Stop to Unlock - Improving the Security of Android Unlock Patterns

Alexander Suchan
SBA Research
alexander.suchan@outlook.com

Emanuel von Zezschwitz
University of Bonn and Fraunhofer FKIE
zezschwitz@cs.uni-bonn.de

Katharina Krombholz
CISPA Helmholtz Center for Information Security
krombholz@cispa.saarland

***Abstract*—Android unlock patterns are among the most common authentication mechanisms on mobile devices. They are fast and easy to use but also lack security as user-chosen gestures are easy to guess and easy to observe. To improve the traditional pattern approach, we propose *Stop2Unlock*, a usable but more secure modification of the traditional pattern lock. *Stop2Unlock* allows users to define nodes where they stop for a limited amount of time before swiping to the next node. We performed a lab study (n=40) and a field study (n=14) to show that this small change in user interaction can have a significant impact on security with a minimal impact on usability. That is, user-selected *Stop2Unlock* patterns are significantly harder to guess while being comparable in terms of usability. Additional analysis showed that users perceived the stop component as a rhythmic and memorable cue which supported the selection of higher entropy patterns.**

I. Introduction

While most commercially available smartphones come with built-in fingerprint sensors or sophisticated face recognition features, PINs and unlock patterns are still indispensable fallback mechanisms whenever probabilistic biometric authentication fails. From a usability perspective however, unlock patterns have many benefits as the sheme is optimized for touch-based devices and fits well in the mobile context (e.g., it is fast and supports eyes-free interaction). At the same time. From a security perspective, low-entropy unlock patterns can pose a major vulnerability to the potentially sensitive user data stored on the device as they are prone to observational and guessing attacks.

As no other mechanism has managed to fully replace unlock patterns, we follow the recommendations by Bianchi et al. [5] to develop iterative improvements of the well-established authentication methods to improve the security of today's smartphone ecosystem, and to build on already optimized usability features. Similar to De Luca's XSide [8] and Krombholz' et al. Force-PINs [18] we propose an improvement for gesture-based unlock patterns as found in commercially available Android phones. However, in contrast to most previous work which mainly focused on the prevention of smudge attacks (e.g., [19]) or observation attacks (e.g., [10], [24]), we aim at increasing the practical password space of such pattern concepts. As primary probabilistic biometric authentication is generally resistant towards shoulder surfing and smudge attacks, it has become increasingly important to harden knowledge-based fallback authentication against guessing attacks [7]. Our mechanism, *Stop2Unlock*, does not require additional hardware and thus works with currently available off-the-shelf smartphones. We show that the concept, which introduces a practically invisible stop component, supports the selection of a diverse set of highly memorable unlock patterns. Figure 1 shows an example of *Stop2Unlock* with one stop node (marked in red).

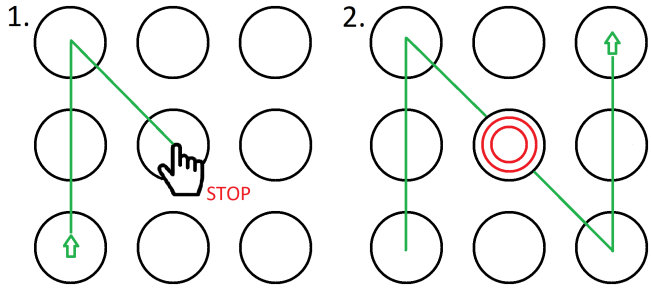

Fig. 1. An exemplary unlock process for *Stop2Unlock*, enhancing the traditional unlock pattern with a *stop* component (marked with a red circle).

To prove the feasibility of our concept, we performed two user studies. The results of our lab study and our field study suggest that *Stop2Unlock* can improve security with only a minimal impact on usability. In particular, we found that the minor modification of adding a time component to the interaction supported users in memorizing *Stop2Unlock* patterns and thus led to a diverse set of high entropy credentials which were actually used over the course of several days.

The contributions of this paper are:

- We propose *Stop2Unlock* as an enhancement of traditional unlock patterns. Our approach enables users to stop at 1 to n nodes in their unlock pattern.
- We implemented a prototype application and performed an evaluation of *Stop2Unlock* in a lab study with 40 participants and a field study with 14 participants.
- The collected data shows that Stop2Unlock nudges users to use high entropy patterns that perform similarly well in terms of usability metrics compared to the less-secure traditional patterns.

Workshop on Usable Security (USEC) 2019
24 February 2019, San Diego, CA, USA
ISBN 1-891562-57-6
https://dx.doi.org/10.14722/usec.2019.23017
www.ndss-symposium.org

II. Related Work

A. Unlocking Behavior and Attacks

According to Harbach et al. [13], participants spend on average 2.9% of their smartphone interaction time authenticating. This result suggests that any security improvement for unlock patterns should keep the additional usability overhead as low as possible. Observation-based attacks have been found to be a major problem in the mobile context. However, Eiband et al. [11] studied actual stories about shoulder surfing on mobile devices from both users and observers and found that such observation attacks mainly occur in an opportunistic, non-malicious way without any serious consequences. In addition to observation attacks (e.g., [26], [21]), prior research has investigated several alternative threat models for mobile unlock mechanisms. E.g., Aviv et al. [4] investigated *smudge attacks* on smartphones with Android unlock patterns and were able to detect 68% of the patterns in their simulated scenarios. Abdelrahman et al. [1] investigated thermal attacks and showed that heat traces left on the screen can be used to successfully guess a secret if the attack was performed immediately after a successful authentication session. Krombholz et al. [17] presented a microbiological attack based on bacterial growth but were unable to show that the attack can be conducted successfully.

In addition to practical interaction-based attacks, several studies have found that pattern selection is highly predictable and thus easy to guess (e.g., [2], [22]).

