Using Grids as Password Entry Devices

Karol Lejmbach
Marquette University

USING GRIDS AS PASSWORD ENTRY DEVICES

by

Karol Lejmbach, B.S.

A Thesis submitted to the Faculty of the Graduate School,
Marquette University,
in Partial Fulfillment of the Requirements for
the Degree of Master of Science

Milwaukee, Wisconsin
May 2021

ABSTRACT
USING GRIDS AS PASSWORD ENTRY DEVICES

Karol Lejmbach, B.S.

Marquette University, 2021

The classic text-based password has been around for a very long time. A lot of security research has been conducted on it. A set of best practices has been available for many years stressing the use of longer and more complex passwords. The issue with this approach is that humans have a hard time recalling long complex sequences of characters. Worse, the more complex the string of characters the more prone it is to being written down which is the most detrimental security threat.

The goal of this paper is to introduce and provide an introductory analysis of a grid-based password system. This system allows weaker passwords to still have the potential security of longer more complex passwords. At the same time the system leverages the human ability to better recall visual patterns to aid in the memorization process. This thesis will discuss the mathematical maxima that may be achieved by using this password system. Compare it against conventional graphical passwords, and finally discuss the human factor in using this password schema.

ACKNOWLEDGEMENTS

Karol Lejmbach, B.S.

I would like to thank my parents, professors, and friends for their support and motivation. Without them I would not be where I am today.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Introduction

It is a well-documented fact that people choose easy to remember passwords. The research that has analyzed this property points to a few reasons for this. One of the reasons is the number of accounts people have, in the paper by Rick Wash et al. they observed users authenticating around 3.2 times per day (Rick Wash et al. 1-10). Since authentication is a requirement for completing a task for many people, choosing weaker passwords means less time is spent on the authentication process (Rick Wash et al. 1-10). The other burden of having so many accounts is password reuse between the accounts.

In the research conducted by Rick Wash et al. they found that majority of reused passwords were more complex and were used frequently. One potential solution to this problem is using password managers; however, only 26 participants said they used this technology (Rick Wash et al. 1-10). During their research Rick Wash et al. found that the average password length of their participants was around 9 characters.

With these facts in mind, a new user authentication system is being proposed that uses a grid of cells into which the user enters their password. Each character that makes up the user password will appear somewhere in the grid provided. The system will then authenticate the user based on if the characters match and if they were entered in the correct order into the grid.

There will be to forms of analysis run on this system. A mathematical analysis that will look to see the gains over conventional text-based passwords. As well as a human factors analysis that will look to see how these mathematical gains are reflected in the real world. The main focus will be to see if there is sufficient variation in how the

grid is used to warrant further research and eventual adoption of the new password system.

Literature Review

The main problem with text-based passwords is human memory. In the research done by Inglesant and Sasse people tended to create passwords based on objects found around them; the main goal being to facilitate the recollection of passwords as needed (qtd. in Rick Wash et al. 1-10). In the research done by Stobert and Kibble people tend to save login credentials to their browsers; however, this practice is still less common than memorizing passwords (qtd. in Rick Wash et al. 1-10). Reliance on memory has other implications.

Passwords that are easy to remember are less complex; moreover, a password that was created in accordance with password policies is still relatively simple (Rick Wash et al. 1-10). This was observed in the research led by Ur's team where the respondents did not fully understand how to increase password complexity (qtd. in Rick Wash et al. 1-10). Ideally people should use password managers to handle passwords; however, the adoption of these tools is very low even saving passwords to browser is not as widely used as memorizing passwords (Rick Wash et al. 1-10). Other password schemes have been developed that try to account for human reliance on memory, one such password scheme is Draw a Secret.

Draw a Secret, DaS, technology has been around for quite some time. It has been employed in many authentication processes from drawing a pattern on a PDA screen to altering an image. One of the advantages of using this form of password authentication is that it leverages the human ability to remember patterns. Today, one of the most commonly encountered forms of DaS is the android swipe pattern that utilizes a grid of nodes.

This grid is used to store a continuous path connecting the nodes that is used to authenticate the user. Security analysis that has been performed on the merit of this authentication scheme has yielded interesting results that in some respects will also be reflected in the grid scheme being proposed in this thesis.

One of the first criticism of the Android swipe pattern is that the default grid is too small. The pattern also must meet specific criteria to be valid. Such pattern consists of at least 4 nodes, all nodes can only be used once, the pattern must be continuous, and finally if a path traces over a node that node is automatically included in the pattern (Aviv, Budzitowski, & Kuber, Dec 7, 2015). Considering these constraints there is a maximum of 389,112 possible patterns; this means that a 3x3 grid has the same magnitude of complexity as a 4-character password that uses all printable symbols (Aviv et al., Dec 7, 2015). These observations do not translate favorably once human factors are added.

In the process of researching this topic two most relevant papers were chosen that analyze the Android swipe pattern in varying degrees. The paper “Complexity Metrics and User Strength Perceptions of the Pattern-Lock Graphical Authentication Method” by Panagiotis Andriotis et al. focuses primarily on the current 3x3 grid present on Android devices. Their paper presents the continuation and further analysis of the initial pilot study done by Panagiotis.

