Strength and Predictability of Graphical Passwords

Alex Tanasescu - Computer Science - 30041538,
Matthew Newton - Computer Science - 30094756
Delara Shamanian Esfahani - Computer Science - 30089408
Ramez Halasah - Computer Science - 30094242

Group 4

April 2023

***Abstract* - Patterns are a part of everyday life, which are subconsciously picked up by people. Can we predict the thought process of certain people based on common visual patterns found in modern society? In this paper, we aim to explore various human factors that come into place when setting graphical passwords. Our approach involves conducting tests and surveys on the authors of the paper in order to analyze and learn more about the predictability of graphical passwords. We will focus on three kinds of graphical passwords: the Android style 3x3, pick points on a picture, and a color pattern-based system. Despite our small sample size, we believe that this study offers valuable insights regarding the predictability of different graphical passwords and can be used as a stepping stone for further analysis on this topic.**

1 INTRODUCTION

As passwords become more and more notable throughout our lives, one must also question their security and predictability. In our research paper, we conduct multiple experiments on a sample size of 100 different passwords of three different graphical password types in order to analyze and measure their security and predictability. In these experiments, we investigate the thought process of why certain passwords were chosen and how an attacker could take advantage of these factors in order to gain knowledge about a user's password.

1.1 Proposed Work

Our proposed work consists of each of our group members generating 100 passwords of each of the three mentioned password types within a given time frame. We would do this using password generators coded up by one of our group members. At the end of the given time frame, we compile our passwords in a processable format and perform a quantitative analysis such as creating heat maps or bar graphs based on our results. After examining the results, we would perform a more qualitative analysis and reflect on why those results came out the way they did. This would then help us understand the reasoning behind why certain passwords are chosen more over others.

2 RELATED WORK

There have been quite a few papers discussing the predictability of graphical passwords, such as Løge and Røstad[3] where similar to us, they conducted a study that analyzed the predictability of unlock patterns, specifically the android 3x3 grid pattern. Their findings, like ours, concluded that certain patterns were a lot more common than others and could be used by attackers to hack into devices. Another study conducted by Dan Goodin[2], also on Android passwords, talked about how even the most seemingly complex Android patterns can easily be cracked since there are a limited number of combinations of lines and nodes. He also talked about how quite a few people base their lock password on a pattern that resembles the initial of a loved one such as a spouse or a pet, using this information one can easily filter and brute force to crack into the victim's phone. In our paper, we plan to further explore these ideas for our generated set of passwords and also try to find the reasoning behind why people choose certain passwords over others.

3 IMPLEMENTATION AND EXPERIMENTS

In order to test the strength of graphical passwords the following three password systems were devised:

- Andriod Password
- Colour-Based Password
- Picture Based Password

Over the course of this project, each member of the team generated 100 passwords of each password kind which we then compiled into a dataset to run our analysis on.

3.1 Android Password

The first password system is one many people are familiar with; the standard Android swipe-style password. In this password, a user will create a pattern by swiping their finger along

the 9 dots that are presented to them on the screen. Familiar Android password rules apply:
1. Can only swipe to cells using straight lines
2. Can not skip intermediate cells. For example, a user can not skip from cell 1 to cell 3 and must first pass through cell 2.
Since there are 9 dots it is natural that the maximum length password will be 9 dots. The minimum length of a pattern is 3 dots.

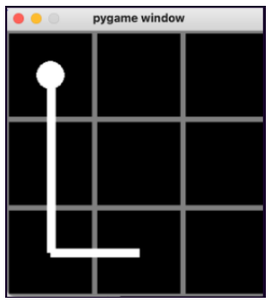

Figure 1: Android Password

3.2 Colour Based Password

The second password system works in a similar way to the Android password. Users can now select any cell they want without the restrictions that are present within the Android system. Namely, the sequence of colors would allow for any order as well as repetition of colours. Similarly to Android, users are limited to length 9 passwords.

Figure 2: Colour Based Password

3.3 Picture Based Password

The final system uses pictures for users to create a password. Each picture is broken down into a 3x3 grid. A user will create a password by clicking on a point on the image that they choose. The point that was chosen will fall within one of the cells of the grid that is under the image. Finally, to create a full password, 3 pictures are strung together so a user must click the correct spot on those 3 images that are presented to them. The 3 images are randomly presented to the user from a set of 6 pictures. Essentially a user will have to correctly choose 3 points on 3 separate 3x3 grids to be authenticated.

