Evaluating the Effect of Tolerance on Click-Draw Based Graphical Password Scheme

Yuxin Meng[1] and Wenjuan Li[2]

[1] Department of Computer Science,
City University of Hong Kong, Hong Kong, China
yuxinmeng@ieee.org
[2] Computer Science Division, Zhaoqing Foreign Language College,
Guangdong, China,
wenjuan.anastatia@gmail.com

Abstract. To enhance graphical passwords, we have developed a system of click-draw based graphical password scheme (named CD-GPS) that combined current graphical password techniques and evaluated its performance with human users. In real settings, we identify that the effect of tolerance is a key factor affecting the usability of our scheme, however, we have not explored its specific effect in our previous work. In this paper, we therefore conduct a user study to investigate the effect of tolerance on creating and confirming the click-draw based graphical passwords. In particular, we conduct two experiments with totally 60 participants in the user study. The results show that accurate memory and reproduction for the CD-GPS scheme can be significantly reduced when the tolerance is greatly decreased (e.g., 12×12 pixels). In the end, we further discuss how to select an appropriate tolerance for the scheme of CD-GPS in real deployment.

Keywords: Graphical Password, Authentication, Usable Security, Click Draw, Tolerance Evaluation, Human Factors.

1 Introduction

User authentication is the process to verify whether a user is allowed to access to a particular system or resource. Traditionally, alphanumeric passwords (or called *text-based passwords*) is a widely used method in authenticating users, however, the alphanumeric passwords have some drawbacks with regard to both usability and security [1,2]. For instance, currently, a secure text-based password should be 8 characters or longer, random with upper-case, lower case characters and special characters. This kind of passwords is meaningless and it is hard for human users to remember so that users are more likely to choose simple and short password instead [1]. These usability problems can be translated directly into security problems [3].

To mitigate the drawbacks of the alphanumeric passwords, graphical passwords have been proposed as an alternative to the text-based passwords. The psychology studies [4,6] showed that human brain was better at remembering and recognizing images than text. Moreover, several studies [8,5] also reported positive results that users were

T.W. Chim and T.H. Yuen (Eds.): ICICS 2012, LNCS 7618, pp. 349–356, 2012.

able to remember their graphical passwords accurately after long periods of time. Generally, graphical passwords can be classified into three categories in terms of the input types [15]: click-based graphical passwords, choice-based graphical passwords and draw-based graphical passwords. The click-based graphical passwords (e.g., [9], [11]) require a user to click on an object or element of an image, the choice-based graphical passwords (e.g., [12], [13]) require a user to select a group of images in an ordered sequence, and the draw-based graphical passwords (e.g., [16], [20]) require a user to draw some secrets on an image for authentication.

However, there are still some intrinsic limitations regarding to each category of the above graphical password schemes (e.g., 'hot-spot' issue [17]). Relevant security studies about graphical passwords can be found in [7], [14] and [19]. In our previous work [10], we proposed and developed a click-draw based graphical password scheme (called *CD-GPS*) with the purpose of better enhancing the graphical passwords by combining the above three techniques. In real applications, we notice that the effect of tolerance is a key factor affecting the usability of the scheme.

In this paper, we therefore conduct another user study to investigate the effect of tolerance on creating and confirming the *CD-GPS* passwords. In the scheme of *CD-GPS*, with different values of N, an image can be divided into different tolerance sizes. In particular, we performed two experiments with totally 60 participants. By analyzing the experimental results, we find that accurate memory and reproduction for the scheme of *CD-GPS* can be significantly reduced when the tolerance size is greatly decreased. For a pixel tolerance of about 12×12 (the corresponding value of N is 32), it is extremely hard for users to reproduce their graphical passwords accurately since a lot of click errors are occurred. Based on the results of user study, we point out that an appropriate value of N should be smaller than 32.

The rest of this paper is organized as follows: in Section 2, we introduce background information of the *CD-GPS* scheme; Section 3 details our experimental methodology; the user study and relevant results are described in Section 4; finally, Section 5 concludes our work.

2 Background

In our previous work [10], we proposed and developed a prototype of click-draw based graphical password scheme (called *CD-GPS*) aiming to better enhance the graphical passwords in both usability and security. There are mainly two steps in the scheme: *image selection* and *secret drawing*.

