A Comparative Long-Term Study of Fallback Authentication

Work in Progress

Philipp Markert
Ruhr University Bochum
philipp.markert@rub.de

Maximilian Golla
Ruhr University Bochum
maximilian.golla@rub.de

Elizabeth Stobert
National Research Council of Canada
elizabeth@stobert.ca

Markus Dürmuth
Ruhr University Bochum
markus.duermuth@rub.de

***Abstract*—Fallback authentication, the process of recovering access to an account if the primary authenticator is forgotten or lost, is of significant importance in real-world applications. A variety of mechanisms are deployed, ranging from secondary channels (such as email and SMS), over personal knowledge questions (such as the "mother's maiden name") to social authentication (such as vouching-based approaches). One central difference with primary authentication is that the elapsed time between enrollment and authentication can be much longer, typically in the range of years. However, few of the mechanisms used today have been studied over such long time-spans, making claims about their usability difficult to generalize to real-world applications. Additionally, most past studies have considered one or two mechanisms only, and deriving a meaningful comparison of a relevant number of mechanisms from the individual datapoints is not easy. In this work in progress paper, we report on the design of a usability study that we will use to study the usability of authentication mechanisms over a more realistic time-frame of up to 18 months, and will provide a fair comparison of the four most widely used fallback authentication schemes. We present results of a pre-study with 74 participants that ran over 4 weeks and indicates that schemes based on email and SMS are more usable. Mechanisms based on designated trustees and personal knowledge questions, on the other hand, fall short, both in terms of convenience and efficiency.**

I. Introduction

Fallback authentication is the mechanism for recovering access to an account after the primary authenticator is lost (the literature also uses the terms backup authentication, emergency authentication, recovery authentication, last-resort authentication, or account recovery). It plays an important role for real-world applications of authentication, as administrators have to deal with forgotten passwords and lost security tokens on a regular basis [17], [34], [29], and manual account recovery can be expensive [9].

The security requirements are equal to those of primary authentication systems, which are used on a daily basis, as fallback authentication forms another means by which the system can be accessed. A weak fallback authentication mechanism compromises the security of the overall system in the same way a weak primary authentication scheme does, as illustrated by the prominent "Celebgate" hack in 2014 [35] or other prominent account takeover attempts [16].

In contrast to this, the requirements for fallback authentication regarding usability are different from the requirements for primary authentication:

i. *Long-Term memorability* is more critical, as the period between enrollment and authentication is almost always longer for fallback authentication. Fallback authentication is also not supported by the rehearsal that results from frequent entry and supports the memorability of more frequently used forms of authentication. The usage of multi-factor authentication (MFA) would support this learning process as users are required to provide additional codes from electronic mails (emails), short message service (SMS), or mobile applications on a regular basis. However, the adoption rates of MFA are minimal [20] which is why it is not recommendable to take these authentication forms as a given.

ii. *Authentication time* is less critical to fallback authentication, because it is a relatively rare action, and is not intended to take place frequently.

iii. *Rate-limiting* can be much stricter than for primary authentication (e. g., in the order of a few authentication attempts per day).

Probably the most typical example of a fallback authentication process is a reset link which is sent via email to a user. By clicking on the link the user is directed to a page where a new password can be set. Other approaches require the user to provide a reset code that was sent via SMS or to answer previously set personal knowledge questions (PKQs).

Research on fallback authentication is complicated by the required long spans between enrollment and authentication. Work on Google's security questions gives some details about the usage of fallback authentication in the wild [2], and reports a nearly linear relation between the time passed and the number of users who used the fallback system. After approximately 150 days (4.9 months) 30 % of the users initiated a reset while 50 % did after 330 days (10.9 months) and 70 % after 540 days (17.8 months).

Several papers studied the usability of fallback authentication schemes. However, very few measure memorability after

Workshop on Usable Security and Privacy (USEC) 2019
24 February 2019, San Diego, CA, USA
ISBN 1-891562-57-6
https://dx.doi.org/10.14722/usec.2019.23030
www.ndss-symposium.org

more than 6 months, or test more than a single scheme. Based on those observations we developed the following research questions that we like to answer by conducting a long-term user study:

- **RQ1:** Given the same conditions, how do the considered fallback authentication schemes perform in terms of success rate, time required, user sentiment, dropout rate, and user burdening?
- **RQ2:** Which fallback authentication scheme performs best given realistic recall times of 6/12/18 months?
- **RQ3:** Which usability issues arise when using fallback authentication after 6/12/18 months?

II. Background

As different fallback authentication schemes have been proposed within the last years, many researchers have also analyzed the usability of those schemes. One of the first contributions came from Zviran and Haga [39]. In their paper from 1990, they introduced and analyzed the concept of a *cognitive password* which is nowadays known under the term *PKQ*. They demonstrated a higher recall rate of this novel approach compared to a conventional password as well as a low recall rate by even closely related persons. However, this conclusion originates from a time when social media platforms with users indirectly posting the answers to their security questions did not exist yet.

