Hollow-Pass: Hide Your Pattern Password From Shoulder Surfers

JIAYI TAN, University of Twente, The Netherlands

This paper introduces Hollow-Pass, a new approach to combat shoulder surfing attacks on pattern passwords, a type of graphical password (GP) schema. Hollow-Pass does not necessitate external devices and renders the grid and pattern imperceptible to distant shoulder surfers. The usability and security of Hollow-Pass were evaluated through two types of user tests, online and onsite. Results showed that Hollow-Pass is successful in resisting shoulder surfing for simple patterns at various viewing angles (front, left-front, and right-front) and distances (1.0m, 1.5m, and 2.0m).

Additional Key Words and Phrases: Pattern password, Graphical password, Shoulder surfing attack, Global Precedence, Color perception, Image blurring.

1 INTRODUCTION

Graphical password (GP) systems are becoming increasingly popular in the authentication. GP can be categorized into four types: recognition-based, recall-based, cued-recall-based, and hybrid schemes (as shown in Appendix A)[1, 2]. One of the commonly used GP is Android Pattern Unlock (depicted in Fig.1)[3]. This is a specific example of the Pass-Go scheme[4], a recall-based GP that is inspired by the Asian game Go. During registration and authentication, users are prompted to draw a pattern in a grid as their credentials[5]. According to a study conducted in 2014[6], 40% of Android users use patterns as their credentials instead of PINs.

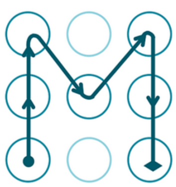


Fig. 1. Android pattern unlock

Several studies have highlighted that shoulder surfing is a significant security concern for GP systems[7–9]. This type of attack involves an adversary observing a user's device screen, keyboard, or mouse in a public place to steal their login credentials[10]. To combat this issue, various countermeasures have been proposed, including increasing the complexity of patterns and hiding part of the pattern using external hardware. However, a systematic literature review in 2018[11] found that 84 countermeasures had been proposed against shoulder surfing for pattern locking on smartphones and only 10 out of 84 techniques were pattern-based. The prevalent pattern-based techniques were increasing the complexity of the patterns while merely one technique was hiding part of the pattern by using external hardware. Since then, new approaches such as eyes-free[12], SysPal[13], Pass–O[14], TinPal[15], gaze tracking[16, 17], and using a swipe behavior-based mechanism[18] have been proposed, but these are typically designed for small-screen devices or require specialized hardware. This research aims to address the issue of shoulder surfing on both mobile devices and monitors without the use of specialized hardware by focusing on processing the pattern password itself.

The use of pattern passwords that incorporate typical features such as colored pictures with a grid structure can be effective in resisting shoulder surfing attacks, as mentioned in a previous study[19]. Researchers [20, 21] have found that the human visual system processes objects over time, with the recognition of general or global objects preceding that of detailed or local features. This phenomenon, known as global precedence, can be leveraged in the design of pattern passwords. By displaying the pattern password to users at a local level, but obscuring it at a global level, it can be protected from being observed by shoulder surfers.

This research aims to develop a new mechanism for pattern passwords, called Hollow-Pass, by integrating the techniques of View manipulation and Image degradation, as identified in a review of anti-shoulder-surfing techniques by Aris and Yaakob [11]. The usability and security of Hollow-Pass will also be evaluated through a small-scale user test.

1.1 Research Questions

Research question 1: To what extent does a pattern drawn on a distorted grid layout, as illustrated in Fig.2, prevent shoulder surfers from identifying a correct authenticated pattern at distances of 1.0m, 1.5m, and 2.0m?

Research question 2: To what extent do color contrast and global precedence prevent shoulder surfers from identifying the pattern at distances of 1.0m, 1.5m, and 2.0m?

Research question 3: To what extent do color contrast between the pattern password and background image, and the distorted grid layout affect the usability and security of Hollow-Pass?

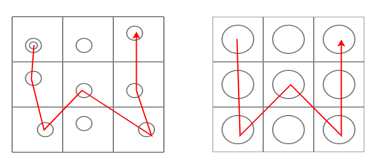


Fig. 2. Left: distorted grid layout; Right: normal grid layout

1.2 Structure

The structure of this research is as follows: Section 2 provides an overview of previous studies that are relevant to the research. Section 3 details the methodology that will be employed in the research.

TScIT 38, February 3, 2023, Enschede, The Netherlands

Section 4 analyzes the results and limitations of the research. The research concludes with a section on Acknowledgements, References, and an Appendix.

2 RELATED WORK

2.1 Pattern grid layout

3x3 is a conventional Android pattern grid size. A previous study has shown that the 3x3 pattern is easy to be guessed and attacked[22]. To increase security, larger grid sizes are encouraged. However, researchers have found that there is little improvement in security from changing the grid size to 4x4[23]. Therefore, instead of increasing the grid size, another alternative is to increase pattern complexity by adapting the 3x3 grid based on 9 points.

Various patterns can be created based on a 9-point layout, but a random layout may be difficult for users to memorize, which makes users prone to choose simple patterns that are vulnerable to attack. Researchers have designed new grid layouts to improve security, such as trapezium[4], circle[4, 14], and house[4]. The password space size of circle and that of house are larger than the original Android grid layout, while their overall recall success rates do not have significant differences[4]. This implies different grid layouts can improve security while maintaining good usability.

2.2 Spatial frequency

Spatial frequency (SF) [24] refers to the number of cycles or oscillations of a pattern that occur within a given distance, typically measured in cycles per unit distance, such as cycles per degree (cpd). In the context of image processing and computer vision, spatial frequency is often used to describe the amount of detail or texture present in an image. High spatial frequencies correspond to fine details and textures, while low spatial frequencies correspond to larger, coarser structures[25].

2.3 Visual acuity

Visual acuity (VA) is the ability of the human visual system to clearly perceive the details of an object. It is commonly measured using the Snellen chart, which is typically viewed at a distance of 6 meters (20 feet). A normal VA is typically represented as 6/6 or 20/20, and corresponds to a line of letters on the chart that subtend an angle of 5 minutes of arc. The Snellen "E" letter, which is made up of 3 strokes and 2 gaps, with each stroke and gap subtending 1 minute of arc, is often used as an example (as shown in Fig.3). When considering each stroke (1 minute of arc) as a peak of a sine wave and the white gap between strokes as a trough, a normal eye visual is equivalent to 30 cycles per degree (cpd) (as shown in Fig.4)[26].

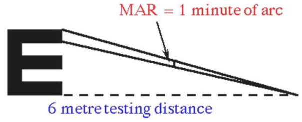


Fig. 3. For visual acuity of 6/6 (or 20/20), every stroke of the letter in the corresponding line subtends to 1 minute of arc. This is called the minimum angle of resolution(MAR)., For the whole letter, the angular resolution is 5 minutes of arc [26].

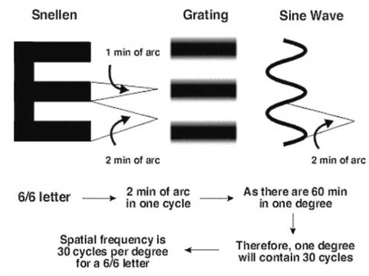


Fig. 4. Conversion of normal visual acuity(in minute of arc) to cycle per degree(cpd) [26].