Generally, in the first step of *image selection*, users are required to select some images from an image pool in a story ordered sequence. Story memorization can help users to better remember their selected images and ordered sequence [12]. Then, users are required to further select one or several images for drawing their secrets. In the step of *secret drawing*, the *CD-GPS* scheme divides an image into a $N \times N$ table with appropriate pixel tolerance. Users are required to click-draw their secrets, that is, using a series of clicks to construct their secrets (e.g., a digital number, a letter).

In our developed example system, the image pool contains 10 everyday images (e.g., images of cartoon characters, images of landscape, etc).

Table 1. The values of N and the relevant smallest tolerance sizes

Value of N	Tolerance size (pixel)	Size in cm^2
20	18.75×18.75	0.45
24	15.6×15.6	0.39
32	11.7×11.7	0.27

- In the first step, users are required to select 4 images out of the image pool and remember the sequence like a story, then users should further choose 1 image for click-drawing their secrets.
- In the following step, the example system set $N = 16$ and divided an image into a 16×16 table with 256 clickable squares. Thus, the smallest pixel tolerance is 23×23. Users can click-draw their secrets by clicking a series of clickable squares within the 16×16 table.

The previous user study showed that participants preferred our scheme with respect to both security and usability, and satisfied with this pixel tolerance of 23×23 with a high success rate of creation, confirmation and login respectively.

3 User Study

In this section, we first introduce several tolerance sizes that are evaluated in the user study and we then give an in-depth description of the experimental methodology.

3.1 Tolerance Size

The example system used in this user study is the same as in [10] so that all the images have the pixel size of 500×375. As described above, the system will divide an image into a $N \times N$ table in the step of *secret drawing* (the system used a technique of floating tolerance in balancing the table). Thus, the value of N can greatly affect the tolerance size of a clickable square. For instance, in our previous work, we set $N = 16$, therefore, the pixel size of the smallest clickable square of an image is 23×23. Our previous user study showed that participants were comfortable with this tolerance size.

In the user study, referred to the work [18], we set N to three different values such as 20, 24 and 32. In general, a bigger value of N means a smaller tolerance size of clickable squares. The values of N and corresponding tolerance sizes are described in Table 1. Three values of N (e.g., 20, 24 and 32) are selected because they are respectively increased by one-quarter, one-half and one compared to the value of 16 that we used in our previous work. With the increase of N, the pixel tolerance is respectively decreased to 18.75×18.75, 15.6×15.6 and 11.7×11.7.

3.2 Experimental Methodology

We conducted an in-lab user study which consisted of two experiments (named *Experiment1* and *Experiment2*) with totally 60 participants those who were interested in our

Table 2. Participants' information in the two experiments respectively

Demographic	Male	Female
Experiment1	12	8
Experiment2	23	17

work. All participants are university students with diverse backgrounds (e.g., 20 undergraduate and 40 graduate students) and 20 participants (8 females and 12 males) joined our previous studies. In total, 25 participants are female while the remaining 35 participants are male. In addition, 28 of them are from the department of computer science (not security related major) and the others are from other science and management majors. All the participants are regular computer and web users, and ranged in age from 19 to 35 years.

Before the experiments, we gave an in-depth description of the *CD-GPS* scheme, introduced our objectives of the user study and showed them how to use the example system. Each participant could finish 2 practice trials to get familiar with the example system before the real trials. To investigate the effect of tolerance on creating and confirming the click-draw based graphical passwords, we divided the participants into two experiments as below:

- *Experiment1*: This experiment involved 20 participants (by means of random selection) and only required all participants to click-draw their secrets on the same image (an image of cartoon characters that was very popular in our previous study). Each participant should create and confirm up to 5 *CD-GPS* passwords corresponding to the three tolerance sizes respectively.
- *Experiment2*: This experiment involved the remaining 40 participants and all these participants were required to regularly use the two steps (image selection and secret drawing) to create and confirm their *CD-GPS* passwords. Similar to the *Experiment1*, each participant should complete 5 *CD-GPS* passwords for the three tolerance sizes respectively.

The detailed participants' information is shown in Table 2. For the *Experiment1*, we attempt to explore the effect of tolerance to users on creating and confirming their graphical passwords on the same image. On the other hand, in the *Experiment2*, we aim to investigate the effect of tolerance on the regular use of the *CD-GPS* scheme and identify the minimum affordable pixel tolerance.

