Usability and Strength in Click-based Graphical Passwords

Elizabeth Stobert
Department of Psychology
Carleton University
Ottawa, Canada
estobert@connect.carleton.ca

Abstract
Click-based graphical passwords have attractive usability properties, such as cueing and good memorability. However, image size and number of click-points in each password significantly affect their security. We investigated the usability of such a graphical password system when its parameters were adjusted to provide security equivalent to (or better than) that of text passwords. We found that manipulating different parameters resulted in similar usability. This suggests that the preferred method for adjusting security can be dictated by the constraints of devices and preferences of users. For example, mobile devices might use smaller image sizes and more click-points.

Keywords
Authentication, graphical passwords, usable security

ACM Classification Keywords
K.6.5 Computing Milieux: Security and Protection – Authentication

General Terms
Experimentation, Human Factors, Security

Introduction
Usable authentication concerns increasing the usability of authentication schemes, such as password systems, while still maintaining the security they support [4]. Poor usability

This work was part of E. Stobert's undergraduate honours thesis at Carleton University.

CHI 2010, April 10–15, 2010. Atlanta, Georgia, USA
ACM 978-1-60558-930-5/10/04.

figure 1. The PCCP interface for creating passwords.

can affect security because people may use the system in an insecure manner, such as selecting predictable passwords or reusing passwords across different accounts [5]. Although text passwords should be both memorable and secure, in practice, most passwords are either memorable but easy-to-guess or secure but difficult-to-remember [8]. Furthermore, as keyboard-less devices become more popular, text passwords may become even less practical.

Alternative approaches, such as graphical passwords [1, 9], seek to have passwords that are both memorable and secure. Graphical passwords use images instead of text, and have two distinct advantages over text passwords. First, the *picture superiority effect* [7] identifies the human ability to remember images better than text, which indicates that graphical passwords may have a memorability advantage. Secondly, some schemes include *cueing* [10], where a memory retrieval cue is provided to help people remember and distinguish their passwords. While several graphical password systems have been shown to be usable, their intent has not been to replace text passwords [1]. Our goal is to investigate the usability of a graphical password system when its parameters are adjusted to provide security equivalent to (or better than) that of text passwords.

One category of graphical passwords with the potential to be more secure is *click-based* graphical passwords [1, 11, 6], where users select click-points on one or more images. We investigated one such system, Persuasive Cued Click-Points [2], shown to have good usability at security settings approximately equivalent to a 6-character random text password. To achieve security levels comparable to 8- or 10-character text passwords, we manipulated two parameters: the image size and the number of click-points selected by users. Our results show that for the same level of security, both manipulations have similar usability and memorability effects. This suggests that the constraints of devices and preferences of users may be also taken into account when adjusting security. For example, mobile devices might use smaller images and more click-points.

Background

Persuasive Cued Click-Points (PCCP) [2] is a click-based graphical password system in which a user is presented with a number of images in sequence, and is asked to choose one click-point on each image (Figure 1). The first image is assigned by the system, but each subsequent image in the sequence is determined by the user's previous click. This means that clicking in different places on an earlier image leads the user to different next images. This provides users with a clue towards the correctness of their password entry attempt — if they see the correct image, they know they have selected the correct click-point. As with other click-based graphical passwords [1, 11], the user cannot be expected to repeat exact pixel selections. Thus an invisible *tolerance square* is defined around each click-point so that any of the enclosed pixels are considered acceptable. To help create more secure passwords, PCCP assists users during password creation by providing a *viewport* that highlights part of the image and asks users to choose a click-point within the viewport, thus resulting in more randomized choices. If users are unable to select a memorable point in the current viewport, they may press the shuffle button, which randomly repositions the viewport. The random viewport, together with the shuffle button, has been shown to ensure that click-points are randomly distributed, addressing a problem seen in earlier schemes. PCCP has been shown to be more secure than other click-based password systems [2] while maintaining login times and success rates comparable to text passwords. However, to be used as a replacement for text passwords, PCCP needs to be at least as secure as standard text passwords. We can adjust the security of PCCP by manipulating its parameters, which in turn affect the size of the theoretical password space.

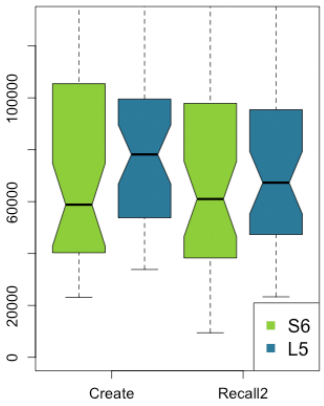

figure 2. Durations in ms: S6 vs L5.