2.4 Global Precedence

Researchers[20, 21] have shown that lower spatial frequencies (SF) tend to facilitate global perception, while higher SFs tend to facilitate local perception. When people begin to extract low and high SFs from an image simultaneously, the visual processing time follows a "coarse-to-fine" strategy, in which they identify global or general objects faster and more accurately than local or detailed features. For example, one might first identify a tree before recognizing its leaves and branches. Another well-known example is Navon's experiment [21], in which participants were quicker to identify the overall shape of an "H" made up of smaller "X"s, as shown in Fig.5. This effect, known as global precedence, suggests that humans can choose to perceive the global level of a scene alone, but cannot skip over global perception to local perception in a single action.

For this research, we won't go into further psychological details but utilize this concept to process the background image.

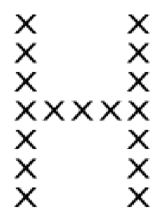


Fig. 5. Navon's stimulus: it has a global feature: an H, and local feature: an X. People tend to identify an H faster than an X [27].

2.5 CIE Color System and ΔE

Color is one of the important characteristics of digital images and affects human visual perception. Therefore, one of the research goals is to improve the global precedence effect by adjusting the image's color difference.

The CIE color system, developed by the International Commission on Illumination, provides a numerical way to describe all colors that are visible to the human eye. Unlike the RGB color model, the color definitions in the CIE color system are absolute, unambiguous, and not dependent on the device or display specifications. The CIE LAB (L*a*b*) model, published in 1976, is widely accepted as a way to quantitatively measure perceived color. This model has three components [28]:

- L* represents the lightness (from 0 to 100), where 0 means black and 100 is white.
- a* represents the green to the red axis (from -128 to +127)
- b* represents the yellow to the blue axis (from -128 to +127)

To understand how the human eye perceives color difference, CIE color system uses the metric ΔE [28, 29]. For CIE LAB, $\Delta E \approx 10$ represents the color difference visible at glance.

3 METHODOLOGY

In this section, we outline the procedures we employed to address the research questions. To accomplish this, we utilized a web development framework[30] and a pattern unlock grid template[31] to develop a website specifically for the user test, and made it available to all participants by deploying it on a web hosting platform.

Strength meter. We additionally created a metric to assess the usability and security of Hollow-Pass concerning each research question (See Table.1).

RQ	Usability	Security
1	• Key space	• Key space • Pattern complexity
2	• Background spatial frequency • Color difference: ΔE	• Background spatial frequency • Color difference: ΔE
3	• Recall error rate • System usability scale	• Observed pattern accuracy

Table 1. Usability and security meter

Pattern Drawing Rule. We developed the pattern drawing rules based on the design of Tupsamudre et al.[14].

(a) the pattern should be drawn using straight lines and without lifting the hand,

(b) the pattern should connect at least 4 nodes and a maximum of 9 nodes,

(c) a node cannot be connected more than once,

(d) Unlike the conventional 3x3 patterns, an unconnected node can be skipped if it lies along the path of a pattern. For example, the user can draw a line segment $1 \longrightarrow 3$ without visiting 2.

Notations. For convenience, we define the following representations for node label and pattern shape.

(a) Node label: For convenience, all nodes in the original 3x3 grid are labeled from 1 to 9 in row-major order. The upper-left node is labeled as 1 and the bottom-right node is labeled as 9.

(b) Pattern representation: A pattern can be written as an ordered sequence of nodes: e.g., 1475963 (See Fig.2).

Key space. The key space refers to the number of valid patterns in Hollow-Pass. The number of valid r-node patterns in a 3x3 grid can be determined by using a mathematical formula:

$$P(n,r) = \frac{n!}{(n-r)!} = \frac{9!}{(9-r)!} \tag{1}$$

where n is the node count of the grid. For a 3x3 grid, n is 9. To sum up valid patterns from 4 to 9 nodes, we used the formula:

$$\sum_{r=4,9} P(n,r) = \sum_{r=4,9} \frac{n!}{(n-r)!} = \sum_{r=4,9} \frac{9!}{(9-r)!} \tag{2}$$

Pattern Complexity. The visual complexity of a pattern is closely related to various factors, such as the number of connected nodes, pattern length, pattern intersections, and pattern overlaps (examples of which are provided in Fig.6). To quantify this complexity, we used the formula proposed by Sun et al.[32]:

$$PS_P = S_P \times log_2(L_P + I_P + O_P) \tag{3}$$

where PS_P is the pattern strength score, S_P is the node number of pattern, L_P is the pattern length, I_P is the intersection number of pattern, O_P is the overlap number of pattern. By dividing the range of scores into five equal segments, patterns can be categorized into five corresponding levels: very weak, weak, medium, strong, and very strong.

(a) Intersection example 125846: one intersection point at node 5. **(b) Overlap example 2564: one overlapped part 56.**

Fig. 6. Examples of intersection and overlap

3.1 Answer Research Question 1

The approach for addressing Research Question 1 (RQ1) is to enhance the randomness of the grid layout. Even though shoulder surfers may observe the pattern, they would be unable to utilize the observed pattern to authenticate due to the varying distribution of nodes and grid layout on each occasion. To accomplish this, we implemented the following two steps:

(a) Divided the 3x3 grid into nine equal patches and randomly placed each node within its patch.

(b) Rotated the grid by 45 degrees clockwise or counterclockwise.

(a) Divided the 3x3 grid into nine equal patches and randomly placed each node within its patch.

Excessive randomness can have a negative impact on legitimate users by making it difficult for them to recall the pattern and potentially encouraging them to create simple, easily memorable patterns that are vulnerable to attack. To strike a balance between usability and security, we scattered the nodes within their designated grid squares while displaying grid borders to aid users in recognizing the node locations. As the nodes were randomly scattered and did not align with the grid borders, users can easily access non-adjacent nodes without connecting through intermediate nodes (as demonstrated in Fig. 7). This increased the range of reachable nodes and the complexity of patterns.

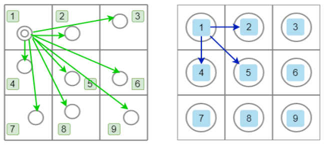
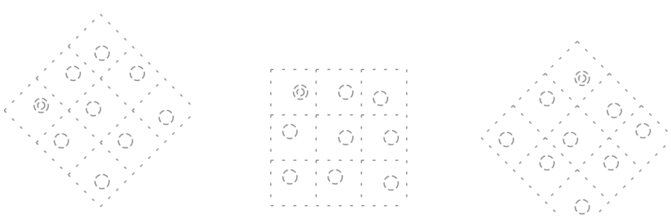


Fig. 7. Left: distorted grid layout. For node 1, it can reach any other node directly and skip the intermediate node. Pattern 17 is allowed. Right: conventional grid layout. For node 1, it can only reach its adjacent nodes: 2, 4, and 5. Pattern 17 is not allowed, it should be 147.

(b) Rotated the grid by 45 degrees clockwise or counterclockwise.

Based on the assumption that users may have difficulty recognizing patterns that have been rotated more than 45 degrees, we have defined three possible grid layout rotations in total (as shown in Fig.8): a 45-degree counterclockwise rotation, the original layout, and a 45-degree clockwise rotation. The system would randomly display one of these three variations each time. To assist users in identifying the rotation direction, node 1 was circled as an indicator. For legitimate users, they drew their patterns in the same ordered sequence of nodes as the original grid with no difference. However, for shoulder surfers, due to the random rotation of the grid, it would be more challenging for them to identify the correct pattern without knowing the current grid orientation.



(a) 45° **counter-clockwise** **(b) Original** **(c)** 45° **clockwise**

Fig. 8. Three rotations of grid layout

3.2 Answer Research Question 2

The approach for addressing Research Question 2 (RQ2) is to create and strengthen the global precedence effect in Hollow-Pass. We accomplished this by implementing the following two steps:

(a) Developed the global precedence effect by converting the foreground grid into a dashed line pattern and adjusting the spatial frequency of gratings in the background;

(b) Reinforced the global precedence by adjusting the color difference (ΔE) between the grid and the background.

