Graphical Authentication Resistance to Over-the-Shoulder-Attacks

Ashley A. Cain
Old Dominion University
Norfolk, VA 23510, USA
acain001@odu.edu

Steffen Werner
University of Idaho
Moscow, ID 83844, USA
swerner@uidaho.edu

Jeremiah D. Still
Old Dominion University
Norfolk, VA 23510, USA
jstill@odu.edu

CHI'17 Extended Abstracts, May 06-11, 2017, Denver, CO, USA
ACM 978-1-4503-4656-6/17/05.
http://dx.doi.org/10.1145/3027063.3053236

Abstract
Graphical passwords offer advantages for memorability over conventional alphanumeric passwords, but in some cases they have been vulnerable to over-the-shoulder-attacks (OSA). Thus, many second-generation graphic based schemes are specifically designed to be resistant to OSA. This is often achieved by not having users select targets directly, but by adding cognitive operations to create seemingly random response patterns. This study takes the first step to directly compare three prototypical graphical password schemes to determine their relative resistance to OSAs employing a within-subjects design. We found that schemes requiring cognitive operations in response to target patterns were superior to direct selection of targets. Convex Hull Click was most secure, followed by What You See is What You Enter, while Use Your Illusion showed high vulnerability to OSA. In addition, we discuss a diversity of previous measurements, which are meant to examine security strength of new approaches. We highlight the need for standard OSA resistance measures depending on threat model needs.

Author Keywords
Graphical password; over the shoulder attack; security.

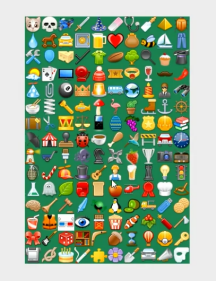

Figure 1: Prototype of CHC

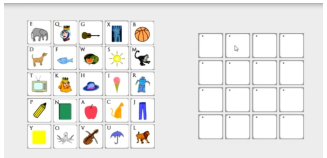

Figure 2: Prototype of WYSWYE

Figure 3: Prototype of UYI

ACM Classification Keywords
D.4.6, K.6.5.

Introduction

Statement of the Problem

Currently, alphanumeric passwords are the most commonly used method of authentication [15]. Alphanumeric passwords offer greater security against brute force attacks when they are long and complex [4]. However, users have difficulty remembering this kind of information. Users tend to deal with difficulties remembering passwords by writing them down [1] or keeping them in a "password file," which weakens security.

To help improve the usability of conventional authentication, researchers have developed graphical schemes [5; 9; 19]. They leverage the picture superiority effect taking advantage of our ability to dual-encode objects visually and semantically [13]. Users can successfully remember graphical passwords after a long delay [12]. However, graphical passwords have been criticized for being vulnerable to over-the-shoulder-attacks (OSA), which happen when an observer peeks at and steals a passcode in a shared space. To solve this security vulnerability, many graphical schemes have been developed to be resistant to OSAs [5; 9; 19]. This is often achieved by grouping targets with distractors [5; 16; 18], mentally translating targets to another location [9; 20; 14], or disguising targets [2; 11; 19].

Previous researchers have investigated the OSA resistance of many approaches [2; 3; 16; 18]. However, a diversity among measurement approaches makes interpretations and comparisons across proposed methods challenging. For example, participants who take on the role of attacker may have opportunities to view a passcode once [5] or multiple times [9]. They may be allowed to make a single attempt [11; 19] or multiple attempts to identify a passcode [9; 12]. They may or may not be motivated by a reward for correctly identifying a passcode [7; 14]. There is a need to for clarification between method approaches to determine relative resistance to OSA.

Graphical Authentication's Resistance to OSA

Previous researchers have experimentally assessed their graphical scheme's resistance to OSA (see table 1). We review the graphical authentication schemes that were designed to resist OSA by grouping targets among distractors, translating targets to another location, and disguising targets.

The following 3 schemes confuse attackers by grouping targets with distractors. In other words, an observer would be unsure if objects are distractors or targets representing the passcode. Goa and colleagues' [5] graphical authentication scheme asked users to indirectly select their passcode by clicking on a row that contains a target icon. Van Eekelen and colleagues' [16] scheme has users select a shape of a certain color, and their target passcode could be either a shape or a color. Yamamoto and colleagues' method [18] prompts the user to select targets among distractors temporally in sequential slide shows. Users could authenticate by selecting the slide show that contained their target among distractors.