B. Improvements

Currently there are different suggestions for improving the security of the authentication process for mobile devices. Most concepts change the way of interaction to protect user input from observation attacks [5]. For example, De Luca et al. [9] presented *ColorPIN*, an authentication mechanism that uses indirect input to provide security-enhanced PIN entry. *ColorPIN* is a combination of a standard PIN with a color, for example 1 (black) - 2 (red) - 3 (white) 4 - (black). *ColorPIN* was significantly stronger than standard PINs while enabling good authentication speed. De Luca et al. [8] also presented *XSide*, an authentication mechanism which uses the front and the back of a smartphone to enter patterns. *XSide* lets the user draw simple shapes or gestures, i.e. is a system that can be used eyes-free and provides increased protection against smudge or shoulder surfing attacks. They also performed a user study to show the effects of switching sides during authentication on usability and security. Song et al. [20] proposed a strength meter to help users to find a stronger pattern and found that about 10% of the patterns generated without the meter could be compromised through 16 guessing attempts. Zezschwitz et al. [24] presented *SwiPIN*, an approach that combines traditional PINs with simple touch gestures. They showed that *SwiPINs* were significantly more secure against shoulder surfing attacks, while being perceived as easy-to-use as PINs. Krombholz et al. [18] developed a new type of PIN with the use of Apples' *3D-Touch*. These so-called *Force-PINs* let users select higher entropy PINs with a pressure component. However, Khan et al. [15] re-implemented Force PINs and conducted a comparative study on countermeasures to shoulder surfing attacks and found that most improvements offer limited protection. Our work extends the state of the art by adding a tactile component to unlock patterns. In this paper, we show that user-selected *Stop2Unlock* patterns are more resistant to guessing attacks but remain usable in the wild.

III. Concept and Prototype

Stop2Unlock (Figure 1) was designed to improve the traditional unlock pattern provided by the standard Android environment. Our method is designed to motivate users to select higher-entropy patterns that are harder to predict while keeping them usable and memorable. Our approach does not require additional hardware. The user can stop at each node on the grid for a short time or run through it quickly when entering the unlock pattern. For better memorability and to train muscle memory (cf. [16], [14]), the user receives a subtle vibration feedback when the application recognizes a *stopping* point. Besides that, the rules for user-selected patterns remain the same, as described by Uellenbeck et al. [23] (i.e. at least four connected nodes and already connected nodes cannot be connected again). *Stop2Unlock* offers a larger pattern space by design. For every existing pattern, the stop component expands the space by 2^k, whereby k is the length of the pattern. Considering the smallest pattern with four visited nodes, the number of possible combinations grows by 2^4=16. We developed a prototype application on Android which simulates a login screen with *Stop2Unlock* to collect data for our evaluation. The application lets users select a new pattern and presents an unlock screen. The user has up to three attempts to unlock the phone with the correct pattern. For comparison with the traditional unlocking method, we also implemented a simulation for standard Android pattern unlock. Additionally to the lab study application we developed a slightly different application for our field study which offers notifications to remind study participants to enter their patterns.

A. Threat Model

We assume that the attacker has physical access to a phone which is protected by a biometric authentication mechanism. The attacker activates the fallback authentication mechanism (i.e., unlock patterns) and tries to unlock the device by guessing the used pattern. As most users select simple and easy-to-predict secrets with low entropy [20], an attacker is likely succeed within a small number of guesses. To improve security in such a scenario, we propose *Stop2Unlock*. Due to the introduced stop elements, *Stop2Unlock* offers a higher theoretical pattern space by design. Additionally, we expect that the time element will be actively used and thus the effective password space will be increase compared to traditional patterns. Please note that we deliberately excluded observation-based attacks from our threat model. We argue that the rise of biometric authentication concepts as a primary unlock mechanisms provides a usable solution which can effectively reduce the risk of such attacks in the wild. However, we would assume that the vulnerability of *Stop2Unlock* would be comparable to traditional patterns.

B. Pilot Study

We conducted a small pilot study with five participants to determine the optimal stop time. The goal was to find a stop time that only minimally compromises usability. All

participants used Android unlock patterns before and had to enter three pre-defined patterns with different stopping times. We selected three different patterns with distinct difficulties. For each pattern and each stopping time we measured failed attempts and the mean authentication time for correct and failed attempts.

Our first tests revealed that the best stopping time is between 200ms and 400ms. We therefore repeated the study with stopping times of 450ms, 350ms and 250ms. Almost all participants reported that they were most comfortable with a stopping time of 350ms. Reasons for that were the increased error rate at the edges of the pattern with slower stopping times, and the overall increased authentication time with slower stopping times. This was confirmed via our collected data: out of all tested stopping times, 350ms had resulted in the fewest failed attempts. As expected, the authentication time increases in accordance with increasing the stopping time and was for all three tested times around 2000ms.

Demographic	Number	Percent
Gender		
Male	21	52.5%
Female	18	45%
No Information	1	2.5%
Age		
Min.	20	
Max.	57	
Median	27	
Mean	29.73	
Used Smartphone		
Android	23	57.5%
iPhone	16	40%
Other	0	0%
None	1	2.5%
Used Authentication Method		
4-digit PINs	25	62.5%
Password (character and digit)	3	7.5%
Unlock Pattern	9	22.5%
Fingerprint Sensor	19	47.5%
Face Unlock	0	0%
None	1	2.5%

TABLE I. PARTICIPANT DEMOGRAPHICS (LAB STUDY, N=40)

IV. LAB STUDY

We conducted a lab study to compare *Stop2Unlock* and traditional Android unlock patterns. Our comparative evaluation is based on the following metrics: (i) authentication time, (ii) error rate, (iii) memorability, and (iv) entropy. These metrics have been used in related work, e.g., [18], [24], [8], [25]. We did not assign unlock patterns but, similar to real-world scenarios allowed participants to choose their own *Stop2Unlock*, which we then used for our security evaluation including entropy calculations.

A. Design and Procedure

Every participant was exposed to two conditions in random order to minimize learning effects. The conditions were as follows: (C1) traditional unlock pattern with the same rule set as in the standard Android environment, and (C2) Stop2Unlock. Participants (with academic and non-academic backgrounds) were recruited in front of the university cafeteria and via mailing lists and social networks. All experiments were conducted in Vienna, Austria. In total, our sample has 40 participants (Table I). The participants were mostly familiar with PINs and unlock patterns. We refrained from studying memorability in the lab study as the setting is not suitable to provide ecologically valid results. Participants were informed that the study was about authentication. We did not use the words usability and security in combination with the authentication methods in order to keep the results unbiased as far as possible. Participation was voluntarily and participants were not compensated. Each session started with a short briefing about the purpose of the study. Our application generated a unique ID for every participant. We did not provide any details on benefits and drawbacks of the two methods to keep the results unbiased. Then, the participants had time to familiarize themselves with the two authentication methods. This short training session, was used to reduce the effects of an unfair comparison: a well-known and well-trained condition against a newly introduced system. The order of conditions was randomized and for each of the conditions participants had to define a new pattern/*Stop2Unlock* pattern; we instructed the participants to select a pattern memorable yet secure pattern. Finally we simulated three authentication sessions during which the user had to correctly authenticate. The metrics were (i) authentication speed and (ii) error rate as defined by De Luca et al. [8]. Concerning *error rate* we distinguished between basic and critical errors. A *basic error* is an overall successful authentication session in which it took the user one or two tries to enter the pattern correctly. A *critical error* on the other hand is a completely failed authentication session, i.e. three erroneous attempts. We measured the *authentication speed* from the first to the last touch of an authentication session. An authentication session was initiated by a participant pressing a button. The session ended after the correct pattern was successfully entered or after three failed attempts.