The majority of their research is conducted through an Android app that is published on the play store and asks respondents to submit an Android swipe pattern and answer a few demographic questions about themselves. The app then provides feedback

regarding the complexity of the submitted pattern and offers the opportunity to submit a different pattern.

The complexity of the pattern is calculated by looking at the number of occurrences of what they call knight moves, following the pattern that the knight piece moves in chess, which happen when a user jumps to a node that is diagonally positioned from the previous node. The presence of an overlapping node which is a node that is hit once then jumped across to continue the pattern further.

The next part of the metric is the length, number of nodes used, which only contributes to the scored if a length of over 5 nodes was submitted. Because of the length constraints placed on the pattern, there will be at least one direction change on a 3x3 grid. The complexity score is increased by one if the pattern submitted has more than one direction change.

The last pattern characteristic that is considered for the complexity score is the starting location of the pattern. If the pattern starts in the upper-left hand node, then there is no contribution to the score. These values are then summed and a verdict of weak, medium, and strong are assigned given the presence of the aforementioned characteristics of the submitted pattern.

The results described in their research reflect the results present in the pilot study done by Panagiotis Andriotis et al which was published in the sixth ACM conference (Andriotis et al., 2014). During the analysis of the patterns, a list of most common sequences of nodes that appear in longer patterns was created.

These sub-patterns consisted of binary, ternary, and quartic node patterns. The presence of these patterns can be leveraged when building patterns that mimic how

people navigate grids. Looking at the lengths of the patterns submitted, the majority fall in the five to seven node range. The upper half yielding an extra point in the complexity calculation being over six nodes long. The results gathered by Androitis et al. show that the most popular starting node was the one in the upper-left hand corner.

Interestingly, majority of the patterns submitted were classified as medium or strong. I would argue that given the way of calculating the complexity metric this does not provide much feedback on what is really happening; nevertheless, there is a tendency to generate stronger patterns with 88 submissions falling in the medium and strong classification (Andriotis et al., 2014). This is in stark contrast to the findings gathered by Rick Walsh where there was a tendency for choosing weaker passwords (Rick Wash et al. 1-10).

The second paper, “Is Bigger Better? Comparing User-Generated Passwords on 3x3 vs. 4x4 Grid sizes for Android’s Pattern Lock” written by Adam J. Aviv et al. dives deeper into the topics present Andriotis’ et al. research. Their research focuses on analyzing the benefits of increasing the grid size to a 4x4. At the same time, Aviv et al. provide a much more robust method of assigning a strength metric to a submitted swipe pattern.

The first form of data collection done by Aviv et al. was a pen-and-paper experiment. The goal of this experiment was to generate patterns and at the same time replicate the results published by Uellenbeck et al. The process would ask the participant to fill out a total of six defensive patterns three for the 3x3 grid as well as for the 4x4 grid. Next the participant would fill out ten offensive patterns as well as a demographic survey.

Aviv et al. define defensive patterns as those owned and created by a respondent; an offensive pattern would be a pattern that the respondent creates as a guess of someone else's pattern (Aviv et al., Dec 7, 2015). At the end of the session, they would be asked to reproduce their original defensive patterns rewarding those participants that successfully recalled their original patterns. The second form of data collection that they performed was by gathering 3x3 patterns from a website utilizing Amazon's Turk to generate the data.

Andriotis et al. define embeds as inscribing a 3x3 pattern into a 4x4 grid. During their analysis of the pen and paper experiment they found that 33% of the 4x4 patterns generated were embeds of 3x3 patterns (Aviv et al., Dec 7, 2015). Symmetries on the other hand are patterns that can be reproduced by rotating or reflecting a pervious pattern. The inclusion of these two metrics in the evaluation of the strength of a pattern is critical because these can be leveraged by guessers to accelerate the cracking process of patterns. Another metric that was employed was guessability. Guessability looks at how easy it is to generate a pattern by randomly connecting nodes together.

Looking at the results present in the paper by Aviv et al. the introduction of another row and column has little impact on the starting and ending positions of patterns. Also, their findings replicated those that were done by Andriotis et al. With the majority of the patterns starting in the upper-left hand corner and terminating in the bottom-right. Since both offensive and defensive patterns were collected from the pen-and-paper respondents, there were significantly less correct guesses of 4x4 patterns compared to 3x3 patterns. It is noted by the researchers that these results may be inflated because of the fact that a single guess would compromise multiple defensive patterns. The most

important part of the results is looking at the recollection rate of the 4x4 pattern compared to the 3x3 pattern were very comparable with the 4x4 grid being only 5% behind in correct recollection to the 3x3 grid.

With these findings the researchers began developing a brute force algorithm that would be run on the patterns gathered from both the online submission and the pen-and-paper experiment. The primary design goal of the algorithm was to maximize the speed of cracking by exploiting and targeting the characteristics such as embeds and symmetries in patterns. The design of the algorithm followed Markov's Model Likelihood Estimates. This model would build patterns based on probabilities of nodes being connected by an edge. There was a modification introduced into the Markov's Model that would also account for the probability of the length of the pattern being generated which would be multiplied to the probability generated by Markov's Model.

With this brute force algorithm in place the researchers began assigning strength metrics to the submitted patterns based on their guessability. This is a much better approach compared to the one presented in the previous paper, providing a much more exacting analysis of the pattern, and yielding more useful feedback about the pattern.