Figure 3: Picture Based Password

Figure 3 is an example of what a user could be presented with.

4 ANALYSIS

For each password system, a qualitative and quantitative analysis was performed in order to determine their relative strengths and predictability.

4.1 Heat Maps

For each password type a heat map highlighted the heavily trafficked cells. The heat maps were generated by overlaying every password that was generated and counting the number of times each cell appeared in a password. The more a cell was clicked, the darker the color of the heatmap.

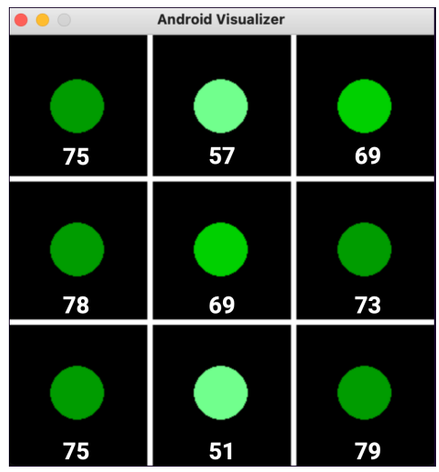

Figure 4: Android Heat Map

In Figure 4 we see that the participants of the study tended towards the outside of the grid with cells 2 and 8 having significantly fewer uses than the rest of the cells

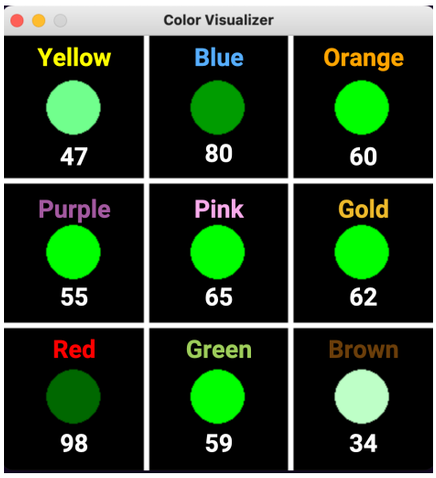

Figure 5: Colour Password Heat Map

As shown in figure 5 red is by far the most used colour from all the passwords with 98 occurrences among 100 passwords generated while brown only appeared 34 times. This heat map depicts the clear bias that users have had towards certain colours.

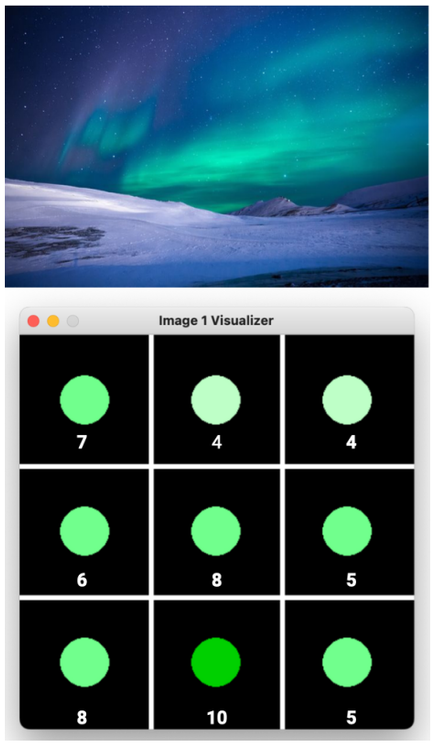

Figure 6: Picture 1 Heat Map

Picture 1 in Figure 6 was the one image that had the best entropy. This is attributed to the wide variety of objects to click on and very few recognizable landmarks, this in turn creates a somewhat uniform distribution as demonstrated in the heat map.

