Usability and Security of Gaze-Based Graphical Grid Passwords

Majid Arianezhad[1], Douglas Stebila[2], and Behzad Mozaffari[2]

[1] School of Engineering Science, Simon Fraser University, Burnaby, B.C., Canada; arianezhad@sfu.ca

[2] Science and Engineering Faculty, Queensland University of Technology, Brisbane, Australia; stebila@qut.edu.au, behzad.mozaffari@connect.qut.edu.au

Abstract. We present and analyze several gaze-based graphical password schemes based on recall and cued-recall of grid points; eye-trackers are used to record user's gazes, which can prevent shoulder-surfing and may be suitable for users with disabilities. Our 22-subject study observes that success rate and entry time for the grid-based schemes we consider are comparable to other gaze-based graphical password schemes. We propose the first password security metrics suitable for analysis of graphical grid passwords and provide an in-depth security analysis of user-generated passwords from our study, observing that, on several metrics, user-generated graphical grid passwords are substantially weaker than uniformly random passwords, despite our attempts at designing schemes to improve quality of user-generated passwords.

Keywords: graphical passwords; eye-tracking; usable security

1 Introduction

Graphical password schemes have the potential to improve user authentication due to easier memorability and use. Typically, a user indicates various regions of the screen or draws some pattern using mouse, touch, or gaze input methods. Gaze-based password input is promising due to its resistance to shoulder surfing and because it may be easier for people with disabilities to use. The usability of graphical password has been extensively studied, but there has been very little investigation into the quality of user-generated graphical passwords. We investigate several variants of graphical grid passwords to determine if variations can improve the quality of user-generated passwords—in terms of point and stroke distribution and symmetry—while maintaining usability

An extensive survey of the vast literature on graphical passwords was recently given by Biddle, Chiasson, and van Oorschot [4]. They describe three main categories of schemes: *recall-based schemes*, such as Draw-A-Secret [14], where the user must recall and enter a secret drawing or pattern from memory; *recognition-based schemes*, where a user must recognize a few personal objects from a set of objects, either images (Passfaces [20], Faces [7]) or text [28]; and *cued-recall*

schemes, such as PassPoints [24,25] or Cued Gaze-Points [12], where the user is given an image cue and must recall and enter certain points or a pattern.

Recall-based schemes can be divided into two main subcategories. In *free-form drawmetric schemes*, such as Draw-A-Secret [14] or Pass-Go [23], the user draws an arbitrary image on a blank canvas. *Grid schemes* restrict the valid target points to a grid; some, such as PassShapes [26] and the gaze-based EyePassShapes [8], restrict moves to adjacent points in the grid or use limited patterns [10], whereas others, such as GridSure [5] and the popular 'pattern lock' 3×3 grid screens for Android and other [22] mobile phones allow users to enter arbitrary patterns of grid points. A few schemes [15] have users enter text-based passwords using *on-screen keyboards*.

Shoulder-surfing, where an attacker watches a user enter their password, is a well-known problem for graphical password schemes [11,29]. Grid schemes on mobile phones can be vulnerable to smudge attacks [3], though shoulder-surfing and smudge attacks can be mitigated using biometric characteristics from entering the password [9]. Magnetic entry schemes also resist shoulder-surfing attack [21]. *Gaze-based passwords* may be more resistant to such attacks, since no visual feedback of the user's entry is displayed on-screen, and may also be suitable as an input method for users with disabilities. (Gaze-based entry is not a security panacea, however: video cameras or attackers surreptitiously watching a user's eye movement may still be able to gain enough information to attack passwords with some success [8].)

A well-known weakness of traditional text-based passwords is that human-generated passwords are not truly random. While an eight-character mixed-case alphanumeric password may be chosen from a large theoretical password space ($(26 + 26 + 10)^8 = 62^8 \approx 2^{47.6}$), humans pick passwords from a non-uniform distribution with much lower entropy.

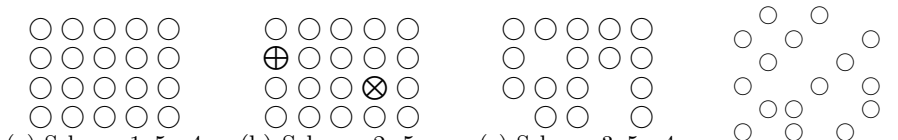
Unfortunately, most papers on graphical password schemes only mention the theoretical password space with no analysis of user-generated passwords, though some research has been done on the password security of some schemes. A line of research by van Oorschot and Thorpe has analysed the space of human-generated passwords in free-form drawmetric schemes [18] as well as the prevalence of image hot spots in cued-recall graphical passwords [17,19], though hot spots can be reduced using masking [6]. User-generated passwords in recognition-based schemes can also have poor entropy and be susceptible to educated guess attacks based on demographic information [7] or personal knowledge [13].

We focus on recall-based graphical grid schemes using eye-tracking for data entry. From the usability perspective, we aim to determine if gaze-based entry of graphical grid passwords, which have no recall cues, can achieve comparable success rates and entry times to cued-recall schemes. On the security side, we aim to provide metrics for the security of human-generated grid passwords, as previous security analyses do not directly carry over to grid schemes. We hypothesize that human-generated passwords will have more symmetry and not use uniformly distributed points and strokes, so we test several variants to see if they improve password quality.

