scale field study of browser security warning effectiveness", *Proceedings of the 22nd USENIX Security Symposium*, pp. 257–272, 2013.

[AG07] C. Alexander and I. Goldberg, "Improved user authentication in off-the-record messaging", in *Proceedings of the 2007 ACM workshop on Privacy in electronic society - WPES '07*, ACM, 2007, p. 41, ISBN: 9781595938831. DOI: 10.1145/1314333.1314340. [Online]. Available: http://portal.acm.org/citation.cfm?doid=1314333.1314340.

[ASB+17] R. Abu-Salma, M. A. Sasse, J. Bonneau, A. Danilova, A. Naiakshina, and M. Smith, "Obstacles to the Adoption of Secure Communication Tools", *Proceedings - IEEE Symposium on Security and Privacy*, pp. 137–153, 2017, ISSN: 10816011. DOI: 10.1109/SP.2017.65.

[BST01] F. Boudot, B. Schoenmakers, and J. Traoré, "A fair and efficient solution to the socialist millionaires' problem", *Discrete Applied Mathematics*, vol. 111, no. 1-2, pp. 23–36, 2001, ISSN: 0166218X. DOI: 10.1016/S0166-218X(00)00342-5.

[DDO+16] A. De Luca, S. Das, M. Ortlieb, I. Ion, and B. Laurie, "Expert and Non-Expert Attitudes towards (Secure) Instant Messaging", in *the Symposium On Usable Privacy and Security (SOUPS)*, 2016, pp. 147–157, ISBN: 978-1-931971-31-7. [Online]. Available: https://www.usenix.org/conference/soups2016/technical-sessions/presentation/deluca.

[DSAB17] S. Dodier-Lazaro, M. A. Sasse, R. Abu-Salma, and I. Becker, "From Paternalistic to User-Centred Security: Putting Users First with Value-Sensitive Design", in *CHI 2017 Workshop on Values in Computing*, 2017.

[DSB+16] S. Dechand, D. Schürmann, K. Busse, Y. Acar, S. Fahl, and M. Smith, "An Empirical Study of Textual Key-Fingerprint Representations", *USENIX Security '16*, pp. 193–208, 2016.

[EHM17] K. Ermoshina, H. Halpin, and F. Musiani, "Can Johnny Build a Protocol? Co-ordinating developer and user intentions for privacy-enhanced secure messaging protocols", in *Proceedings of the 2nd European Workshop on Usable Security (EuroUSEC)*, Internet Society, 2017, pp. 1–14, ISBN: 1891562487. DOI: 10.14722/eurousec.2017.230016. [Online]. Available: https://www.internetsociety.org/sites/default/files/eurousec2017%7B%5C_%7D16%7B%5C_%7DErmoshina%7B%5C_%7Dpaper.pdf.

[EJP+14] S. Egelman, S. Jain, R. S. Portnoff, K. Liao, S. Consolvo, and D. Wagner, "Are You Ready to Lock?", in *Proceedings of the 2014 ACM SIGSAC Conference on Computer and Communications Security - CCS '14*, ACM, 2014, pp. 750–761, ISBN: 9781450329576. DOI: 10.1145/2660267.2660273. [Online]. Available: http://dl.acm.org/citation.cfm?doid=2660267.2660273.

[FLK+13] M. Farb, Y.-H. Lin, T. H.-J. Kim, J. McCune, and A. Perrig, "SafeSlinger", in *Proceedings of the 19th annual international conference on Mobile computing & networking - MobiCom '13*, 2013, p. 417, ISBN: 9781450319997. DOI: 10.1145/2500423.2500428. [Online]. Available: http://devd.me/papers/pwdmgr-usenix14.pdf%7B%5C%%7D5Cnhttps://www.usenix.org/conference/usenixsecurity14/technical-sessions/presentation/li%7B%5C_%7Dzhiwei%20http://dl.acm.org/citation.cfm?doid=2500423.2500428.

[Gan17] M. Ganguly, *WhatsApp design feature means some encrypted messages could be read by third party*, 2017. [Online]. Available: https://www.theguardian.com/technology/2017/jan/13/whatsapp-design-feature-encrypted-messages (visited on 11/24/2017).

[GG14] D. Gelles and V. Goel, *Facebook Enters $16 Billion Deal for WhatsApp*, 2014. [Online]. Available: https://dealbook.nytimes.com/2014/02/19/facebook-to-buy-messaging-start-up/ (visited on 11/24/2017).

[Gib14] S. Gibbs, *Six alternatives to WhatsApp now that Facebook owns it*, 2014. [Online]. Available: https://www.theguardian.com/technology/2014/feb/20/six-alternatives-whatsapp-facebook (visited on 11/24/2017).

