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

Alexa Schlegel

September 15, 2015

Abstract

***Android Unlock Pattern* is a graphical password scheme, which has been widely adopted for unlocking the screen on Android smart phones. Instead of using a PIN number or textual password, the user can set up an unlock pattern by connecting dots in a 3×3 grid. In theory the security of Android Unlock Pattern is more secure than a 5-digit PIN scheme. Several studies have shown that users tend to pick easy to guess passwords, so the security of user chosen patterns is close to a 3-digit PIN scheme. To overcome the problem of weak passwords textual password schemes integrated *password composition policies*. In general this leads to more security, but sometimes can have a negative effect on usability, when HCI principles are disregarded. The following research proposal aims to introduce password creation policies to graphical passwords with having no negative impact on usability. The user will be guided and will get constructive feedback during the password creation process to encourage stronger passwords. A user study will be conducted.**

1 Introduction

* authentication on smart phones needs to be done on a regular basis for unlocking
* different manufacturers implemented various authentication methods
* text-base passwords and PIN numbers are widely used[TODO/cite]
* smart phones are widely distributed now and Android phones dominating the market with about 78.0% marked share[1]
* Android Unlock Pattern, a *recall-based* graphical password scheme, it is the default authentication choice and therefor used very often, one user study with 51% using Android unlock pattern - [20] another on stating 48% to 56%[21]
* like textual password, graphical passwords are *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 4-9 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. at least 4 points must be chosen
2. no point can be used twice
3. only straight lines are allowed
4. one cannot jump over points not visited before

* text-base passwords are typically difficult to remember and are predictable if user choice is allowed, so graphical password schemes have been proposed as a alternative to overcome those usability and security issues, "reduced memory burden will facilitate the selection and use of more secure or less predictable passwords", "It is now clear that the graphical nature of schemes does not, by itself, avoid the problems typical of text password systems."[5]
* *theoretical password space (TPS)* is the total number of unique passwords that could be generated according to the given rules, *effective password space (EPS)* is the number of passwords in the TPS that are likely to

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

be chosen by real-world users[9]
* the TPS of the 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.[19] 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
* 50% of the patterns were able to be guessed with only 1000 guesses, this correspond to an EPS of 3-digit PIN scheme for half of the Android Users
* building on this study, extended research has been done related to Android Unlock Pattern[18][17][3][1] , which will be discussed in detail later
* as there are no password leaks, like from text-based passwords, Android Unlock Pattern are only collected from in-lab studies (with real devices and/or pen& paper studies) or from web-based studies (self-reporting or web applications)[TODO/howToCiteAPoster][2], with participants from Amazon Mechanical Turk[17][3] or university students[19][18]
* security of Android Unlock Patterns can be improved either by (a) increase TPS or (b) expand EPS, different methods are applied and evaluated, like password meters (week, medium strong indication during creation process)[18][17], random starting point[17], alternative patterns (e.g. circle) [19] and increasing grid size[TODO/howToCiteAPoster][3]
* To the best of my knowledge there is no guidance and user feedback during the creation process evaluated or tested yet. Password meters are giving real time feedback about the underlying security, but provide no advice on what and why to change something to accomplish a more secure pattern.
* password composition policies have been studied for text-based passwords a little bit[11][13], resulting in stronger password, but when to strong resulting in 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[15]
* I want to transfer password policies to graphical passwords and increase security, while not decreasing usability.

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 patterns that are 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

TODO - some sentence here

3.1 Graphical Passwords

* first graphical password by Blonder (1996), Draw A Secret (DAS) by Jermyn et al. (1999), improving DAS by using background images to make user create more complex passwords, called BDAS, Tao and Adams in 2008 Pass-Go, Yet another graphical password (YAGP) proposed by Gao et al. in 2008, read [18], [14] as a summary or have a look at the work of Biddle et al. [5] for an extensive overview on graphical password during the last 12 years.
* Android Unlock Patterns are very similar to Pass-Go
* also existing research on improving DAS using *password complexity factors* and recommending password rules for DAS (like symmetry, stroke count)[very important!][14]

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.

Uellenbeck et al. [19] (2013) was conducting the first study about security of Android Unlock Patterns
several study on university campus with in total 584

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

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

participants generating 2.900 patterns, including a pen & paper study with 105 participants to collect data about users "real world" pattern
bias in starting point towards corners 75%
people tend to chose adjacent points
found often used typical sub-patterns
pattern strength is measured using partial guessing entropy[6], which measures the average number of guesses that the optimal attack needs in order to find a correct password (or just fraction of accounts)
using a Markov Model for cracking passwords
found out that entropy is in between 2-digit PIN scheme and 3-digit PIN scheme
alternative patters were evaluated in a second study with 366 participants
a drawback is here that the underlying security model allows no conclusion on how to create a stronger password
only starting point problem can be addressed and n-grams, which would lead to dictionary checks
no usability aspects are considered at all

Andriotis et al. [1] (2013) 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 choose a what they think secure and easy to remember pattern
the following parameter were analyzed for creating heuristics: average pattern length, number of direction changes, start and end points, sub-patterns with length 1-4, based on Shannon's entropy
the secure pattern was longer and included more direction changes
the mixed attack was tested with 22 participants, resulting in cracking 20 of 22 patterns.