(a) Developed the global precedence effect by converting the foreground grid into a dashed line pattern and adjusting the spatial frequency of gratings in the background.

Human visual perception system is more responsive to low spatial frequencies for global processing and more responsive to high spatial frequencies for local processing, as per the research in [20]. According to Kalloniatis and Luu[26], sinusoidal gratings can be represented in terms of SF and vice versa. We used the Python open-source package PsychoPy to generate sinusoidal gratings as the background image. The background contains four layers of gratings corresponding to four orientations: 0°, 45°, 90° and 135°, to cover every border of the foreground grid (see Fig.9). Spatial frequencies in cycle per degree (cpd) of four layers were represented in the list, respectively.

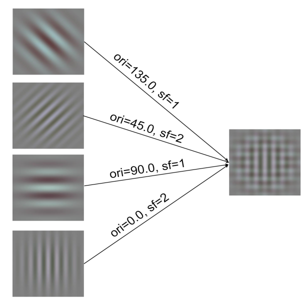


Fig. 9. An example of background gratings: every background is composed of four layers of gratings, corresponding to four orientations: 0°, 45°, 90° **and** 135°**. The spatial frequency of every layer is in the range of (1,3). The spatial frequency of this example is sf=[1,2,1,2]**

To determine the threshold of the auto-generated spatial frequency, we experimented with spatial frequencies ranging from 0 to 30 (the normal visual acuity, see Fig.4). Examples of comparison results are shown in Appendix B. Based on the results of our perceptual performance evaluations, we chose a lower bound of 1 and an upper bound of 3 (exclusive upper bound) for the spatial frequency of each layer of gratings. All layers of gratings were masked with a default 2-D Gaussian filter (sd=3) in PsychoPy.

Additionally, to obscure the foreground contour at the global level, we broke down the pattern and grid into local features through the use of a dashed line drawing technique.

(b) Reinforced the global precedence by adjusting ΔE between grid and background.

We assumed that if the color of the foreground grid is significantly different from the background, shoulder surfers may easily observe the grid orientation and the positions of each node. To prevent this, we limited the color contrast between the foreground grid and the background gratings to a specific level. To achieve this, we converted the grid's RGB color to CIE LAB using an algorithm written by Manoj Pandey[33], and then calculated the color difference (ΔE) between the grid and the background image (see Equation 4):

$$\begin{aligned}\Delta E_{Lab} &= \sqrt{(L^*_{grid} - L^*_{bg})^2 + (a^*_{grid} - a^*_{bg})^2 + (b^*_{grid} - b^*_{bg})^2} \\ &= \sqrt{(\Delta L^*)^2 + (\Delta a^*)^2 + (\Delta b^*)^2}\end{aligned} \quad (4)$$

where *bg* represents the background image.

To minimize the generation time of background images during user testing, five pre-generated images were utilized. The mean difference in color, as measured by ΔE, between the grid color and the background image was 10.817 (sd = 3.998). Similarly, the mean difference in color between the pattern line color and the background image was 31.841 (sd = 1.646). During the test, a random selection of the five pre-generated images was displayed on the user interface each time the users created a pattern password on the grid. This approach is expected to have the desired visual effect, as illustrated in Appendix B: Fig.18b and Fig.18c.

3.3 Answer Research Question 3

The approach for addressing Research Question 3 (RQ3) is to evaluate the usability and security of Hollow-Pass. Usability is measured through the recall error rate and the System Usability Scale (SUS), while security is measured through observed pattern accuracy.

Online test. There were seven sections in the online user test (See Fig.10) [34]. Participants were first asked to provide consent and were given instructions on how to use Hollow-Pass. They were then asked to register a pattern on the grid for the first time, with only patterns containing at least 4 nodes and no more than 9 nodes being considered valid. Demographic information such as gender, age, quality of eyesight, testing device, and pattern password experience were also collected.

To evaluate the anti-shoulder surfing effect, participants were asked to identify six pattern passwords that had been simulated to reflect the perceptual effect at viewing distances of 1.0m, 1.5m, and 2.0m. The accuracy of their answers was used to evaluate shoulder surfing resistance. Participants were then asked to recall and redraw their first-time patterns, with their recall error rate being used to verify password memorability and usability. Finally, a set of SUS questions were asked to further evaluate usability. On average, the online test took between 10-15 minutes to complete.

(1) Purpose of informed consent: Participants were informed about (and consented to) the user test at the beginning of the test. This included the purpose of the user test, alternatives of the procedure or intervention, and potential risks. (see Appendix C.1).

(2) Instructions: Participants were instructed to draw a pattern on either the desktop website or the mobile version of the site during the test (see Appendix C.2).

(3) 1st time drawing (Registration): Participants were instructed to create a username and create a Hollow-Pass pattern, which they then confirmed in the registration phase (see Appendix C.3).

(4) Demographic questions: Participants were asked to provide demographic information, including gender, age, quality of eyesight, testing device, and prior experience with pattern passwords (see Appendix C.4).

(5) Identify pattern passwords: Participants were asked to identify six pattern passwords that had been simulated to reflect the perceptual effect at viewing distances of 1.0m, 1.5m, and 2.0m. These

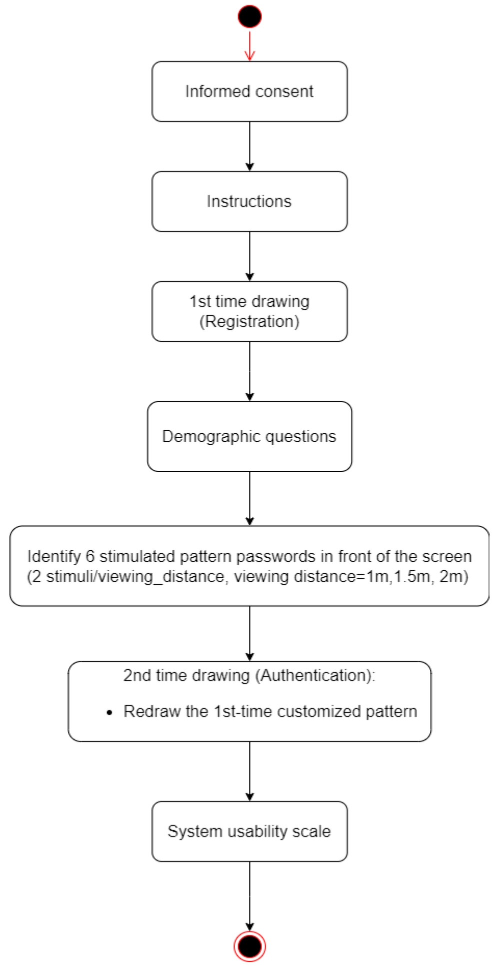


Fig. 10. Online user test procedure

stimuli were created to virtually test the perceptual effect at different viewing distances, taking into consideration factors such as perceived resolution and perceived size (see Appendix C.5).

- **Perceived resolution.** The human visual system has been found to closely resemble a low-spatial-frequency pass filter. Research by Pappas and Neuhof[35] has determined that the impulse responses of the 1-D eye filters closely match a Gaussian shape with appropriate standard deviation. Specifically, at a resolution of 300 dots per inch (dpi) and a viewing distance of 30 inches (0.76m), the impulse response of the 1-D eye filter is identical to that of a Gaussian filter with a standard deviation (σ) of 1.5 and spacing of the dots (τ) of 0.0095 degrees. Based on this understanding, we utilized a Gaussian filter to simulate the perceived resolution at various viewing distances by adjusting the standard deviation. The procedure for perceived resolution simulation is described in detail in the appendix (see Appendix D).