Newer studies [25], [30], [19], [23] came to a different conclusion: in today's presence of personal information on the Internet and especially in social networks PKQs do not provide the initially intended level of security and usability.

Simson L. Garfinkel wrote an article promoting the use of emails for authentication in 2003 [11]. He argues that this approach is highly deployable as the required infrastructure to send emails already exists and the known security risks are manageable considering the widespread usage by various services.

Another approach uses SMS instead of email. In 2015, Bonneau et al. [2] analyzed the account recovery processes at Google. There, a SMS-based scheme had the highest success rates followed by email recovery and PKQs. Nevertheless, the researchers also identified a disadvantage of this approach. The scheme requires having possession of the phone thus a user who has not got the phone within reach, cannot use SMS for account recovery.

Brainard et al., in turn, proposed a novel group of social authentication schemes in 2006 [5] and involved designated trustees of a user in the recovery process. While they were researchers at RSA Security and based their scheme on the proprietary hardware authentication token SecurID, Schechter et al. generalized the idea using email addresses [31]. Although the results suggested that using designated trustees is not as efficient as other schemes, the new approach showed a high success rate. In their experiment, 17 out of 19 participants who contacted their designated trustees were able to successfully complete the recovery process.

The presented works depict the diversity of different fallback authentication schemes. The subsequent task for service providers to agree on one mechanism is challenging and requires comparative, long-term analyses. Bonneau et al. [3] compared several fallback authentication schemes, including email, SMS, designated trustees, and PKQs, but only synthesized individual analyses [5], [30], [36].

In a subsequent work Bonneau et al. [2] directly compared different schemes and demonstrated a higher recovery rate for SMS (81 %), and email (75 %), than for PKQs (61 %). However, they did not consider the elapsed time between account creation and recovery claim.

III. Fallback Authentication Schemes

Several approaches are used for fallback authentication, and many more have been proposed in the literature. We discuss the most relevant in the following.

A. Secondary Channel

One of the most common techniques is to use a secondary channel that has been set up while the user still had access to the account.

a) Email: When using email as secondary channel, one registers an email address (often specifically for the recovery) while still having access to the account. In case access to the account is lost, the account recovery can be initiated using the account name only. The service provider sends an email containing a link or a reset code to the person who initiated the process, and by clicking on the link or typing in the code on a special page, the user is able to set a new password. Alternatively, a temporary password can be sent, with a forced reset after using it.

b) SMS: Using SMS as a secondary channel is very similar to email-based recovery, but instead of an email address, a phone number is linked with the account. For account recovery, an SMS is sent that contains a reset code.

As for the email scenario, it is also possible to send links although this only makes sense if the user has a smartphone. Otherwise, the usability is negatively affected as the link needs to be copied manually into a web browser. If password-based authentication is used, a third option is to send a temporary password.

B. Social Authentication

Social authentication describes a class of mechanisms that rely on "who you know," i. e., information about one's social graph.

a) Designated Trustees: For using designated trustees, the user selects several contacts while still having access to the account. The initial proposal by Schechter et al. [31] from 2009 used email addresses entered by the user to identify the designated trustees, while later implementations by Facebook allowed users to select the trustees from their friend list. For account recovery, the trustees will receive reset codes, and a subset of the reset codes is required to regain access to the account.

TABLE I: The considered fallback authentication schemes as well as the security assumption they rely on.

Scheme	Description	Security Assumption
Email	Click on reset link sent to registered email account	Secrecy of the channel and access to the email
SMS	Provide reset code sent via SMS to a registered phone number	Secrecy of the channel and access to the phone
Designated Trustees	Provide reset codes sent to registered trusted contacts	Ability of trusted contacts to only share the reset code with the user
PKQ	Answer security questions referring to personal knowledge	Difficulty to correctly guess the answers (targeted and trawling attacks)
Browser Fingerprint	Fingerprint including IP address, geolocation, and user agent	Difficulty to obtain and mimic the user's precise browser fingerprint

b) Friends Selected at Recovery Time: Another social recovery scheme, called *Trusted Friends*, was introduced by Facebook in October 2011 [7]. It allowed the selection of trusted friends, from a list of active friends, after the access was lost. In May 2013, Facebook introduced a redesign, called *Trusted Contacts* [8], which is constrained in the selection of friends in an attempt to avoid attacks that selected recently-added fake friends under the control of an attacker [18]. Similar to before, each of those contacts receives instructions on how to obtain a reset code, and by providing three of those codes, the user can perform a password reset.

c) Knowledge About the Social Graph: A third approach is to generate questions about a user's social graph. The idea was first proposed by Yardi et al. [38] in 2008, who based a prototype implementation on the social information that is provided by Facebook. The system uses the social network graph as well as auxiliary information like photos with tags of the shown persons to authenticate users. This is done by presenting photos from the user's database and asking questions, for example about the names of the photographed persons or the date the photo was taken. In 2011 Facebook adopted the idea to provide an additional barrier in case a suspicious login is detected. The underlying idea is that an attacker who might have been able to obtain the password for an account does not know the corresponding social graph and is therefore not able to answer the questions correctly.