2 Schemes

In a pre-trial phase, we had a handful of users try out a basic grid scheme, and noticed that the passwords they created tended to being symmetric and have poor distribution of the first and last points; in particular, a significant proportion of users chose the top-left point as their first point. Pre-study results were similar to the results for Scheme 1; see Appendix B for distribution of points during the main study. This motivated us to design several variants to see if we could improve the quality of user-generated passwords.

We propose four gaze-based graphical grid password schemes as shown in Figure 1. Scheme 1 is a basic 5×4 grid, a generalization of the 'pattern lock' screen popular on Android devices. In Scheme 2, we cued the user to start and end at the specified points, visually displayed with different colours; by picking the first point for the user, we hypothesize that the second point (i.e., the first user-selected point) might have better distribution; this also eliminates perfectly symmetric shapes that use the same start and end point. In Scheme 3, we removed a few random grid points: users may be less likely to pick symmetric shapes since not all the points were available to them. We also wanted to know if a bigger, sparser, less grid-like scheme induced more random passwords: Scheme 4 was a much sparser subset of a larger, 6x6 grid. Note that while we designed schemes 2 through 4 by selecting/removing points at random, we did this randomization once: all users used the exact same fixed grids in Figure 1.



(a) Scheme 1: 5×4 grid

(b) Scheme 2: 5×4 grid with cued start $\oplus$ and end $\otimes$ points

(c) Scheme 3: 5×4 grid with holes

(d) Scheme 4: 6×6 sparse grid

Fig. 1. Gaze-based graphical grid password schemes in our study

To enter passwords, users gaze at the first point in the password, press the space bar to tell the system to begin recording, gaze at each subsequent point for at least 0.5 seconds, then press the space bar again to stop recording. Note in particular that users do not have to press the space bar at each point, just gaze at it for at least 0.5 seconds. No visual feedback is displayed to the user while entering their password — no indicated of points gazed or even when a gaze is registered; the only visual feedback comes after they press the space bar to stop recording, which results in a dialog box indicating successful or failed entry. Subsequent points have no restriction for adjacency; the same point cannot be gazed at twice in a row, though can be later used again.

The entry grid was displayed on a 19" monitor running at a resolution of 1920×1080 pixels. Gaze points were displayed as circles of radius 65 pixels with a 11×11 pixel 'cross' (+) displayed in the centre of the circle to help users focus on a target. The user did not need to gaze directly at the circle: we took the closest circle to their gaze fixation.

3 Experiment Design

We conducted a within-subjects lab study. Participants were approached through personal contacts and received no compensation; the study was approved by the university's ethics board.

A standard Windows 7 desktop PC with a 19" monitor was equipped with a Mirametrix S2 Eye Tracker, placed just below the monitor. The device has a data rate of 60 Hz with infrared binocular tracking. The accuracy range is 0.5° to 1° and the drift range is less than 0.3°. Our gaze-based password scheme was a custom-written C# program.

Each participant was assigned to use three of the four schemes: all participants used Schemes 1 and 2, and were randomly assigned to either Scheme 3 or 4. First, participants were introduced to the system and ran the eye-tracker's 9-point calibration routine. We told users to gaze at points for at least 1 second, even though the system would register a gaze after just 0.5 seconds. For each of the three schemes assigned (1, then 2, then either 3 or 4) participants were directed to (a) *create* a new password "of at least 6 points that would be easy for [them] to remember but hard for others to guess"; (b) *confirm* the password; (c) answer three short survey questions[3]; and (d) *login* using the password. After doing this for the three assigned schemes, the participant did (e) a *final login* using the password from Scheme 1. During confirmation and login sub-tasks, participants could keep trying until successful, skip the task, or restart the task (recreate and reconfirm a new password).

The login (d) after sub-task (c) and the final login at the end were designed to test recall after a passage of time. The login (d) after distraction task (c) typically occurred approximately 1 minute after completing steps (a)–(b). This is similar to the 30-second distraction task of Forget et al. [12].

The final login (e) in Scheme 1 typically occurred approximately 10 minutes after completing steps (a)–(d) for Scheme 1. In fact, in our study we also emailed participants two days after their participation, asking them to reply with the scheme 1 password, but not enough participants responded for us to report results.

It should be noted that, by having all subjects proceed sequentially through the tasks, a potential learning effect is introduced in which users find the later schemes easier to use: thus *usability* results may not be fully comparable between schemes. However, studies have found that *security* behaviour can change if the user has been "primed" for security (for example, Whalen and Inkpen [27] found

[3] Survey questions in Appendix A. User password dataset and Java code for metrics available at http://eprints.qut.edu.au/58524/

that no users looked for web browser security indicators before being asked to do so). In our context, this means that a user who sees scheme 2 before scheme 1 may choose different start/end points in scheme 1 than had she seen scheme 1 before scheme 2. To compare password quality consistently across schemes and to avoid priming subjects to choose more random or asymmetric passwords in scheme 1, we used a fixed sequence of tasks. This tradeoff between learning effects in usability or in password security seems to be inherent to any study where subjects use multiple variants.

Participants were randomly assigned to either scheme 3 or scheme 4 before they arrived; time constraints prevented us from having participants use both schemes.

Some of our survey questions, regarding security and computer expertise, are a subset of the survey questions of Arianezhad et al. [2].

