Improving the Security of Android Unlock Patterns by giving Feedback to the User during the Password Creation Process

Alexa Schlegel

September 18, 2015

Abstract

An Android Unlock Pattern is a graphical password scheme, widely adopted for unlocking the screen on Android smartphones. Instead of using a PIN or a textual password, the user can set up an unlock pattern by connecting dots in a 3×3 grid.

In theory, an Android Unlock Pattern is more secure than a five-digit PIN. According to several studies, users tend to pick easy-to-guess patterns so the security of user-chosen patterns is close to a three-digit PIN scheme. To overcome the problem of weak passwords, textual password schemes have integrated password composition policies. This, in general, strengthens security but sometimes can negatively affect usability when HCI[1] principles are disregarded.

The following research proposal aims to introduce password creation policies into graphical passwords having no negative impact on usability. The user will receive constructive feedback, in form of hint, as regards creating a stronger pattern during password creation. A user study will be conducted.

1 Introduction

Authentication on smartphones needs to be done on a regular basis for unlocking the device. Different manufacturers implemented various authentication methods. Widely used and well known are textual password and PIN numbers. Looking at the distribution of smartphones today, Android phones are dominating the market with about 78.0%[2] marked share. The Android Unlock Pattern, a *recall-based* graphical password scheme, is the default authentication choice and therefor used very often. According to recent studies about 50% of Android users are using Android Unlock Pattern [25, 26].

Like textual passwords, a graphical password scheme is a *knowledge-based* authentication mechanism in which users enter a shared secret as evidence of their identity. During enrollment, the user has to choose a pattern with four to nine dots and during the authentication phase, needs to recall the pattern and draw the path on the screen.

The user can select a path according to the following rules:

1. four to nine dots have to be used
2. one dot can be used only once
3. dots are connected with a straight line
4. one cannot jump over dots not visited before

Textual passwords are typically difficult to remember, also depending on length, and are predictable if users are allowed to chose passwords. Graphical password schemes have been proposed as a alternative to overcome those usability and security issues. "The reduced memory burden will facilitate the selection and use of more secure or less predictable passwords, but It is now clear that the graphical nature of schemes does not, by itself, avoid the problems typical of text password systems." [5]

The *theoretical password space* (TPS) is the total number of unique passwords that can be generated

[1] Human Computer Interaction

[2] `http://www.idc.com/prodserv/smartphone-os-market-share.jsp`, 31.08.2015 - 12:52PM

according to the given rules, in contrast the *effective password space* (EPS) is the number of passwords in the TPS that are likely to be chosen by real-world users. [11] The TPS of Android Unlock Pattern contains 389.112 possible patterns [4], which makes it in theory more secure than a 5-digit PIN scheme.

Uellenbeck et al. [24] was the first to demonstrate the skewed distribution of Android Unlock Patterns, e.g. bias in starting point and n-grams, that make user chosen patterns guessable. They showed that 50% of the patterns were able to be guessed with only 1.000 guesses, this correspond to an EPS of a 3-digit PIN scheme for half of the Android Users. Building mainly on Uellenbeck's study, research has been done related to Android Unlock Pattern focusing on various topics: [22, 21, 3, 1]. Those papers will be discussed later in detail.

In contrast to textual passwords, which made available via password leaks, patterns are only collected from in-lab studies (using devices and/or pen&paper) or from web-based studies (self-reporting or web applications).[3] Some studies [21, 3] acquired participants from from Amazon Mechanical Turk (MTurk) or did user tests with university students [24, 22]

The security of Android Unlock Patterns can be improved either by (a) increasing the TPS or (b) expanding the EPS. Different methods have been applied and evaluated, like password strength meters[4] [22, 21], random starting point [21], alternative patterns (e.g. circle) [24] and increasing grid size[5].

To the best of my knowledge, there is no user feedback during the creation process evaluated or tested yet. Password strength meters are giving real time feedback about the underlying security measurement, what depends on the used mathematical model calculating a strength score, but it provides no advice for the user on what and why to change something, to accomplish a more secure pattern.

Password composition policies have been studied for textual passwords (e.g., [13, 16]). Applying policies usually results in stronger passwords, but when to strong causing bad usability and strange user behavior. Real-time feedback has an positive impact on usability and can help users create strong passwords with fewer errors [19].

I want to transfer password-composition policies to graphical passwords in order to increase security, while not decreasing usability. This would be an contribution towards more secure mobile devices.