In addition to the data collected through the Android application we asked the participants a series of questions about the tested authentication system. Apart from demographic data about the participant (age and gender) we did not collect any personal information. In order to link the technical data to the questionnaire, we assigned each participant an id. The questionnaire can be found in the appendix A.

Ethical Considerations: Our institute located in central Europe does not have a formal IRB process but a set of ethical guidelines to follow. Our data collection and processing was compliant with these guidelines and the GDPR. We limited the collection of personal information to a minimum. All participants signed consent forms and agreed to participate voluntarily.

B. Results

Our sample includes 40 participants who under two conditions had to complete three successful authentication sessions. Hence our quantitative results are based on 40 * 2 * 3 = 240 authentication sessions.

1) Authentication Speed: We only considered successful authentication sessions, i.e. all sessions with a maximum of two failed authentication attempts. In order to reduce the impact of measurement errors, we removed two outliers from our sample with authentication times larger than 14 and 21 seconds, since they occurred because participants were distracted from the study task. All collected *Stop2Unlock* patterns were unique in our sample (i.e. 40 different patterns); concerning

traditional patterns, *1-2-3-5-7-8-9* was selected four times and *1-4-7-8-9* twice (i.e. we have 36 different patterns in total).

Both samples (C1/C2) are normally distributed and have a homogeneity of the variances. Further, by study design the two conditions are independent from one another. In Table II we show the mean authentication speed and error rates for both conditions. An independent sample t-test suggests significant effects regarding the difference in authentication time between *Stop2Unlock* and the traditional unlock pattern (t(40) = 5.4117, p ¡ 0.05)).

Authentication Speed	**Mean**	**SD**
StopUnlockPattern	2.79	1.80
Traditional Unlock Pattern	1.32	0.76
Errors	**Basic**	**Critical**
StopUnlockPattern	9	1
Traditional Unlock Pattern	5	0

TABLE II. MEAN AUTHENTICATION TIME IN SECONDS, BASIC AND CRITICAL ERRORS (LAB STUDY).

As our study compares a novel modification of unlock patterns to well-established traditional patterns, our participants were not sufficiently trained to provide an ecologically valid comparison. Assuming training effects over time, we estimate a lower bound for authentication speed. According to Harbach et al. [12] the average duration of a successful unlocking session for traditional Android patterns is 0.91 seconds. Our solution adds 350ms for every stop node of a selected pattern. Hence, our approach adds an average of 0.76 seconds solely due to the additional stopping component. Therefore, the lowest possible authentication time for *Stop2Unlock* is approx. 1.67 seconds.

2) Error Rate: As shown in Table II, out of all 240 authentication sessions only one session with *Stop2Unlock* failed. Furthermore 9 (3.75%) failed attempts (basic errors) occurred with *Stop2Unlock* and 5 (2.08%) with traditional unlock patterns.

3) Perceived Usability and Security: To find out how the participants perceived this new authentication method, we asked them a series of questions. We asked our participants if they would like to use *Stop2Unlock* on their smartphones. As expected, Android users were more interested in using the method in comparison to iOS users. More precisely, 70% of the participants with Android devices stated that they would use *Stop2Unlock* (9% don't want to use it, and 22% said maybe). In comparison, only 56.25% of participants with iPhones or other smartphones stated that they would use *Stop2Unlock*, 18.75% don't want to use it, and 25% maybe).

To determine how participants perceived *Stop2Unlock* compared to other authentication methods, we asked them to rank several methods, including *Stop2Unlock*, according to authentication speed and security (Table III). Interestingly, our participants perceived the fingerprint sensor as the securest and fastest authentication method. However, some users stated that the fingerprint sensor is only the fastest if the recognition works on the first try; often the users need two or more tries to unlock the phone, and in that cases the PIN was perceived as faster. Regarding security, most of the users thought that *Stop2Unlock* is more secure than traditional unlock patterns and 4-digit PINs. When considering authentication speed, participants perceived *Stop2Unlock* as slower than traditional unlock patterns and the fingerprint sensor. Finally, most users are undecided about the unlock speed regarding *Stop2Unlock* and 4-digit PINs.

	Participant Votes			
Method	First	Second	Third	Fourth
Security				
Fingerprint Sensor	**26**	6	4	4
StopUnlockPattern	7	**21**	12	0
4-digit PIN	7	11	**13**	9
Unlock Pattern	0	2	11	**27**
Auth. Speed				
Fingerprint Sensor	**30**	5	2	3
Unlock Pattern	4	**21**	13	2
StopUnlockPattern	2	5	**14**	**19**
4-digit PIN	4	9	**11**	**16**

TABLE III. USER-BASED ORDERING OF AUTHENTICATION METHODS ACCORDING TO SECURITY AND AUTHENTICATION SPEED.

All this votes are of course highly subjective; the questions were intentionally asked in a very open way to give the participants room to interpret the security and speed based on their own observations and experiences. Nevertheless, we think this question gives valuable insights into perceptions of authentication methods.

4) Informal Participant Statements: Ten participants explicitely mentioned to like the haptic (vibration) feedback when the system recognizes a stop at one node and perceived the stop component as security feature. 20 participants thought that this system is not easily observable by other persons.

"This new pattern includes a somehow rhythmic component, this maybe helps some people to memorize their patterns." (ID-94)

We also received negative feedback and gathered suggestions for improvement; "*People who already have problems using their smartphone (or are overwhelmed with the functionality), would have problems with this unlocking method.*" (ID-103)

V. FIELD STUDY

In order to see how *Stop2Unlock* performs in a real-world setting, we conducted a field study. We installed a prototype app on the participants' private Android phones smartphones to measure authentication time and error rate over time. In contrast to previous work which investigated password choice in online studies [], we opted to assess selection strategies in a longitudinal field study. We argue that this approach supports the ecological validity of the findings as selected patterns had to be actually used multiple times over a longer period of time. In addition, we studied the memorability of *Stop2Unlock*.