[Gle13] L. P. Glenn Greenwald, Ewen MacAskill, *Edward Snowden: the whistleblower behind the NSA surveillance revelations | US news | The Guardian*, 2013. [Online]. Available: https://www.theguardian.com/world/2013/

jun/09/edward-snowden-nsa-whistleblower-surveillance (visited on 04/09/2018).

[GM05] S. L. Garfinkel and R. C. Miller, "Johnny 2: A User Test of Key Continuity Management with S/MIME and Outlook Express", *Proceedings of the 2005 symposium on Usable privacy and security*, vol. 6, pp. 13–24, 2005. DOI: 10.1145/1073001.1073003. [Online]. Available: http://portal.acm.org/citation.cfm?id=1073001.1073003.

[GPM17] K. Gallagher, S. Patil, and N. Memon, "New Me: Understanding Expert and Non-Expert Perceptions and Usage of the Tor Anonymity Network", in *Thirteenth Symposium on Usable Privacy and Security (SOUPS 2017)*, USENIX Association, 2017, pp. 385–398, ISBN: 978-1-931971-39-3. [Online]. Available: https://www.usenix.org/conference/soups2017/technical-sessions/presentation/gallagher.

[GS67] B. G. Glaser and A. L. Strauss, *The Discovery of Grounded Theory: Strategies for Qualitative Research*, 4. Aldine Publishing Company, Chicago, 1967, vol. 1, p. 271, ISBN: 0202302601. DOI: 10.2307/2575405. arXiv: 9809069v1 [arXiv:gr-qc]. [Online]. Available: http://www.amazon.com/dp/0202302601.

[HL16] A. Herzberg and H. Leibowitz, "Can Johnny Finally Encrypt ? Evaluating E2E- Encryption in Popular IM Applications", in *ACM Workshop on Socio-Technical Aspects in Security and Trust (STAST)*, 2016, ISBN: 9781450348263. DOI: 10.1145/3046055.3046059.

[KDFK15] R. Kang, L. Dabbish, N. Fruchter, and S. Kiesler, ""My data just goes everywhere": User mental models of the internet and implications for privacy and security", in *Symposium on Usable Privacy and Security (SOUPS) 2015*, 2015, pp. 39–52, ISBN: 978-1-931971-249.

[KMSW17] K. Krombholz, W. Mayer, M. Schmiedecker, and E. Weippl, ""I Have No Idea What I'm Doing" – On the Usability of Deploying HTTPS", *USENIX Security*, pp. 1–18, 2017. [Online]. Available: https://www.usenix.org/system/files/conference/usenixsecurity17/sec17-krombholz.pdf%20https://www.usenix.org/conference/usenixsecurity17/technical-sessions/presentation/krombholz.

[KTW09] C. Karlof, J. D. Tygar, and D. Wagner, "Conditioned-safe ceremonies and a user study of an application to web authentication", in *Symposium on Usable Privacy and Security (SOUPS) 2009*, 2009, p. 20, ISBN: 9781605587363. DOI: 10.1145/1572532.1572578. [Online]. Available: http://portal.acm.org/citation.cfm?doid=1572532.1572578.

[LZR17] A. Lerner, E. Zeng, and F. Roesner, "Confidante: Usable Encrypted Email: A Case Study with Lawyers and Journalists", *Proceedings - 2nd IEEE European Symposium on Security and Privacy, EuroS and P 2017*, pp. 385–400, 2017. DOI: 10.1109/EuroSP.2017.41.

[Mar17] M. Marlinspike, *There is no WhatsApp 'backdoor'*, 2017. [Online]. Available: https://signal.org/blog/there-is-no-whatsapp-backdoor/ (visited on 11/24/2017).

[MBB+15] M. S. Melara, A. Blankstein, J. Bonneau, E. W. Felten, and M. J. Freedman, "CONIKS : Bringing Key Transparency to End Users", *USENIX Security*, pp. 383–398, 2015.

[MCHR15] S. E. McGregor, P. Charters, T. Holliday, and F. Roesner, "Investigating the Computer Security Practices and Needs of Journalists", *USENIX Security Symposium*, 2015.

[MWA+17] S. E. McGregor, E. A. Watkins, M. N. Al-Ameen, K. Caine, and F. Roesner, "When the weakest link is strong: Secure collaboration in the case of the Panama Papers", *26th USENIX Security Symposium (USENIX Security)*, pp. 505–522, 2017. [Online]. Available: https://www.usenix.org/conference/usenixsecurity17/technical-sessions/presentation/mcgregor.