- **Perceived size.** According to Emmert's law, the perceived image size changes proportionally with its distance from the observer controlling for the visual angle[36]. Assuming that the observer had normal visual acuity and the default viewing distance was 0.5m from the screen to the observer's eye, we scaled the original stimuli size (340x340 px) for viewing distances of 1.0m, 1.5m, and 2.0m to simulate the perceived size.

To assess the security of the system, participants were presented with six pattern identification questions. Each question involved two Hollow-Pass patterns (with complexity levels of weak or medium) that were stimulated at three different viewing distances: 1.0m, 1.5m, and 2.0m. The pattern complexity was estimated using Equation 3. For each question, participants were asked to identify the correct pattern among four options, and the accuracy of their responses was used to evaluate the Hollow-Pass's security.

(6) 2nd time drawing (Authentication): Participants were instructed to redraw their initial pattern on a grid. Should a participant be unable to recall their initial pattern, they were permitted to move on to the next question. The accuracy of the participant's redrawn pattern in comparison to their initial pattern was utilized to determine the usability of the password.

(7) System usability scale: A 5-point Likert scale assessment, utilizing various levels of agreement, was administered to gauge participants' perceptions of the pattern, their experience with Hollow-Pass, and their affinity towards it. Subsequently, participants were requested to provide comments regarding the system (as outlined in Appendix C.6).

On-site test. To ensure reliability, a small-scale sampling test was conducted on-site. The onsite test was conducted using an Acer TravelMate P2 laptop, which has a screen size of 14 inches and a pixel density of 164.64 pixels per inch. The purpose of utilizing a laptop as the testing device in our onsite test was to examine the effectiveness of the Hollow-Pass system in resisting shoulder surfing on a large display.

The laptop screen was utilized as the center of a circle, and the viewing distance was calculated as the radius extending from the screen to the observer's position. The semicircle was divided into four equal sections, with each quarter being considered as a distinct viewing angle position: left-front, front, and right-front, as illustrated in Fig.11.

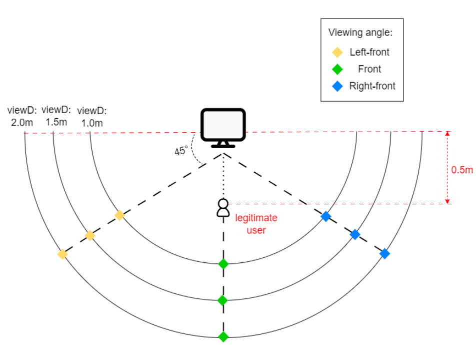


Fig. 11. Visualisation of on-site user test

The researcher, acting as a legitimate user, sat at a distance of 0.5m, and utilized a mouse to draw six weak patterns on the screen. The password strength of these patterns was evaluated using Equation 3. For each pattern, participants were instructed to stand at three distinct viewing angles (front, left-front, right-front) and a specified viewing distance from the screen, and to then draw what they observed on a test form. This measure aimed to investigate the resistance of the mechanism for identical patterns to shoulder surfing at various viewing angles. The test included three viewing distances (1.0m, 1.5m, and 2.0m) and was intended to investigate the resistance of Hollow-Pass for weak patterns to shoulder surfing at different viewing distances. Participants were permitted to request that the researcher redraw the pattern.

4 RESULTS & CONCLUSION

4.1 Password Strength

Key space. We conducted a comparison of the key space between a conventional 3x3 grid (a typical Android pattern unlock grid, see Fig.1) proposed by Sun et al.[32] and that of Hollow-Pass, which was calculated using Equations 1 and 2. The results, as presented in Table 2, indicates that the key space of Hollow-Pass is larger than that of the conventional method, increasing from 389,112 to 985,824, particularly in the cases of node size 4, 5, 6, and 7. This suggests that users can create more diverse patterns, making it more challenging for shoulder surfers to recognize them, without necessarily increasing the number of nodes.

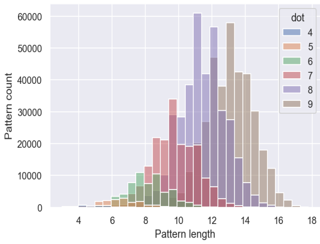
# of nodes	Conventional GP	Hollow-Pass
4	1,624	3,024
5	7152	15,120
6	26,016	60,480
7	72,912	181,440
8	140,704	362,880
9	140,704	362,880
Total	389,112	985,824

Table 2. Key space comparison between conventional GP and Hollow-Pass

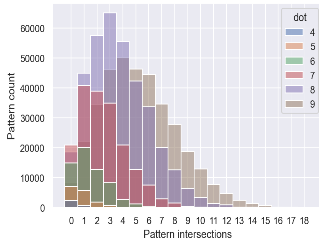
Pattern complexity. The password strength score of all valid patterns was calculated using Equation 2, taking into account pattern length (Fig. 12a), pattern intersections (Fig. 12b), and pattern overlaps (Fig. 12c). The distribution of length, intersections, and overlaps demonstrate that, as more nodes were used, the pattern became longer, and had more intersections and overlaps, making the pattern more perceptually complex. The score ranged from 6.34 to 46.81 (Fig. 12d).

4.2 Demographic information of participants

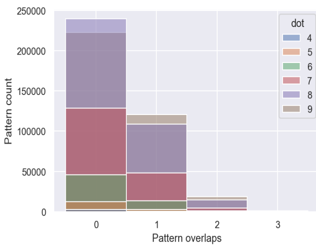
30 participants (15 female) between the ages 20-80 participated in the online user test, and 19 undergraduates (10 female) between the ages 20–23 from the University of Twente participated in the onsite user test. All participants provided informed consent, which was approved by the University of Twente's committee for ethical concerns (as shown in Table 3). The majority of the participants (21 out of 30) were in the 20-29 age group, while 2 participants were in



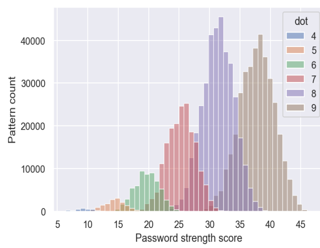
(a) Pattern length



(b) intersection



(c) overlap



(d) password strength score

Fig. 12. Distribution of valid patterns attributes.

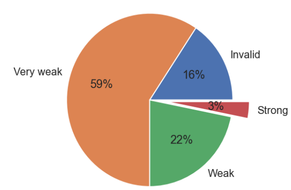
the 30-49 age group, 6 were in the 50-59 age group, and 1 was over 60. The visual acuity of the participants was recorded in decimal Snellen notation, with 11 participants having an eyesight result lower than 0.8, 6 between 0.8-1.0, 5 higher than 1.0, 6 uncertain of their result but wearing glasses, and 2 uncertain and not wearing glasses. Most of the participants were familiar with pattern passwords, with 17 having used it before and 11 having knowledge of it but not having used it. Only 2 participants were unfamiliar with pattern passwords before the test. The participants also indicated the device used for the test, which could be a desktop, phone, or tablet.