C. Personal Knowledge Questions

Personal knowledge questions are a form of knowledge-based authentication, that tests already known information, by answering questions about past experiences. The user typically selects the questions during the initial account setup from a predefined list of questions. Some services allow users to create security questions on their own as well. For account reset, the questions need to be answered whereby a certain variation may be allowed to tolerate different spellings.

D. Browser Fingerprinting

Browser fingerprinting is a strategy often used to improve overall account security. The underlying idea is to collect information about the user's browser and location each time the user visits the website and compare it to the data from previous sessions. Assuming an attacker is not able to precisely mimic the formerly used system, it is thus possible to distinguish legit and malicious login attempts. The collected information may include, among other things, IP, user agent, and referrer, as well as, client-side features like language settings, window size, time zone, canvas-, local storage-, and WebGL support, hashes of installed browser plugins and fonts, screen resolution, used OS, and browser version. Due to its intended use case, i. e., being part of an "arms race" between attackers and defenders, not much about browser fingerprinting-based security systems is known [13]. In 2018, Google disclosed not only to collect browser fingerprints but to also monitor user behavior to drive their authentication decision [20].

E. Helpdesk

As a last resort, services sometimes offer an option for users who forgot their password to contact or visit a helpdesk [9]. The high costs of employing support personnel and maintaining helpdesks can lead to reconsiderations of deployed security mechanisms [28], [17]. The authentication via support employees often happens via *soft factors*. This includes personally identifiable information like the name, the address, the date of birth and account usage questions like the date of account registration or parts of the credit card number registered with the account. There are many examples of *targeted* attacks that exploited this particularly insecure way of fallback authentication [16].

IV. Study Design

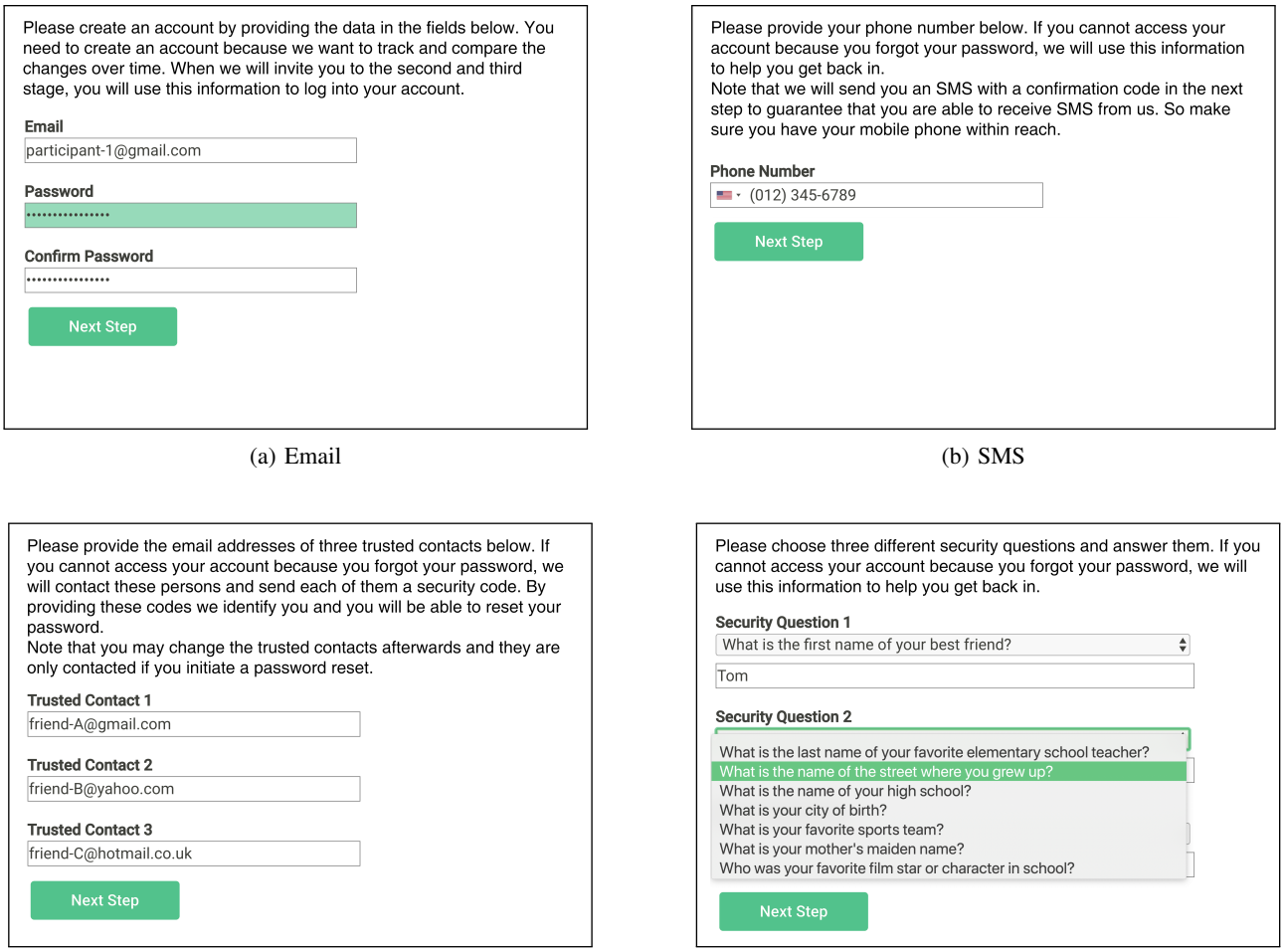
We plan to conduct a long-term user study in order to find answers to our proposed research questions. Subsequently, we explain the design of this study.