[RAH+16] S. Ruoti, J. Andersen, S. Heidbrink, M. O'Neill, E. Vaziripour, J. Wu, D. Zappala, and K. Seamons, ""We're on the Same Page": A Usability Study of Secure Email Using Pairs of Novice Users", in *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems - CHI '16*, 2016, pp. 4298–4308, ISBN: 9781450333627. DOI: 10.1145/2858036.2858400. arXiv: 1510.08554. [Online]. Available: http://arxiv.org/abs/1510.08554%7B%5C%%7D0Ahttp://dx.doi.org/10.1145/2858036.2858400%20http://dl.acm.org/citation.cfm?doid=2858036.2858400.

[RAZS15] S. Ruoti, J. Andersen, D. Zappala, and K. Seamons, "Why Johnny Still, Still Can't Encrypt: Evaluating the Usability of a Modern PGP Client", 2015, ISSN: 2010-1945. DOI: 10.1142/S2010194516601320. arXiv: 1510.08555. [Online]. Available: http://arxiv.org/abs/1510.08555.

[RKB+13] S. Ruoti, N. Kim, B. Burgon, T. van der Horst, and K. Seamons, "Confused Johnny: When Automatic Encryption Leads to Confusion and Mistakes", *Proceedings of the Ninth Symposium on Usable Privacy and Security - SOUPS '13*, 5:1–5:12, 2013. DOI: 10.1145/2501604.2501609. [Online]. Available: http://doi.acm.org/10.1145/2501604.2501609.

[RS16] S. Ruoti and K. Seamons, "Standard Metrics and Scenarios for Usable Authentication", *Symposium on Usable Privacy and Security (SOUPS)*, pp. 1–2, 2016.

[RVR14] K. Renaud, M. Volkamer, and A. Renkema-Padmos, "Why Doesn't Jane Protect Her Privacy?", *Privacy Enhancing Technologies*, vol. 8555, pp. 244–262, 2014. DOI: 10.1007/978-3-319-08506-7_13. [Online]. Available: http://dx.doi.org/10.1007/978-3-319-08506-7%7B%5C_%7D13.

[SBHK06] S. Sheng, L. Broderick, J. J. Hyland, and C. A. Koranda, "Why Johnny still can't encrypt: evaluating the usability of email encryption software", *Symposium On Usable Privacy and Security*, pp. 3–4, 2006. [Online]. Available: http://chariotsfire.com/pub/sheng-poster%7B%5C_%7Dabstract.pdf.

[SHC17] R. Statica, C. A. Howell, and K. L. Coppa, *IN-BAND IDENTITY VERIFICATION AND MAN-IN-THE-MIDDLE DEFENSE*, 2017.

[SHWR16] S. Schröder, M. Huber, D. Wind, and C. Rottermanner, "When SIGNAL hits the Fan : On the Usability and Security of State-of-the-Art Secure Mobile Messaging", in *European Workshop on Usable Security (EuroUSEC)*, 2016, ISBN: 1891562452. DOI: 10.14722/eurousec.2016.23012.

[Spi05] C. Spinuzzi, "The Methodology of Participatory Design", *Technical Communication*, vol. 52, no. 2, pp. 163–174, 2005, ISSN: 00493155. DOI: 10.1016/j.infsof.2008.09.005. arXiv: 0402594v3 [arXiv:cond-mat]. [Online]. Available: http://www.ingentaconnect.com/content/stc/tc/2005/00000052/00000002/art00005.

[SYG08] R. Stedman, K. Yoshida, and I. Goldberg, "A user study of off-the-record messaging", *the 4th Symposium on Usable Privacy and Security*, pp. 95–104, 2008. DOI: 10.1145/1408664.1408678. [Online]. Available: http://portal.acm.org/citation.cfm?id=1408664.1408678%7B%5C%%7D5Cnpapers2://publication/uuid/F0AFE0B9-8535-4626-949D-09891C11FBA0.

[TBB+17] J. Tan, L. Bauer, J. Bonneau, L. F. Cranor, J. Thomas, and B. Ur, "Can Unicorns Help Users Compare Crypto Key Fingerprints?", in *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems - CHI '17*, 2017, pp. 3787–3798, ISBN: 9781450346559. DOI: 10.1145/3025453.3025733. [Online]. Available: http://dl.acm.org/citation.cfm?doid=3025453.3025733.

[The13] The Guardian, *The NSA files*, 2013. [Online]. Available: https://www.theguardian.com/us-news/the-nsa-files (visited on 04/09/2018).

[TV 18] TV Tropes Wiki, *Spy Speak*, 2018. [Online]. Available: http://tvtropes.org/pmwiki/pmwiki.php/Main/SpySpeak (visited on 03/29/2018).