4 Results

We had 25 participants total, though the eye-tracking equipment only recorded results for 22 of them due to astigmatisms. Participants ranged in age from 19–41 with an average age of 26.1. Most had a high degree of computer expertise; only 1 reported using Android pattern lock.

To help the reader understand what types of passwords are entered by users, we include in Appendix C the points for the passwords entered by our users in Scheme 1.

4.1 Security

Since passwords in our scheme are user-generated, not randomly generated, it is not appropriate to assume that all possible passwords are equally likely. Table 1 reports several measures of password randomness; cells in sections (b)–(d) of the table are of the form a/b, where a is the value of the metric for passwords our users created and b is the value for passwords generated uniformly at random, computed either algebraically (for (b)) or on a sample of 100000 passwords of length 7 generated uniformly at random (for (c) and (d)).

Password length. Users were directed to create a new password "of at least 6 points that would be easy for [them] to remember but hard for others to guess". As reported in Table 1, the average length of passwords in all schemes around $7\frac{1}{3}$ characters. Note that in Scheme 2, users seemed to interpret this instruction for length of at least 6 as including the cued start and end points, hence in the table we report only the number of user-selected—and hence secret—points.

Point frequency. For all four schemes, the frequency of points selected by users is quite close to random: for example, in Scheme 1, the entropy of user-selected points is 4.11 bits, compared to the maximum 4.32 bits for random points.

	Scheme 1 Grid	**Scheme 2** Grid with cued start/end	**Scheme 3** Grid with holes	**Scheme 4** Sparse grid
(a) User-generated password length				
Mean* (SD)	7.59 (2.42)	5.36 (2.01)	7.33 (2.69)	7.23 (1.64)
(b) Binary entropy of points				
All	$^{4.11}/_{4.32}$	$^{3.87}/_{4.17}$	$^{3.75}/_{4.00}$	$^{3.95}/_{4.00}$
First†	$^{2.18}/_{4.32}$	$^{2.54}/_{4.25}$	$^{2.50}/_{4.00}$	$^{2.78}/_{4.00}$
Last†	$^{3.54}/_{4.32}$	$^{2.63}/_{4.25}$	$^{2.50}/_{4.00}$	$^{2.14}/_{4.00}$
(c) Binary entropy of stroke direction & length‡				
	$^{3.47}/_{5.65}$	$^{3.05}/_{5.54}$	$^{3.20}/_{5.64}$	$^{3.73}/_{6.33}$
(d) Symmetry score‡ (higher = more symmetry)				
Vertical	$^{0.71}/_{0.58}$	$^{0.70}/_{0.55}$	$^{0.66}/_{0.57}$	$^{0.48}/_{0.47}$
Horizontal	$^{0.66}/_{0.57}$	$^{0.68}/_{0.59}$	$^{0.63}/_{0.56}$	$^{0.43}/_{0.46}$
(e) Search estimate for 7-point passwords				
Theoretical	$2^{30.2}$	$2^{29.4}$	$2^{28.0}$	$2^{28.0}$
Point entropy	$2^{28.8}$	$2^{27.1}$	$2^{26.3}$	$2^{27.7}$
First+strokes	$2^{23.0}$	$2^{20.8}$	$2^{21.7}$	$2^{25.2}$

Table 1. Security metrics for user-generated passwords versus uniformly random passwords

* For Scheme 2: excluding cued start/end points.
† First & last *user-selected* points. Thus, for Scheme 2: second & second-last.
‡ Values for uniformly random passwords calculated from 100000 uniformly randomly generated samples of length 7.

However, first and last user-selected points are not very random. In Scheme 1, the entropy of user-selected first points was just 2.18 out of 4.32 bits; in fact 50% of user-generated passwords started in the top-left corner. Scheme 2 was no better: the second and second-last points (i.e., the first and last *user-selected* points) were clustered around the cued points and had low entropy (2.54 and 2.63 out of 4.25 bits). Frequency tables for all, first, and last points of all schemes are given in Appendix B.

Strokes. We next consider the distribution of "strokes", meaning the direction and length between subsequent points. For example, a password where the first point was $(1, 1)$ and the second point was $(2, 3)$ corresponds to the stroke $(1 \downarrow, 2 \rightarrow)$. We observed that the entropy of strokes in user-generated passwords is quite poor, in all cases between 55% and 62% of the entropy of strokes in randomly generated passwords. Frequency tables for stroke distribution for all schemes are given in Appendix B.

Symmetry. We observed that many users entered passwords that looked to be quite symmetric. For example, consider the fifth password (second row, second column) in Appendix C that was entered by participant #5 entered in Scheme 1.

We devised a metric to measure the symmetry present in a graphical grid password based in part on symmetry analyses of free-form drawmetric schemes [16,18]. The vertical (respectively, horizontal) symmetry score is computed as follows: for each possible vertical (horizontal) axis (axes exist either in between or along columns (rows) of points), fold along the axis, count the number of password points that match on both sides of the fold, and divide by the total number of password points; the vertical (horizontal) symmetry score is the maximum over all possible axes. (Note that both previous works on symmetry analyses of drawmetric schemes include at least some off-centre axes [16] or maximize over all possible axes [18].)