Participants received training on Gao and van Eekelen and colleagues' schemes, and then they took on the role of attacker. Gao and colleagues found that when

participants observed the researcher log in one time, none guessed the passcode given three attempts. Van Eekelen and colleagues found that when participants only viewed a video once of a researcher logging in, none were able to guess the passcode given one attempt to guess. Yamamoto and colleagues simulated an attacker. The researchers used the assumption that the attacker would be able to remember every image viewed in the slide show. Given a single observation of the passcode, the simulated attacker would not be able to reliably guess the passcode, but given four observations, an attacker would be able to narrow down the choices and identify the passcode.

The following 3 schemes allow users to translate targets spatially elsewhere instead of directly clicking on them. Participants were trained on a method in which they translated the locations of target faces on a grid to bars protruding from the side of the grid [9]. They took on the role of attacker and observed 3 logins. After a distractor task, 21% of participants were able to steal the passcode. Participants were trained on a method in which they typed text associated with target images into a text box [20]. They viewed videos of a single login, and 3 out of 10 attackers were able to identify the passcode. Participants were trained on a method in which they used a scroll bar on the side of an image to select points on the image indirectly [14]. As attackers, participants viewed sets of screenshots from 2 logins. They were allowed 10 attempts and 2 weeks to steal the passcode. If they guessed right, they received a gift. None guessed the passcode.

Other methods provide resistance to OSA by disguising targets [2; 6; 11; 19]. Use Your Illusion (UYI) [6] and Rapid Serial Visual Presentation schemes allow users to log in by selecting degraded versions of targets. Degrading targets interferes with object recognition of attackers who are unfamiliar with the original targets.

Schemes	Number of Attempts	Reward
Liu, Gao, Wang, & Chang (2011)	At least 5	No
De Luca, Hertzschuch, & Hussmann (2010)	1	N/A (simulation)
Gao, Liu, Dai, Wang, & Chang (2009)	3	No
Zakaria, Griffiths, Brostoff, & Yan (2011)	1	Yes
Sun, Chen, Yeh, & Cheng (2016)	10	Yes
Yamamoto, Kojima, & Nishigaki (2009)	1	N/A (simulation)
Kim et al. (2010)	3	No
Lin, Dunphy, Olivier, & Yan (2007)	1	No
Zangooei, Mansoori, & Welch (2012)	1	No
Jenkins, McLachlan, & Renaud (2014)	1	No
van Eekelen, van den Elst, & Khan (2013)	1	No
Cain & Still (2016)	1	No

Table 1: Differences among measurement practices.

The Current Research

Previous investigations have considered each graphical scheme individually, and there has been much diversity among the ways OSAs resistance was measured. In a within-subjects study, we directly compare 3

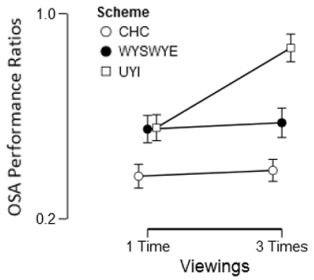

Figure 4: Mean OSA performance by scheme and number of viewings. Error bars represent 95% confidence intervals.

prototypical examples of OSA-resistant graphical passwords. Our prototypes were based on Weidenbeck's [17] Convex Hull Click (CHC), Khot's [8] What You See is What You Enter (WYSWYE), and Hayashi's [6] Use Your Illusion (UYI). They represent passcodes that group targets with distractors, translate targets to another location, and disguise targets. CHC is an example of a method that uses grouping to obscure a passcode. Instead of clicking target icons, users click in a region that is formed by them. WYSWYE is a representation of methods that require translating the passcode to another location. Users view their passcode on a grid, they mentally delete the row and column that does not contain their passcode, and they click the resulting locations of their targets on a smaller, blank grid. UYI is a prototypical example of a method that uses disguising. Targets are presented among distractors on a grid. Users click their targets to authenticate, but the images are degraded by removing detail. Because we used measurements that were consistent among 3 prototypes, we offer a direct comparison between graphical schemes designed to be OSA resistant.

Method

Participants

Twenty undergraduate students volunteered to participate and were compensated with course research credit. There were 9 males ranging from 18 to 53 years old (M = 23.05).