[UDB+15] N. Unger, S. Dechand, J. Bonneau, S. Fahl, H. Perl, I. Goldberg, and M. Smith, "SoK: Secure messaging", *Proceedings - IEEE Symposium on Security and Privacy*, vol. 2015-July, pp. 232–249, 2015, ISSN: 10816011. DOI: 10.1109/SP.2015.22.

[VWH+16] E. Vaziripour, J. Wu, S. Heidbrink, K. Seamons, and D. Zappala, "Social Authentication for End-to-End Encryption", in *Symposium on Usable Privacy and Security (SOUPS) 2016*, 2016, pp. 1–2. arXiv: arXiv:1510.08555.

[VWN+17] E. Vaziripour, J. Wu, M. O. '. Neill, R. Clinton, J. Whitehead, S. Heidbrink, K. Seamons, and D. Zappala, "Is that you, Alice? A Usability Study of the Authentication Ceremony of Secure Messaging Applications", *Soups*, no. Soups, 2017. [Online]. Available: https://www.usenix.org/conference/soups2017/technical-sessions/presentation/vaziripour.

[WHS15] S. Weber, M. Harbach, and M. Smith, "Participatory Design for Security-Related User Interfaces", in *USEC*, 2015, ISBN: 1891562401.

[Wor14] J. Wortham, *WhatsApp Deal Bets on a Few Fewer 'Friends'*, 2014. [Online]. Available: https://www.nytimes.com/2014/02/22/technology/whatsapp-deal-bets-on-a-few-fewer-friends.html (visited on 11/24/2017).

[WT99] A. Whitten and J. Tygar, "Why Johnny can't encrypt: A usability evaluation of PGP 5.0", *Proceedings of the 8th USENIX Security Symposium*, pp. 169–184, 1999. DOI: 169-184. [Online]. Available: http://scholar.google.com/scholar?hl=en%7B%5C&%7DbtnG=Search%7B%5C&%7Dq=intitle:Why+Johnny+can?t+encrypt:+A+usability+evaluation+of+PGP+5.0%7B%5C#%7D0.

Glossary

Five Eyes An alliance between Australia, Canada, New Zealand, the United Kingdom, and the United States of America for exchanging signals intelligence. 1

GCHQ The Government Communications Headquarters (GCHQ) is a British foreign intelligence agency providing signals intelligence. 1

HTTPS HTTP connection with SSL protection. Original it was only used for banking websites or e-commerce but nowadays is often expected by default. 11

NSA The National Security Agency (NSA) is an US American foreign intelligence agency providing signals intelligence. 1

perfect forward secrecy A property of encryption schemes that protects data encrypted in the past against a key compromise in the future. Users do not reveal their entire message history if one key is compromised, which leads to a reduced incentive for attackers to compromise those keys. 1

PGP Software that uses the OpenPGP standard to email end-to-end encryption. Uses a Web-of-Trust for Trust Establishment between users and only works with a plugin in most email clients. 1, 9, 10, 12, 37

QR Code Two-dimensional graphical code with automatic error-correction. Commonly used to encode links to websites, but can also be used to encode other information such as WiFi connection details or payment details. 3, 4, 18, 35, 51, 52

repudiation A property of encryption schemes that does not allow associating a message with a specific person, i.e. more than one person had the ability to generate the message. 1

S/MIME End-to-end email encryption standard developed in 1995 and is integrated in all common email clients. Uses authority-issued X.509 certificates for Trust Establishment. 1, 10

SSL Protocol that provides point-to-point encryption and uses authority-issued X.509 certificates for Trust Establishment. Has been superseded by the TLS protocol, but SSL is still used as a colloquial term. 10, 11, 21, 22, 57

Tor Network that uses Onion routing to anonymize its users. Tor-Browser users can anonymously use websites, circumvent censorship, and access Tor hidden services (sometimes also referred to as Dark Web). 15

Acronyms

E2EE end-to-end encryption. 1–3, 6, 13, 16, 18

GUI Graphical User Interface. 9

HCI Human-Computer Interaction. 16, 21

IM Instant Messaging. 1, 10, 13, 14

KCM Key Continuity Management. 10

MitM Monster-in-the-Middle. ix, xi, 2–5, 16, 19, 36, 41, 49, 52, 56

OTR Off-the-Record. 1, 18, 49

SAS Short Authentication String. 15, 36, 56

SIM Secure Instant Messaging. 1, 2, 13, 14

SMP Socialist Millionaire Protocol. 15, 49

SUS System Usability Scale. 10, 11

TOFU Trust On First Use. 10, 15

VSD Value-Sensitive Design. xiii, 5, 12, 13, 21, 45, 57, 62, 63