For example, for the firth password in Appendix C, the vertical symmetry score is 1.0, since by folding along the optimal vertical axis (through the third column), we have perfect overlap, whereas the horizontal symmetry score is $7/8 = 0.875$, since by folding along the optimal horizontal axis (through the second row), we have 7 of 8 points overlapping.

Note that although schemes 2–4 are somewhat asymmetric by design, the symmetry score does not become obsolete. Rather, the question becomes: are user-generated passwords more or less symmetric that randomly generated passwords in the same scheme?

As seen in Table 1(d) Schemes 1, 2, and 3 had higher vertical and horizontal symmetry scores than randomly generated passwords, suggesting that Schemes 2 and 3 did not introduce much "asymmetry". However, user-generated passwords in Scheme 4 were as asymmetric as random passwords.

Password space estimate. We used the above password metrics to estimate an upper-bound on the amount of work to search for a 7-point password using three different strategies: the *theoretical* search space (computed as $7 \cdot$ (ideal entropy of all points)); based on the *point entropy* of user-generated passwords ($7 \cdot$ (user entropy of all points)); and *first+strokes*, based on the entropy of the first point and subsequent strokes of user-generated passwords ((user entropy of 1^{st} point) $+ 6 \cdot$ (user entropy of strokes)). For Schemes 1–3, these techniques show decreases in search space of at least 7 bits compared to the theoretical space.

Limitations. The symmetry measures we employ do not address rotational symmetry or reflection on non-vertical/horizontal axes. While such symmetries are natural [16] for free-form drawmetric schemes, it is not clear how to correctly define them for grid schemes.

4.2 Usability

Table 2 (following reporting techniques of Forget et al. [12]) reports a wide variety of metrics on the usability of our 4 schemes and compares with 3 other gaze-based schemes. We use non-parametric tests due to small sample size.

	Scheme 1 Grid	**Scheme 2** Grid with cued start/end	**Scheme 3** Grid with holes	**Scheme 4** Sparse grid	**CGP T-51** Cued-recall [12]	**EyePassShapes** Grid with adjacent movements [8]	**EyePassword** On-screen keyboard (Dwell-QWERTY) [15]
# of participants	22	22	9	13	25	24	18
# of trials	22	22	9	13	169	–	–
Mean creates per participant	1.09	1.55	1.89	2.77	–	–	
Successful confirms on 1st try	91%	64%	67%	38%	67%	–	–
Successful confirms on $\leq$ 3 tries	95%	91%	78%	69%	82%	–	–
Successful logins on 1st try	73%	91%	89%	54%	73%	86%	97%
Successful logins on $\leq$ 3 tries	91%	95%	100%	77%	93%	–	–
Successful final logins on 1st try	45%*	–	–	–	–	57%*	–
Successful final logins on $\leq$ 3 tries	55%*	–	–	–	–	–	–
Mean confirm errors per trial	0.18	0.68	0.89	1.54	1.21	–	–
Mean login errors per trial	0.55	0.45	0.22	0.64	0.51	–	–
Mean (SD) total create time	14.6 (6.4)	19.2 (12.4)	24.2 (18.5)	38.9 (41.1)	44.2 (22.0)†	–	–
Mean (SD) total confirm time	17.9 (22.5)	21.6 (16.5)	26.3 (16.5)	36.4 (42.3)	47.1 (78.5)†	–	–
Mean (SD) total login time	21.4 (21.9)	18.0 (20.2)	19.5 (24.0)	17.4 (12.2)	36.7 (35.9)†	–	–
Mean (SD) succ. create time/point	1.78 (0.50)	1.76 (0.59)	1.94 (0.87)	1.75 (0.58)	–	–	–
Mean (SD) succ. confirm time/point	1.70 (0.51)	1.80 (0.62)	2.04 (0.98)	1.52 (0.65)	–	–	–
Mean (SD) succ. login time/point	1.68 (0.66)	1.60 (0.65)	1.98 (1.29)	1.26 (0.93)	–	1.56	1.08
Ease of use (4-point Likert scale: very easy-easy-hard-very hard)	5-14-3-0	3-17-2-0	1-4-3-1	3-4-5-1	See [12] Fig. 2	mid-scale (2.67/5)	–

Table 2. Usability metrics of our schemes and other gaze-based schemes

* Scheme 1 final login performed 10 minutes after initial use; EyePassShapes final logins performed 5 days after initial study.

† CGP T-51 times included username entry and calibration time; other results do not.

Successes and errors. As described in Table 2, the mean number of password creation operations per participant in Scheme 1 was (mostly) significantly smaller than Schemes 2 (Wilcoxon signed-rank $V = 3$, $p = 0.037$), 3 ($V = 0$, $p = 0.098$), and 4 ($V = 0$, $p = 0.034$). Scheme 1 also required fewer tries for confirmation, but the difference was significant only versus Scheme 4 ($V = 0$, $p = 0.021$).

For number of tries for successful login after the short distraction task (3 survey questions, ∼45 seconds), Schemes 1, 2, and 3 all performed well, and better than Scheme 4, though the difference was not statistically significant. However, the success rate for final logins to Scheme 1, which participants did at the end of the study after doing the Scheme 2 and Scheme 3 or 4 tasks (∼10 minutes later), was quite poor (≤ 3 tries: 55%). This suggests users may forget grid passwords quickly, may become confounded when working with several grid passwords, or did not have enough repetition to ensure memorability. Our recall rate is not far off that of EyePassShapes [8], though theirs was after a much longer period (5 days vs. ∼10 minutes). A Spearman rank correlation test observed no statistically significant correlation between password length and number of confirmation or login errors.