The results of applying the cracking algorithm show the massive improvement resulting from increasing of the grid size to 4x4. With the total number of guessed patterns dropping to below 40% after 10,000 attempted guesses; compared to over 80% being cracked by that point on the 3x3 grid. The other important finding present here is that offensive patterns are much weaker than defensive patterns. This could be justified by the fact that people guess common patterns or simpler patterns when trying to guess an unknown pattern.

The password entry device being proposed in this work can be considered a variant of the Android Swipe pattern. The main difference being is that there is no longer a binary system of either using a node or not using a node. The cells, nodes, can be used as often as the user wishes. At the same time, other constraints such as: the pattern requiring at least four nodes, there can be no reuse of previous nodes, the pattern being drawn in one stroke, and not avoiding previously un-selected nodes are not enforced (Aviv et al., Dec 7, 2015). Since these constraints are not being adopted in the new password schema, many more possible paths through the grid will have to be considered increasing the complexity and slowing down the brute force attack. The biggest contribution of these two papers is the analysis of how grids are navigated by humans which is something that will need to be considered in the new system being proposed.

The demonstration of the attack results provides an invaluable way of further analyzing and testing the grid password system. Being able to remove some of the inherent limitations and simplifications of the pattern system will undoubtedly make it harder for attacks to take place. At the same time understanding the limitations introduced by symmetries and embeds may be used to create password policies that mitigate the risks introduced by their presence.

Considering the findings of the two papers there is clear evidence that combining text-based passwords with the recollection benefits of Draw a Secret will allow an additional level of security when users have weaker passwords that are easier to remember. There is strong potential in building upon the swipe pattern using the grid-based system and provide an additional level of complexity that mitigates the risks of short passwords.

The scheme being introduced will be more involved than a simple swipe pattern that is employed on Android devices. Questions regarding the human ability to recall their passwords will have to be addressed. The article “Working Memory Training: Assessing the Efficiency of Mnemonic Strategies” written by Di Santo et al. looks at memory training and compares how the trainings impact the ability of humans to improve their working memory. The part that is most relevant to this research is looking at the differences between sequential memorization and grid-based memorization.

The experiment conducted by Di Santo et al. separated a group of forty-three participants into a control and experimental group. Both groups would perform the same tasks meant to exercise their working memory. These tasks included: N-Back task, seminar, and a memory test. The main difference between the control and the experimental group was in the seminar. The experimental group were given a seminar on how to improve their working memory whilst the control group were given a seminar on memory. After the seminar both groups were given a three-hour break after which the memory tests were administered.

The memory test consisted of five pieces: word sequence, image sequence, digits, word grid, image grid. During all these tests the people had to commit to memory at their own pace the information being presented. The main difference being that the experimental group could employ the tactics that were covered during their seminar while the control group had to memorize the information in whatever way they knew. Some of the outliers in the control group later told the researchers that they were taught similar strategies as those that were taught to the experimental group which increased their performance on the tests (Di Santo, De Luca, Isaja, & Andreetta, 2020).

The results of the experiment show that there is improvement to using grids over sequential memorization; the experimental group had the best scores with the fraction of correct answers being between 80-100% (Di Santo et al., 2020). Comparing the results of the control group looking at the experiment of sequential words to words presented in a grid there is improvement with the sequential average being around 20% correct jumping to an average of around 70% when the information was presented in a grid (Di Santo et al., 2020). The group of researchers justify this spike in performance in the control group as to the fact that grids can be used as visualizations making it easier to internalize information versus creating a chain of information that is being presented in order. In the end the performance gains in working memory demonstrate the potential that grids have when it comes to remembering information.

Mathematical Analysis

For an Android swipe pattern to be legal the following requirements must be met (Aviv, Budzitowski, and Kuber 301-310):

- The pattern must consist of at least 4 nodes
- Nodes cannot be reused
- The pattern must be continuous
- Nodes cannot be skipped in the pattern. If the pattern line passes over a node, that node is added to the pattern.

With these restrictions in place the maximum number of legal patterns available on a 3x3 grid is only 389,112 (Aviv, Budzitowski, and Kuber 301-310). By removing these restrictions, the proposed password grid scheme will allow for more complex password without necessarily increasing the length of the password needed to be memorized by the user. At the same time, being able to provide basic protection when written down.

The new system introduces a grid into which the user enters their password by placing characters present in their password into the cells of the grid. Once complete, the system authenticates the user by checking: if the correct characters are present as well as if they were entered into the grid in the correct position and order. Upon fulfilling these two factors the user is granted access. The system does not have any requirements for how the grid needs to be filled out; however, password policies can be introduced to enforce specific levels of complexity as need be.

In order to maximize adoption of this system, current user password storing solutions will be used. To accomplish this, as the user is entering the password into the grid the computer is translating the input into a sequence of characters. For example:

a,1,2 would mean that the letter “a” appears second in the first cell. The user only must recall is that their password has the letter “a” appear in the first cell after another character in that cell. Since the order in which the grid was filled out is preserved by the sequence of characters and their ordered pairs, finding a grid written down somewhere will not necessarily compromise the security because the order in which the characters appear will have to be figured out. The presence of repeated characters in the password can aid in disguising how the grid was filled out.