A. Design and Procedure

We recruit 14 Android users from sample from our lab study and installed a dedicated app on their smartphones and explained the goals of our study, i.e., to measure the performance of *Stop2Unlock* different situations over a longer period of time. We asked the participants to enter their *Stop2Unlock* pattern at least three to five times a day. We did not mention that we were going to study memorability in the course of the field study. To remind the participants to use *Stop2Unlock* multiple times throughout the day, we issued notifications three times a day (morning, midday, and afternoon) and encouraged

them to enter their pattern as often as possible. After installing the app, our participants chose a new pattern. We instructed them to select a pattern that is secure, easy to remember and has at least one stop node. The field study was conducted over a period of two weeks. We expected our participants to complete between 100 and 200 successful authentication sessions. In case the participant forgot their patterns, or wanted to change it for whatever reason, they were provided an option to do so. We also gave them an option to e-mail us in case they had any problems or comments. We opted to deploy a separate app and refrained from modifying the standard lock screen as we didn't want to change the preferred security settings of our participant's phones as some participants were concerned about this at the beginning of the study. These participants would not have taken part if the application had overridden their selected unlocking method. For the evaluation, we used the same metrics as in our lab study(Section IV-A).

After the study period, we interviewed about their experiences during the field study to learn about problems in specific everyday situations, for example when using public transportation. We also asked them how often and why they decided to change their patterns. We gathered feedback about the stopping time, i.e., if the stopping time is perceived as too long after a training period of two weeks. All interview questions can be found in appendix B. Additionally we conducted a memorability experiment where we asked participants approx. two months after completion of the field study if they could remember their selected pattern.

B. Results

Overall, the 14 participants (P1-P14) completed 3,223 authentication sessions. Every participant completed 230.21 authentication sessions on average, i.e. about 16 sessions per day. In total, the participants produced 405 basic errors and 30 critical errors (Table IV). Furthermore, we included the error rate for basic and critical errors, pattern length, the stops from every selected pattern and the completed study days, i.e. the days on which a participant entered the pattern at least one time. Participants P5, P11 and P13 changed their pattern during the field study.

As requested, most participants distributed their authentication sessions as evenly as possible over the period of two weeks. Figure 2 shows the distribution of successful authentication sessions per day. Some participants exceeded the requested study time as well as the requested number of authentication sessions, whereas others skipped a few days. In the following, we list the participants who did not enter the pattern on more than one day in a row. Subject P3 and subject P12 skipped six days at the end of the study time due to failed notifications. Subject P5 skipped four days and subject P7 three days. We removed 14 authentication sessions due to measurement errors. In order to detect these measurement errors, we calculated the authentication time without the predefined time for every stop and removed outlying authentication sessions, i.e. those longer than 10 seconds. The mean authentication speed over all successful authentication sessions was 2.41 seconds (median=2.06, SD=1.56), which is an improvement over the results from the lab study and suggests training effects over time. The shortest authentication session was 0.696 seconds, the longest was 11.59 seconds.

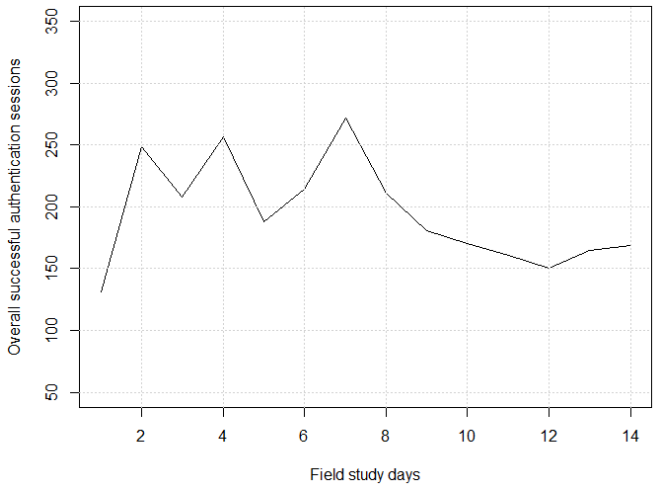

Fig. 2. Successful authentication sessions over time.

To visualize the development of authentication time and error rate, we grouped the results in 16 bins (one bin per anticipated study day, the average study duration was 15.5 days) to simulate a trend and make the results comparable. Three participants changed their patterns halfway through the field study and increased the number of stops, which yields an increased overall authentication time. Consequently, we split the results according to participants who changed the pattern (interrupted lines) and participants who did not change the pattern (solid line). In the first sample (participants who changed patterns) we considered every pattern change as a new start of the study. Despite the user being already familiar with the new unlocking system, a newly selected pattern requires additional learning time. We are confident that this solution utilizes both samples in a comparable and clean way without loosing the data from participants who changed their patterns. Figure 3 shows the development of the authentication time

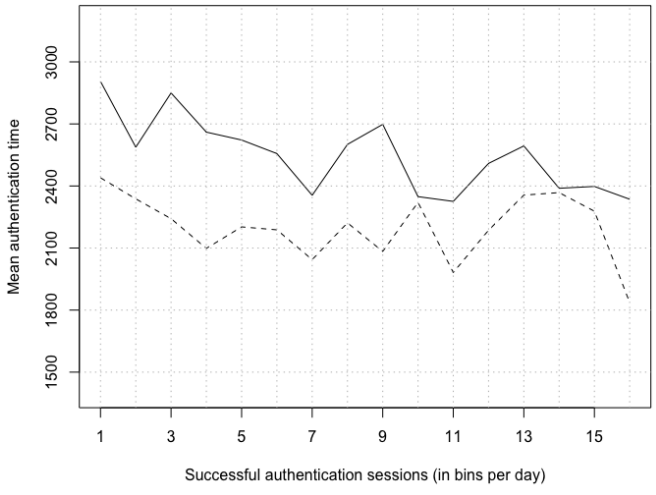

Fig. 3. Authentication time development based on the successful authentication sessions across all participants.

across 16 bins. Our results suggest that the mean authentication time decreases over time. Figure 4 shows the error rate across the study duration. For participants who did not change their