Times. Table 2 reports for creation, confirmation, and login. Note that we report two different types of times:

- *total* time required for creation, confirmation, or login, which includes time elapsed during errors and re-tries, but does not include eye-tracker calibration
- time per point for *successful* creation, confirmation, or login, which includes only the time elapsed during the entry that actually succeeded, and is averaged on a per point basis.

We report both times to allow meaningful comparison with other schemes, some of which (Cued Gaze Points (CGP) T-51) reported total time and some of which (EyePassShapes, EyePassword) reported successful time. Note CGP T-51 [12] times also include time for a 1-point calibration and keyboard-based username entry; all other times do not include calibration or username entry. At the start of the study, we used our device manufacturer's 9-point calibration, which requires ∼20 seconds.

Times required for Schemes 1, 2, and 3 were fairly similar, whereas Scheme 4 had higher creation and confirmation times. Due to high standard deviation, only a few of the differences in means were statistically significant: creation time, Scheme 1 vs. 4 (Wilcoxon signed-rank $V = 16$, $p = 0.043$); confirmation time, Scheme 1 vs. 2 ($V = 63$, $p = 0.041$) and 1 vs. 4 ($V = 6$, $p = 0.006$).

Since we allowed users to choose the length of their password, we separately report times for just the successful creation/confirmation/login operations, averaged over the number of points in the password. Mean time per point is relatively consistent across all schemes and tasks: on average, users require 1.77 seconds per spot. Hence, an experienced user who makes no errors should be able to login with a 7-point password in around 12 seconds or less.

Our times are generally comparable with other schemes. In particular, per-point time during successful logins (ranging from 1.26–1.98 seconds per point) is

on par with that for EyePassShapes (1.56 seconds), although a bit higher than EyePassword (1.08 seconds per point).

User perception. For each scheme, participants rated "how difficult it was to complete the task" on a 4-point Likert scale (very easy, easy, hard, very hard). Nearly all participants rated Schemes 1 and 2 easy or very easy (slightly lower than Cued Gaze-Points [12]; higher than EyePassShapes [8]), but only about half did for Schemes 3 and 4.

5 Conclusion

We have studied the usability and security of various recall-based graphical grid password schemes when used with gaze-based user interfaces. Though it can be difficult to precisely compare usability results across studies, in general our success rates and entry times are comparable with existing gaze-based cued-recall schemes. We give the first thorough treatment of the quality of passwords generated by users in graphical grid password schemes.

Assessing the strength of user-generated passwords on a variety of metrics is essential. User-generated graphical passwords may perform well on some metrics but poorly on others. Thus, for user-generated passwords, a simple password-space calculation in which all potential passwords are considered equally likely is overly optimistic. We have proposed the first metrics for assessing randomness of grid password schemes, which can be applied to all grid schemes, including for example Android pattern lock. In all of our schemes, the distribution of the first and last user- points was quite poor. The distribution of strokes between subsequent points in a password was also quite poor. Our attempt in Scheme 4 at increasing asymmetry in user-generated passwords worked, but at the cost of significantly longer creation and confirmation time and significantly lower confirmation and login success rates. Of the four schemes we proposed, the basic grid scheme, Scheme 1, seems to provide the best ease-of-use (high success rates, small time), with password distribution quality comparable to the other schemes.

Larger-scale real-world studies testing gaze-based graphical grid password scheme would provide insight into several open questions, such as the usability of gaze-based authentication in a non-laboratory setting, generalization to other user populations, suitability for users with disabilities, long-term recall rates, whether use of multiple grid passwords has a confounding effect, and the relative security of human-generated grid passwords in settings with more realistic risks.

Acknowledgements

The authors gratefully acknowledge helpful discussions with Margot Brereton and Daniel Johnson.

