Do Graphical Authentication Systems Solve the Password Memorability Problem?

Soumyadeb Chowdhury, Ron Poet, and Lewis Mackenzie

School of Computing Science, University of Glasgow
soumc@dcs.gla.ac.uk,
{Ron.Poet,Lewis.Mackenzie}@glasgow.ac.uk

Abstract. Passwords are the most common form of authentication. The password memorability problem is magnified with increasing number of systems users have to access. Graphical authentication systems (GASs) have received significant attention as one potential alternative to alphanumeric passwords to provide more usable authentication. In this paper we review all the existing work which had explored the memorability of multiple graphical passwords. The review reveals that human memory capabilities should not be overestimated and the password memorability problem remains unsolved, even when graphical passwords are employed. Hence we propose a novel graphical authentication system with certain new security features which could solve the problem. This paper will be of interest to Human Computer Interaction-Security researchers investigating approaches to usable and secure authentication techniques.

Keywords: graphical authentication, memorability, password problem.

1 Introduction

In the current practice, alphanumeric passwords are the most widely used mechanism to authenticate users. According to Adams and Sasse, as the number of passwords per user increases, the rate of forgetting them also increases [1]. In order to cope with multiple passwords, users tend to adopt unsafe strategies, which include writing them down, reusing the same passwords and sharing them with others [1, 2].

An alternative approach that has received significant attention is that of graphical authentication [3-11], which uses images to form passwords. The motivating idea is that humans can supposedly remember images better than alphanumeric text [12], so use of the former may be a way of devising more memorable passwords. In this context, GASs can be categorized as follows.

- *Cognometrics*: During registration users can either choose their password images from a collection presented by the system [4,5 and 6] or the passwords are issued by the system [7] to the users. The users can also provide their own images to be used as password too [9 and 10]. Each password is a combination of certain number of target images. During authentication, users must recognize each target image among a collection of decoys (figure 1). The work reported in [12] suggests that human beings have exceptional ability to recognize images that they have previously seen, even if the image has been viewed for a very short period of time.

T. Tryfonas and I. Askoxylakis (Eds.): HAS 2014, LNCS 8533, pp. 138–148, 2014.

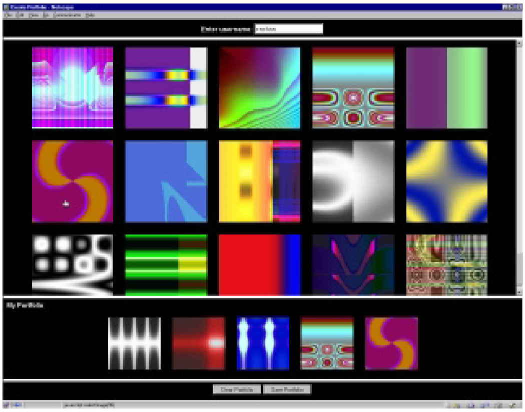


Fig. 1. Authentication screen for Dejavu system [6]

- *Locimetrics:* In these systems, specific points in an image that is either selected by the user or issued by the system form the password (figure 2). An example of such a system is Passpoints [3]. These systems are often referred to as cued recall based systems. The cognitive studies in the past have explained that items in human memory may be available, but not accessible for retrieval at a later time [13]. Ideally, a cue should be helpful only to the legitimate users and not to intruders to break into the system. In graphical authentication systems using this approach, the users don’t have to remember the image, but remember specific points in the image that has been selected by them as their password.

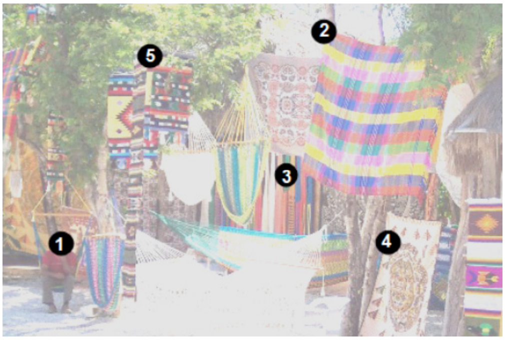


Fig. 2. Authentication screen for Click-based password [3]

- Drawmetrics: In these systems, users must draw an image during the password creation stage and they have to reproduce that same image during authentication (figure 3). This is same as pure recall, where the users are asked to retrieve their password from memory, which they have used or chosen in the past without any cues. Unaided recall is considered to be the least accurate type of memories because the accuracy would decay after a considerable amount of time, if the password is not used frequently [14]. In case of graphical passwords, the users have to reproduce their passwords without any cues. It is a difficult memory task and the users may sometimes use the interface as the cues, even if it is not intended as such. Examples of systems include DAS [15] and Pass go [11]

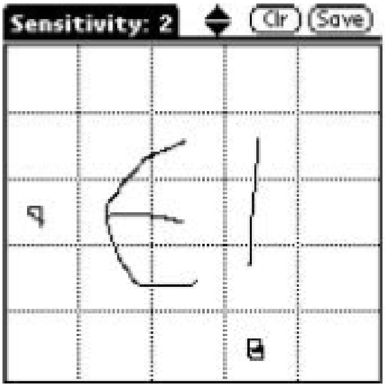


Fig. 3. Authentication screen for DAS [15]

Given the need for more usable authentication and the existing interest of the research community in graphical passwords as a potential solution, an important limitation of the existing work is: most studies in the field of GASs have focused on the use of single password. We believe that people will need to remember and use graphical passwords in the same way as they currently use alphanumeric passwords. In this paper we will review all the studies that have explored the memorability of multiple graphical passwords. This review to our knowledge is first of its kind which will help to understand, whether graphical passwords in their current form had been able to solve the issue of remembering multiple passwords.