A. Selected Schemes & Implementations

Based on the overview given in Section III, we selected five fallback authentication schemes that are tested in our study. An overview is given in Table I, and parts of the different enrollment phases are displayed in Figure 1.

1) Email: The schemes using a secondary channel are among the most widely used fallback authentication schemes, as they can be implemented universally, and typically have reasonable usability, albeit their security can be problematic. The email scheme is probably the most common form of fallback authentication [4]. It is often easy to implement as users provide an email address during account registration, which then can be used for fallback authentication as well.

We use a straightforward implementation of the idea. We inform participants that the email address will be used to invite them for future rounds of the study, so they do not use throwaway email addresses, but we do not send verification emails, and we implemented a feature to update the stored email address. For recovery, the participant needs to state the email address, which is then matched against the database. If the email exists, we send a link to a password reset page.



(a) Email

(b) SMS

(c) Designated Trustees

(d) PKQ

Fig. 1: The pages on which the information for the different fallback schemes is provided. The form of the email scheme shown in Figure 1a is the standard form that all participants have to complete to create their account.

2) SMS: While this approach is also very frequently used, it differs from the email scheme in that a different channel with different security properties is used. More importantly, SMS are typically received on another device with potentially increased security, but decreased usability as another device needs to be accessed. The NIST currently discourages the use of SMS as a *second factor for primary authentication* [14]. Still, it is widely used for fallback authentication in practice.

During setup we ask for the user's mobile phone number, explicitly giving account recovery as the reason. Users may be hesitant giving out mobile numbers to websites for privacy reasons. We expect this to be true for both our study as well as real-world use. For account reset, we send a confirmation code via SMS, which the participant needs to enter in a form to reset the password. We do not disclose the phone number during the reset for privacy reasons.

3) Designated Trustees: We selected the designated trustee scheme for our study in a variant very similar to the one studied by Schechter et al. [31]. During account creation, participants are asked to give email addresses of three friends, where we explicitly state account recovery as the reason. We offer an interface to update the list of trustees. For account recovery, the trustees receive an email with a reset code, and instructions to relay the code to the owner of the account. We explicitly ask them only to pass the code once they verified the participant's identity. For a successful recovery, two out of three codes are required.

Many forms of social fallback authentication can only be implemented by social networks. This form by Schechter et al. [31] can reasonably be implemented by almost any website, and we expect it to share many usability properties with related implementations by social networks. We assume, the main difference being the much less comfortable selection of trustees during account creation and the potential to insert non-existing email addresses into the form. Thus, we verify the existence of the provided address by sending an email to it. In the case the email cannot be delivered, we ask participants to provide other trusted contacts.

4) Personal Knowledge Questions: Due to severe security issues [30], [25], [12] the popularity of this classical form of fallback authentication that was widely used has diminished. Nevertheless, we wanted to understand the usability aspects,

as it is still in use today. During account creation, participants select answers for three personal knowledge questions. We have selected four "classical" questions that were in use for a long time, but are rather easy to guess (cf. [12], [25]):

- "What is your mother's maiden name?"
- "What is your city of birth?"
- "What is your favorite sports team?"
- "What is the name of your high school?"

In addition, we have selected four questions with presumably slightly better security properties, following previous findings [2].

- "What is the name of the street where you grew up?"
- "What is the first name of your best friend?"
- "Who was your favorite film star or character in school?"
- "What is the last name of your favorite elementary school teacher?"

For account recovery, two out of the three registered questions are randomly selected, and the user has to provide correct answers for both of them. We accept answers that have an edit distance of at most one, we ignore capitalization, and special characters, as well as spaces, are removed.

5) Browser Fingerprinting: The browser fingerprinting approach is different from the previous approaches in that it does not require user interaction, and that it is typically used in addition to other factors in the background (cf. [1], [24], [10], [15]). In this study, we collect browser fingerprints from all participants in all steps, and will later evaluate how much additional security the fingerprint will likely give us in a fallback authentication scenario. We collect features including the IP, various client-side browser features, and information about the operating system using a JavaScript library developed as part of a browser fingerprinting study by Pugliese et al. [24].

B. Recruitment

We will recruit participants online, using a crowdsourcing platform such as Amazon Mechanical Turk (MTurk) or Prolific. While MTurk is widely used, Prolific is explicitly set up for academic research studies. Analyses have shown that MTurk yields reasonable results, even though not adhering to strict standards of experimental design [26]. Research has shown that participants from Prolific are more equally distributed across the world, and tend to answer more honest than people from MTurk while the data quality is comparable [22]. An online study is advisable as it provides an environment close to when real fallback authentication occurs.