Stimuli

Videos were taken using screen capture software of a researcher logging into prototypes of graphical schemes built in Paradigm©. The prototypes were based on CHC [17], WYSWYE [8], and UYI [6]. See figure 1, 2 and 3 for screenshots of prototypes. The prototypes were designed to provide low-security similar to a 4-digit PIN. Our implementations resulted in a range of password strength between 9.5 bits (UYI) to 10.8 bits (WYSWYE). CHC's strength varies significantly depending on the size of the convex hull but averaged in between the other two schemes.

Procedure

Participants read and signed a consent form while sitting in front of a desktop computer. They received instructions for each scheme before attempting to authenticate. To assure that participants were familiar with each scheme, they attempted to authenticate 10 times using our prototype of CHC, 10 times using WYSWYE, and 9 times using UYI. Prototypes were counterbalanced across participants. After this training with each scheme, the participants took on the role of attacker. They viewed a video of the researcher logging in 1 time using a mouse pointer, and they were asked to identify the passcode just observed on a sheet of paper. The video of the login allowed for the participants to have an ideal view. This represented a best case scenario for an attacker performing an OSA. Real world situations would be expected to be less vulnerable. Then participants viewed a video of the researcher logging in 2 more times. Targets and distractors were randomly rearranged for each video. Each passcode was the same as was entered on the first viewing. Participants made a second guess to identify the passcodes after the third video.

Results

With full knowledge of a passcode, it would be possible for attackers to identify 3 out of 3 icons for CHC, 4 out of 4 images for WYSWYE, and 3 out of 3 images for

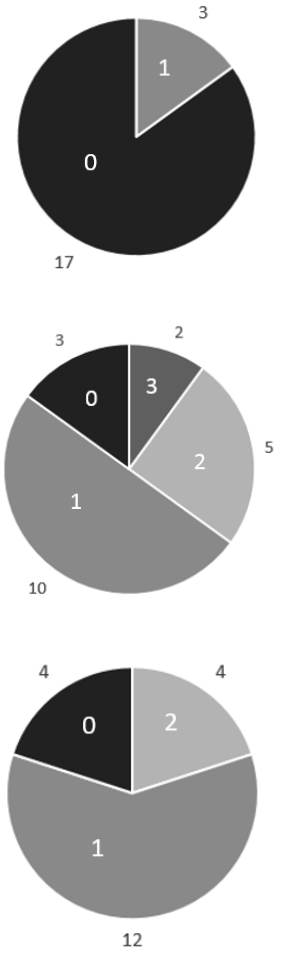

Figure 5: CHC, WYSWYE, and UYI after 1 viewing. The whole represents all participants. The numbers on the pie represent numbers of targets identified.

UYI. We measured OSA performance as ratios of each passcode that were identified. For example, if a participant identified 1 out of 3 targets correctly, they would have a score of .33. A 3 (authentication scheme: CHC, WYSWYE, UYI) x 2 (number of viewings: 1, 3) repeated measures ANOVA was conducted to evaluate OSA performance. There was a main effect of scheme and of viewings, $p < .001$. The main effects were accounted for by the 2 way interaction. There was an interaction between number of times viewed and scheme, $F(2, 38) = 25.87$, $p < .001$, partial $\eta^2 = .577$. The interaction can be explained by an increase in OSA performance for UYI but not CHC or WYSWYE after 3 viewings compared to 1 viewing (see figure 4). Simple effects were investigated using a Bonferroni correction. CHC was more resistant to OSA than WYSWYE, $t(38) = -6.99$, $p < .001$, and UYI, $t(38) = -12.61$, $p < .001$. WYSWYE was more resistant than UYI, $t(38) = -5.62$, $p < .001$. Allowing 1 viewing of the video provided more resistance to OSA than allowing 3 viewings, $t(38) = -6.51$, $p < .001$.

After viewing 1 login, none of the participants were able to identify a full passcode. Higher ratios of partial passcodes were identified for UYI and WYEWYE than CHC. See figure 5 for targets identified after 1 viewing at the participant level. After viewing 3 logins, no participant guessed the full passcode for CHC or WYSWYE. However, after viewing 3 videos, 9 out of 20 participants identified the full passcode for UYI. See figure 6 for targets identified after 3 viewings at the participant level. For CHC, chance performance after 3 viewings was 1.4 for 1 icon and 18.6 for none. For WYSWYE, chance performance was 0.1 for 3 images, 2.0 for 2 images, 8.4 for 1 image, and 9.5 for none. Chance performance for UYI was 0.03 for 3 images, 0.7 for 2, and 5.3 for 1.