For each experiment, we later gave a set of questions to all participants and collected their feedback about the tolerance sizes. Ten-point Likert scales were used in each question where 1-score indicates strong disagreement and 10-score indicates strong agreement. We denoted 5-score as the statement "It is hard to say" for a participant.

4 Results and Analysis

In this section, we describe the results of the two experiments (*Experiment1* and *Experiment2*) and analyze the participants' feedback.

Table 3. The success rate of *CD-GPS* creation and confirmation in the *Experiment1* for the three tolerance sizes

Success Rate	18.75×18.75 pixels	15.6×15.6 pixels	11.7×11.7 pixels
Creation	88/100 (88%)	80/100 (80%)	73/100 (73%)
Confirmation	85/100 (85%)	75/100 (75%)	65/100 (65%)

Table 4. Several questions and relevant scores in the *Experiment1*

Questions	Score (average)
1. I could easily create a password with the pixel tolerance of 18.75×18.75	8.5
2. I could easily create a password with the pixel tolerance of 15.6×15.6	8.0
3. I could easily create a password with the pixel tolerance of 11.7×11.7	7.1
4. I could easily confirm my password with the pixel tolerance of 18.75×18.75	8.5
5. I could easily confirm my password with the pixel tolerance of 15.6×15.6	7.8
6. I could easily confirm my password with the pixel tolerance of 11.7×11.7	7.1

4.1 Experiment1

In this experiment, each participant should create and confirm 5 *CD-GPS* passwords. Totally 100 trails have been recorded during the experiment. The success rates for these two phases are shown in Table 3. The success rate in the phase of *Creation* means that participants created their passwords without restarting while the success rate in the phase of *Confirmation* means that participants confirmed their passwords without restarting and failed attempts. In Table 3, it is easily visible that the success rate is greatly dropped down with the decrease of the tolerance size (i.e., from 88% to 73% in the phase of *Creation*, from 85% to 65% in the phase of *Confirmation*). The main reason is that by reducing the tolerance sizes, participants were hard to accurately remember and reproduce their secrets. For instance, click errors could be significantly increased.

After the experiment, we gave relevant questions to participants and collected their feedback. The questions and average scores are presented in Table 4. The average scores are simply average values calculated by the feedback of all participants. In Table 4, it is easily visible that participants were satisfied with the pixel tolerance of 18.75×18.75 since the questions of No.1 and No.4 received a high average score of 8.5 respectively. With regard to the other two tolerance sizes of 15.6×15.6 and 11.7×11.7, the average scores were greatly decreased. During the experiment, most participants reflected that they considered the pixel tolerance of 15.6×15.6 was still fine but it was very difficult for them to use the CD-GPS scheme if the pixel tolerance was only 11.7×11.7. Therefore, the experimental results show that by increasing the value of N, users could suffer from the problem of drawing reproduction especially when the pixel tolerance is decreased to 11.7×11.7.

4.2 Experiment2

The *Experiment2* involved the remaining 40 participants and each participant should complete 5 *CD-GPS* passwords. Each trial contains two phases: *Creation* and *Confirmation*. In this experiment, we attempt to investigate the effect of tolerance on affecting the regular use of *CD-GPS*.

Table 5. The success rate of *CD-GPS* creation and confirmation in the *Experiment2* for the three tolerance sizes

Success Rate	18.75×18.75 pixels	15.6×15.6 pixels	11.7×11.7 pixels
Creation	188/200 (94.0%)	173/200 (86.5%)	156/200 (78.0%)
Confirmation	179/200 (89.5%)	166/200 (83.0%)	140/200 (70.0%)

Table 6. Several questions and relevant scores in the *Experiment2*

Questions	Score (average)
1. I could easily create and confirm a password in the tolerance of 18.75×18.75.	8.3
2. I could easily create and confirm a password in the tolerance of 15.6×15.6.	7.2
3. I could easily create and confirm a password in the tolerance of 11.7×11.7.	6.5
4. I prefer the pixel tolerance of 18.75×18.75 in CD-GPS scheme.	9.4
5. I prefer the pixel tolerance of 15.6×15.6 in CD-GPS scheme.	8.2
6. I prefer the pixel tolerance of 11.7×11.7 in CD-GPS scheme.	5.7