C. Procedure

The study takes place in three stages: *registration*, a *short-term callback* after about 2 weeks, and a *long-term callback*. Participants are assigned round-robin to one of the four fallback authentication schemes, described in Section IV-A. As stated above, browser fingerprints are collected from all of the contributors, regardless of their assignment. Additionally, participants are assigned to one of three *duration groups* of 6 months, 12 months, and 18 months. These groups differ in the timing of the recall of the fallback authentication scheme and are chosen following findings from Bonneau et al. [2], who measured that 33 %, 50 %, and 75 % of users had started an account recovery after the mentioned periods.

The study is framed as being about long-term performance trends in a mental rotations test (MRT) (cf. Section IV-D). Participants are debriefed after the third stage. Before the study, participants are made aware of the length of the commitment. Participants are compensated for each stage individually, with increasing compensations for the later stages in an attempt to mitigate attrition.

Stage 1: Registration At the beginning of the first stage, participants create an account on the website for which we ask them to provide an email address and a password. Depending on the assigned fallback authentication scheme additional information needs to be provided as part of the registration process. We explain that the long running time of the study may make fallback authentication necessary. Afterward, the contributors complete the primary task for the first time. A demographic questionnaire concludes the first stage.

Stage 2: Two-week callback All contributors are emailed after two weeks to return and complete the primary task a second time. They have to log in using their email address and password combination but are also able to reset their password using the respective fallback authentication mechanism.

A critical challenge in our study is gauging dropout rates. This step is included to remind participants about the study, to select participants who will be more likely to come back after an extended time, and to give further incentives to follow through the entire study. Additionally, it will give us another data point after 2 weeks.

Stage 3: 6/12/18-month callback Depending on their assigned condition, contributors are emailed to return after 6 months, 12 months, or 18 months. When logging in, we force a password reset using the fallback authentication scheme by telling contributors that their password is incorrect regardless of correctness. With this approach, we are able to measure how many contributors successfully log in using their fallback authentication mechanism. Additionally, we can measure how many people got their password right. After logging in, contributors complete the primary task a third time, before we disclose the real purpose of our study. At the end, we ask the contributors to complete a usability questionnaire regarding the reset process.

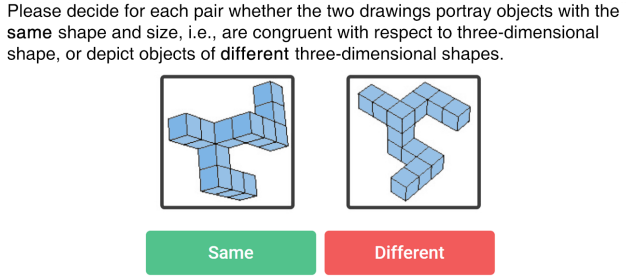


Fig. 2: An example of a mental rotations test (MRT) which is used as a distractor task in the study.

D. Primary Task

We use an MRT [32], [37] as "primary task" in the study (cf. Figure 2). The purpose of the primary task is to distract the contributors from the real purpose of the study and to increase the ecological validity of the authentication task. Framing the long-term nature of the study as being a study of the cognitive ability over time allows us to warn people about the length of the commitment without revealing our interest in the authentication step. The advantage of the MRT is that it gives us a measure of how much attention participants were devoting to the primary task. The MRT is also a strong cognitive distractor, and should suitably prevent participants from remaining focused on the authentication task. The results from our pilot study suggest that this assumption also holds in practice.

E. Metrics

To answer the proposed research questions, we collect various data points. This includes quantitative information such as the success rate, the overall time required, dropout rates, number of profile updates, and issues during the process.

Furthermore, we use a usability questionnaire to gather qualitative data about the fallback process such as user sentiment and perceived effort. This includes questions about the perceived time required to recover their account. Moreover, we will ask questions regarding the perceived effort that was required to reset the account, e. g., copying the reset code from their phone or answering the personal knowledge questions. In the case of a dropout, we also email the participants to ask for additional feedback to learn more about the reasons for not completing the recovery process. This could enable us also to discuss issues related to privacy and trust.

F. Ethical Considerations

In order to be able to receive objective results for our comparison between the different fallback authentication schemes, we try to drive the focus of the participants to the MRT. As our institution does not have a review board, we minimize any potential negative effects by following the ethical principles laid out in the Belmont report [21]. This includes an informed consent and a debriefing that explains the real purpose of the study. Thus, we also conform to earlier research [33] that proved the critical role of debriefings in pretextual studies.

V. Pilot Study

Between December 2017 and January 2018 we conducted a pilot study for which we recruited students from our university. As testing our implementation was the primary purpose of the study, we reduced the time span between the first and second stage to 1 week, as well as the period between stage two and three to 3 weeks. Still, the results enable us to draw some first conclusions even though they are of restricted validity due to the limited extent of this pilot study. The pilot study deviated from the description above in one detail only: The designated trustees were not informed at enrollment time. Thus, we have not checked the existence of the provide email addresses.