References

1. Proc. 30th International Conference on Human Factors in Computing Systems (CHI) 2012. ACM (2012)
2. Arianezhad, M., Camp, L.J., Kelley, T., Stebila, D.: Comparative eye tracking of experts and novices in web single sign-on. In: Proc. 3rd ACM Conference on Data and Application Security and Privacy (CODASPY) 2013 (2013), to appear. Available at http://www.douglas.stebila.ca/research/papers/acks13/
3. Aviv, A.J., Gibson, K., Mossop, E., Blaze, M., Smith, J.M.: Smudge attacks on smartphone touch screens. In: USENIX WOOT 2010 (2010), https://www.usenix.org/conference/woot10/smudge-attacks-smartphone-touch-screens
4. Biddle, R., Chiasson, S., van Oorschot, P.C.: Graphical passwords: Learning from the first twelve years. ACM Computing Surveys 44(4), 19:1–19:41 (September 2012)
5. Bond, M.: Comments on Gridsure authentication. Online (March 2008), http://www.cl.cam.ac.uk/~mkb23/research/GridsureComments.pdf
6. Bulling, A., Alt, F., Schmidt, A.: Increasing the security of gaze-based cued-recall graphical passwords using saliency masks. In: Proc. 30th International Conference on Human Factors in Computing Systems (CHI) 2012 [1], pp. 3011–3020
7. Davis, D., Monrose, F., Reiter, M.K.: On user choice in graphical password schemes. In: Proc. 13th USENIX Security Symposium. pp. 151–164 (2004), http://static.usenix.org/event/sec04/tech/davis.html
8. De Luca, A., Denzel, M., Hussmann, H.: Look into my eyes!: can you guess my password? In: Proc. 5th Symposium on Usable Privacy and Security (SOUPS) 2009. pp. 7:1–7:12. ACM (2009)
9. De Luca, A., Hang, A., Brudy, F., Lindner, C., Hussmann, H.: Touch me once and I know it's you!: implicit authentication based on touch screen patterns. In: Proc. 30th International Conference on Human Factors in Computing Systems (CHI) 2012 [1], pp. 987–996
10. De Luca, A., Weiss, R., Drewes, H.: Evaluation of eye-gaze interaction methods for security enhanced PIN-entry. In: Proc. 19th Australasian Conf. on Computer-Human Interaction (OZCHI) 2007. pp. 199–202. ACM (2007)
11. Dunphy, P., Heiner, A.P., Asokan, N.: A closer look at recognition-based graphical passwords on mobile devices. In: Proc. 6th Symposium on Usable Privacy and Security (SOUPS) 2010. pp. 3:1–3:12. ACM (2010)
12. Forget, A., Chiasson, S., Biddle, R.: Shoulder-surfing resistance with eye-gaze entry in cued-recall graphical passwords. In: Proc. 28th International Conference on Human Factors in Computing Systems (CHI 2010). pp. 1107–1110. ACM (2010)
13. Hayashi, E., Hong, J., Christin, N.: Security through a different kind of obscurity: evaluating distortion in graphical authentication schemes. In: Proc. of the 29th International Conference on Human Factors in Computing Systems (CHI 2011). pp. 2055–2064. ACM (2011)
14. Jermyn, I., Mayer, A., Monrose, F., Reiter, M.K., Rubin, A.D.: The design and analysis of graphical passwords. In: Proc. 8th USENIX Security Symposium (1999), http://static.usenix.org/events/sec99/full_papers/jermyn/jermyn.pdf
15. Kumar, M., Garfinkel, T., Boneh, D., Winograd, T.: Reducing shoulder-surfing by using gaze-based password entry. In: Proc. 3rd Symposium on Usable Privacy and Security (SOUPS) 2007. pp. 13–19. ACM Press (2007)
16. Nali, D., Thorpe, J.: Analyzing user choice in graphical passwords. Technical Report TR-04-01, School of Computer Science, Carleton University (May 2004), http://www.cs.carleton.ca/research/tech_reports/2004/TR-04-01.pdf

17. van Oorschot, P.C., Salehi-Abari, A., Thorpe, J.: Purely automated attacks on passpoints-style graphical passwords. IEEE Transactions on Information Forensics and Security 5(3), 393–405 (September 2010)
18. van Oorschot, P.C., Thorpe, J.: On predictive models and user-drawn graphical passwords. ACM Transactions on Information and System Security 10(4), 5:1–5:33 (January 2008)
19. van Oorschot, P.C., Thorpe, J.: Exploiting predictability in click-based graphical passwords. Journal of Computer Security 19(4), 669–702 (December 2011)
20. Passfaces: The science behind Passfaces. Online (September 2001), http://www.passfaces.com/enterprise/resources/white_papers.htm
21. Sahami Shirazi, A., Moghadam, P., Ketabdar, H., Schmidt, A.: Assessing the vulnerability of magnetic gestural authentication to video-based shoulder surfing attacks. In: Proc. 30th International Conference on Human Factors in Computing Systems (CHI) 2012 [1], pp. 2045–2048
22. Tafasa: PatternLock. Online, http://www.tafasa.com/patternlock.html
23. Tao, H.: Pass-Go, a new graphical password scheme. Master's thesis, University of Ottawa (2006), http://site.uottawa.ca/~cadams/papers/HaiTaoThesis.pdf
24. Weidenbeck, S., Waters, J., Birget, J.C., Brodskiy, A., Memon, N.: Authentication using graphical passwords: Basic results. In: Proc. Human-Computer Interaction International (HCII) 2005 (July 2005), http://clam.rutgers.edu/~birget/grPssw/susan3.pdf
25. Weidenbeck, S., Waters, J., Birget, J.C., Brodskiy, A., Memon, N.: Authentication using graphical passwords: effects of tolerance and image choice. In: Cranor, L.F., Zurko, M.E. (eds.) Proc. Symposium on Usable Privacy and Security (SOUPS) 2005. pp. 1–12. ACM (2005)
26. Weiss, R., De Luca, A.: Passshapes: utilizing stroke based authentication to increase password memorability. In: Proceedings of the 5th Nordic Conference on Human-Computer Interaction (NordiCHI) 2008. pp. 383–392. ACM (2008)
27. Whalen, T., Inkpen, K.M.: Gathering evidence: use of visual security cues in web browsers. In: Inkpen, K.M., van de Panne, M. (eds.) Proceedings of Graphics Interface 2005. Graphics Interface, vol. 112, pp. 137–144. Canadian Human-Computer Communications Society (2005), http://portal.acm.org/citation.cfm?id=1089532
28. Wright, N., Patrick, A.S., Biddle, R.: Do you see your password?: applying recognition to textual passwords. In: Cranor, L.F. (ed.) Proc. 8th Symposium on Usable Privacy and Security (SOUPS) 2012. pp. 8:1–8:14. ACM (2012)
29. Zakaria, N.H., Griffiths, D., Brostoff, S., Yan, J.: Shoulder surfing defence for recall-based graphical passwords. In: Cranor, L.F. (ed.) Proc. 7th Symposium on Usable Privacy and Security (SOUPS) 2011. pp. 6:1–6:12. ACM (2011)