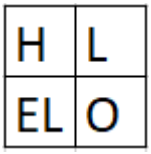


Figure I Password "HELLO" stored in a 2x2 grid.

Figure II Password "HELLO" stored in a 3x3 grid

Figure 1 demonstrates this property. The two L’s in HELLO do not automatically yield a pattern in which the gird was filled out. Two possible passwords would need to be tested one in which the “L” is entered after the “E” or the other case where it is entered after the “E” in the upper-right hand corner. One of these passwords could be: *H,1,1,E,3,1,L,2,1,L,3,2,O,4,1*. Entering the password left to right, top to bottom would yield *H,1,1,E,3,1,L,3,2,L,2,1,O,4,1* which clearly are not the same password. This means that finding the order in which the Ls were entered is mandatory. The same can be observed if the grid is large enough to not need cell reuse as shown in figure 2. These two examples also demonstrate the feature that this new scheme has. Namely, knowing what

characters make up the password may not be sufficient to gaining access to a user's account.

A list of potential text-based passwords is generated from a collection of available characters. These lists can be adjusted according to password policies, but in the end the maximum number of possible passwords, M_p, can be calculated using the following equation.

$$C = cardinality\ of\ the\ character\ set$$

$$n = number\ of\ characters\ in\ the\ password$$

$$M_p = C^n$$

The practice of using more complex character sets can be justified by the fact that increasing the cardinality of the character set creates many more potential passwords than increasing the password length of the same character set.

For an eight-character lower-case letter password the calculation would be as follows.

$$M_p = 26^8$$

$$M_p = 2.09 * 10^{11}$$

Increasing the password by one character the original value is multiplied by a factor of 26.

$$M_p = 26^9$$

$$M_p = 5.43 * 10^{12}$$

If a larger cardinality is chosen, then the number of potential passwords may have a larger increase. For example, including upper-case letters will yield:

$$M_p = 52^8$$

$$M_p = 5.35 * 10^{13}$$

In this case keeping the password the same length but introducing upper-case letters yielded a result that is almost an order of magnitude larger than increasing the length of the lower-case password by one.

This does not always hold true for example if instead of upper-case letters digits 0-9 were introduced then the following result would be obtained.

$$M_p = 36^8$$

$$M_p = 2.82 * 10^{12}$$

This value is smaller than the value obtained when the password was increased by one lower-case letter.

Table I Table that shows the number of potential passwords and at which point increasing the cardinality is better than adding another letter.

Length of password	26	36
5		6.0E+07
6	3.1E+08	2.18E+09
7	8.0E+09	7.84E+10
8	2.1E+11	2.82E+12
9	5.4E+12	1.02E+14
10	1.4E+14	3.66E+15
11	3.7E+15	1.32E+17
12	9.5E+16	4.74E+18
13	2.5E+18	1.71E+20
14	6.5E+19	6.14E+21
15	1.7E+21	

The table shows at which point increasing the password length provides less passwords than increasing the cardinality. The password length can be increased up to a length of 11 characters yielding more passwords than keeping the same length and increasing the cardinality. From that point going forward it is better to increase cardinality than increasing the length of the password. As mentioned previously people tend to choose shorter passwords this means that something has to be introduced to increase the number of potential passwords that are shorter. This is where the password grid comes into play. The idea being that by introducing a grid the number of potential passwords is increased.

When the grid is introduced the equation for finding the maximum number of potential passwords is multiplied by the number of ways there are to choose the cells from the grid.

$$M_p = C^n * (n + M - 1 \; choose \; n)$$

In this case the M stands for the number of available cells and the C and n preserve their meaning from the original equation.

If the password set contains two letters for example A and B and a 2-cell grid is provided for entering the password, then the maximum number of passwords will be calculated as follows.

$$M_p = 2^2 * (3\ choose\ 2)$$

$$M_p = 12$$

Writing out all the arrangements the following list can be created.

Table II Shows all possible arrangements for a two-letter password in a 2-cell grid.

AA|	BB|	AB|	BA|
A|A	B|B	A|B	B|A
|AA	|BB	|AB	|BA

It is clear that the complexity introduced by filling out the grid has significant improvements even on this very simple example. Looking at a similar table from the previous regular passwords it can be seen the impact of a 9-cell grid and a 16-cell grid.

Table III Shows how many characters are needed to match or exceed the number of potential passwords available from a 3x3 and 4x4 grid.

Length	Lower-case	With 3x3 grid	With 4x4 grid
8	2.09E+11	2.69E+15	1.02E+17
9	5.43E+12	1.32E+17	7.10E+18
10	1.41E+14	6.18E+18	4.61E+20
11	3.67E+15	2.77E+20	2.84E+22
12	9.54E+16	1.20E+22	1.66E+24
13	2.48E+18	5.05E+23	9.29E+25
14	6.45E+19	2.06E+25	5.00E+27
15	1.68E+21	8.22E+26	2.60E+29
16	4.36E+22	3.21E+28	1.31E+31
17	1.13E+24	1.23E+30	6.41E+32
18	2.95E+25	4.61E+31	3.06E+34

In order to exceed the minimum number of passwords available with a 3x3 grid a lower-case character password with at least 11 characters would have to be chosen. At the same time in order to match the 4x4 grid an additional two characters would have to be chosen.