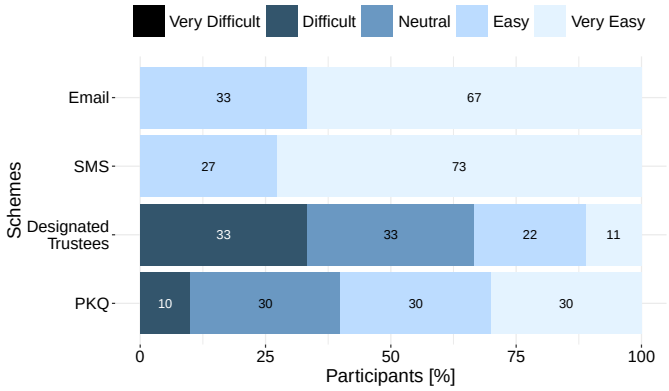


Fig. 3: A comparison that shows how participants of all four fallback authentication schemes would finish the sentence *"Resetting my password was . . . "*. This question is part of the usability questionnaire at the end of the third stage.

A. Results

From 74 participants who started the first stage, 44 completed all three stages. During the first stage, we recorded 12 dropouts, all of them after the assignment to a fallback authentication scheme. While no one from the email group backed out from the study, one participant of the SMS group, four of the PKQ group, and seven persons who were assigned to the designated trustees scheme did. For the recovery process, we observed similar results.

All participants who received a reset link via email completed the account recovery. This is followed by the SMS where 11 out of 12 participants successfully set a new password; in the PKQ case also 11 out of 12 did. From 12 persons of the trustee scheme who initiated the password reset only 9 finished the account recovery.

The best results regarding efficiency were achieved by the email scheme. The average reset took only 34 seconds, and 75 % of the participants described the duration of the password reset as a "Very short time."

The SMS scheme ranked second. The time-span for a reset was 98 seconds on average which is nearly 3 times as long as for the email scheme. The subjective ratings of the participants still suggest that the scheme is efficient to use who describe the time as "Very short" or "Short."

The third most efficient scheme was the one based on PKQ. In this case, two outliers lead to an average reset time of 16 minutes. The ratings from the participants differ as well. Some stated that the password reset took a "Very short time," while the participant whose reset took 2 hours 37 minutes used the term "Long time."

The mean reset time in case of the designated trustees scheme was around 7 hours which is by far the longest time. Here, we need to differentiate two approaches. Participants who acted in the intended way and stated other persons as their trusted took around 14 hours on average to reset the password. The other group used three of their email accounts, and an average reset took around 3 hours due to two outliers. The

rating of the participants is negative as even some participants who stated only to have used their email addresses, instead of email addresses from trustees, needed several hours to reset their passwords.

The subjective impressions of the participants are very much alike. Figure 3 depicts the answers to the question that asked the participants to rate the convenience of the password reset. The email and SMS group describe the process as "Very easy" or "Easy" whereas the answers for the other two schemes tend to be more negative.

B. Discussion

One insight from the pilot study is the importance to reduce the use of made-up email addresses for the designated trustees scheme. Thus, we decided to send enrollment emails to the trustees to check the existence of the address and ask for a new contact if we were not able to deliver the email by the beginning of the second stage. At this point, our implementation differs from previous work by Schechter et al. [31], who has not sent emails to trustees during enrollment.

Another difference with Schechter et al. is the way we interact with the trustees. As described in Section IV-A3 we send the codes in the email along with a warning message explaining potential misuse. Schechter et al. required the designated trustees to complete several steps before they receive a code, among others a pledge, in order to minimize the risk of an account takeover.

For our pilot study, we simplified the complicated procedure which Schechter et al. were able to test in the context of Microsoft's "Live" services. Still, the preliminary results of our pilot study indicate that even with this simplification the designated trustee scheme is the most user burdening and time-consuming.

We also want to stress that we solely focus on the usability of the schemes. Security analyses were carried out in the past [3], [5], [6], [27], [30] and showed that different aspects need to be considered which is why a comparison is difficult.

Furthermore, it needs to be taken into account that the email-based fallback scheme might not be deployable in every case (i. e., email account recovery), although it performs reasonably well regarding usability and security.

VI. Conclusion

In this work, we outline a long-term study that explores the usability of different fallback authentication schemes given realistic recall times of 6/12/18 months. In the study, we like to compare four different schemes, based on email, SMS, designated trustees, as well as PKQs. A preliminary pilot study suggests that the email and SMS scheme are the most usable ones. PKQs and a scheme based on designated trustees fell short, in both efficiency and convenience.

Acknowledgment

This research was supported by the research training group "Human Centered Systems Security" sponsored by the state of North Rhine-Westphalia, Germany.

References

[1] F. Alaca and P. C. Van Oorschot, "Device Fingerprinting for Augmenting Web Authentication: Classification and Analysis of Methods," in *Annual Conference on Computer Security Applications*, ser. ACSAC '16. Los Angeles, California, USA: ACM, Dec. 2016, pp. 289–301.