4.3 User Pattern

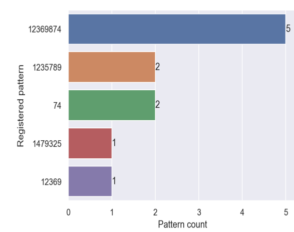
In the online user test, participants were encouraged (but not mandatory) to register a pattern that contained a minimum of four nodes and a maximum of nine nodes, as Hollow-Pass is a novel pattern password that some participants may find difficult to use at first. 59% of participants used a very weak pattern as their credential, 22% of participants used a weak pattern, and only 3% of participants used a strong pattern (as depicted in Figure 13a). Patterns with three or fewer connected nodes were considered invalid. The strength of a pattern was estimated using Equation 3. Upon examining the frequency of registered patterns (as shown in Figure 13b), the most commonly used valid pattern was 12369874, with a strength of weak (as depicted in Figure 13c), followed by 1235789, with a strength of very weak (as depicted in Figure 13d). The majority of participants used 5-6 nodes to create their patterns. To evaluate participants' recall error rate, the accuracy of their registered patterns (first-time drawing) and authenticated patterns (second-time drawing) was calculated. 3 out of 30 participants among all the age groups failed to redraw their patterns in the authentication phrase (as shown in Figure 13e).

Demographic information	Desktop	Phone	Tablet	Total
Gender				
Male	8	7	0	15
Female	7	7	1	15
Non-binary	0	0	0	0
Prefer not to say	0	0	0	0
Age				
20-29	10	10	1	21
30-49	1	1	0	2
50-59	4	2	0	6
60+	0	1	0	1
Prefer not to say	0	0	0	0
Eyesight quality				
Below 0.8	4	6	1	11
0.8-1.0	4	2	0	6
Over 1.0	3	2	0	5
unsure, wears glasses	2	4	0	6
unsure, does not wear glasses	2	0	0	2
Pattern experience				
used	7	9	1	17
known but not used	8	3	0	11
not known	0	2	0	2
Total	15	14	1	30

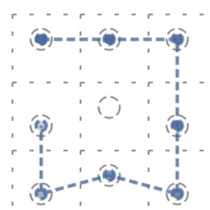
Table 3. Demographic Information of the Participants



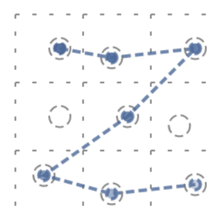
(a) percentage of registered patterns



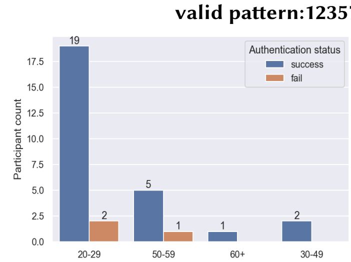
(b) frequency of registered pattern.



(c) the most frequently valid pattern:12369874



(d) the second most frequently valid pattern:1235789



(e) Authentication status of each age group

Fig. 13. User pattern in the online test.

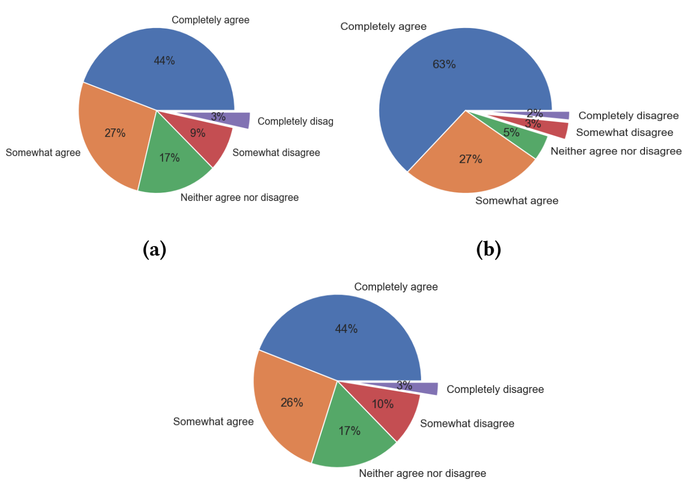
4.4 System Usability

In the online user test, we asked participants to rate the system usability ranging from Completely disagree(1) to Somewhat disagree(2), Neither agree nor disagree(3), Somewhat agree(4), and Completely agree(5). We measured usability in three aspects: mechanism feasibility, mechanism reliability, and user affinity.

Mechanism feasibility. More than 70% of participants provided positive feedback on the mechanism feasibility (44% completely agreed, 27% somewhat agree, see Fig.14a). Specifically, they can see the grid and nodes without difficulty, they can see my pattern while drawing without difficulty and they can draw the pattern without difficulty.

Mechanism reliability. 90% of participants provided positive feedback on the mechanism reliability (63% completely agreed, 27% somewhat agree, see Fig.14b). Participants felt more secure in Hollow-Pass compared with the existing pattern password, and they agreed that it would make shoulder surfers difficult to observe the credentials afar.

User affinity. 70% of participants provided positive feedback on the grid layout and background design (44% completely agreed, 26% somewhat agree, see Fig.14c). Participants felt easy to learn and use this new mechanism in the testing. The participants agreed that the automatically generated background design was more secure compared to a customizable background design. However, they expressed a desire to have the option of customizing their background image, if such a feature was available.



(a) (b)

(c)

Fig. 14. User pattern in online test:(a) Usability scale on mechanism feasibility; (b) Usability scale on mechanism reliability; (c) Usability scale on mechanism affinity.

4.5 System Security

Online test-Identify stimuli. Every participant identified two different patterns per viewing distance(30*2=60 patterns/viewing distance). The mean accuracy of identifying stimuli that simulate perceptual effects at different viewing distances (1.0m, 1.5m, 2.0m) was 57.78% in the online user test, see Fig.15. The correctness slightly increased as the viewing distance increased. In general, viewing distance did not have a significant effect on the accuracy of stimuli identification.

Onsite test-Observe pattern password in practice. In the onsite test, every participant observed six patterns per viewing distance and viewing angle, respectively. We recruited 19 participants, thus, we collected 114 patterns in total. The accuracy in the onsite test, as shown in Fig.16a, was significantly lower than that of the online test, as shown in Fig.15, with a mean accuracy of 19.59% compared to 57.78%. It can be seen from Fig.16a that the accuracy decreased as the viewing distance increased from 1.0m to 2.0m. In terms of viewing angle, as depicted in Fig.16b, participants had the highest accuracy of observed patterns when they were positioned directly in front of the screen (41.23%), while the least accurate patterns were observed when they were positioned at the right-front of the screen (24.56%). The mean accuracy of observing Hollow-Pass at different viewing angles was 31.87%.

The discrepancy between the results of the online and onsite tests may be due to a bias in the online test, where participants tended to guess the options (a) or (b), and most of the correct options were located in these options. Another factor contributing to the difference is that online participants were presented with multiple-choice questions, while onsite participants were required to draw the full patterns without any options. Additionally, the onsite participants reported difficulties in recognizing the grid orientation as the edge and the top-left node indicator were hard to discern from a distance, and it was also challenging to determine if a node was skipped, as the node positions changed dynamically each time.

Interestingly, it was discovered that the accuracy of the observed patterns decreased when participants asked the researchers to redraw the pattern. This result aligns with the findings of Adam et al. [37] as shown in Fig.17.