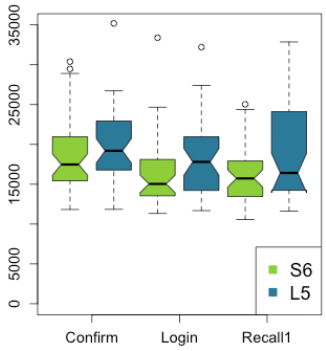

figure 3. Durations in ms: S6 vs L5.

table 1. Theoretical password space for different length text passwords, and PCCP passwords with varying parameters.

chars	n	space
95	6	2^{39}
95	8	2^{53}
95	10	2^{66}

	w	h	c	space
S5	451	331	5	2^{44}
S6	451	331	6	2^{53}
S7	451	331	7	2^{61}
L5	800	600	5	2^{52}
L6	800	600	6	2^{63}
L7	800	600	7	2^{73}

The theoretical password space for a password system is the number of possible passwords that could be generated according to the system specifications. A larger theoretical password space indicates a lower likelihood that any particular password may be guessed. For text passwords, the theoretical password space is typically reported as 95^n, where n is the length of the password, and 95 is the number of typeable characters on the keyboard. For PCCP, the theoretical password space is calculated as: $((w \times h)/t^2)^c$ where the number of places that the user could click (the width (w) multiplied by the height (h) of the image) divided by the size of the tolerance square (t^2, commonly set to 19^2, but may vary) is all raised to the power of the number of click-points (c). Table 1 shows the theoretical password space for several lengths of text passwords and the theoretical password space for PCCP with different parameters. As shown in Table 1, the theoretical password space for PCCP can be adjusted to approximate the space of text passwords of varying lengths. For example, an 8-character text password has approximately the same password space as a PCCP password with a small image size (451 by 331) and 6 click-points (S6), or a large image size (800 by 600) and 5 click-points (L5).

Memorability is an important issue for a password scheme. While a password system should be easy to understand and quick to use, neither of these features is relevant if users cannot remember their passwords. PCCP takes advantage of both the picture superiority effect and memory cueing when it gives the user a distinct image for each click-point. We are unaware of any work examining whether image size affects memorability, and while we suspect that having more click-points in a password will negatively affect memorability, there is no published work on the topic.

We expect that targeting points on a larger image will take longer. Based on Fitts' Law, a linear increase in distance (with fixed target size) should result in a logarithmic increase in time. Locating points on an image also involves aspects of visual search and visual memory. We expect the time for these additional tasks to increase with the area of the image. We also expect that adding another click-point will increase the time because it adds a visual search on an additional image and requires extra time to target.

Experiment

With PCCP, we can manipulate the image size and the number of click-points per password. We were interested in seeing which manipulation resulted in better usability and memorability. We hypothesized that for conditions with approximately comparable theoretical password spaces, the condition with the higher number of click-points would have lower usability, because the extra physical task of entering another click-point would be dominant.

There were four experimental conditions: *S6*: small image, 6 click-points; *S7*: small image, 7 click-points; *L5*: large image, 5 click-points; and *L6*: large image, 6 click-points. This allowed for two comparisons between conditions with similar theoretical password spaces: one between S6 and L5, and one between S7 and L6 (Table 1). We hypothesized that S6 would have lower usability than L5, and S7 would have lower usability than L6.

We conducted a two-part lab study, with sessions scheduled approximately two weeks apart. In the first session, participants created and re-entered PCCP passwords. In the second session, participants re-entered these same

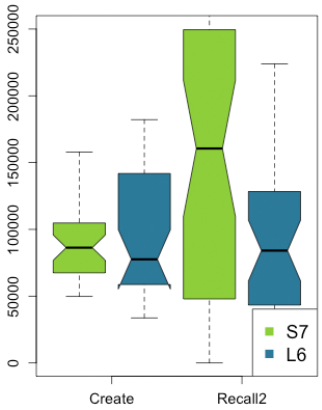

figure 4. Durations in ms: S7 vs L6.

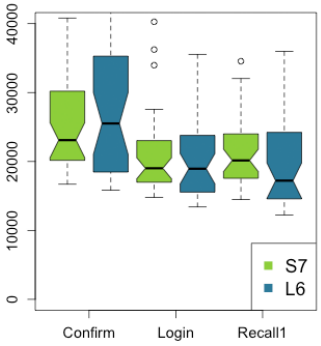

figure 5. Durations in ms: S7 vs L6.

passwords. This created five experiment phases over the two sessions: *create*, *confirm*, *login*, *recall-1* and *recall-2*. Create, confirm, login and recall-1 occurred in the first session and measured participants' ability to successfully create and confirm PCCP passwords, as well as short-term memorability. Recall-2 was conducted in the second session and tested memorability two weeks later, to observe results where memorability was difficult.