[2] J. Bonneau, E. Bursztein, I. Caron, R. Jackson, and M. Williamson, "Secrets, Lies, and Account Recovery: Lessons from the Use of Personal Knowledge Questions at Google," in *International World Wide Web Conference*, ser. WWW '15. Florence, Italy: ACM, May 2015, pp. 141–150.

[3] J. Bonneau, C. Herley, P. C. Van Oorschot, and F. Stajano, "The Quest to Replace Passwords: A Framework for Comparative Evaluation of Web Authentication Schemes," in *IEEE Symposium on Security and Privacy*, ser. SP '12. San Jose, California, USA: IEEE, May 2012, pp. 553–567.

[4] J. Bonneau and S. Preibusch, "The Password Thicket: Technical and Market Failures in Human Authentication on the Web," in *Workshop on the Economics of Information Security*, ser. WEIS '10. Cambridge, Massachusetts, USA: ACM, Jun. 2010.

[5] J. Brainard, A. Juels, R. L. Rivest, M. Szydlo, and M. Yung, "Fourth-Factor Authentication: Somebody You Know," in *ACM Conference on Computer and Communications Security*, ser. CCS '06. Alexandria, Virginia, USA: ACM, Oct. 2006, pp. 168–178.

[6] S. Bugiel, A. Dmitrienko, K. Kostiainen, A.-R. Sadeghi, and M. Winandy, "TruWalletM: Secure Web Authentication on Mobile Platforms," in *International Conference on Trusted Systems*, ser. INTRUST '10. Beijing, China: Springer, Dec. 2010, pp. 219–236.

[7] Facebook Security, "Facebook: Introducing Trusted Friends," Oct. 2011, https://www.facebook.com/notes/facebook-security/national-cybersecurity-awareness-month-updates/10150335022240766/, as of September 8, 2023.

[8] ——, "Facebook: Introducing Trusted Contacts," May 2013, https://www.facebook.com/notes/facebook-security/introducing-trusted-contacts/10151362774980766/, as of September 8, 2023.

[9] S. Foo, S. C. Hui, P. C. Leong, and S. Liu, "An Integrated Help Desk Support for Customer Services Over the World Wide Web – A Case Study," *Computers in Industry*, vol. 41, no. 2, pp. 129–145, Mar. 2000.

[10] D. M. Freeman, S. Jain, M. Dürmuth, B. Biggio, and G. Giacinto, "Who Are You? A Statistical Approach to Measuring User Authenticity," in *Symposium on Network and Distributed System Security*, ser. NDSS '16. San Diego, California, USA: ISOC, Feb. 2016.

[11] S. L. Garfinkel, "Email-Based Identification and Authentication: An Alternative to PKI?" *IEEE Security & Privacy*, vol. 1, no. 6, pp. 20–26, Nov. 2003.

[12] M. Golla and M. Dürmuth, "Analyzing 4 Million Real-World Personal Knowledge Questions (Short Paper)," in *International Conference on Passwords*, ser. PASSWORDS '15. Cambridge, United Kingdom: Springer, Dec. 2015, pp. 39–44.

[13] M. Golla, T. Schnitzler, and M. Dürmuth, ""Will Any Password Do?" Exploring Rate-Limiting on the Web," in *Who Are You?! Adventures in Authentication Workshop*, ser. WAY '18. Baltimore, Maryland, USA: USENIX, Aug. 2018.

[14] P. A. Grassi, J. L. Fenton, and W. E. Burr, "Digital Identity Guidelines – Authentication and Lifecycle Management: NIST Special Publication 800-63B," Jun. 2017.

[15] C. Herley and S. Schechter, "Distinguishing Attacks from Legitimate Traffic at an Authentication Server," Microsoft, Technical Report MSR-TR-2018-19, Jun. 2018.

[16] M. Honan, "How Apple and Amazon Security Flaws Led to My Epic Hacking," Aug. 2012, http://www.wired.com/2012/08/apple-amazon-mat-honan-hacking/, as of September 8, 2023.

[17] P. G. Inglesant and M. A. Sasse, "The True Cost of Unusable Password Policies: Password Use in the Wild," in *ACM Conference on Human Factors in Computing Systems*, ser. CHI '10. Atlanta, Georgia, USA: ACM, Apr. 2010, pp. 383–392.

[18] A. Javed, D. Bletgen, F. Kohlar, M. Dürmuth, and J. Schwenk, "Secure Fallback Authentication and the Trusted Friend Attack," in *International Distributed Computing Systems Workshops*, ser. ICDCSW '14. Madrid, Spain: IEEE, Jun. 2014, pp. 22–28.

[19] M. Just and D. Aspinall, "Personal Choice and Challenge Questions: A Security and Usability Assessment," in *Symposium on Usable Privacy and Security*, ser. SOUPS '09. Mountain View, California, USA: ACM, Jul. 2009, pp. 8:1–8:11.

[20] G. Milka, "Anatomy of Account Takeover," in *USENIX Enigma Conference*, ser. Enigma '18. Santa Clara, California, USA: USENIX, Jan. 2018.