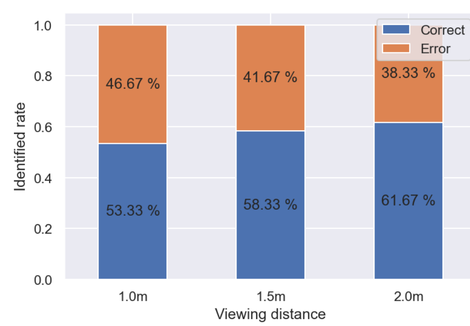
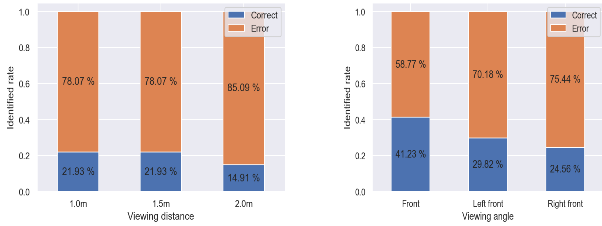


Fig. 15. Accuracy of identifying stimuli in online user test



(a) **(b)**

Fig. 16. Accuracy of observed pattern password in onsite user test: (a) Accuracy at different viewing distances; (b) Accuracy at different viewing angles.

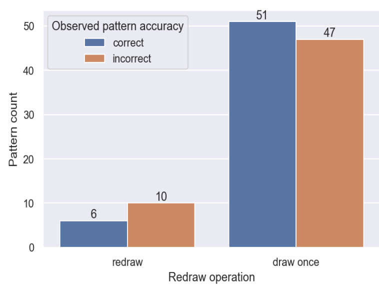


Fig. 17. Accuracy of observed patterns comparing drawing once and redrawing

4.6 Limitation

There are several limitations to Hollow-Pass mechanism:

(1).The default display pixel density of 164.54 pixels per inch and a viewing distance of 0.5m were used to create the background image, which may lead to a biased perceptual effect when users draw patterns on displays with different specifications or at a greater viewing distance.

(2).In the online user test, participants were not directly observed, and it was their first time using Hollow-Pass, which may result in inconsistent responses compared to real-life use. This could occur due to participants not following instructions properly, or opting for weaker patterns that were easy to recall and draw. To mitigate this limitation, we advised participants to use at least four nodes in drawing the pattern.

(3).The online user test stimuli were created based on average visual acuity and a 0.5m viewing distance, not taking into account human eye adjustment. Human eye adjustments refer to the changes in the shape of the lens inside the eye that allow it to focus on objects at different distances, which is known as accommodation. This adjustment is important for seeing objects clearly, whether they are near or far[26]. And participants may engage in random guessing when they were unable to recognize the stimuli in the form of options a or b, which could lead to biased results depending on the position of the correct option. The simulated effect and the design of the online test need to be re-evaluated to better align with human visual perception in reality.

(4). The scope of the user test is limited, and the onsite user test only involved undergraduate students in the age range of 20-29. Younger adults tend to be more receptive to new technology compared to middle-aged and older adults[38]. As a result, the impact of the background image and color contrast on other age groups should be further explored through a larger-scale study that encompasses all age groups.

5 CONCLUSION

In this study, we introduce a new pattern password mechanism, Hollow-Pass, that utilizes global precedence and color difference, ΔE, allowing users to draw patterns on a dynamic 3x3 grid as their authentication credentials. An online user test was conducted with $n = 30$ participants aged between 20-80, using desktop, phone, and tablet devices. To assess the practicality of the mechanism, an onsite user test was also conducted with $n = 19$ undergraduate students, viewed from three distances (1.0m, 1.5m, 2.0m) and three angles (front, left-front, right-front).

The results show that Hollow-Pass enhances the security of weak pattern passwords against shoulder surfing attacks while maintaining usability. The simulated shoulder surfer observed 19.59% of tested patterns on average in the onsite user test. Over 70% of online participants also gave positive feedback after using the mechanism. These findings suggest that Hollow-Pass effectively resists shoulder surfing attacks at close distances and viewing angles of front, left-front, and right-front, and balances security and usability.

Future work. In the future, we plan to conduct a comprehensive and targeted onsite usability study to assess any potential tradeoff between usability and security that may arise from the implementation of Hollow-Pass.

6 ACKNOWLEDGEMENT

I would like to express my gratitude to my supervisor, Dipti K. Sarmah, who guided me throughout this project, and to all participants who voluntarily joined the user test online or onsite. Last but not least, I would like to thank my family and friends who supported me all the time.

REFERENCES

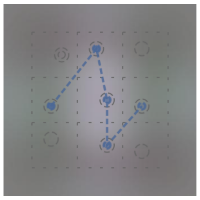
[1] Tahmina Islam Shammee, Taslima Akter, Muthmainna Mou, Farida Chowdhury, and Md Sadek Ferdous. A systematic literature review of graphical password schemes. *Journal of Computing Science and Engineering*, 14(4):163–185, 2020.
[2] Arti Bhanushali, Bhavika Mange, Harshika Vyas, Hetal Bhanushali, and Poonam Bhogle. Comparison of graphical password authentication techniques. *International Journal of Computer Applications*, 116(1), 2015.
[3] 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, page 161–172, New York, NY, USA, 2013. Association for Computing Machinery.
[4] Lei Zhang, Yajun Guo, Xiaowei Guo, and Xiaowei Shao. Does the layout of the android unlock pattern affect the security and usability of the password? *Journal of Information Security and Applications*, 62:103011, 2021.
[5] Hai Tao and Carlisle M. Adams. Pass-go: A proposal to improve the usability of graphical passwords. *Int. J. Netw. Secur.*, 7:273–292, 2008.
[6] Dirk Van Bruggen. *Studying the Impact of Security Awareness Efforts on User Behavior.* PhD thesis, University of Notre Dame, Apr 2014.