The experiment was conducted using a Windows desktop computer, running a custom stand-alone C# application. A set of 465 images was used in the experiment, and no images were repeated between passwords. The smaller images were lower resolution versions of the large images. There were 28 participants across these conditions, mainly university undergraduates from various degree programs. No participants were majoring in computer security, and none had previous experience with graphical passwords. Participants were assigned randomly to one of the four experiment conditions, and each participant created and used six distinct passwords. To help the user distinguish their six passwords, each password was attached to a unique fictional account. For each password, participants created the password by selecting their click-points, confirmed the password to make sure they remembered it, and were asked to login using that password. At the end of the first session, participants were asked to log in to the same accounts in shuffled order. In the second session, participants returned to the lab and were asked to once again log in to their accounts.

For each of the study phases (create, confirm, login, recall-1 and recall-2) we measured usability in three ways: the time it took participants to complete each phase of the study; the number of errors they made in entering their passwords; and their success rates when logging in. Conditions that took less time, had fewer errors, and had higher success rates were judged to have better usability.

Results

In this section, we report on our pairs of conditions with comparable password spaces (S6 vs. L5 and S7 vs. L6). Several figures show boxplots to illustrate distributions. Boxplots show the median, the inner quartiles (as a box), the outer quartiles (as whiskers) and any outliers (as circles). The notches indicate the 95% confidence interval around the median. Statistical tests showed no significant differences in the distributions illustrated by these figures.

Times: Durations were calculated as the time from when the first image appears on the screen until the user presses the login button (which includes entering their username). Conditions with comparable theoretical password spaces took similar lengths of time to complete. We found that having more click-points did not result in longer times for participants than having a larger image. This was surprising, because having more click-points means that the user has to cope with finding extra click-points on new images.

S6/L5: As seen in Figures 2 and 3, S6 and L5 durations were similar in each of the phases (create, confirm, login, recall-1 and recall-2). t-tests between the two conditions for each phase showed no significant differences.

S7/L6: As seen in Figure 4 and 5, for S7 and L6, durations were again similar for most phases, and t-tests showed no significant differences for create, confirm, login and recall-1 durations. A difference was suggested by the box plots for recall-2 durations, but with $t(10.678) = 1.382, p < 0.171$, the difference was not significant.

Errors: An error was counted any time a participant restarted their password attempt, or pressed login with the wrong password. The median number of errors was zero for the confirm, login and recall-1 phases in all four examined conditions. Figure 6 shows the median number of errors for the recall-2 phase in all four examined conditions. Wilcoxon tests (used because distributions were non-normal) showed that there were no significant

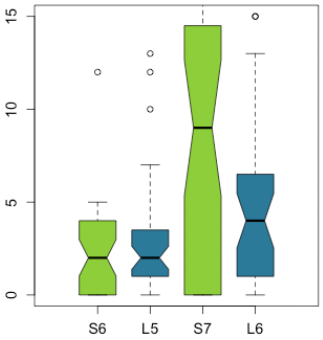

figure 6. Recall2 Errors: S6 vs L5 - S7 vs L6.

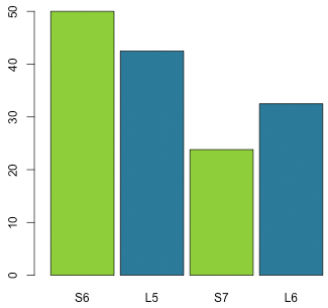

figure 7. Recall2 Success %: S6 vs L5 - S7 vs L6.

differences in number of errors between S6 and L5 or between S7 and L6 in any of the 5 phases.

Success Rates: A password entry attempt was considered successful any time the entire correct password for an account was entered, with no mistakes or restarts. Success rates were obtained by dividing a user's number of successful password logins by their total number of password entry attempts. Success rates were high for the confirm, login and recall-1 phases, with medians near 100% for those phases. As expected, success rates were considerably lower for the recall-2 phase (Figure 7). It was difficult for users to remember six passwords after two weeks with no interim rehearsal.

S6/L5: In the recall-2 phase, success percentages were approximately 50% and 40% for S6 and L5 respectively. A Chi-squared test found $\chi^2 = 0.211, p < 0.646$, and the difference in percentages was not significant.

S7/L6: In the recall-2 phase, success percentages were approximately 25% for S7 and 35% for L6. The difference was not statistically significant ($\chi^2 = 0.666, p < 0.415$).

Discussion

Our hypothesis was that in conditions with similar theoretical password spaces, usability would be better in the conditions with fewer click-points. However, there were no significant differences in duration, number of errors, or success rates in either of the pairs of conditions with comparable theoretical password spaces. We found no evidence that increasing the number of click-points has a greater effect on usability than increasing the size of the image.