This system can be thought of as an extension of the Android swipe pattern. In the proposed system each node no longer stores a binary value reflecting its use in the pattern. Instead, each cell can store any number of characters present in the password. Next there is no requirement for how many cells must be used by the user. At the same time, the pattern in which the grid is filled out does not have to be continuous, where consecutive cells need to be adjacent. An example of this would be filling out the

password using only the outside corners of the grid. These generalizations introduce new levels of complexity that must be addressed when attempting to crack passwords.

Human Factors Analysis

Mathematical complexity does not show the whole picture. Human tendencies can severely impact the potential that any security system has. In order to analyze these impacts, an IRB approved survey was performed to answer the following questions:

- Does the respondent use a password manager?
- How do they unlock their phone?
- How would they fill out a grid with a predetermined password?
- Where they see this password scheme being adopted?

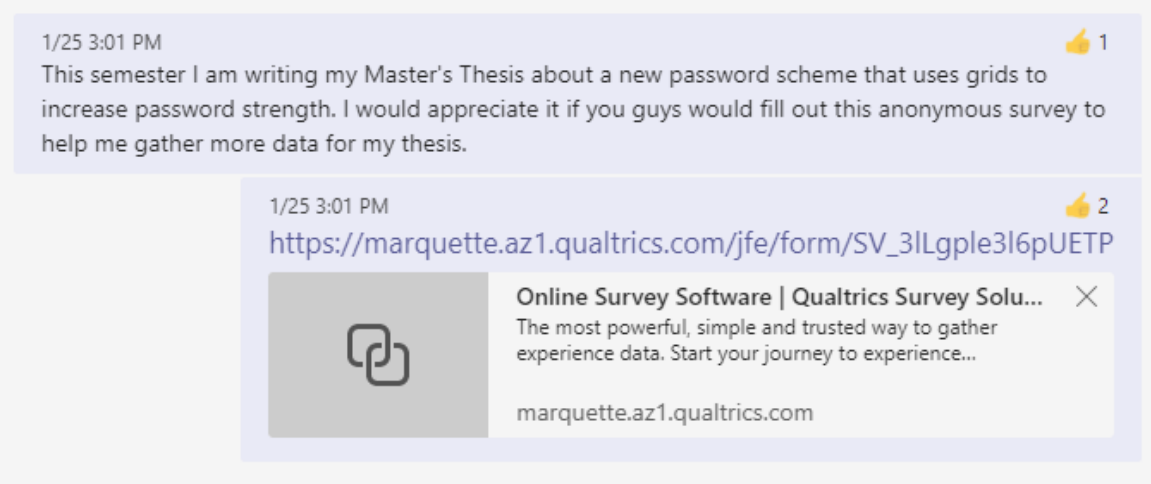


Figure III Example of the invitation to the survey.

The survey was built in Qualtrics and an anonymous link was posted in group chats for people to click on and fill out the survey. Figure III shows an example of one of the recruitment messages sent along with the link to the survey. There were two posts performed.

The first link was posted in a group chat where CS college students would hang out and discuss classes and topics. The second recruitment was posted on a group chat for graduate students in the Computing program at Marquette. There was a total of 14 responses collected of which two were removed because of technical issues. Of the remaining 12 one person chose to only submit the grids. The results of the survey reflect some of the findings present in the research described by Andriotis et al. and Aviv et al.

In the survey only 50% of the respondents replied “yes” to using a password manager. This is like the results gathered by Rick Wash et al. where 24% of their participants claimed they used a password manager (Rick Wash et al. 1-10). These two results reflect the fact that using password managers is not yet widely adopted. It is worth noting that the research done by Rick Wash et al. was focused mainly on college students outside of the computing and engineering fields (Rick Wash et al. 1-10). The next question in the survey asked about how people unlock their phones.

18% of the respondents claimed they use a swipe pattern to unlock their phone. 36% of the respondents claimed that they used a pin. 27% claimed they use biometrics. 9% claimed they use a password. And finally, 9% claimed they do not use a lock on their phone. Andriotis et al. had a similar question in their research but only asked for Pattern Lock, Pin, and other. The results they got 53.3%, 39.2% and 7.5% respectively. Adjusting the groupings to match those done by Andriotis et al., yields the following results: 20%, 40%, and 40% excluding the one person that did not use a lock. It is important to note that Andriotis et al. focused only on the android users gathering their data using an Android app published on the Google App Store (Andriotis, Tryfonas, and Oikonomou 115-126).

Next came the first 4x4 grid into which the respondent had to place the password “lazydog123” into. The reason for providing a 4x4 grid was to ensure that cell reuse was optional. The instructions did make it clear that cells could be reused as well as one of the examples provided along with the prompt also had cells reused. The results gathered reflected the research done by both Andriotis et al. as well as Aviv et al. Figure IV shows the results of the cells used by the respondents and during which time they were used.