Up to 200 trials have been recorded during this experiment. The success rates for these two phases are shown in Table 5. It is easily visible that participants can achieve a high success rate (94% for the phase of *Creation* and 89.5% for the phase of *Confirmation*) with the pixel tolerance of 18.75×18.75. The same as the *Experiment1*, the success rate was quickly decreased when the pixel tolerance was reduced to 15.6×15.6 and especially to 11.7×11.7. For the pixel tolerance of 11.7×11.7, the succuss rates were only 78% and 70% for the phases of Creation and Confirmation respectively. Most patricians indicated that it was very difficult for them to click on the correct clickable squares with the pixel tolerance of 11.7×11.7 so that many click errors were occurred. These click errors cost most participants a lot of time in click-drawing their secrets again, which was unaffordable for a regular user.

After the experiment, we also gave relevant questions to all participants and collected their feedback. The questions and average scores are described in Table 6. For the No.1 question, the average score of 8.3 showed that most participants were satisfied with the pixel tolerance of 18.75×18.75. The scores of the No.2 question was 7.2 while the No.3 question only obtained a score of 6.5 which indicated that the pixel tolerance of 11.7×11.7 was not suitable in real deployment.

By comparing the scores in the No.4, No.5 and No.6 questions, we can find that most participants preferred the pixel tolerance of 18.75×18.75 (obtaining a very high score of 9.4) in that they can easily and accurately create and confirm their *CD-GPS* passwords. In addition, participants also gave a score of 8.2 for the No.5 question which indicated that they felt the pixel tolerance of 15.6×15.6 was still affordable in actual application. The low score of 5.7 for the No.6 question showed that the pixel tolerance of 11.7×11.7 was not appropriate in real settings.

On the whole, the experimental results indicate that the value of N should be smaller than 32. The usability of *CD-GPS* scheme will be greatly reduced when using a small pixel tolerance of around 11.7×11.7.

4.3 Discussion

Based on the two experiments, we can find that determining an appropriate value of N in the *CD-GPS* scheme is very crucial. To select an appropriate value, we should make a balance between usability and security.

Usability. This factor is very important according to our experimental results. In addition, some previous work [3,18] has shown that a lot of usability problems tended to translate directly into security problems. Therefore, in the scheme of *CD-GPS*, we should first ensure that users can use the scheme comfortably. For a comfortable and convenient graphical password scheme, users are more likely to create more secure passwords. The above two experiments showed that it was not comfortable to users if the value of N reached 32, so that a smaller value of N (i.e., smaller than 32) should be used in real settings.

This is another important factor for a graphical password. By safeguarding the usability, a more secure graphical password scheme is desirable. In terms of our previous work [10], the password space of the *CD-GPS* scheme can be greatly enlarged by increasing the value of N. Take $N = 16$ as an example, the password space of *CD-GPS* is 5.34×10^{18} with 6 clicks on the selected image, which is very larger than the text-based passwords with 8 characters over a 64-character alphabet (the password space is 2.81×10^{14}). The calculation of the password space can be referred to our previous work [10]. Therefore, we believe that the value N of 16, 20 and 24 can provide large enough password space in real settings.

Overall, the value of N should be smaller than 32 and can be selected according to the specific work environment. For example, if a very high security level is desirable (i.e., in a bank), we can select the value of N to 24. On the other hand, in a regular environment, we can choose the value of N to 16 or 20.

5 Concluding Remarks

In this paper, we mainly attempt to investigate the effect of tolerance on creating and confirming the click-draw based graphical passwords. We conducted two experiments (*Experiment1* and *Experiment2*) with totally 60 participants. From the experimental results, we find that accurate memory and reproduction for the scheme of *CD-GPS* can be significantly reduced when the pixel tolerance is greatly decreased (e.g., 11.7×11.7 pixels). With this small tolerance size, users cannot accurately click on the right clickable squares. By balancing both usability and security, we find that an appropriate value of N should be smaller than 32. The value of 16, 20 and 24 can all provide suitable properties in the aspect of both usability and security.

Our work is mainly conducted by means of the click-draw based graphical password scheme. Future work could include conducting another user study with larger and more varied participants to validate our results, and exploring more specific values of N. In addition, future work could also include investigating the effect of tolerance on other click-related graphical password schemes and identifying the relationships.