Increasing the number of click-points clearly adds to the cost of memorability, visual search, and targeting associated with each additional click-point. However, a larger image will add to the cost of memorability, visual search, and targeting associated with each individual click-point. It appears that these two kinds of increase may balance each other out.

This suggests that in increasing the security of PCCP, there is no particular advantage to one approach over the other. There may be other reasons for favouring one approach. For example, shoulder-surfing could potentially be more of a problem on a larger image, and larger images may be difficult to display on small screens (such as mobile devices). Although success rates were low and error rates were high for the recall-2 condition, we were not overly concerned about the apparent lack of usability after two weeks. We chose this design to emphasize differences among conditions, and understood that this did not accommodate ecological validity. In real life, users are unlikely to ever create six passwords in a row, and then wait two weeks to try and log in using them all at one time. This issue points to a need for future work involving field trials, in which more ecologically valid data could be collected.

We can compare the results of this study to the results of a study on multiple password interference [3] using text passwords and PassPoints [11], an earlier click-based password system, that followed the same methodology. Whereas Passpoints uses five clicks on one image, PCCP uses five images and one click per image. We found comparable success rates to their PassPoints condition. Chiasson et al. [3] suggested that the usability and memorability of PassPoints was superior to that of text passwords, so it is thus reasonable to speculate that PCCP is also superior to text passwords. Adding more click-points to Passpoints would appear to create a harder task for the user, since not only would they have to remember more points without individual cues, but also the order of the points. That PCCP's security may be increased without adding additional memory tasks is a strength of the PCCP system, and a possible advantage of PCCP over PassPoints.

Conclusion

We asked how increasing the number of click-points or the image size in a PCCP graphical password scheme would affect usability, and found that while both increases have effects on usability, neither produces a markedly different effect. We compared conditions with similar theoretical password spaces, and found that when the password spaces are comparable, changing the number of click-points or image size had similar demands on usability. This suggests that there is no advantage to one kind of increase in security over another. This result might allow other considerations to be taken into account when making modifications to the system, such as the size of the screen on a mobile device, which would favour the use of smaller images. Future work in this area might include a field study to investigate these modifications in a more ecologically valid situation and more investigation into how the cost to usability is similarly affected by both click-points and image size.

Acknowledgments

Thanks to Robert Biddle, Sonia Chiasson and Alain Forget for their guidance. This work was supported by the Natural Sciences and Engineering Research Council of Canada.

References

[1] R. Biddle, S. Chiasson, and P. C. van Oorschot. Graphical passwords: Learning from the first generation. Technical Report TR-09-09, Computer Science, Carleton University, `http://www.scs.carleton.ca/research/tech_reports`, 2009.

[2] S. Chiasson, A. Forget, R. Biddle, and P. C. van Oorschot. Influencing users towards better passwords: Persuasive Cued Click-Points. In *Human Computer Interaction (HCI), The British Computer Society*, 2008.

[3] S. Chiasson, A. Forget, E. Stobert, P. C. van Oorschot, and R. Biddle. Multiple password interference in text and click-based graphical passwords. In *ACM Conf. on Computer and Communications Security (CCS)*, 2009.

[4] L. Cranor and S. Garfinkel. *Security and Usability: Designing Systems that People Can Use*. O'Reilly Media, edited collection edition, 2005.

[5] D. Florencio and C. Herley. A large-scale study of WWW password habits. In *16th ACM International World Wide Web Conference (WWW)*, May 2007.

[6] W. Jansen, S. Gavrila, V.Korolev, R. Ayers, and R. Swanstrom. Picture password: A visual login technique for mobile devices. Technical Report NISTIR 7030, National Inst. of Standards and Tech, 2003.

[7] D. Nelson, V. Reed, and J. Walling. Pictorial Superiority Effect. *Journal of Experimental Psychology: Human Learning and Memory*, 2(5):523–528, 1976.

[8] M. A. Sasse, S. Brostoff, and D. Weirich. Transforming the 'weakest link' – a human/computer interaction approach to usable and effective security. *BT Technology Journal*, 19(3):122–131, July 2001.

[9] X. Suo, Y. Zhu, and G. Owen. Graphical passwords: A survey. In *Annual Computer Security Applications Conference (ACSAC)*, December 2005.

[10] E. Tulving and Z. Pearlstone. Availability versus accessibility of information in memory for words. *Journal of Verbal Learning and Verbal Behavior*, 5:381–391, 1966.

[11] S. Wiedenbeck, J. Waters, J. Birget, A. Brodskiy, and N. Memon. PassPoints: Design and longitudinal evaluation of a graphical password system. *Int. Journal of Human-Computer Studies*, 63(1-2):102–127, 2005.