2 Research Questions

The purpose of this research is to find out, if password-creation policies applied to graphical passwords (e.g. Android Unlock Patterns) lead to stronger user chosen passwords, or rather extend the effective password space, with no negative impact on usability. The research questions can be formulated as follows:

Q1 Are user chosen patterns, created using a password-composition policy, stronger (more secure) than patterns created in the conventional way?

Q2 Can password-composition policies be applied to graphical passwords with similar implications on security and usability?

3 Related Work

The following section gives an overview about graphical passwords in general, its possible security attacks and discusses relevant studies about Android Unlock Patterns and its limitations. In addition latest research about password-composition policies and its impact on security and usability is summarized.

3.1 Graphical Passwords

The first graphical password scheme was introduced by Blonder [6] in 1996, followed by Draw A Secret (DAS) by Jermyn et al. [14] in 1999. DAS was improved by using background images to make users create more complex passwords, called BDAS [10]. Tao and Adams in 2008 invented Pass-Go [23], which is very similar to Android Unlock Patterns. For an extensive overview on graphical passwords read the work of Biddle et al. [5] or Oorschot and Thorpe [18] or for a short summary Sun et al. [22] is worth reading.

Oorschot and Thorpe [18] improved DAS by recommending password rules based on password complexity properties:

[3] Aviv et al. "Comparisons of Data Collection Methods for Android Graphical Pattern Unlock" poster at SOUPS 2015

[4] Indicating password strength during creation process using visual or textual representations for week, medium and strong.

[5] Aviv et al. "Do bigger grid sizes mean better passwords? 3x3 vs. 4x4 Grid Sizes for Android Unlock Patterns", poster at SOUPS 2015

1. Require a stroke count of at least [...],

2. Disallow passwords having global reflective (mirror) symmetry [...],

3. Require at least one stroke of length 1.

3.2 Studies about Android Unlock Pattern

The results and methods of recent studies are summarize and their limitations are explained, in order of publication time.

2013–Uellenbeck et al. [24] They conducted the first study about security of Android Unlock Patterns. Several studies on university campus with in total 584 participants generating 2.900 patterns was carries out. It including a pen&paper study with 105 participants, to collect data about users "real world" pattern. They observed a bias in starting point towards corners with 75% probability and a tendency that people tend to chose adjacent points. They also found often used typical sub-patterns.

Pattern strength is measured using *partial guessing entropy* [7], which measures the average number of guesses that the optimal attack needs in order to find a correct password (or just fraction of accounts). They trained a Markov-chain model for cracking passwords. They found out that the entropy is in between a 2-digit PIN scheme and a 3-digit PIN scheme. Based on those findings alternative patters (e.g. circle, random) were evaluated in a second user study with 366 participants.

A drawback is here, that the underlying security model allows no conclusion on how to create a stronger password. Only the starting point problem is addressed here and n-grams, which would lead to dictionary checks. As the focus lies on security improvement, no usability aspects are considered at all.

2013–Andriotis et al. [1] This paper is resulting in a mixed attack combining physical (trace of fingers on screen, replicating Aviv et al. [4]) and psychological (heuristics about how user set unlock patterns) attacks. A user study with 144 participants was conducted. User had to create a what they think "secure" and "easy to remember" pattern. The following parameters were analyzed for creating heuristics: average pattern length, number of direction changes, start and end points, sub-patterns with length one to four. The evaluation based on Shannon's entropy. The secure pattern was longer and included more direction changes. In the end the mixed attack was tested with 22 participants, resulting in cracking 20 of 22 patterns.

2014–Sun et al. [22] They started with a detailed statistical analysis on all possible pattern looking at characteristic (number of dots, physical length, intersections, overlaps) and its distributions. Two different pattern strength meters were evaluated during a user study conducted on university campus with 81 participants. They showed that a password strength meter, which gives instant feedback to the user during the creation process, had an positive effect on security. People using the password strength meter, created passwords with more dots, longer length and more intersections. They state that pattern strength is largely determined by its visual complexity. For calculating entropy, Burr's formula [9] was modified based the characteristics.

Although they analyzed characteristic, they don't transform those findings into constructive feedback, of how to create a strong pattern. Also a distribution analysis of user chosen passwords with respect to characteristics is missing. The study lacks to measure the memorability (usability) of the created passwords.

2014–Aviv et al. [3] This study focuses on visual perception of usability and security. 384 participants (from MTurk) had to choose between two passwords (pairwise preference) the one who looked (a) more secure and (b) more usable. They found out that usability and security are inversely related. Visual features that can be attributed to complexity indicated a stronger perception of security, spatial features (shifts up/down, left/right) are not so strong indicators for security or usability. They built an logistic model to predict perception preference by training on features used in the survey and other related work. They achieved 70% of predicted preference. The strongest feature, they found out, is password length (sum of all euclidean length). They measured the following features: number of points, crosses (and exes), non-adjacent, knight-moves, height, side.