Subjects	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	P11	P12	P13	P14
Completed Unlock Sessions	183	235	212	271	155	157	200	104	180	190	502	100	571	163
Basic Errors	47	25	23	25	43	25	30	15	15	5	48	14	38	52
Basic Error Rate (%)	25.7	10.6	10.8	9.2	27.7	15.9	15.0	14.4	8.3	2.6	9.6	14.0	6.7	31.9
Critical Errors	3	0	1	1	4	4	0	1	0	0	2	0	1	13
Critical Error Rate (%)	1.64	0.00	0.57	0.37	2.58	2.55	0.00	0.96	0.00	0.00	0.40	0.00	0.18	7.98
Pattern Length	14	8	8	9	6 8	5	9	8	7	10	6 12 6	12	7 10	11
Stops in Pattern	5	1	1	2	1 3	1	2	1	2	3	1 4 2	5	1 2	4
Completed Study Days	17	15	9	16	11	16	12	18	12	16	12	11	16	13

TABLE IV. DETAILED RESULTS FOR ALL FIELD STUDY PARTICIPANTS. N=14

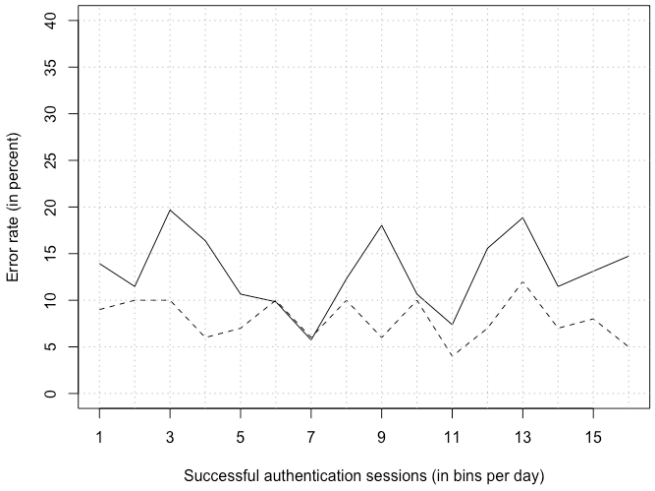

Fig. 4. Error rate development based on the successful authentication sessions across all participants.

patterns, the error rate varies between 20% and 5%. For participants who changed their patterns the error rate varies between 11% and 4%.

1) Debriefing Questionnaire: All field study participants also completed the debriefing questionnaire consisting of six questions. Most of these questions where open-ended to encourage them to express their opinion.

We again asked the participants if they would use *Stop2Unlock* to secure their own smartphones after having used the method on a daily basis in a real-world setting. Seven users said they would use this unlocking method, four said maybe , and three claimed that they would not like to use it. Eight participants reported that they perceived a decrease in authentication time; the remaining six did not perceive any difference compared to the beginning of the study.

Three participants changed their pattern during the field study, and we asked them why they decided to do so. Two participants stated that they wanted to try out a pattern with more than one stop node; another participant just wanted to try something different. One participant tried to change the pattern, but immediately changed it back without having it entered correctly but did not provide reasons for this behavior.

We also asked about problems occurred during the authentication process in the real world (for example when using public transportation). A few users stated that they sometimes had problems hitting the stopping point, i.e. they did not stop exactly on the point and the stop was not recognized by the system. Six people reported that they experienced these problems more often while they were on public transport or busy with something else.

"*Sometimes I had difficulties when entering the pattern with a wet display, gloves, or in jerky environments (fast walk, bus ride,...) in comparison to the traditional version. I think I make more mistakes in these situations.*" (P4)

"*Like with traditional patterns, I often visited the wrong nodes, but I had no problems with the additional stopping feature.*" (P1)

We furthermore asked our participants if they thought that the stopping time should be reduced, increased or remain as is after the two-week trial. Most of the participants stated they where more or less satisfied with the selected stopping time. One participant requested a feature to select an individual stopping time. More precisely, eight users thought the stopping time is well selected, five said the stopping time could be shorter, and one person wanted a longer stopping time.

C. Memorability

To assess the memorability of the user-chosen *Stop2Unlock* patterns, we sent out emails approx. two months after completion of the field study and asked the participants whether they still remembered their patterns. In order to prevent the participants from writing their pattern down, we did not tell them that we would contact them again in this regard.

13 out of 14 participants immediately replied to our request and sent us their selected patterns; nine participants correctly remembered their complete *Stop2Unlock* patterns. Four participants only partially remembered it. One user was not able to recall the last two nodes, but managed to draw the rest of the pattern, including all stopping points. Another user drew one pattern in a mirrored way and the other pattern correctly, but marked the wrong stopping point. Finally, two participants were not able to draw the pattern from memory, but were able to locate the stopping nodes when shown the pattern. Our results suggest that *Stop2Unlock* secrets are memorable, even after two months not using them. Even complex patterns with up to five stopping points were remembered even two months after they stopped using them.

VI. Security Evaluation

Based on the collected data from our lab and field study, we evaluated the security of this new authentication method with respect to our threat model (Section III-A) This includes the evaluation in relation to the practical and theoretical entropy of *Stop2Unlock* as well as the basic pattern space. In this section, we additionally want to discuss the impact of pattern length on mean authentication time and error rate.

A. Entropy

Password entropy measures the probability distribution of passwords over the entire search space. These measures are based on different mathematical models, for example the guessing entropy measures the average number of guesses that an optimal attack needs in order to find the correct password [23].

Obviously, the password space of *Stop2Unlock* is larger than for traditional patterns, by design. One can simply enumerate all possible combinations, resulting in 389.112 2^{19} [23], [4] possible patterns. For every existing pattern of length k the stopping component expands these possibilities by 2^k, whereby k is the length of the pattern. Considering the shortest pattern with four visited nodes, the possibilities grow by factor 2^4=16.

As explained by Cherapau et al. [7] zero-order entropy measures the entire search space of all possible secrets of a given length and the size of a given alphabet. Zero-order entropy assumes that each character is randomly selected and represents the effort an attacker needs to spend on guessing it. Zero-order entropy is measured in bits and calculated as $L * log_2(N)$, whereby L is the length of the secret and N the size of the character set. For *Stop2Unlock* we have a length between 4 and 9 connected nodes and a character set of 18 (9 times 2, for stop/no stop). Therefore, the lower bound for zero-order entropy is 16.68 bits, and the upper bound is 37.53 bits. While traditional unlock patterns have a zero-order entropy between 12.68 bits and 28.52 bits, 4-digit PINs have an entropy of 13.28 bits [7].