[21] National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research, "The Belmont Report: Ethical Principles and Guidelines for the Protection of Human Subjects of Research," Sep. 1978, https://www.hhs.gov/ohrp/regulations-and-policy/belmont-report/index.html, as of September 8, 2023.

[22] E. Peer, L. Brandimarte, S. Samat, and A. Acquisti, "Beyond the Turk: Alternative Platforms for Crowdsourcing Behavioral Research," *Journal of Experimental Social Psychology*, vol. 70, no. 5, pp. 153–163, May 2017.

[23] J. L. Pinchot and K. L. Paullet, "What's in Your Profile? Mapping Facebook Profile Data to Personal Security Questions," *Issues in Information Systems*, vol. 13, no. 1, pp. 284–293, Mar. 2012.

[24] G. Pugliese, Z. Benenson, and F. Freiling, "Study on Browser Fingerprinting," Jul. 2016, https://browser-fingerprint.cs.fau.de/, as of September 8, 2023.

[25] A. Rabkin, "Personal Knowledge Questions for Fallback Authentication: Security Questions in the Era of Facebook," in *Symposium on Usable Privacy and Security*, ser. SOUPS '08. Pittsburgh, Pennsylvania, USA: ACM, Jul. 2008, pp. 13–23.

[26] E. M. Redmiles, S. Kross, , and M. L. Mazurek, "How Well Do My Results Generalize? Comparing Security and Privacy Survey Results from MTurk, Web, and Telephone Samples," in *IEEE Symposium on Security and Privacy*, ser. SP '19. San Francisco, California, USA: IEEE, May 2019, pp. 227–244.

[27] R. Reeder and S. Schechter, "When the Password Doesn't Work: Secondary Authentication for Websites," *IEEE Security & Privacy*, vol. 9, no. 2, pp. 43–49, Mar. 2011.

[28] M. A. Sasse and I. Flechais, *Usable Security: Why Do We Need It? How Do We Get It?*, 1st ed. Sebastopol, California, USA: O'Reilly and Associates, 2005, ch. 2, pp. 13–30.

[29] M. A. Sasse, M. Steves, K. Krol, and D. Chisnell, "The Great Authentication Fatigue – And How to Overcome It," in *International Conference on Cross-Cultural Design*, ser. CCD '14. Heraklion, Crete, Greece: Springer, Jun. 2014, pp. 228–239.

[30] S. Schechter, A. J. B. Brush, and S. Egelman, "It's No Secret. Measuring the Security and Reliability of Authentication via "Secret" Questions," in *IEEE Symposium on Security and Privacy*, ser. SP '09. Oakland, California, USA: IEEE, May 2009, pp. 375–390.

[31] S. Schechter, S. Egelman, and R. W. Reeder, "It's Not What You Know, But Who You Know: A Social Approach to Last-Resort Authentication," in *ACM Conference on Human Factors in Computing Systems*, ser. CHI '09. Boston, Massachusetts, USA: ACM, Apr. 2009, pp. 1983–1992.

[32] R. N. Shepard and J. Metzler, "Mental Rotation of Three-Dimensional Objects," *Science*, vol. 171, no. 3972, pp. 701–703, Feb. 1971.

[33] S. S. Smith and D. Richardson, "Amelioration of Deception and Harm in Psychological Research: The Important Role of Debriefing," *Journal of Personality and Social Psychology*, vol. 44, no. 5, pp. 1075–1082, May 1983.

[34] E. Stobert and R. Biddle, "The Password Life Cycle: User Behaviour in Managing Passwords," in *Symposium on Usable Privacy and Security*, ser. SOUPS '14. Menlo Park, California, USA: USENIX, Jul. 2014, pp. 243–255.

[35] L. Vaas, "Former High School Teacher Pleads Guilty to Hacking Celebrities," Oct. 2018, https://nakedsecurity.sophos.com/2018/10/24/former-high-school-teacher-pleads-guilty-to-hacking-celebrities/, as of September 8, 2023.

[36] T. W. van der Horst and K. E. Seamons, "Simple Authentication for the Web," in *Conference on Security and Privacy in Communication Networks*, ser. SecureComm '07. Nice, France: IEEE, Sep. 2007, pp. 473–482.

[37] S. G. Vandenberg and A. R. Kuse, "Mental Rotations, a Group Test of Three-Dimensional Spatial Visualization," *Perceptual and Motor Skills*, vol. 47, no. 2, pp. 599–604, Oct. 1978.

[38] S. Yardi, N. Feamster, and A. Bruckman, "Photo-Based Authentication Using Social Networks," in *Workshop on Online Social Networks*, ser. WOSN '08. Seattle, Washington, USA: ACM, Aug. 2008, pp. 55–60.

[39] M. Zviran and W. J. Haga, "User Authentication by Cognitive Passwords: An Empirical Assessment," in *Jerusalem Conference on Information Technology*, ser. JCIT '90. Jerusalem, Israel: IEEE, Oct. 1990, pp. 137–144.