Perceived security is a good indicators, but not identical to standard metrics for password strength. A conclusion could also be that users need to be educated about security. Also perceived usability needs to be

evaluated, if it hold ins practice (memorability, error rates, and so on).

2015–Siadati et al. [21] They increased the strength of user chosen patterns by using a *persuasive security framework* [11, 12], a set of principles to get users to behave more securely. They conducted a user study with 270 participants form MTurk to evaluate and test two improvements expanding EPS: (1) BLINK (suggested starting point) and (2) EPSM (continues visual feedback during creation). With creating 60% strong passwords with BLINK and 77% strong passwords with EPSM, they raised security noticeable. They used the same Markov-chain model as Uellenbeck et al. [24]

Also here is no feedback given of how to create a stronger password.

> EPSM does not provide any hint on how to create a better pattern because users are already aware of which patterns are more secure. [...] strength of pattern 2(b) and 2(d) is almost same, where their strength is not the same in reality [...]

This is definitely a conflict, which needs to be further investigated and is in line with findings from Aviv et al.[3].

3.3 Attacks on graphical Passwords

Knowledge factor attacks on graphical passwords can be divided into:

(1) guessing or psychological attacks e.g. bias in patterns like skewed distribution, to limit search space, dictionary attacks like textual passwords, brute force guessing

(2) capture or physical attacks also called side-channel attacks, e.g., smudge attack, shoulder surfing

Smudge attacks are studied by Aviv et al. [4] in detail. Using photographs, they showed that it was possible to full or partial recover patterns. Generating more complex pattern makes it automatically more resistant to (1) and (2), but would need to be investigated further.

3.4 Password-Composition Policies

Looking at textual passwords a study by Shay et. al. shows that real-time feedback during password creation helps the user to create stronger passwords with fewer errors. But also password policies may cause usability problems. [19]. Usability and security of passphrases[6] is studied by Keith et al. [15], stating that "passphrases lead to more typographical errors, are perceived as more difficult to use, but are actually no more difficult to remember than other password methods.". "Password policies requiring length lead to more usability, and in some cases more security, than those requiring only a comprehensive mix of character classes and a dictionary check." [20]. Inglesant and Sasse [13] conclude that "rather than focusing password policies on maximizing password strength and enforcing frequency alone, policies should be designed using HCI principles to help the user to set an appropriately strong password in a specific context of use.". Komanduri et al. [16] found out that commonly held beliefs about password composition and strength are inaccurate:

> 1. Adding numbers to passwords is thought to add little entropy; we found, by contrast, a lot of entropy in numbers.
> 2. Dictionary checks, although other- wise useful, add much less entropy than expected.
> 3. Unexpectedly, users typically create passwords that exceed minimum requirements, thus increasing password entropy

Password-composition policies do have an positive effect on security and with keeping HCI principles in mind, not effecting usability in a negative way.

4 Security Measurement

The *security* or *strength* of a graphical password describe how hard it is for an attacker to guess or crack the pattern [15]. What people think or perceive as secure is not in line with what really is secure [3]. Different approaches to measure pattern strengths can be found in recent work:

Guessing Entropy Can be used to measure strength of password distribution. Measures the average

[6] Passphrases are longer passwords consisting of multiple words.

number of guesses that the optimal attack needs in order to find the correct password. [17]

Partial Guessing Entropy (α-guesswork) by Bonneau [8]. Finds the minimal number so that the guesses cover at least a fraction α of the passwords (used in [24]).

Shannon's entropy monograms, start and end points, entropy is calculated based on probability of point X being selected in the pattern or being at start (or end), for n-grams conditional entropy is calculated (used in [3]).

Modified entropy formula Burr's [9] entropy formula is modified for graphical password adding a score for visual complexity, based on pattern characteristics (used in [22]).

Score Function MM-score score function $f(X)$ based on the probabilistic of a given pattern X, using the Markov model by Uellenbeck [24], so a more likely pattern gets a lower score, and a less likely one get a higher score (used in [21]).

5 Pattern Characteristic

Pattern characteristics or visual complexity features were looked at when doing pattern analysis. Those characteristics are very promising in relation to giving user feedback. The following part summarizes used features in recent literature about Android Unlock Patterns and graphical passwords in general:

- start point
- end point
- size (number of connected dots)
- length (sum over all euclidian distances between dots)
- Intersections (two groups: X crossings with angle 90 and others)
- overlaps (no crossing but touching)
- non-adjacent
- knight moves
- height
- side
- sub-patterns with different number of dots
- direction changes
- symmetry (vertical/horizontal)

In 1957, Attneave [2] studied the judged complexity problem of shapes, and concluded that the complexity is related to the number of turns in the contour of the shape, the symmetry of the shape, and the variability of angular change between successive turns.