Acknowledgments. We thank all the participants for their hard work in our user study and all anonymous reviewers for their valuable comments.

References

1. Brown, A.S., Bracken, E., Zoccoli, S., Douglas, K.: Generating and Remembering Passwords. Applied Cognitive Psychology 18, 641–651 (2004)
2. Yan, J., Blackwell, A., Anderson, R., Grant, A.: Password Memorability and Security: Empirical Results. IEEE Security and Privacy Magazine 2(5), 25–31 (2004)
3. Klein, D.: Foiling the Cracker; A Survey of, and Improvements to Unix Password Security. In: Proceedings of the USENIX Security Workshop, pp. 83–86 (1990)
4. Shepard, R.N.: Recognition Memory for Words, Sentences, and Pictures. Journal of Verbal Learning and Verbal Behavior 6, 156–163 (1967)
5. De Angeli, A., Coventry, L., Johnson, G., Renaud, K.: Is a Picutre Really Worth a Thousand Words? Reflecting on the Usability of Graphcial Authentication Systems. International Journal of Human Computer Studeies 63(2), 128–152 (2005)
6. Nelson, D.L., Reed, U.S., Walling, J.R.: Picture Superiority Effect. Journal of Experimental Psychology: Human Learning and Memory 3, 485–497 (1977)
7. Gołofit, K.: Click Passwords Under Investigation. In: Biskup, J., López, J. (eds.) ESORICS 2007. LNCS, vol. 4734, pp. 343–358. Springer, Heidelberg (2007)
8. Paivio, A., Rogers, T.B., Smythe, P.C.: Why are Pictures Easier to Recall than Words? Psychonomic Science 11(4), 137–138 (1976)
9. Blonder, G.E.: Graphical Passwords. United States Paten 5559961 (1996)
10. Meng, Y.: Designing Click-Draw based Graphical Password Scheme for Better Authentication. In: Proceedings of 7th IEEE International Conference on Networking, Architecture, and Storage (NAS 2012), pp. 39–48 (2012)
11. Wiedenbeck, S., Waters, J., Birget, J.C., Brodskiy, A., Memon, N.: PassPoints: Design and Longitudinal Evaluation of a Graphical Password System. International Journal of Human-Computer Studies 63, 102–127 (2005)
12. Davis, D., Monrose, F., Reiter, M.K.: On User Choice in Graphical Password Schemes. In: Proceedings of USENIX Security Symposium, pp. 151–164. USENIX Association, Berkeley (2004)
13. Passfaces (accessed by May 20, 2012), `http://www.realuser.com/`
14. Nali, D., Thorpe, J.: Analyzing User Choice in Graphical Passwords. Technical Report. Careton University (2004)
15. Jali, M., Furnell, S., Dowland, P.: Quantifying the Effect of Graphical Password Guidelines for Better Security. In: Camenisch, J., Fischer-Hübner, S., Murayama, Y., Portmann, A., Rieder, C. (eds.) SEC 2011. IFIP AICT, vol. 354, pp. 80–91. Springer, Heidelberg (2011)
16. Jermyn, I., Mayer, A., Monrose, F., Reiter, M.K., Rubin, A.D.: The Design and Analysis of Graphical Passwords. In: Proceedings of USENIX Security Symposium, pp. 1–14. USENIX Association, Berkeley (1999)
17. Thorpe, J., van Oorschot, P.C.: Human-Seeded Attacks and Exploiting Hot-Spots in Graphical Passwords. In: Proceedings of 16th USENIX Security Symposium, pp. 1–16. USENIX Association, Berkeley (2007)
18. Wiedenbeck, S., Birget, J.C., Brodskiy, A., Memon, N.: Authentication using Graphical Passwords: Effects of Tolerance and Image Choice. In: Proceedings of Symposium on Usability Privacy and Security (SOUPS), pp. 1–12 (2005)
19. Dirik, A.E., Memon, N., Birget, J.C.: Modelling User Choice in the Passpoints Graphical Password Scheme. In: Proceedings of Symposium on Usability Privacy and Security (SOUPS), pp. 20–28 (2007)
20. Dunphy, P., Yan, J.: Do Background Images Improve "Draw A Secret" Graphical Passwords? In: Proceedings of ACM Conference on Computer and Communiation Security (CCS), pp. 36–47 (2007)