In theory *Stop2Unlock* is an improvement over traditional unlock patterns because of the extended pattern space and the higher zero-order entropy. However, this statement assumes that user-chosen patterns are equally distributed across the pattern space, which barely holds in practice. Real-world patterns are distributed over a much smaller search space and are therefore easier to guess. To provide a rough estimate, we calculated the partial guessing entropy gain of our stopping component based on the collected patterns (n = 58) from our field and lab study. None of our participants used the same pattern for both the lab and the field study, which is why we use these datasets together.

1) Partial Guessing Entropy: In contrast to the guessing entropy, the partial guessing entropy [6] (or α-guesswork) calculates the average number of guesses that an attacker needs to break a certain fraction of passwords. We calculate the partial guessing entropy according to by Uellenbeck et al. [23]. As our dataset is also too small to assess the practical entropy, we focus on the engropy gain of the stop component isolated from the respective nodes.

Let $\mu_\alpha = \min\left\{j \mid \sum_{i=1}^{j} p_i \geq \alpha\right\}$ be the minimal number of guesses to cover at least a α fraction of the patterns (i=1 is the pattern with the highest probability). While $\lambda_\alpha = \sum_{i=1}^{\mu_\alpha} p_i$ represents the actual fraction covered, which is greater or equal to α. The partial guessing entropy is defined as follows:

$$G_\alpha(X) = (1 - \lambda_\alpha) \cdot \mu_\alpha + \sum_{i=1}^{\mu_\alpha} i \cdot p_i \tag{1}$$

In equation 1 the first term represents the values that were not guessed in the given fraction, and the second term represents those that were guessed [23]. We express entropy in *bits* to make this estimate comparable with other measurements. This can be achieved with equation 2.

$$\tilde{G}_\alpha(X) = \log\left(\frac{2 \cdot G_\alpha(X)}{\lambda_\alpha} - 1\right) + \log\frac{1}{2 - \lambda_\alpha} \tag{2}$$

Table V shows the probability of user-selected *Stop2Unlock* patterns from the lab- and field study (n = 58). We omitted patterns that were selected only once within our sample. We included patterns from both studies, participants who completed both the lab and field study were allowed to re-use the pattern from the lab study for the field study.

Stop Pattern	Number	Probability
- - - S - - -	4	0.069
- - S S S - -	4	0.069
- - S - S - -	3	0.052
S - - - -	3	0.052
S - - - S	3	0.052
- - - - S - -	2	0.034
- - - - S - S	2	0.034
- - S - -	2	0.034
- - S S S	2	0.034
S - - - - - S	2	0.034
S - S - S - S	2	0.034
S S S S - - - - S	2	0.034
...	...	...

TABLE V. Stop patterns as chosen by the participants in the lab- and field study. N = 58 user-selected patterns, whereby S represents the stopping node. For reasons of brevity, all patterns with less than 2 occurrences were omitted.

Based on the probabilities in Table V, we calculated the partial guessing entropy. The results of our calculations for $\alpha = 0.50$ are as follows: $\mu_\alpha = 12$, $\lambda_\alpha = 0.532$, $G_\alpha(X) = 2.973$, $\tilde{G}_\alpha(X) = 1.3423$ bit. This means that the stop component offers an additional partial guessing entropy of 1.34 bit. According to Aviv et al. [3], the self-reported real-world 3x3 patterns offer an entropy of 9.94 bits (with $\alpha = 0.50$). Together with our estimated entropy gain of 1.34 bits, we are close to the entropy of 4x4 patterns which is 11.61 bits, according to Aviv et al. [3].

B. Pattern Length

In this section, we discuss the impact of pattern length on mean authentication time and error rate. *Stop2Unlock* adds overhead due to the stopping nodes. E.g.,3 stops add a minimum stopping time of 1.05 seconds. Due to this timing component, we evaluate at which point the selection of multiple stopping nodes becomes infeasible.

Figure 5 shows the mean authentication time with respect to pattern length. Our pattern length considers all selected

nodes including stopping nodes, e.g.,*1-2L-3L* equals a length of 5. The mean time is based on all 58 selected patterns from both studies and the first three successful authentication attempts (n = 173). Harbach et al. [12] found that the average pattern

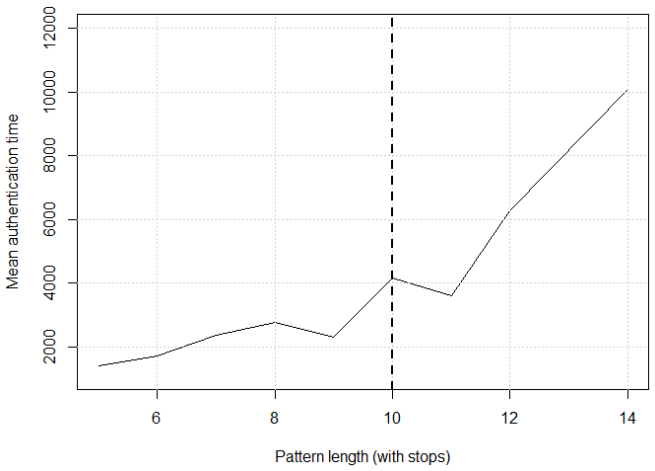

Fig. 5. Authentication time development based on the length of the selected pattern (n = 173).

length for traditional unlock patterns is 5.9 nodes and the average authentication time is 0.91 seconds. They also found that every additional node increases the successful login time by 147ms on average. We argue that a pattern length larger than 10 (including stopping points) is therefore infeasible. In this case the mean authentication time reaches 4 seconds which is way over the average authentication time for traditional patterns. The average number of selected nodes in our sample is 6.2, which is slightly more than for traditional patterns. Figure 6 shows a similar representation of the error rate. With respect to pattern length we not measure higher error rates. The observed error rate varies between 0.0% and 5.0%, whereby the patterns with a length of 4 to 7 nodes have an error rate under 2.5%.