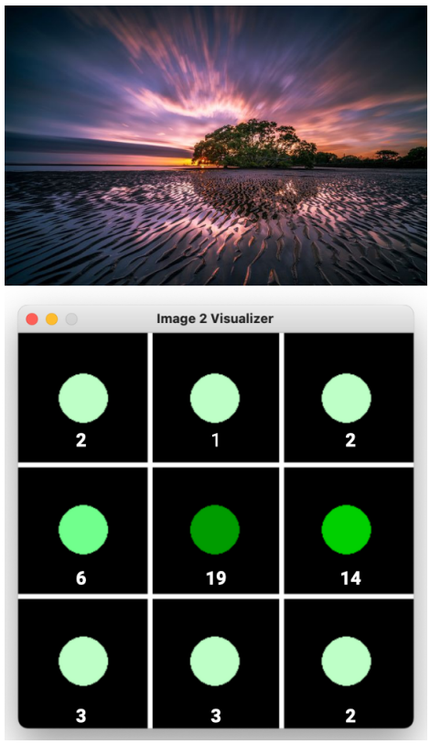

Figure 7: Picture 2 Heat Map

We can see from the heat map in picture 2 that a lot of people identified with the tree in the middle and clicked on that the most, this of course lowers the entropy of the password system that features this image. This image eventually had the lowest entropy among all the images.

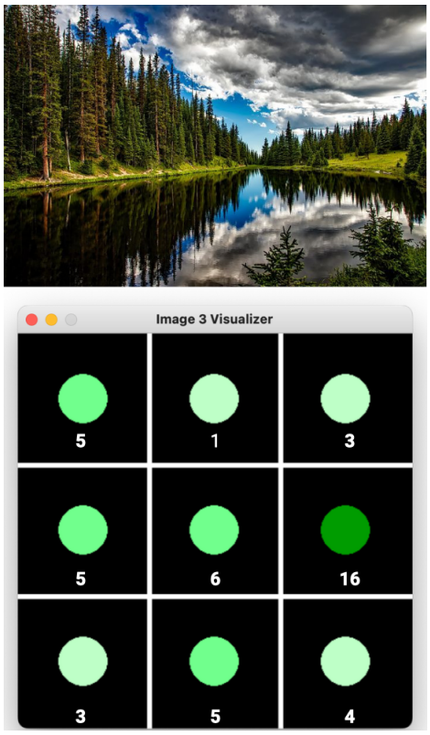

Figure 8: Picture 3 Heat Map

Figure 8 shows a very similar situation to figure 7. Since the right river bank sticks out, people are more likely to click on it and thus skewing the distribution which in turn reduces the entropy of the password system involving this picture.

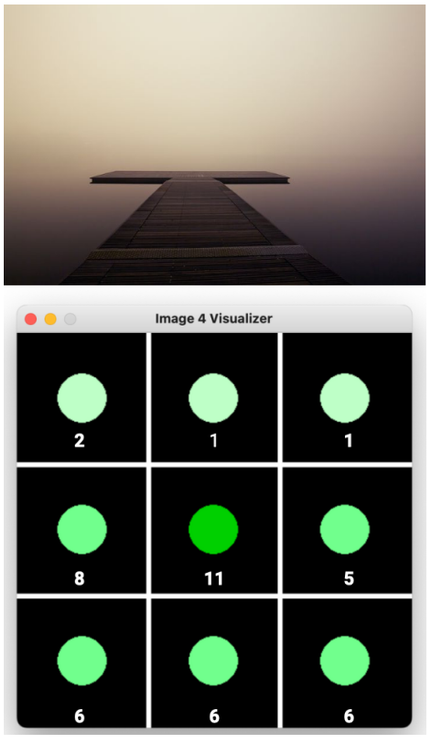

Figure 9: Picture 4 Heat Map

The results for the heat map for image four were quite a bit surprising, we thought that since the only object featured in the image was the dock, that would be the most clicked-on image, and while this was partially the case the clicks were a lot more spread out than we initially thought.

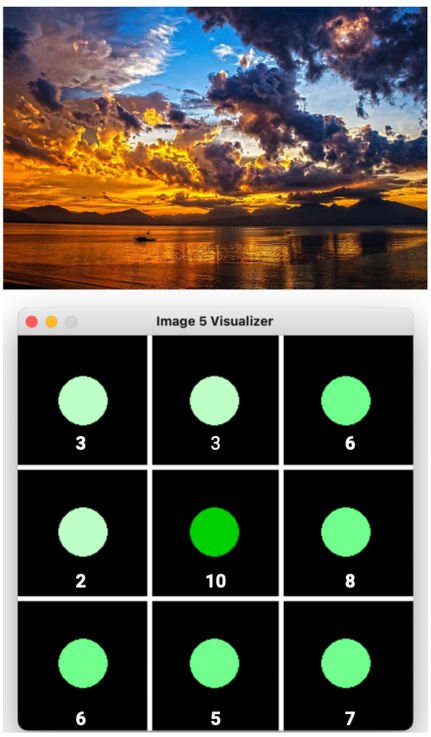

Figure 10: Picture 5 Heat Map

The image in figure 9, like image one, also presented us with high entropy. This is once again due to the large variety of clouds scattered across the picture and the fact that there isn't one set object that stands out for the user to click on the most.