6 Usability Measurement

The term *usability* describes how easy a password is for a user to both remember and correctly enter into a login prompt [15]. Usually usability and security are seen as counterparts, but the goal should be to increase usability and security simultaneously. For extensive recommendations regarding methods for evaluation of usability (e.g., login success rates, login times, password creation times) and also security see [5]. Recommendations are given for lab, field and web-based studies. The most important task regarding usability testing related to passwords are short and long term memorability tests.

7 Methods to Improve Security

This part summarizes different methods to improve security found in literature. In general there are two different approaches: (1) expand effective password space, (2) increase theoretical password space, in order to indirectly expand effective password space.

- manipulate pattern (e.g., remove top left dot)
- rearrange pattern (e.g., random – no three points lie on one line, circle pattern)
- blacklisting (e.g., forbidding frequently used or easy to guess patterns, similar to dictionary checks at textural passwords)
- random assignment (e.g., random assigned pattern, random starting point)
- user education (e.g., password strength meter)
- change pattern size (e.g. from $3x \times 3$ to 4×4)

8 Research Idea

The aim is to improve the security of Android Unlock pattern, while keeping usability in mind. Improving security can be achieved in different ways (see section 7), but providing feedback to the user during the password creation process has not been evaluated yet. The password strength meter evaluated by Sun [22] and Siadati [21] is a good starting point, but should be studied further. The underlying characteristics of patterns (see section 5) can be used to derive constructive feedback, which will help the user to understand what to do, to generate a stronger password. Research has shown that feedback has an positive effect to security of textual password (see section 3.4), but this has to be evaluated for graphical password as well, only transfering is not sufficient.

Furthermore, I think people are not completely aware of the rules on how to create a pattern. Especially I think not all users know about crossings and overlying lines. This would need to be investigated as well.[TODO/studie wiederfinden]

9 Research Plan

My research plan consist of the following steps:

1. Conduct a pen & paper study on pattern construction rules
2. Evaluate summarized security measurements (see section 4)
3. Invent a new measurement or use existing security measurements
4. Rate existing password characteristics (see section 5) based on some criteria (which needs to be defined first)
5. Derive different policies $P_1, P_2, \ldots P_n$ from most promising characteristics
6. Create prototype for feedback interface
7. Test invented policies and feedback interface with pen & paper
8. Improve policies and feedback interface
9. Conducting a user study[7]

The most important and at the same time challenging part is to "find" a measurement for evaluating the influence of invented policies. The Markov Chain model using partial guessing entropy by Uellenbeck [24] seems promising and will be used as a first try.

References

[1] Panagiotis Andriotis, Theo Tryfonas, George Oikonomou, and Can 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*, WiSec '13, pages 1–6, New York, NY, USA, 2013. ACM.

[2] Fred Attneave. Physical determinants of the judged complexity of shapes. *Journal of Experimental Psychology*, 53(4):221, 1957.

[3] Adam J. Aviv and Dane Fichter. Understanding visual perceptions of usability and security of android's graphical password pattern. In *Proceedings of the 30th Annual Computer Security Applications Conference*, ACSAC '14, pages 286–295, New York, NY, USA, 2014. ACM.

[4] Adam J. Aviv, Katherine Gibson, Evan Mossop, Matt Blaze, and Jonathan M. Smith. Smudge attacks on smartphone touch screens. In *Proceedings of the 4th USENIX Conference on Offensive Technologies*, WOOT'10, pages 1–7, Berkeley, CA, USA, 2010. USENIX Association.

[5] Robert Biddle, Sonia Chiasson, and P.C. Van Oorschot. Graphical passwords: Learning from the first twelve years. *ACM Comput. Surv.*, 44(4):19:1–19:41, September 2012.

[6] Greg E Blonder. Graphical password, September 24 1996. US Patent 5,559,961.

[7] Joseph Bonneau. The science of guessing: analyzing an anonymized corpus of 70 million passwords. In *Security and Privacy (SP), 2012 IEEE Symposium on*, pages 538–552. IEEE, 2012.

[7] with following recommendations by Biddle et al. [5]

[8] Joseph Bonneau, Cormac Herley, Paul C. van Oorschot, and Frank Stajano. The quest to replace passwords: A framework for comparative evaluation of web authentication schemes. In *Proceedings of the 2012 IEEE Symposium on Security and Privacy*, SP '12, pages 553–567, Washington, DC, USA, 2012. IEEE Computer Society.