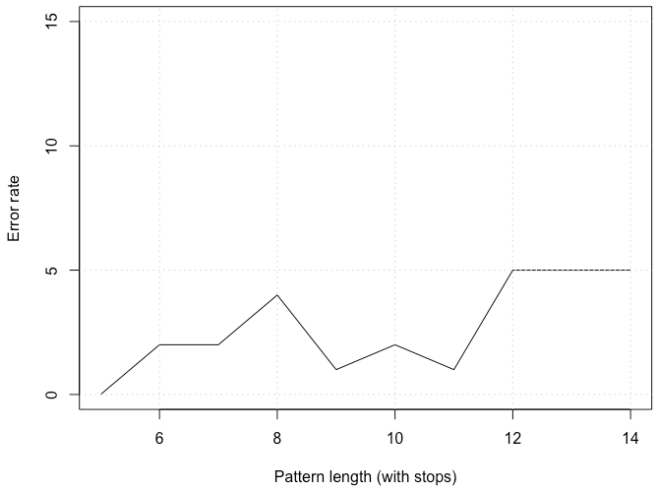

Fig. 6. Error rate development based on the length of the selected pattern (n = 173).

In summary, pattern length has an impact on authentication time which should be reflected in length restriction of *Stop2Unlock*.

VII. Discussion and Limitations

In the following, we discuss key observations and limitations for *Stop2Unlock* based on our studies. During the lab- and field study no participant forgot their pattern, which suggests that *Stop2Unlock* secrets are memorable. This is a particularly important feature as we expect that *Stop2Unlock* will mainly serve as a fallback mechanism for biometric solutions. Another interesting observation is that participants perceived the stop nodes as "rhythmic components" which improved user experience and additionally tackled muscle memory and therefore supported implicit learning.

Stop2Unlock allows users to select 1 to n stopping points per pattern. Some participants selected very long patterns with many stopping points resulting in exceedingly long authentication times and increased error rates. While this is a major shortcoming of *Stop2Unlock* when it is used as primary authentication mechanism, the use of long and complex patterns is a desired user behavior for fallback mechanism (which are used less frequently). Since *Stop2Unlock* might make the login process more difficult in specific situations, for example when the user is exposed to environmental constraints; e.g., authenticating while driving a the car or on a hands-free device, we generally recommend this concept as a secondary fallback mechanism.

As already pointed out, we did not measure the risk of observation-based attacks like shoulder surfing or smudge attacks as we expect that users will rely on observation-resistant biometric approaches in their daily lives. While we assume that *Stop2Unlock* are as vulnerable to observation attacks as traditional unlock patterns, this needs to be shown in future work. Another limitation is that our dataset of 58 patterns is relatively small and therefore not sufficent to fully determine the practical entropy of *StopUnlock*. Our results should thus be treated as a rough indicator.

Finally, our sample of study participants is heavily biased in terms of education and economic status. Thus, our results might not be generalizable to other populations.

VIII. Conclusions

In this paper we presented *Stop2Unlock*, a concept which adds a practically hidden stopping component to traditional Android unlock patterns. This timing concept enables the user to select higher entropy patterns with minimal impact on usability. We evaluated *Stop2Unlock* through a lab study and a field study in terms of usability metrics such as authentication time, error rate and memorability, and security metrics with respect to our threat model (i.e., guessing attacks). Fast authentication times (2.97 seconds), a low error rate (3.75% for basic errors) combined with good memorability suggest that our design is a feasible concept for everyday fallback authentication which nudges participants to select more diverse and harder to guess patterns.

Acknowledgment

The competence center SBA Research (SBA-K1) is funded within the framework of COMET Competence Centers for

Excellent Technologies by BMVIT, BMDW, and the federal state of Vienna, managed by the FFG. This work was partially funded by the ICT of the future program of the FFG and the bmvit (IoT4CPS).

APPENDIX

A. Questionnaire

1) What was your ID during the lab experiments?
2) Gender (male/female/no information)
3) Age
4) What kind of smartphone are you currently using? (single- choice: iPhone, Android, Other, I don't use a smartphone)
5) What methods are you currently using to unlock your smartphone? (multiple-choice: 4-digit PINs, password with character and digit, traditional unlock pattern, fingerprint sensor, face unlock, none)
6) Order the following methods according to their security. (4-digit PINs, fingerprint sensor, *StopUnlock Pattern*, traditional unlock pattern)
7) Order the following methods according to their authentication speed. (4-digit PINs, fingerprint sensor, *StopUnlock Pattern*, traditional unlock pattern)
8) Would you use the *StopUnlock Pattern* on your own smartphone to unlock it? (yes/no/maybe)
9) What did you like about *StopUnlock Pattern*? (optional, open-ended)
10) What did you NOT like about *StopUnlock Pattern*? (optional, open-ended)

B. Questionnaire

1) Would you use the *StopUnlock Pattern* on your own smartphone to unlock it? (yes/no/maybe)
2) Would you say you could enter the *StopUnlock Pattern* faster and with fewer errors after a few days of using our app? (yes/no/no difference)
3) Did you at one point change the *StopUnlockPattern* and why? (open-ended)
4) What problems did occur when entering the *StopUnlock Pattern* in the real word, for example on public transport? (open-ended)
5) What do you think of the stopping-time of the *StopUnlock Pattern*, should it be faster or slower? (open-ended)
6) Is there anything else you would like to let us know? (optional, open-ended)

REFERENCES

[1] Y. Abdelrahman, M. Khamis, S. Schneegass, and F. Alt, "Stay cool! understanding thermal attacks on mobile-based user authentication," in *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*. ACM, 2017, pp. 3751–3763.

[2] P. Andriotis, T. Tryfonas, G. Oikonomou, and C. Yildiz, "A pilot study on the security of pattern screen-lock methods and soft side channel attacks," in *Proceedings of the Sixth ACM Conference on Security and Privacy in Wireless and Mobile Networks*, ser. WiSec '13. New York, NY, USA: ACM, 2013, pp. 1–6. [Online]. Available: http://doi.acm.org/10.1145/2462096.2462098

[3] A. J. Aviv, D. Budzitowski, and R. Kuber, "Is bigger better? comparing user-generated passwords on 3x3 vs. 4x4 grid sizes for android's pattern unlock," in *Proceedings of the 31st Annual Computer Security Applications Conference*, ser. ACSAC 2015. New York, NY, USA: ACM, 2015, pp. 301–310. [Online]. Available: http://doi.acm.org/10.1145/2818000.2818014

[4] A. J. Aviv, K. L. Gibson, E. Mossop, M. Blaze, and J. M. Smith, "Smudge attacks on smartphone touch screens." *Woot*, vol. 10, pp. 1–7, 2010.

[5] A. Bianchi, I. Oakley, and D.-S. Kwon, "Open sesame: Design guidelines for invisible passwords," *Computer*, vol. 45, no. 4, pp. 58–65, 2012.