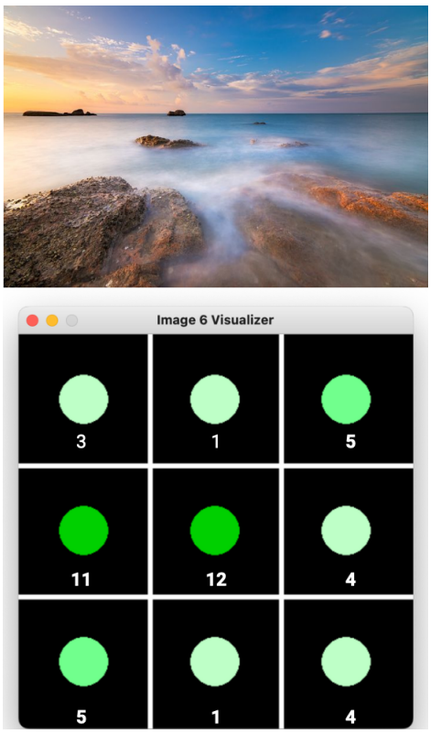

Figure 11: Picture 6 Heat Map

Image six features our second worst entropy, which is because the rocks on the left-hand side are very much in the foreground and stick out very evidently to the user, encouraging them to click on the rocks. This is demonstrated on the heat map as we can see that cells four and five are clicked on the most.

Although each image has a unique distribution of clicks compared to each other the user tends to click on the more recognizable objects. For example, in Image 2, the majority of clicks were toward the center cell which had a lone tree as a landmark. however, a picture such as Image 1 had a more smooth distribution since there were many places in the image that the user could click on (stars, rocks, mountains).

4.2 Entropy Estimation

For each password system, a maximum entropy was calculated assuming uniform distribution of all password combinations.

4.2.1 Android Password Entropy

The possible combinations of an Android password system which can be lengths 3 to 9 is 389436. This was found using a combination generator[]. From this, we can now calculate the max entropy using Hartley's Information as follows:

$$\log_2 N = \log_2 389436 = 18.571 bits$$

This leaves the result that an Android password system has max 18.571 bits of entropy.

4.2.2 Colour Based Password Entropy

Unlike the Android system, the colour password system has no input restrictions and the password lengths can be 1 to 9. In order to determine the total combinations we will sum all possibilities for each length:

$$\sum_{n=1}^{9} 9^n = 435848049$$

Once again we calculate the max entropy using Hartley's Information: $\log_2 N = \log_2 435848049 = 24.108 bits$

Clearly, the colour-based system has a higher maximum entropy than the Android system due to there being no restrictions on how a user can create their password.

4.2.3 Picture Based Password Entropy

Finally, for the picture-based password, 3 pictures are combined together to create a password. We know that each picture is split into a 3x3 grid giving 9 choices per picture. Thus the total combinations of passwords are:
$9 \times 9 \times 9 = 729$

Calculating max entropy using Hartley's Information we get: $\log_2 N = \log_2 729 = 9.510 bits$

As we can see, this password type has the least amount of entropy which is due to the fact that the password is composed of three pictures which are similar to having only a length three password option for the color password.

4.2.4 Experimental Picture System Entropy

As seen in figure 6 to 11 each picture has a unique distribution of clicks. Applying Shannon-Entropy to these distributions will provide a metric we can use to compare to the maximum entropy calculated earlier and determine the strength of a system.
The following pie charts represent the probability that a cell is clicked for its corresponding image. The bar graphs are a visualization of the amount of information each cell gives for that image.
Information of each cell is calculated by applying Shannon-entropy on the probabilities seen in the pie chart: $\log_2 p(i) \forall$ i $\epsilon [1...9]$
To calculate the entropy of an image we sum the information given by each cell of that image.