[9] William E Burr, Donna F Dodson, and William T Polk. *Electronic authentication guideline*. Citeseer, 2004.

[10] Paul Dunphy, Andreas P. Heiner, and N. Asokan. A closer look at recognition-based graphical passwords on mobile devices. In *Proceedings of the Sixth Symposium on Usable Privacy and Security*, SOUPS '10, pages 3:1–3:12, New York, NY, USA, 2010. ACM.

[11] Alain Forget, Sonia Chiasson, and Robert Biddle. Persuasion as education for computer security. In *World Conference on E-Learning in Corporate, Government, Healthcare, and Higher Education*, volume 2007, pages 822–829, 2007.

[12] Alain Forget, Sonia Chiasson, Paul C van Oorschot, and Robert Biddle. Persuasion for stronger passwords: Motivation and pilot study. In *Persuasive Technology*, pages 140–150. Springer, 2008.

[13] Philip G. Inglesant and M. Angela Sasse. The true cost of unusable password policies: Password use in the wild. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '10, pages 383–392, New York, NY, USA, 2010. ACM.

[14] Ian Jermyn, Alain Mayer, Fabian Monrose, Michael K. Reiter, and Aviel D. Rubin. The design and analysis of graphical passwords. In *Proceedings of the 8th Conference on USENIX Security Symposium - Volume 8*, SSYM'99, pages 1–1, Berkeley, CA, USA, 1999. USENIX Association.

[15] Mark Keith, Benjamin Shao, and Paul John Steinbart. The usability of passphrases for authentication: An empirical field study. *International Journal of Human-Computer Studies*, 65(1):17 – 28, 2007. Information security in the knowledge economy.

[16] Saranga Komanduri, Richard Shay, Patrick Gage Kelley, Michelle L. Mazurek, Lujo Bauer, Nicolas Christin, Lorrie Faith Cranor, and Serge Egelman. Of passwords and people: Measuring the effect of password-composition policies. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '11, pages 2595–2604, New York, NY, USA, 2011. ACM.

[17] James L Massey. Guessing and entropy. In *Information Theory, 1994. Proceedings., 1994 IEEE International Symposium on*, page 204. IEEE, 1994.

[18] P. C. van Oorschot and Julie Thorpe. On predictive models and user-drawn graphical passwords. *ACM Trans. Inf. Syst. Secur.*, 10(4):5:1–5:33, January 2008.

[19] Richard Shay, Lujo Bauer, Nicolas Christin, Lorrie Faith Cranor, Alain Forget, Saranga Komanduri, Michelle L. Mazurek, William Melicher, Sean M. Segreti, and Blase Ur. A spoonful of sugar?: The impact of guidance and feedback on password-creation behavior. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, CHI '15, pages 2903–2912, New York, NY, USA, 2015. ACM.

[20] Richard Shay, Saranga Komanduri, Adam L. Durity, Phillip (Seyoung) Huh, Michelle L. Mazurek, Sean M. Segreti, Blase Ur, Lujo Bauer, Nicolas Christin, and Lorrie Faith Cranor. Can long passwords be secure and usable? In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '14, pages 2927–2936, New York, NY, USA, 2014. ACM.

[21] Hossein Siadati and Nasir Memon. Fortifying android patterns using persuasive security framework. 2015.

[22] Chen Sun, Yang Wang, and Jun Zheng. Dissecting pattern unlock: The effect of pattern strength meter on pattern selection. *Journal of Information Security and Applications*, 19(4âĂŞ5):308 – 320, 2014.

[23] Hai Tao and Carlisle Adams. Pass-go: A proposal to improve the usability of graphical passwords. *IJ Network Security*, 7(2):273–292, 2008.

[24] Sebastian Uellenbeck, Markus Dürmuth, Christopher Wolf, and Thorsten 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*, CCS '13, pages 161–172, New York, NY, USA, 2013. ACM.

[25] Dirk Van Bruggen, Shu Liu, Mitch Kajzer, Aaron Striegel, Charles R. Crowell, and John D'Arcy. Modifying smartphone user locking behavior. In *Proceedings of the Ninth Symposium on Usable Privacy and Security*, SOUPS '13, pages 10:1–10:14, New York, NY, USA, 2013. ACM.

[26] Dirk C Van Bruggen. *Studying the Impact of Security Awareness Efforts on User Behavior*. PhD thesis, University of Notre Dame, 2014.