[6] J. Bonneau, "The science of guessing: Analyzing an anonymized corpus of 70 million passwords," in *2012 IEEE Symposium on Security and Privacy*, May 2012, pp. 538–552.

[7] I. Cherapau, I. Muslukhov, N. Asanka, and K. Beznosov, "On the impact of touch id on iphone passcodes." in *SOUPS*, 2015, pp. 257–276.

[8] A. De Luca, M. Harbach, E. von Zezschwitz, M.-E. Maurer, B. E. Slawik, H. Hussmann, and M. Smith, "Now you see me, now you don't: Protecting smartphone authentication from shoulder surfers," in *Proceedings of the 32nd Annual ACM Conference on Human Factors in Computing Systems*, ser. CHI '14. New York, NY, USA: ACM, 2014, pp. 2937–2946. [Online]. Available: http://doi.acm.org/10.1145/2556288.2557097

[9] A. De Luca, K. Hertzschuch, and H. Hussmann, "Colorpin: securing pin entry through indirect input," in *CHI '10: Proceedings of the 28th international conference on Human factors in computing systems*. New York, NY, USA: ACM, 2010, pp. 1103–1106.

[10] A. De Luca, E. Von Zezschwitz, N. D. H. Nguyen, M.-E. Maurer, E. Rubegni, M. P. Scipioni, and M. Langheinrich, "Back-of-device authentication on smartphones," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, 2013, pp. 2389–2398.

[11] M. Eiband, M. Khamis, E. Von Zezschwitz, H. Hussmann, and F. Alt, "Understanding shoulder surfing in the wild: Stories from users and observers," in *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*. ACM, 2017, pp. 4254–4265.

[12] M. Harbach, A. De Luca, and S. Egelman, "The anatomy of smartphone unlocking: A field study of android lock screens," in *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, ser. CHI '16. New York, NY, USA: ACM, 2016, pp. 4806–4817. [Online]. Available: http://doi.acm.org/10.1145/2858036.2858267

[13] M. Harbach, E. Von Zezschwitz, A. Fichtner, A. De Luca, and M. Smith, "Its a hard lock life: A field study of smartphone (un) locking behavior and risk perception," in *Symposium on usable privacy and security (SOUPS)*, 2014, pp. 9–11.

[14] B. Haslinger, P. Erhard, E. Altenmüller, A. Hennenlotter, M. Schwaiger, H. Gräfin von Einsiedel, E. Rummeny, B. Conrad, and A. O. Ceballos-Baumann, "Reduced recruitment of motor association areas during bimanual coordination in concert pianists," *Human brain mapping*, vol. 22, no. 3, pp. 206–215, 2004.

[15] H. Khan, U. Hengartner, and D. Vogel, "Evaluating attack and defense strategies for smartphone pin shoulder surfing," in *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. ACM, 2018, p. 164.

[16] D.-E. Kim, M.-J. Shin, K.-M. Lee, K. Chu, S. H. Woo, Y. R. Kim, E.-C. Song, J.-W. Lee, S.-H. Park, and J.-K. Roh, "Musical training-induced functional reorganization of the adult brain: Functional magnetic resonance imaging and transcranial magnetic stimulation study on amateur string players," *Human brain mapping*, vol. 23, no. 4, pp. 188–199, 2004.

[17] K. Krombholz, A. Dabrowski, P. Purgathofer, and E. Weippl, "Poster: The petri dish attack - guessing secrets based on bacterial growth," in *Proceedings of the NDSS Symposium 2018*. Internet Society, 2018.

[18] K. Krombholz, T. Hupperich, and T. Holz, "Use the force: Evaluating force-sensitive authentication for mobile devices," in *Twelfth Symposium on Usable Privacy and Security (SOUPS 2016)*. USENIX Association, 2016, pp. 207–219.

[19] S. Schneegass, F. Steimle, A. Bulling, F. Alt, and A. Schmidt, "Smudgesafe: geometric image transformations for smudge-resistant user authentication," in *Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing*. ACM, 2014, pp. 775–786.

[20] Y. Song, G. Cho, S. Oh, H. Kim, and J. H. Huh, "On the effectiveness of pattern lock strength meters: Measuring the strength of real world pattern locks," in *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*. ACM, 2015, pp. 2343–2352.

[21] F. Tari, A. Ozok, and S. H. Holden, "A comparison of perceived and real shoulder-surfing risks between alphanumeric and graphical passwords," in *Proceedings of the second symposium on Usable privacy and security*. ACM, 2006, pp. 56–66.

[22] S. Uellenbeck, M. Duermuth, C. Wolf, and T. Holz, "Quantifying the security of graphical passwords: The case of android unlock patterns," in *Proceedings of the 2013 ACM SIGSAC Conference on Computer & Communications Security*, ser. CCS '13. New York, NY, USA: ACM, 2013, pp. 161–172. [Online]. Available: http://doi.acm.org/10.1145/2508859.2516700

[23] S. Uellenbeck, M. Dürmuth, C. Wolf, and T. Holz, "Quantifying the security of graphical passwords: The case of android unlock patterns," in *Proceedings of the 2013 ACM SIGSAC Conference on Computer & Communications Security*, ser. CCS '13. New York, NY, USA: ACM, 2013, pp. 161–172. [Online]. Available: http://doi.acm.org/10.1145/2508859.2516700

[24] E. von Zezschwitz, A. De Luca, B. Brunkow, and H. Hussmann, "Swipin: Fast and secure pin-entry on smartphones," in *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, ser. CHI '15. New York, NY, USA: ACM, 2015, pp. 1403–1406. [Online]. Available: http://doi.acm.org/10.1145/2702123.2702212

[25] E. von Zezschwitz, P. Dunphy, and A. De Luca, "Patterns in the wild: a field study of the usability of pattern and pin-based authentication on mobile devices," in *Proceedings of the 15th international conference on Human-computer interaction with mobile devices and services*, ser. MobileHCI '13. New York, NY, USA: ACM, 2013, pp. 261–270. [Online]. Available: http://doi.acm.org/10.1145/2493190.2493231

[26] S. Wiedenbeck, J. Waters, J.-C. Birget, A. Brodskiy, and N. Memon, "Passpoints: Design and longitudinal evaluation of a graphical password system," *International journal of human-computer studies*, vol. 63, no. 1-2, pp. 102–127, 2005.