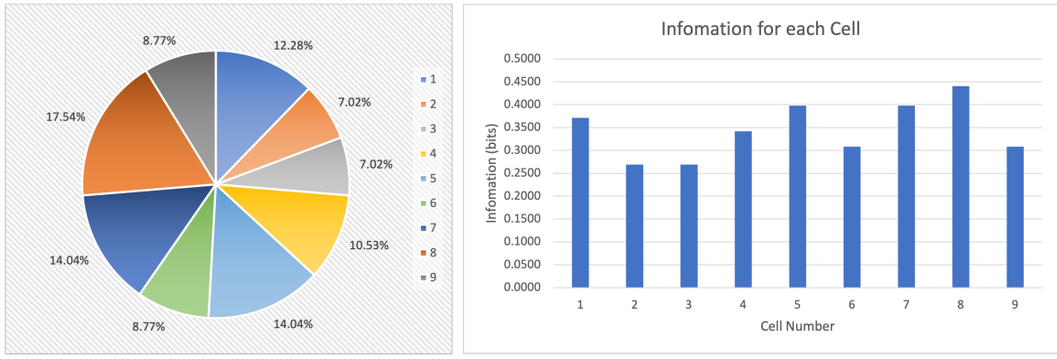

Figure 12: Picture 1 Probability Distribution and Information

Picture 1 displays the most uniform distribution of all the pictures which results in the highest entropy.
Picture 1 Entropy:
$log_2(p(1) = 0.1228) + log_2(p(2) = 0.0702) + log_2(p(3) = 0.0702) + log_2(p(4) = 0.1053) + log_2(p(5) = 0.1404) + log_2(p(6) = 0.0877) + log_2(p(7) = 0.1404) + log_2(p(8) = 0.1754) + log_2(p(9) = 0.0877) = 3.103$ bits.

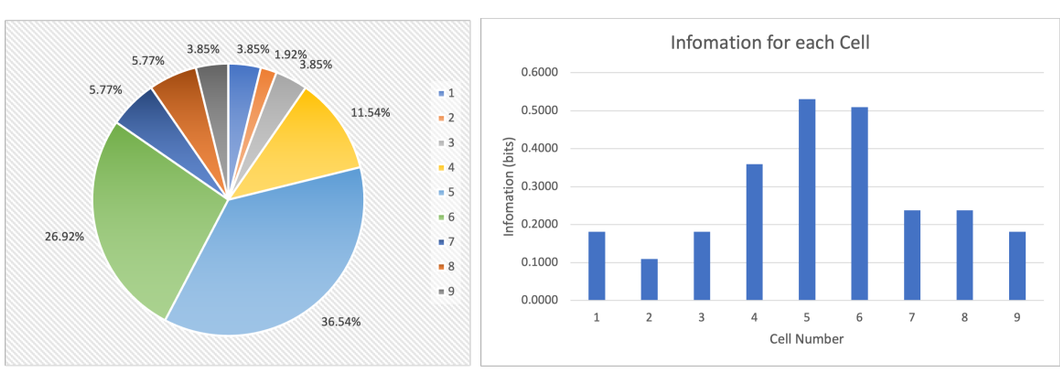

Figure 13: Picture 2 Probability Distribution and Information

Picture 2 displays the least uniform distribution of all the pictures which results in the lowest entropy.

Picture 2 Entropy:
$log_2(p(1) = 0.0385) + log_2(p(2) = 0.0192) + log_2(p(3) = 0.0385) + log_2(p(4) = 0.1154) + log_2(p(5) = 0.3654) + log_2(p(6) = 0.2692) + log_2(p(7) = 0.0577) + log_2(p(8) = 0.0577) + log_2(p(9) = 0.0385) = 2.527$ bits.