A Survey

[In the survey questions reproduced below, we use _____ to indicate that the question allowed a free-form answer, ◯ to indicate that a single choice could be made, and ☐ to indicate that multiple choices could be made. Participants completed questions 1–4 during the distraction during Scheme 1, questions 5–8 during the distractions during Scheme 2, questions 9–12 during the distractions during Scheme 3 or 4, and questions 13–16 at the end of the study.]

You can skip any questions you prefer not to answer.

1. What is your participant number? _____
2. What is your age? _____
3. What is your gender?
   ◯ Male ◯ Female ◯ Prefer not to say
4. What is the highest level of education you have completed?
   ◯ Some high school
   ◯ High school diploma
   ◯ TAFE diploma[4]
   ◯ Some university education
   ◯ Bachelor's degree
   ◯ Master's degree
   ◯ Doctoral degree
   ◯ Other
5. Are you currently a student?
   ◯ Yes ◯ No
   If yes, what are your year and major? _____
6. Are you currently employed?
   ◯ Yes ◯ No
   If yes, what is your occupation? _____
7. Do you use a computer daily for work?
   ◯ Yes ◯ No
8. Do you have a degree in OR are currently studying toward a degree in an IT-related field (e.g., information technology, computer science, electrical engineering, etc.)?
   ◯ Yes ◯ No
9. Have you ever (select all that apply)
   ☐ Designed a website
   ☐ Registered a domain name
   ☐ Used SSH
   ☐ Configured a firewall
   ☐ Created a database
   ☐ Installed a computer program
   ☐ Written a computer program
   ☐ None of the above
10. Have you ever taken or taught a course on computer security?
    ◯ Yes ◯ No
11. Please check all of the following statements that describe your password habits.
    ☐ I use the same password for every website.
    ☐ I have a few passwords that I use interchangeably.
    ☐ I have one password that I use for important sites and another password I use for less important sites.
    ☐ I use different passwords for each site.
    ☐ I use my web browser's password manager to store my passwords.
    ☐ I write my passwords down on a piece of paper.
    ☐ I use a separate program to store my passwords.
12. Please specify the brand and model of your mobile phone. _____
13. If you have an iPhone, which of the following options best describes your passcode lock habits?
    ☐ I have set a numerical passcode to lock/unlock my iPhone

[4] *[In Australia, TAFE stands for Technical and Further Education, and such institutions typically offer vocational tertiary education courses.]*

☐ I have installed a third-party application to simulate Android grid lock screen on my iPhone
☐ I have no lock screen setting on my iPhone
☐ I don't have an iPhone

14. If your mobile supports Android, which of the following options best describes your lock screen habits?
☐ I have set a numerical passcode to lock/unlock my mobile
☐ I use grid lock screen on my mobile phone
☐ I have no lock screen setting on my mobile phone
☐ I don't use an Android mobile phone
15. Please rate each task in the study based on how difficult it was to complete the task (1=very easy, 2=easy, 3=hard, 4=very hard).
(a) T1: Creating password in a grid.
(b) T2: Creating password in a grid with start and end points.
(c) T3: Creating password in a grid with holes, if you did this task.
(d) T4: Creating password in an asymmetric screen, if you did this task.
16. After completing these tasks, would you use this password scheme if your computer was equipped with an eye-tracking device?
◯ Yes ◯ No
Why or why not? _____

B Frequency Tables

B.1 Scheme 1: Grid

Scheme 1: All points

0.0719	0.0479	0.0838	0.0599	0.0599
0.0539	0.0778	0.0958	0.0599	0.0479
0.0060	0.0479	0.0838	0.0539	0.0299
0.0060	0.0240	0.0419	0.0299	0.0180

Scheme 1: First point

0.5000	0.1818	0.0455	0.0000	0.0455
0.0909	0.0455	0.0000	0.0000	0.0000
0.0000	0.0000	0.0000	0.0000	0.0000
0.0000	0.0000	0.0909	0.0000	0.0000

Scheme 1: Last point

0.0455	0.0000	0.0000	0.1364	0.1364
0.0455	0.0455	0.0000	0.0455	0.0455
0.0000	0.0000	0.0909	0.0909	0.0000
0.0000	0.0455	0.0909	0.0455	0.1364