Sun et al. [18] (2014) analyzed characteristic (number of dots, physical length, intersections, overlaps of all available patterns and its distribution)
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 feedback about the strength of password (week, medium, strong), had an positive effect on password strength. People using the password strength meter, created password with more dots, longer length and more intersections.
they state that pattern strength is largely determined by its visual complexity, so based on characteristics the formula calculating entropy[8] is modified
Also they analyzed characteristic, but don't use those findings for giving constructive feedback to the user, of how to create a strong pattern.
also distribution of user chosen passwords with respect to characteristics is missing
The study lacks to measure the memorability (usability) of the created passwords.

Aviv et al. [3] (2014) This study focuses on visual perception of usability and security. Participants (384 from Amazon Mechanical Turk, resulting in 2.000 rated password) had to choose between two passwords (pairwise preference) the one who looks (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 is password length (total length of all lines)
features measured are: 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.
Conclusion could 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).

Siadati et al. [17] (2015) increasing strength of by using a persuasive security framework[9], [10], a set of principles to get user to behave more securely
conducting a user study with 270 participants form Amazon Mechanical Turk
two improvements were tested to expand EPS: (1) BLINK (suggested starting point), EPSM (continues feed during creation: strong, medium, weak)
60% strong passwords with BLINK and 77% strong passwords with EPSM

using same Markov Chain model as Uellenbeck et al.[19]
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." The next sentence is "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 it textual passwords, brute force guessing) and *(2) capture or physical attacks*, also called side-channel attacks (smudge attack, shoulder surfing)
* smudge attacks are studied studied by Aviv et al. [4]
Improving security (generating more complex patterns) makes it automatically more resistant to (1) guessing attacks, because ...

3.4 Password Composition Policies

* Looking at textual password a study by Shay et. al. shows that real-time feedback while password creation helps the user to create stronger passwords with fewer errors. But also password policies may cause usability problems. [15]
usability and security of passphrases[4] is studied by Keith et al.[12], 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.[16].
Inglesant and Sasse[11] conclude that "rather than focussing 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. [13] 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 [12]. 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 number of guesses that the optimal attack needs in order to find the correct password.[TODO/cite]

Partial Guessing Entropy (α-guesswork) by Bonneau [7] finds the minimal number so that the guesses cover at least a fraction α of the passwords (used in [19]).

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 [8] entropy formula is modified for graphical password adding scores for visual complexity, based on pattern characteristics (used in [18]).

[4] passphrases are longer passwords consisting of multiple words.

Score Function MM-score score function $f(X)$ based on the probabilistic of a given pattern X, using the Markov model by Uellenbeck [19], a more likely pattern gets a lower score, and a less likely one get a higher score (used in [17]).

5 Pattern Characteristic

Pattern characteristics or visual complexity features were looked at a lot 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 lines between dots)
- intersections (two groups: X crossings with angle 90 and others)
- overlaps (no crossing but touching)
- non-adjacent
- knigth 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.[12]
Usually usability and security are seen as counterparts, but the goal should be 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.
Memorability needs to be tested, shot and long term

7 Methods to Improve Security

Summary on different methods (some are tested already) to improve security found in literature. Two different approaches:

expand EPS manipulate pattern (e.g., removing top left dot), rearrange pattern (e.g., random &circle patterns), blacklisting (similar to dictionary checks), random assignment (random starting point), user education (password meter)

increase TPS change pattern size (e.g., from $3x3$ to $4x4$

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, as shown in 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[18] and Siadati [17] is a good starting point but should be studied further. The underlying characteristics of patterns 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, not only transferred.

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.

9 Research Plan

My research plan consist of the following steps:

1. Conduct a pen & paper study on construction rules
2. Evaluate summarized security measurements 4
3. Invent a new measurement or use existing security measurements
4. Rate existing password characteristics 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[5]

The most important and at the same time challenging part is to measurement for evaluating the influence of invented policies. The Markov Chain model using partial guessing entropy by Uellenbeck [19] seem 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] 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] 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.

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

[9] 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.

[10] 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.

[11] Philip G. Inglesant and M. Angela Sasse. The true cost of unusable password policies: Password use in the wild. In *Proceedings of the SIGCHI*

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

Conference on Human Factors in Computing Systems, CHI '10, pages 383–392, New York, NY, USA, 2010. ACM.

[12] 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.

[13] 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.

[14] 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.

[15] 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.

[16] 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.

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

[18] 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(4âÂĂÂŞ5):308 – 320, 2014.

[19] 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 &#38; Communications Security*, CCS '13, pages 161–172, New York, NY, USA, 2013. ACM.

[20] 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.

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