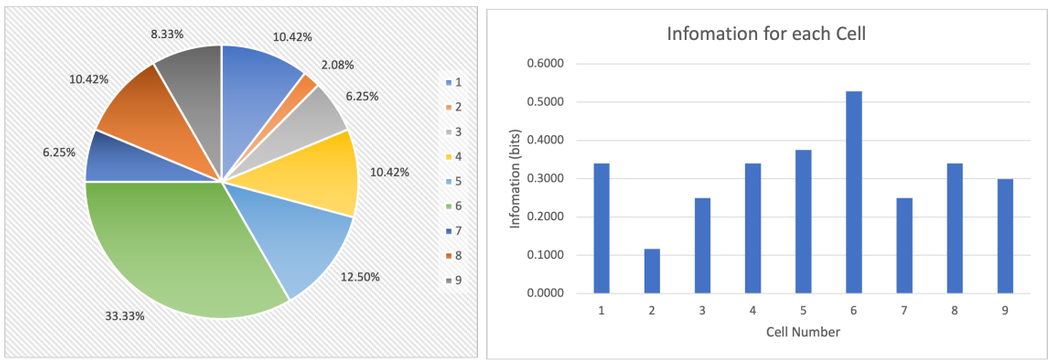

Figure 14: Picture 3 Probability Distribution and Information

Picture 3 Entropy:
$log_2(p(1) = 0.1042) + log_2(p(2) = 0.0208) + log_2(p(3) = 0.0625) + log_2(p(4) = 0.1042) + log_2(p(5) = 0.1250) + log_2(p(6) = 0.3333) + log_2(p(7) = 0.0625) + log_2(p(8) = 0.1042) + log_2(p(9) = 0.0833) = 2.838$ bits.

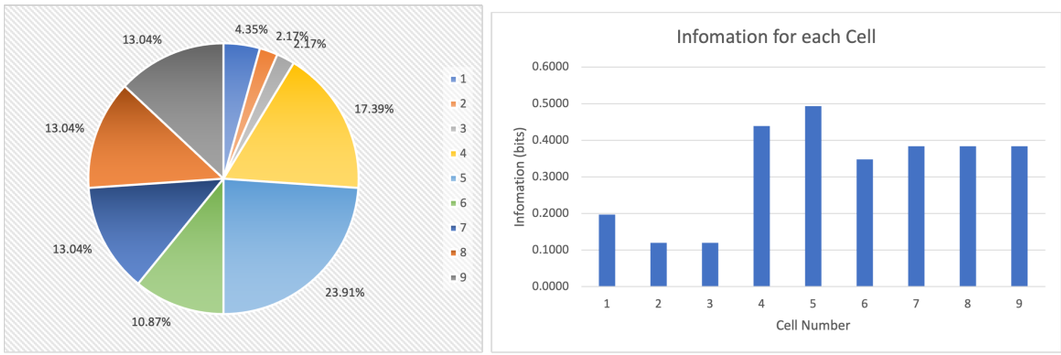

Figure 15: Picture 4 Probability Distribution and Information

Picture 4 Entropy:
$log_2(p(1) = 0.0435) + log_2(p(2) = 0.0217) + log_2(p(3) = 0.0217) + log_2(p(4) = 0.1739) + log_2(p(5) = 0.2391) + log_2(p(6) = 0.1087) + log_2(p(7) = 0.1304) + log_2(p(8) = 0.1304) + log_2(p(9) = 0.1304) = 2.867$ bits.

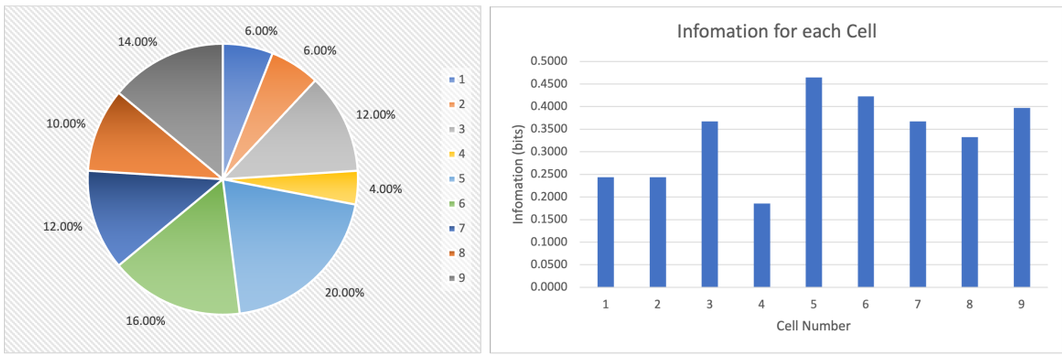

Figure 16: Picture 5 Probability Distribution and Information

Picture 5 Entropy: $log_2(p(1) = 0.06) + log_2(p(2) = 0.06) + log_2(p(3) = 0.12) + log_2(p(4) = 0.04) + log_2(p(5) = 0.2) + log_2(p(6) = 0.16) + log_2(p(7) = 0.12) + log_2(p(8) = 0.1) + log_2(p(9) = 0.14) = 3.024$ bits.