[7] Johannes Vorster and Renier van Heerden. Graphical passwords: A qualitative study of password patterns. 03 2015.
[8] Johannes Vorster. *A Framework for the Implementation of Graphical Passwords*. PhD thesis, 12 2014.
[9] Hung-Min Sun, Shiuan-Tung Chen, Jyh-Haw Yeh, and Chia-Yun Cheng. A shoulder surfing resistant graphical authentication system. *IEEE Transactions on Dependable and Secure Computing*, 15(2):180–193, 2018.
[10] Arash Habibi Lashkari, Samaneh Farmand, Dr Zakaria, Omar Bin, Dr Saleh, et al. Shoulder surfing attack in graphical password authentication. *arXiv preprint arXiv:0912.0951*, 2009.
[11] Hazleen Aris and Wira Firdaus Yaakob. Shoulder surf resistant screen locking for smartphones: A review of fifty non-biometric methods. In *2018 IEEE Conference on Application, Information and Network Security (AINS)*, pages 7–14, 2018.
[12] Flynn Wolf, Adam J. Aviv, and Ravi Kuber. "it's all about the start" classifying eyes-free mobile authentication techniques. *Journal of Information Security and Applications*, 41:28–40, 2018.
[13] Geumhwan Cho, Jun Ho Huh, Junsung Cho, Seongyeol Oh, Youngbae Song, and Hyoungshick Kim. Syspal: System-guided pattern locks for android. In *2017 IEEE Symposium on security and privacy (SP)*, pages 338–356. IEEE, 2017.
[14] Harshal Tupsamudre, Vijayanand Banahatti, Sachin Lodha, and Ketan Vyas. Pass-o: A proposal to improve the security of pattern unlock scheme. In *Proceedings of the 2017 ACM on Asia conference on computer and communications security*, pages 400–407, 2017.
[15] Harshal Tupsamudre, Sukanya Vaddepalli, Vijayanand Banahatti, and Sachin Lodha. Tinpal: An enhanced interface for pattern locks. In *Workshop on Usable Security, ser. USEC*, volume 18, 2018.
[16] Deyashini Chakravorty. What if we used graphical passwords for authentication? https://uxdesign.cc/graphical-passwords-for-authentication-4e716b94eb47, 2020. Accessed: 2022-11-21.
[17] Martin Mihajlov and Borka Jerman. Eye tracking graphical passwords. pages 37–44, 07 2018.
[18] Wenjuan Li, Jiao Tan, Weizhi Meng, and Yu Wang. A swipe-based unlocking mechanism with supervised learning on smartphones: Design and evaluation. *Journal of Network and Computer Applications*, 165:102687, 2020.
[19] Lim Seng, Associate Prof Dr Norafida Ithnin, and Hazinah Km. User's affinity of choice: Features of mobile device graphical password scheme's anti-shoulder surfing mechanism. *International Journal of Computer Science Issues*, 8, 07 2011.
[20] Anastasia V. Flevaris, Antigona Martínez, and Steven A. Hillyard. Attending to global versus local stimulus features modulates neural processing of low versus high spatial frequencies: An analysis with event-related brain potentials. *Frontiers in Psychology*, 5(APR), 2014.
[21] David Navon. Forest before trees: The precedence of global features in visual perception. *Cognitive psychology*, 9(3):353–383, 1977.
[22] Youngbae Song, Geumhwan Cho, Seongyeol Oh, Hyoungshick Kim, and Jun Ho Huh. On the effectiveness of pattern lock strength meters: Measuring the strength of real world pattern locks. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, CHI '15, page 2343–2352, New York, NY, USA, 2015. Association for Computing Machinery.
[23] Adam J. Aviv, Devon Budzitowski, and Ravi Kuber. Is bigger better? comparing user-generated passwords on 3x3 vs. 4x4 grid sizes for android's pattern unlock. In *Proceedings of the 31st Annual Computer Security Applications Conference*, ACSAC '15, page 301–310, New York, NY, USA, 2015. Association for Computing Machinery.
[24] Fergus W. Campbell and Lamberto Maffei. Contrast and spatial frequency. *Scientific American*, 231(5):106–115, 1974.
[25] Keita Takahashi, Madoka Hasegawa, Yuichi Tanaka, and Shigeo Kato. A structural similarity assessment for generating hybrid images. In *2011 Conference Record of the Forty Fifth Asilomar Conference on Signals, Systems and Computers (ASILOMAR)*, pages 240–243, 2011.
[26] Michael Kalloniatis and Charles Luu. Visual acuity. *Webvision: The Organization of the Retina and Visual System [Internet]*, 2007.
[27] Navon task. https://www.psytoolkit.org/experiment-library/navon.html#refs. Accessed: 2023-01-17.
[28] PT Service. Defining and communicating color: The cielab system, 2013.
[29] Zachary Schuessler. Delta e 101. http://zschuessler.github.io/DeltaE/learn/#toc-delta-e-76. Accessed: 2023-01-17.
[30] Welcome to flask. https://flask.palletsprojects.com/en/2.2.x/. Accessed: 2023-01-28.
[31] zhangstar. gesturepassword. https://github.com/zhangstar1331/gesturePassword, 2018.
[32] 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, 11 2014.
[33] Manoj Pandey. rgb2lab.py. https://gist.github.com/manojpandey/f5ece715132c572c80421febebaf66ae. Accessed: 2023-01-17.
[34] Jiayi Tan. Hollow-pass online survey. https://hollow-pass.up.railway.app/. Accessed: 2023-02-02.
[35] T.N. Pappas and D.L. Neuhoff. Least-squares model-based halftoning. *IEEE Transactions on Image Processing*, 8(8):1102–1116, 1999.
[36] Edwin G. Boring. Size constancy and emmert's law. *The American Journal of Psychology*, 53(2):293–295, 1940.
[37] Adam J. Aviv, John T. Davin, Flynn Wolf, and Ravi Kuber. Towards baselines for shoulder surfing on mobile authentication. *CoRR*, abs/1709.04959, 2017.
[38] Monica Anderson and Andrew Perrin. Barriers to adoption and attitudes towards technology. *Tech adoption climbs among older adults*, 2017.

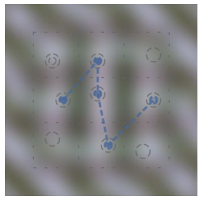
A APPENDIX A : OVERVIEW OF GRAPHICAL PASSWORD CATEGORY

Category	Procedure	Advantage	Disadvantage
Recognition	Users identify and distinguish their pass images from other decoy images.	• Easy to remember	• Time-consuming • Susceptible to shoulder-surfing attacks
Recall	Users create a secret drawing in registration and reproduce it in a grid or blank canvas in the authentication.	• No language restriction • Easy to use	• Recall error • Susceptible to dictionary attack, shoulder-surfing attacks
Cued-recall	Users create passwords by choosing a random set of points in a specific region in a user-chosen image during registration, and identify those points in the correct order during authentication.	• Increase entropy by using random images • Enlarge password space by choosing as many points as possible	• Time-consuming • Difficult to memorize • Susceptible to shoulder-surfing attacks
Hybrid	Combination of two or more different types of GPs or other authentication methods.	Scheme dependent, i.e., • increase password space • better memorability • user-friendly, etc.	Scheme dependent, i.e., • Time-consuming • Difficult to memorize • extra technique required • Susceptible to shoulder-surfing attacks, etc.

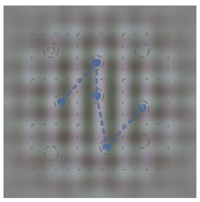
B APPENDIX B : COMPARISON OF DIFFERENT SPATIAL FREQUENCY GRATINGS



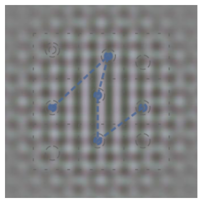
(a) sf=[0.3,0.3,0.3,0.3], pattern is easily observed.



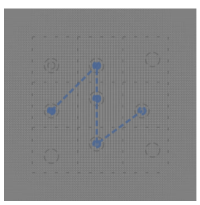
(b) sf=[1,1,1,1], pattern is somehow disguised.



(c) sf=[1,2,1,2], pattern is somehow disguised.



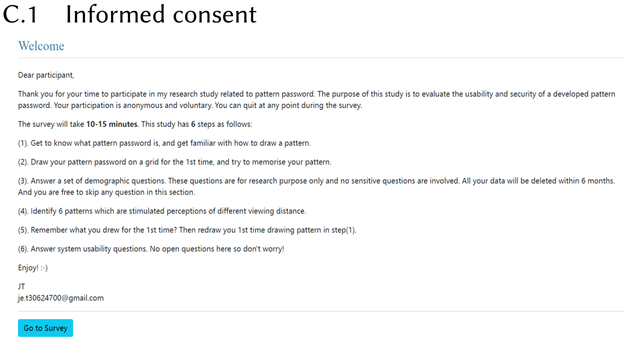
(d) sf=[3,3,3,3], the background starts to have packed of holes, which may cause revulsion.



(e) sf=[30,30,30,30], pattern is easily observed.

C APPENDIX C : ONLINE TEST

C.1 Informed consent



Welcome

Dear participant,

Thank you for your time to participate in my research study related to pattern password. The purpose of this study is to evaluate the usability and security of a developed pattern password. Your participation is anonymous and voluntary. You can quit at any point during the survey.

The survey will take **10-15 minutes**. This study has **6** steps as follows:

(1). Get to know what pattern password is, and get familiar with how to draw a pattern.

(2). Draw your pattern password on a grid for the 1st time, and try to memorise your pattern.

(3). Answer a set of demographic questions. These questions are for research purpose only and no sensitive questions are involved. All your data will be deleted within 6 months. And you are free to skip any question in this section.

(4). Identify 6 patterns which are stimulated perceptions of different viewing distance.