Conclusions

Findings suggest that grouping targets with distractors and translating targets to another location are effective methods for securing graphical passwords. While in all cases more participants than predicted by chance were able to identify elements of the passcodes, this did not present a severe vulnerability within 3 authentication observations in an OSA under ideal conditions for CHC. CHC's original study [17] did not test for OSA vulnerabilities. However, our findings that grouping targets with distractors was an effective defense aligns with previous studies. For example, when participants clicked on a row containing a target instead of clicking targets directly, no participant identified the passcode given 3 attempts [5]. WYSWYE was more vulnerable than CHC but showed a clear security advantage over UYI. Our finding that WYSWYE was resistant to OSA was consistent with the scheme's original study [8] in which no participants could steal a full passcode. Participants had viewed screenshots from 1 login and had 3 attempts to identify the passcode. It is evident that authentication is much more secure when a user does not click directly on their targets because they are grouped with distractors or translated to another location. Clicking directly on degraded targets using UYI did not thwart attackers given 3 viewings. This finding is not aligned with a previous study on the resistance of disguising targets using the RSVP scheme [2]. RSVP presents images serially rather than statically. This temporal presentation style made identifying targets more difficult. It was found that only half of participants could identify 1 out of 4 targets given 1 attempt. UYI's vulnerability to OSA applies generally and likely even

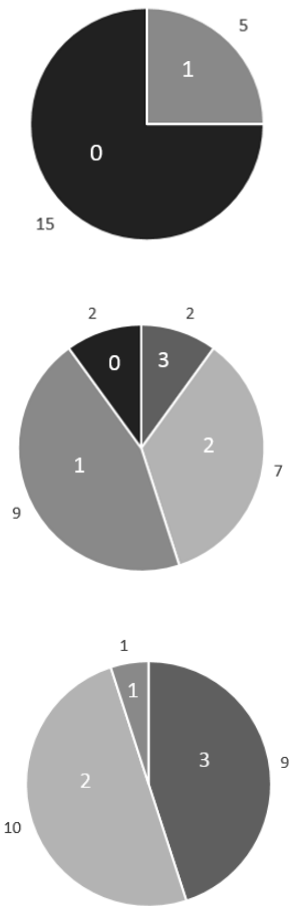

Figure 6: CHC, UYI, and WYSWYE after 3 viewings. The whole represents all participants. The numbers on the pie represent numbers of targets identified.

more so to other direct recognition-based methods of graphical authentication that require the direct, visible selection of static target items.

Future Work

The current research took a step towards ranking the resistance of graphical schemes by using a within-subjects design and uniform measures. We chose to allow for observation of logins using a prerecorded video to ensure that all participants experienced the same conditions. Another approach sometimes used in previous literature is to allow for observation of a live researcher logging in. Both options provide useful information, but a lack of uniformity across the literature makes comparisons among methods more difficult.

Most previous researchers allow for between 1 and 3 viewings of a login before attempts to identify passcodes. Another strategy is to allow for as many playbacks as is desired [10]. Allowing for limited views can be advantageous for capturing behavior representing a casual attacker. Multiple playbacks could also be advantageous to represent attackers who use video recordings.

We did not reward our participants for correctly identifying passcodes. Often in previous literature participants are not rewarded, but other researchers have offered gifts for correctly stealing a passcode [14; 19]. In a within-subject experiment in which we were comparing methods, a lack of motivation likely impacted each method similarly. As a result, rewarding participants would likely not impact results of a comparison.

Moving forward, research needs to be completed that supports a standard practice for exploring OSAs. This ought to facilitate the scientific development of these schemes. In addition, we need to consider characteristics associated with the user, environment, and technology that impact attack performance. Of course, authentication schemes reflect different usability and security strength. Therefore, given a system's security strength needs (e.g., banking vs. home desktop access) an authentication scheme will be selected from an array of options. If the goal is to resist casual attackers not employing a camera, the security requirements are much lighter compared with a system that needs to thwart knowledgeable and resourced attackers. Ideally, a system would protect against both, but that might not offer the best user experience. The literature needs to make a distinction reflecting different design needs, and measurements ought to be standardized.