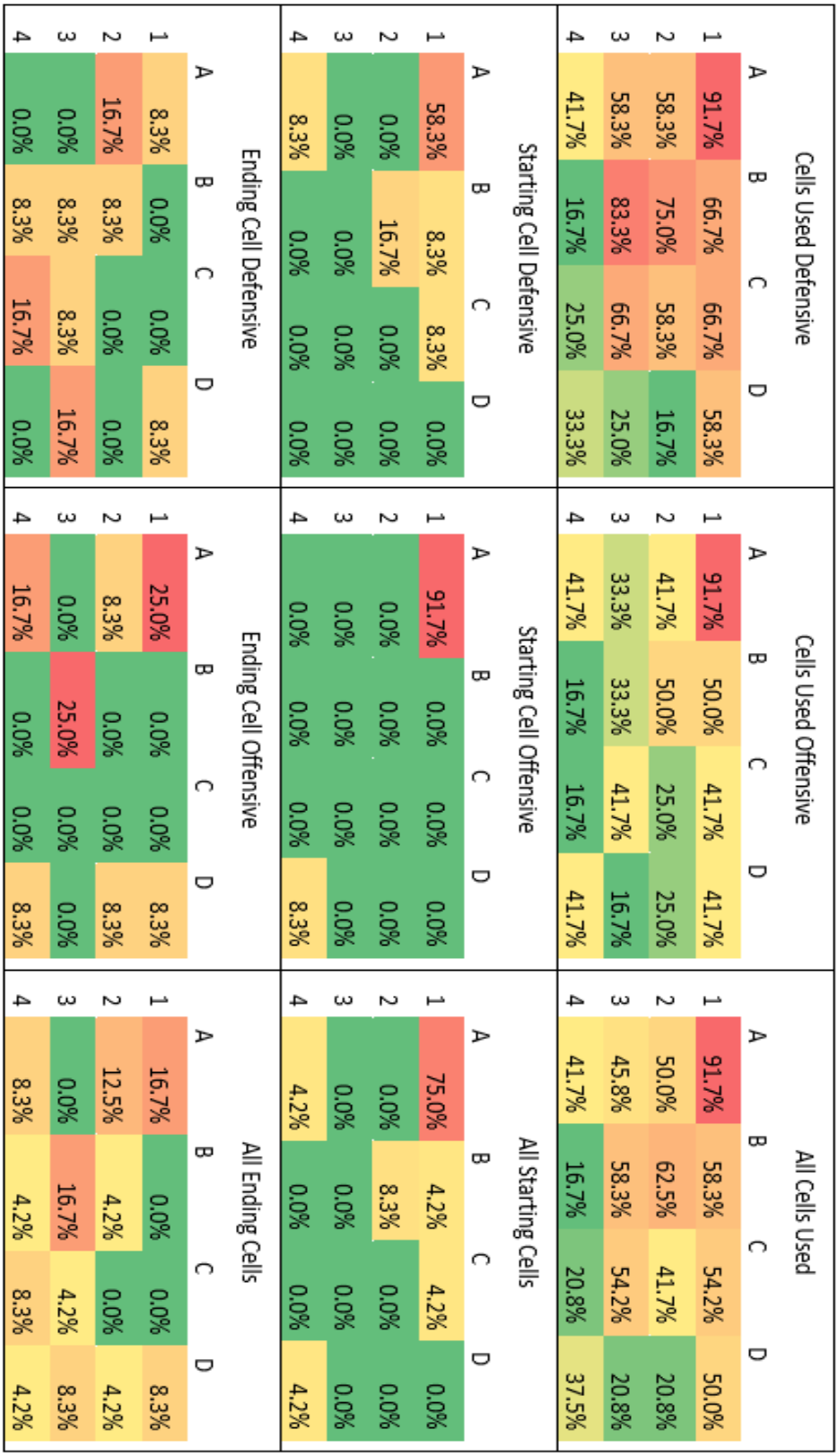


Figure IV Heat maps showing which cells were used most often and at what time.

To better compare to the results gathered by Aviv et al. as well as Andriotis et al. a cell was counted once even if multiple characters appeared in it. The results show that there is a tendency for majority of the patterns to be around the upper-left hand corner at the same time the use of the main diagonal, upper-left to bottom-right, can be observed in the heatmaps. Looking at the results gathered by Aviv et al. there is a much higher distribution of ending positions compared to their research where the bottom-right hand node occurring 21% (Aviv, Budzitowski, and Kuber 301-310).

In their research Aviv et al. also discuss common patterns they call tri-grams and quad-grams. Looking for those patterns in the defensive grid submissions it is hard to find them. Some of the patterns submitted skip around the grid and do not necessarily follow a distinct path.

	A	B	C	D
1	l	y	1	
2			o	2
3		d		3
4	z		g	a

Figure V Example of a more scrambled pattern

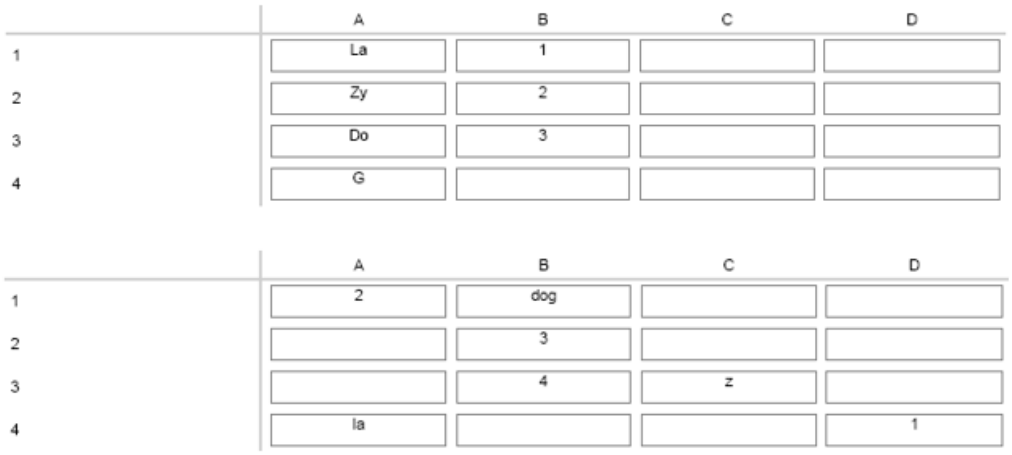
In Figure V it can be observed that the pattern starts in the upper-left hand corner then jumps around filling in cells almost randomly.