(5). Remember what you drew for the 1st time? Then redraw you 1st time drawing pattern in step(1).

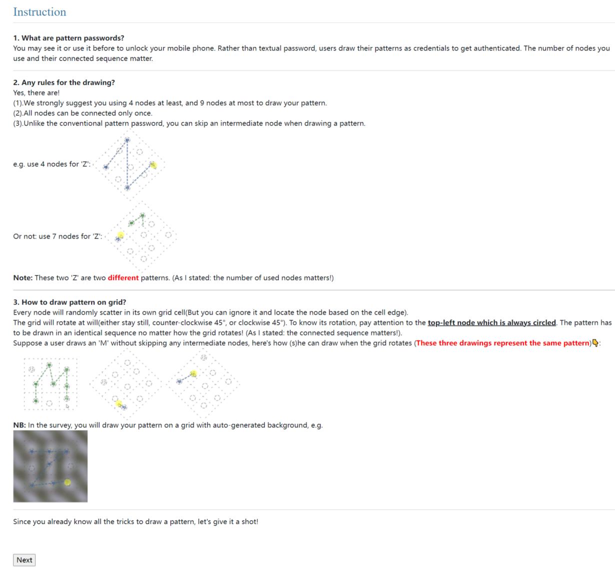
(6). Answer system usability questions. No open questions here so don't worry!

Enjoy! :-)

JT
je.t30624700@gmail.com

Go to Survey

C.2 Hollow-Pass Instructions



Instruction

1. What are pattern passwords?
You may see it or use it before to unlock your mobile phone. Rather than textual password, users draw their patterns as credentials to get authenticated. The number of nodes you use and their connected sequence matter.

2. Any rules for the drawing?
Yes, there are!
(1).We strongly suggest you using 4 nodes at least, and 9 nodes at most to draw your pattern.
(2).All nodes can be connected only once.
(3).Unlike the conventional pattern password, you can skip an intermediate node when drawing a pattern.

e.g. use 4 nodes for 'Z':

Or not: use 7 nodes for 'Z':

Note: These two 'Z' are two **different** patterns. (As I stated: the number of used nodes matters!)

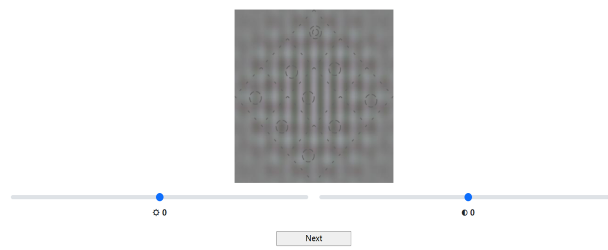
3. How to draw pattern on grid?
Every node will randomly scatter in its own grid cell(But you can ignore it and locate the node based on the cell edge).
The grid will rotate at will(either stay still, counter-clockwise 45°, or clockwise 45°). To know its rotation, pay attention to the **top-left node which is always circled**. The pattern has to be drawn in an identical sequence no matter how the grid rotates! (As I stated: the connected sequence matters!).
Suppose a user draws an 'M' without skipping any intermediate nodes, here's how (s)he can draw when the grid rotates (**These three drawings represent the same pattern**)👇:

NB: In the survey, you will draw your pattern on a grid with auto-generated background, e.g.

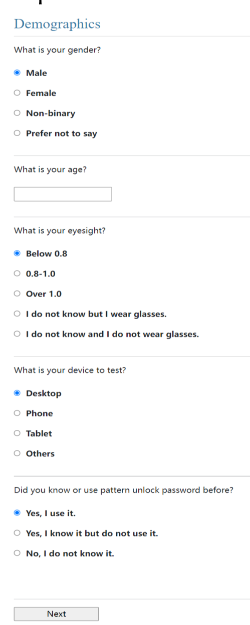
Since you already know all the tricks to draw a pattern, let's give it a shot!

Next

C.3 1st time drawing and 2nd time drawing



C.4 Demographic questions



Demographics

What is your gender?

- Male
- Female
- Non-binary
- Prefer not to say

What is your age?

What is your eyesight?

- Below 0.8
- 0.8-1.0
- Over 1.0
- I do not know but I wear glasses.
- I do not know and I do not wear glasses.

What is your device to test?

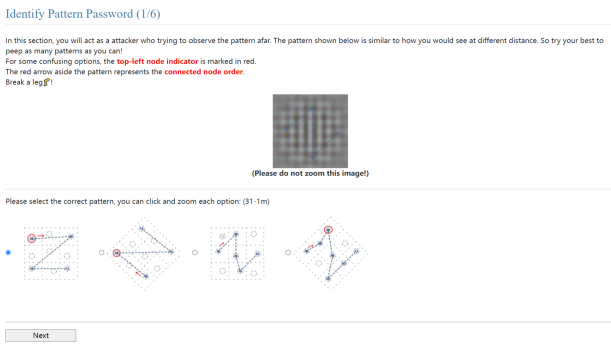
- Desktop
- Phone
- Tablet
- Others

Did you know or use pattern unlock password before?

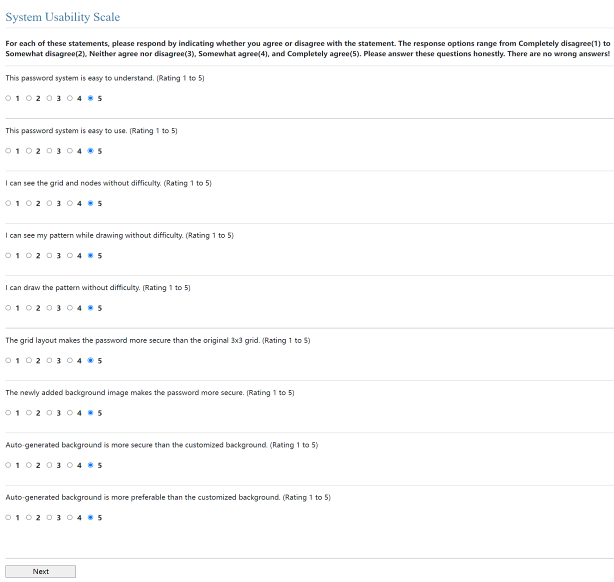
- Yes, I use it.
- Yes, I know it but do not use it.
- No, I do not know it.

Next

C.5 Identify pattern password



C.6 System usability scale



D APPENDIX D : PERCEIVED RESOLUTION SIMULATION

We used the linear relationship between the spacing of the grating period in the image, represented by τ, and the viewing distance in centimeters, represented by $viewD$, as outlined in Equation 5 from the study by Pappas and Neuhof [35]. This relationship is based on the fact that the impulse response of the 1-D eye filter is well-approximated by a Gaussian shape with an appropriate standard deviation. To simulate the perceived resolution at different viewing distances, we scaled the standard deviation of the Gaussian filter using the inverse relationship between the grating frequency, represented by f_s, and the grating period, represented by τ, as outlined in Equation 6.

$$\begin{aligned} \sigma &= coef * \tau \\ coef &= \frac{2.54}{20}\, viewD \end{aligned} \tag{5}$$

where $coef$ is the coefficient that scales standard deviation, $viewD$ is viewing distance in cm.

$$f_s = (2 * arctan(\frac{gp}{2 * sr * viewD}))^{-1} \tag{6}$$

where f_s is the spatial frequency of grating stimuli in cycles per degree (cpd), gp is the grating period in image in pixels, sr is the screen resolution in px/cm and $viewD$ is the viewing distance in cm. In this research, we use screen resolution(sr) 64.82 px/cm in default.