Scheme 1: Stroke frequency

	4 ←	3 ←	2 ←	1 ←	1 →	2 →	3 →	4 →
3 ↑	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
2 ↑	0.0000	0.0000	0.0000	0.0000	0.0207	0.0000	0.0069	0.0069	0.0000
1 ↑	0.0000	0.0000	0.0069	0.0207	0.0690	0.0345	0.0000	0.0069	0.0000
	0.0000	0.0000	0.0138	0.1310	0.0000	0.2276	0.0069	0.0000	0.0000
1 ↓	0.0000	0.0138	0.0276	0.0276	0.2414	0.0621	0.0138	0.0138	0.0000
2 ↓	0.0000	0.0000	0.0000	0.0069	0.0069	0.0138	0.0069	0.0000	0.0000
3 ↓	0.0000	0.0000	0.0000	0.0000	0.0138	0.0000	0.0000	0.0000	0.0000

B.2 Scheme 2: Grid with cued start/end

Scheme 2: All points

0.0085	0.0932	0.0508	0.0593	0.0169
0.0000	0.1017	0.1186	0.1017	0.0424
0.0085	0.1017	0.1102	0.0254	0.0424
0.0085	0.0254	0.0254	0.0254	0.0339

Scheme 2: Second point

0.0455	0.2273	0.0000	0.0000	0.0455
0.0000	0.4091	0.0455	0.0000	0.0455
0.0455	0.0909	0.0000	0.0000	0.0000
0.0000	0.0455	0.0000	0.0000	0.0000

Scheme 2: Second last point

0.0000	0.0000	0.0455	0.0000	0.0000
0.0000	0.0000	0.0455	0.3636	0.0000
0.0000	0.0000	0.2273	0.0455	0.0455
0.0000	0.0455	0.0455	0.1364	0.0000

Scheme 2: Stroke frequency

	4 ←	3 ←	2 ←	1 ←		1 →	2 →	3 →	4 →
3 ↑	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
2 ↑	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
1 ↑	0.0000	0.0071	0.0071	0.0143	0.0357	0.0286	0.0000	0.0000	0.0000
	0.0000	0.0000	0.0071	0.0929	0.0000	0.2357	0.0000	0.0214	0.0000
1 ↓	0.0000	0.0214	0.0071	0.0643	0.3357	0.0643	0.0286	0.0000	0.0000
2 ↓	0.0000	0.0000	0.0000	0.0071	0.0000	0.0000	0.0000	0.0000	0.0000
3 ↓	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000

B.3 Scheme 3: Grid with holes

Scheme 3: All points

0.0455	0.0303	0.0758	0.0303	0.0455
0.0455		0.0152	0.0455	0.0303
0.1061	0.1515	0.1061		0.0303
	0.1212	0.0758		0.0455

Scheme 3: First point

0.2222	0.1111	0.0000	0.0000	0.0000
0.1111		0.0000	0.0000	0.0000
0.2222	0.0000	0.0000		0.0000
	0.1111	0.0000		0.2222

Scheme 3: Last point

0.0000	0.0000	0.1111	0.0000	0.0000
0.0000		0.0000	0.0000	0.2222
0.0000	0.2222	0.1111		0.0000
	0.2222	0.1111		0.0000

Scheme 3: Stroke frequency

	4 ←	3 ←	2 ←	1 ←		1 →	2 →	3 →	4 →
3 ↑	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0175	0.0000
2 ↑	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
1 ↑	0.0000	0.0000	0.0000	0.0877	0.0877	0.0526	0.0175	0.0000	0.0000
	0.0000	0.0000	0.0000	0.1579	0.0000	0.1754	0.0175	0.0000	0.0000
1 ↓	0.0000	0.0000	0.0175	0.0000	0.2632	0.0175	0.0000	0.0000	0.0000
2 ↓	0.0000	0.0000	0.0000	0.0351	0.0000	0.0000	0.0000	0.0000	0.0000
3 ↓	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000

B.4 Scheme 4: Sparse grid

Scheme 4: All points

	0.0532		0.1064	0.0000
0.0532		0.1064		0.0000
	0.0745			0.0957 0.0000
0.0426			0.0851	
		0.0213		0.0213
	0.0000		0.0000 0.0426	

Scheme 4: First point

	0.0769		0.0000	0.0000
0.3077		0.0769		0.0000
	0.0000			0.0000 0.0000
0.0000			0.0000	
		0.0000		0.0000
	0.0000		0.0000 0.0000	

Scheme 4: Last point

	0.0000		0.0769	0.0000
0.0000		0.0000		0.0000
	0.0000			0.0000 0.0000
0.0769			0.0769	
		0.0000		0.0769
	0.0000		0.0000 0.2308	

Scheme 4: Stroke frequency

	5 ←	4 ←	3 ←	2 ←	1 ←		1 →	2 →	3 →	4 →	5 →
5 ↑	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
4 ↑	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
3 ↑	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0123
2 ↑	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
1 ↑	0.0000	0.0000	0.0000	0.0370	0.0370	0.0000	0.0864	0.0000	0.0000	0.0000	0.0000
	0.0000	0.0000	0.0123	0.0370	0.0247	0.0000	0.0617	0.0370	0.0123	0.0247	0.0000
1 ↓	0.0123	0.0000	0.0123	0.0123	0.2840	0.0370	0.0741	0.1111	0.0000	0.0000	0.0000
2 ↓	0.0000	0.0000	0.0000	0.0247	0.0000	0.0123	0.0247	0.0000	0.0000	0.0000	0.0000
3 ↓	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
4 ↓	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
5 ↓	0.0000	0.0123	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000

C Sample User-Generated Passwords — Scheme 1