	A	B	C	D
1	g	1	2	3
2	o	l	a	
3	d	y	z	
4				

	A	B	C	D
1	g	1	2	3
2	o	l	a	
3	d	y	z	
4				

Figure VI Two different defensive submissions that match.

In the submissions two respondents provided the same defensive pattern. In both cases a spiral pattern was chosen. Cell reuse was also not common in the defensive patterns with only 42% of them having cell reuse. Another characteristic that was observed was that some submissions modified the password “lazydog123” by changing the characters.



	A	B	C	D
1	La	1		
2	Zy	2		
3	Do	3		
4	G			

	A	B	C	D
1	2	dog		
2		3		
3		4	z	
4	la			1

Figure VII Two submissions where the given password was altered.

Since the primary focus of the survey was to look at how the grid was filled out, neither of these submissions were excluded.

One defining feature of the offensive submissions is that they are a lot less random. Looking at the two heat maps there are much more distinct paths being followed by the offensive guesses.

Table IV Comparing the heatmaps of defensive and offensive grid patterns.

Cells Used Defensive				
	A	B	C	D
1	91.7%	66.7%	66.7%	58.3%
2	58.3%	75.0%	58.3%	16.7%
3	58.3%	83.3%	66.7%	25.0%
4	41.7%	16.7%	25.0%	33.3%

Cells Used Offensive				
	A	B	C	D
1	91.7%	50.0%	41.7%	41.7%
2	41.7%	50.0%	25.0%	25.0%
3	33.3%	33.3%	41.7%	16.7%
4	41.7%	16.7%	16.7%	41.7%

	A	B	C	D
1	la			zy
2				
3		g	12	3
4				do

Q5. Where do you feel this new password scheme will be easiest to use? (Any that apply)

- [ ] Unlocking a Phone
- [x] Web Login
- [ ] Computer Login

Q6.
If you wanted to gain access to someone's account and you knew their password, "lazydog123" put in your best guess of how that user filled out the grid.

	A	B	C	D
1	la			
2		zy		
3			do	
4				g123

Figure VIII Example of how the pattern changes between a guess and a defensive pattern.

Once again, two submissions guessed each other's password. Both submissions put the password in the upper-left hand corner.

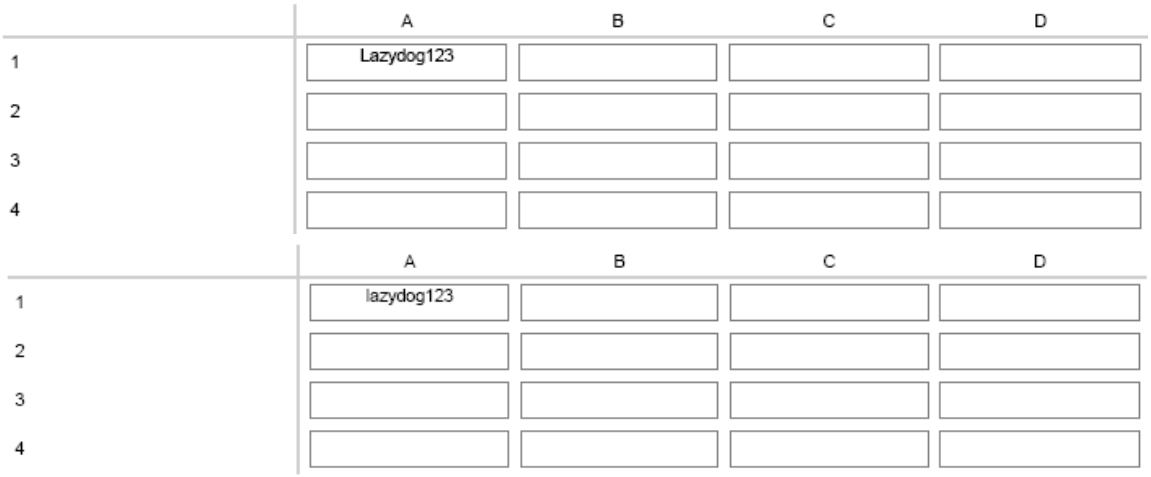


Figure IX Two offensive patterns matching.

As mentioned before the goal was to look at patterns of the submission that is why these two submissions are classified as a match, even though the password itself is altered one using an uppercase letter the other using a lowercase letter. Another interesting observation regarding the difference between the offensive and defensive patterns is the starting location.