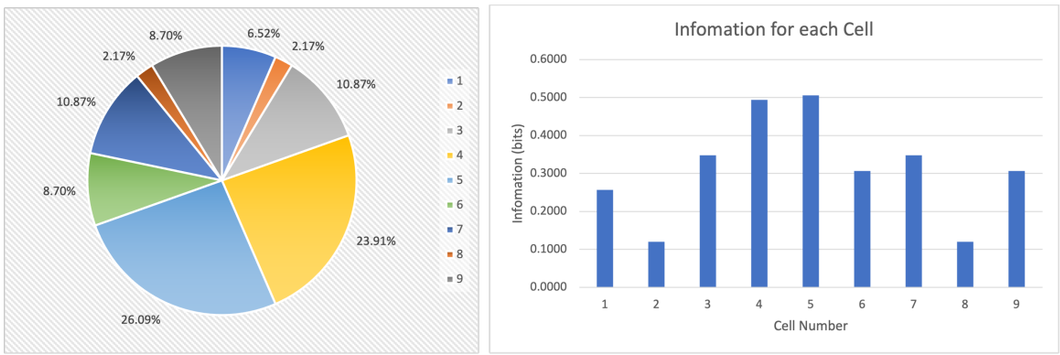

Figure 17: Picture 6 Probability Distribution and Information

Picture 6 Entropy:
$log_2(p(1) = 0.0652) + log_2(p(2) = 0.0217) + log_2(p(3) = 0.1087) + log_2(p(4) = 0.2391) + log_2(p(5) = 0.2609) + log_2(p(6) = 0.0870) + log_2(p(7) = 0.1087) + log_2(p(8) = 0.0217) + log_2(p(9) = 0.0870) = 2.805$ bits.

Finally adding up 3 pictures' entropy will give the entropy for a potential password. Here the best experimental entropy is 8.994 bits which is calculated by summing pictures 1, 4, and 5 entropy together.
The worst experimental entropy is 8.170 bits from summing pictures 2, 3, and 6 together.

4.3 Trends

Each password system displayed certain patterns and trends which are likely due to the biases that a user tends towards when choosing a password. Some factors that may affect how a user creates a password are:
- How relatable a password is to a user
- Ease of input for a password
- Memorability of a password

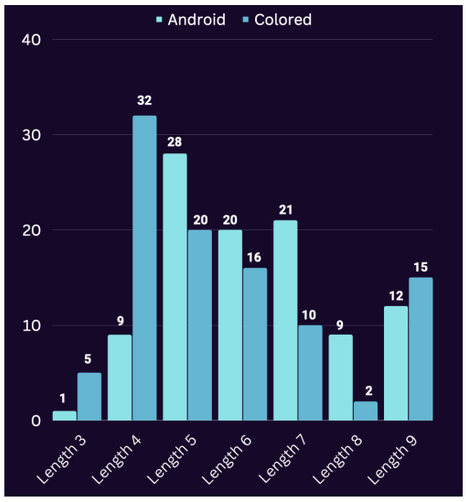

Figure 18: Length Comparison

Here we see that the participants of the study tended to input passwords of length 4 and 5 the most with the colored password having length 4 the most and the android password having length 5 mostly.

Another aspect of the Android and Coloured password systems is the frequencies that the cells are included in a password.
For the Android system, each cell can only be used in a password once so the frequency will reflect the same as what is seen in the heat maps. The Coloured system however can have a colour appear multiple times within one password:

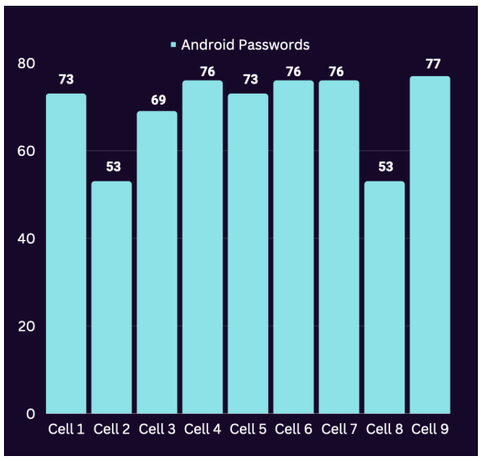

Figure 19: Android Cell Frequency

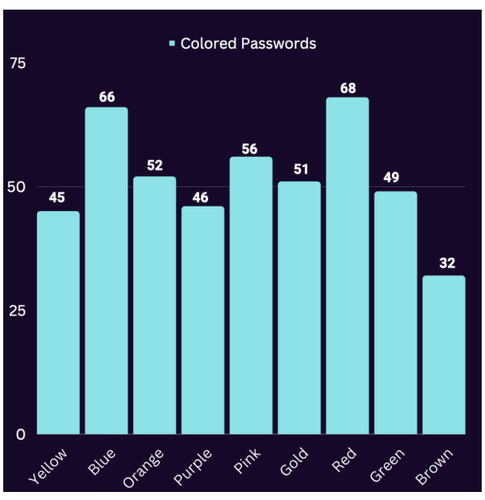

Figure 20: Coloured Password Cell Frequency

Comparing the Coloured cell frequency we notice that red appears in 68 different passwords but appears 98 times in total meaning that red appears in many passwords more than once.

This will have a direct effect on how predictable a password is. If an adversary knows a user will choose red more often than the other colours, the adversary will have a much easier time cracking the password.

5 CONCLUSION

Throughout our research project, we have performed an analysis on three different kinds of graphical password systems: the 3x3 Android grid pattern, the Color-pattern based password, and the picture-based "pick-points" password. Despite the small sample size, we have managed to identify common behaviors and tendencies that humans choose to make when setting up their passwords. When looking into this psychology of graphical passwords, we also looked into what key factors of graphical passwords hackers can exploit or take advantage of.

Our findings suggest that while graphical passwords can be effective, they are not immune to predictability and security issues. For example, when taking a look at our heat map for the color one, we can see how often red was used, and hackers can take advantage of this. This would make a brute-forcing attack considerably easier. We can prevent this by implementing strong graphical passwords alongside other security measures such as two-factor authentication; having these multiple layers of defense can help prevent breaches in security.

Ultimately our results did not differ that vastly from our proposal. We had mentioned the same experiment and implementations that we have in this research paper. One thing we did want to test out, that we were unable to due to time constraints, was memorability. Along with our original implementation, we also wanted to see which of the passwords that we generated we were also able to memorize. Memorability impacts predictability quite a bit since people tend to choose passwords that are easy to remember, however by doing this, the passwords become easier to predict and decrease their strength. This would've been tested over the course of the whole project by trying to recall each week as many passwords as we could until that week. We may approach this in future research.

Overall, our study offers valuable insights into the strength and predictability of different graphical passwords. However, we acknowledge that our sample size was small and limited to a specific population, which may affect the generalizability of our findings. Future research would include doing an experimental entropy estimation on more password systems, getting ethics approval to perform a user study on a much larger population, and generally looking into more factors that could affect password strength and predictability.

6 CONTRIBUTIONS

Matthew - Project Proposal, Generation of Passwords, Analysis - Calculations and Graphs, Three Questions, Presentation and Final Report
Alex - Project Proposal, Generation of Passwords, Coding Password Generators and Analysis Tools, Heat Maps and Quantitative Analysis, Three Questions, Presentation and Final Report
Delara - Project Proposal, Generation of Passwords, Three Questions, Presentation, and Final Report
Ramez - Project Proposal, Generation of Passwords, Three Questions, Presentation, and Final Report

References

[1] Delight-Im. (2014). List of all combinations for the android pattern lock. AndroidPatternLock. Retrieved April 1, 2023, from https://github.com/delight-im/AndroidPatternLock

[2] Dan Goodin - Aug 20, 2015 10:15 am U.T.C. (2015) New data uncovers the surprising predictability of Android Lock Patterns, Ars Technica. Available at: https://arstechnica.com/information-technology/2015/08/new-data-uncovers-the-surprising-predictability-of-android-lock-patterns/ (Accessed: April 2, 2023).

[3] Løge, M., & Røstad, L. (n.d.). "Tell Me Who You Are and I Will Tell You Your Unlock Pattern." Retrieved February 6, 2023, from https://ntnuopen.ntnu.no/ntnu-xmlui/bitstream/handle/11250/2380967/12788_FULLTEXT.pdf?sequence=1&isAllowed=y

[4] Rass, S., Schuller, D., & Kollmitzer, C. (2010, May 31). Entropy of Graphical Passwords: Towards an Information-Theoretic Analysis of Face-Recognition Based Authentication. Hal.science; Springer. https://link.springer.com/chapter/10.1007/978-3-642-13241-4_16