References

1. Avarne, S. (1988). How to find out a password. *Data Processing & Communication Security*, *12*(2), 16-17.
2. Cain, A. A., & Still, J. D. (2016). A rapid serial visual presentation method for graphical authentication. In *Advances in Human Factors in Cybersecurity* (pp. 3-11). Springer International Publishing.
3. De Luca, A., Hertzschuch, K., & Hussmann, H. (2010, April). ColorPIN: Securing PIN entry through indirect input. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 1103-1106). ACM.
4. Eljetlawi, A. M., & Ithnin, N. (2008, November). Graphical Password: Comprehensive study of the usability features of the Recognition Base Graphical Password methods. In *Convergence and Hybrid*

Information Technology, 2008. ICCIT'08. Third International Conference on (Vol. 2, pp. 1137-1143). IEEE.

5. Gao, H., Liu, X., Dai, R., Wang, S., & Chang, X. (2009, September). Analysis and evaluation of the colorlogin graphical password scheme. In *Image and Graphics, 2009. ICIG'09. Fifth International Conference on* (pp. 722-727). IEEE.
6. Hayashi, E., Dhamija, R., Christin, N., & Perrig, A. (2008, July). Use your illusion: Secure authentication usable anywhere. In *Proceedings of the 4th symposium on Usable privacy and security* (pp. 35-45). ACM.
7. Jenkins, R., McLachlan, J. L., & Renaud, K. (2014). Facelock: familiarity-based graphical authentication. *PeerJ*, *2*, e444.
8. Khot, R. A., Kumaraguru, P., & Srinathan, K. (2012, November). WYSWYE: Shoulder surfing defense for recognition based graphical passwords. In *Proceedings of the 24th Australian Computer-Human Interaction Conference* (pp. 285-294). ACM.
9. Kim, D., Dunphy, P., Briggs, P., Hook, J., Nicholson, J. W., Nicholson, J., & Olivier, P. (2010, April). Multi-touch authentication on tabletops. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 1093-1102). ACM.
10. Lee, M. K. (2014). Security notions and advanced method for human shoulder-surfing resistant PIN-entry. *IEEE Transactions on Information Forensics and Security*, *9*(4), 695-708.
11. Lin, D., Dunphy, P., Olivier, P., & Yan, J. (2007, July). Graphical passwords & qualitative spatial relations. In *Proceedings of the 3rd symposium on Usable privacy and security* (pp. 161-162). ACM.
12. Liu, X. Y., Gao, H. C., Wang, L. M., & Chang, X. L. (2011). An enhanced drawing reproduction graphical password strategy. *Journal of Computer Science and Technology*, *26*(6), 988-999.
13. Nelson, D. L., Reed, V. S., & Walling, J. R. (1976). Pictorial superiority effect. *Journal of Experimental Psychology: Human Learning and Memory*, *2*(5), 523.
14. Sun, H. M., Chen, S. T., Yeh, J. H., & Cheng, C. Y (2016). A shoulder surfing resistant graphical authentication system. *Transactions on Dependable and Secure Computing* (pp. 1-14). IEEE.
15. Suo, X., Zhu, Y., & Owen, G. S. (2005, December). Graphical passwords: A survey. In *21st Annual Computer Security Applications Conference (ACSAC'05)* (pp. 10-pp). IEEE.
16. van Eekelen, W. A., van den Elst, J., & Khan, V. J. (2013, April). Picassopass: A password scheme using a dynamically layered combination of graphical elements. In *CHI'13 Extended Abstracts on Human Factors in Computing Systems* (pp. 1857-1862). ACM.
17. Wiedenbeck, S., Waters, J., Sobrado, L., & Birget, J. C. (2006, May). Design and evaluation of a shoulder-surfing resistant graphical password scheme. In *Proceedings of the working conference on Advanced visual interfaces* (pp. 177-184). ACM.
18. Yamamoto, T., Kojima, Y., & Nishigaki, M. (2009, July). A shoulder-surfing-resistant image-based authentication system with temporal indirect image selection. In *Security and Management* (pp. 188-194).
19. Zakaria, N. H., Griffiths, D., Brostoff, S., & Yan, J. (2011, July). Shoulder surfing defense for recall-based graphical passwords. In *Proceedings of the Seventh Symposium on Usable Privacy and Security* (p. 6). ACM.
20. Zangooei, T., Mansoori, M., & Welch, I. (2012, November). A hybrid recognition and recall based approach in graphical passwords. In *Proceedings of the 24th Australian Computer-Human Interaction Conference* (pp. 665-673). ACM.