In almost all offensive patterns the upper-left hand corner was used with only one submission using the lower-right hand corner. This is useful in demonstrating the potential limitations of human tendencies to lowering the mathematical complexity to just a few common patterns.

Table V Heatmaps of the starting locations for the defensive and offensive patterns.

Starting Cell Defensive				
	A	B	C	D
1	58.3%	8.3%	8.3%	0.0%
2	0.0%	16.7%	0.0%	0.0%
3	0.0%	0.0%	0.0%	0.0%
4	8.3%	0.0%	0.0%	0.0%

Starting Cell Offensive				
	A	B	C	D
1	91.7%	0.0%	0.0%	0.0%
2	0.0%	0.0%	0.0%	0.0%
3	0.0%	0.0%	0.0%	0.0%
4	0.0%	0.0%	0.0%	8.3%

These results mirror the results gathered by Aviv et al. where 37.5% of the patterns used the upper-left hand corner.

Interestingly the ending cell positions are much more scattered compared to the results gathered by Aviv's team. They had 21.6% of the patterns end in the lower-right hand corner with the upper-right being the second most used with 14.3%.

Table VI Heat maps for the ending cell.

Ending Cell Defensive				
	A	B	C	D
1	8.3%	0.0%	0.0%	8.3%
2	16.7%	8.3%	0.0%	0.0%
3	0.0%	8.3%	8.3%	16.7%
4	0.0%	8.3%	16.7%	0.0%

Ending Cell Offensive				
	A	B	C	D
1	25.0%	0.0%	0.0%	8.3%
2	8.3%	0.0%	0.0%	8.3%
3	0.0%	25.0%	0.0%	0.0%
4	16.7%	0.0%	0.0%	8.3%

The main reason why this could be the case is that cells are allowed to be reused meaning that the ending position of the pattern is harder to determine.

From all the results gathered it is clear that there is definite potential present in the new password scheme and the aforementioned limitations caused by human tendencies to fill out a grid are present; however, do not greatly impact the overall performance of the grid.

Table VII Heat map of all cells used.

All Cells Used				
	A	B	C	D
1	91.7%	58.3%	54.2%	50.0%
2	50.0%	62.5%	41.7%	20.8%
3	45.8%	58.3%	54.2%	20.8%
4	41.7%	16.7%	20.8%	37.5%

The use of the main diagonal as well as the first row and column can be seen; however, there is a much larger spread of use across majority of the grid focused around the upper-left hand side.

One could argue that using a 3x3 grid and inscribing it inside of the 4x4 grid would make guessing easier. This strategy was looked at by Aviv et al.; however, in the grid-based system cells can be reused and the cells do not have to be adjacent to each other. This means that inscribing 3x3 grid patterns in a 4x4 grid will be less fruitful than doing the same in an android swipe pattern grid.

Finally, looking at the results of where the respondent thinks this password scheme could be used there was a tie for web login and computer login both with 50%. This choice makes sense mainly from the easier input and navigation of using a keyboard and mouse.

Conclusion

The findings gathered in this preliminary research boost confidence in the potential of this system. Further research must be performed using an actual implementation of this password scheme. This would include implementing the login system on computers and performing further analysis.

As part of the analysis, a set of best practices for grids would be created. These would provide a foundation for password policies that could be implemented. The policies would then be tested to see the impact they have the on the performance of the grids. Time would also be devoted to guessing user passwords and analyzing the security similarly to the analysis done by Aviv et al. where they assigned strength based on guessability of the password. The influence of human factors will continue to be studied as well.

One human factor that was analyzed only theoretically was the ability to recall information stored in a grid. Having the users create their own patterns as well as passwords could change what cells are used. At the same time, it is speculated, that the new patterns could resemble the offensive patterns collected in this research. Being that they are more organized and potentially easier to remember. Another human aspect that would need further analysis is the impact of language on how the grid is filled in.

During the research it was speculated, but not tested, how the manor of filling in the grid would change based on the native language of the user. This could be analyzed by looking at the heat maps and seeing if the zone shifts from the upper-left hand corner to the upper-right hand corner, if the system is used by people whose native language is written right to left.

The exploratory research into Grid Passwords has revealed that the system has theoretical and practical potential. The results gathered provide a glimpse into what should be explored further, and at the same time reflect the results gathered by other groups conducting research on similar authentication systems. Focusing on human factors will also be critical. Overall, the system has potential, and warrants further exploration and implementation.

BIBLIOGRAPHY

Andriotis, Panagiotis, Theo Tryfonas, and George Oikonomou. "Complexity Metrics and User Strength Perceptions of the Pattern-Lock Graphical Authentication Method." Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics) 8533 LNCS (2014): 115-26. Web.

Aviv, Adam, Devon Budzitowski, and Ravi Kuber. "Is Bigger Better? Comparing User-Generated Passwords on 3x3 Vs. 4x4 Grid Sizes for Android's Pattern Unlock".ACM , Dec 7, 2015. 301-310. Print.

Di Santo, Serena, et al. "Working Memory Training: Assessing the Efficiency of Mnemonic Strategies." Entropy (Basel, Switzerland) 22.5 (2020): 577. CrossRef. Web.

Rick Wash, et al. "Understanding Password Choices: How Frequently Entered Passwords are Re-used Across Websites." Symposium on Usable Privacy and Security (2016): 1-10. Web.