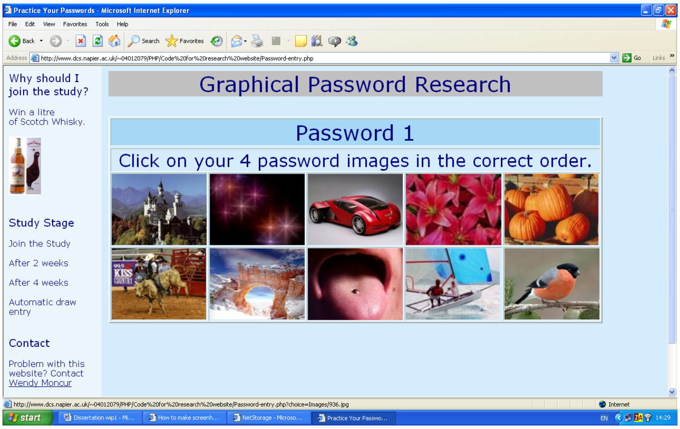
2 Survey of Multiple Graphical Password Studies

In the last fifteen years, only four studies [3, 4, 7 and 8] in the field of GASs had explored the memorability of multiple graphical passwords. We will review them and draw our inferences for each study based upon the results reported by the respective authors. This review will not discuss the conclusions and claims made by the authors.

2.1 Pictures at the ATM: Exploring the Usability of Multiple Graphical Passwords (CHI 2007) –Moncur and Leplatre [8]

The first study with multiple graphical passwords was conducted by Moncur & LePlâtre [8]. They compared the memorability of multiple graphical passwords to multiple PINs. Photographic images of food, music, sports, flowers etc. were used as the visual cue in case of graphical passwords.

System. Each user was assigned five numerical or graphical passwords, depending upon their respective group. The assignment of password to the user was done on a random basis. In case of graphical password, each of the passwords comprised of four colourful and meaningful photographic image. During authentication, a challenge set was displayed to the participants containing 10 images (figure 4). The participants had to select four target images in the correct order, among a collection of 6 decoy images. In case of PINs the 0-9 numerals were displayed on the screen and the participants had to click on the numbers in correct order that formed the digits of their PIN.



Practice Your Passwords - Microsoft Internet Explorer

Address http://www.dcs.napier.ac.uk/~04012079/PHP/Code%20for%20research%20website/Password-entry.php

Graphical Password Research

Why should I join the study?

Win a litre of Scotch Whisky.

Password 1

Click on your 4 password images in the correct order.

Study Stage

Join the Study

After 2 weeks

After 4 weeks

Automatic draw entry

Contact

Problem with this website? Contact Wendy Moncur

http://www.dcs.napier.ac.uk/~04012079/PHP/Code%20for%20research%20website/Password-entry.php?choice=Images/936.jpg

Fig. 4. Challenge set reported in Moncur and Leplatre [8]

User Study. The study examined the memorability of five system-issued passwords with 172 university students, who were assigned randomly to one of the five conditions given below:

- Condition 1: four digit pin;
- Condition 2: graphical passwords;
- Condition 3: graphical password + signature colour background;
- Condition 4: graphical password + mnemonic strategy;
- Condition 5: graphical password + signature colour background + mnemonic strategy.

Three memorability rests (RT1, RT2, and RT3) were conducted, with a gap of two weeks between each one of them. The dropout rate in the user study was 64.91%, which made it difficult to analyze the results.

Results. According to the statistics presented in the paper [13, figure 5], the mean login success percentages are discussed below:

- Condition 1: RT1 was 15%, dropped to almost 5% after 2 weeks and remained almost the same for RT3;
- Condition 2: RT1 was 55%, dropped to 10% after 2 weeks and remained almost the same for RT3;

- Condition 3: RT1 was close to 70%, dropped to almost 10% after two weeks and was slightly more than 10% in RT3;
- Condition 4: RT1 was almost 90%, dropped to 15% after two weeks and was almost 20% in RT3;
- Condition 5: 80% in RT1, dropped to 10% in RT2 and close to 20% in RT3.

However, the mean values reported in the paper are high because the retention just after the passwords were supplied (training session) was taken in to account.

Inferences. The results obtained from the user study clearly demonstrated that the memorability of multiple graphical passwords drops off over time. The mean login success for RT2 and RT3 in case of condition 4 and 5 demonstrated that employing a mnemonic strategy as well as signature background color did not improve the password memorability. The study also revealed that multiple graphical passwords are difficult to remember, when they are issued by the system.

2.2 Multiple Password Interference in Text Passwords and Click-Based Graphical Passwords (ACM CCS 2009) – Chiasson et al. [3]

Chiasson et al. [3] conducted a lab study with 65 university students to compare the memorability of multiple text passwords (MTP) and multiple click-based passwords (MCP). The study investigated the phenomenon of password interference, i.e. whether remembering a password for one system might affect the user's memory of a password for another system.

System. The participants were randomly divided into two groups:

- Members of the first group were required to remember six text passwords created by them during the registration stage;
- Members of the second group created six click-based passwords. Each password comprised of five click points on an image. The users were provided with six distinct images to create each of their passwords.

User Study. The lab study was divided into two sessions:

- Session1: All the participants registered with six passwords depending upon their group. After completing the registration, they were asked to login, once they have performed a distraction task. The distraction tasks were conducted to clear the textual and visual working memory. The login success for each of the participants in the session was obtained and reported as recall1.
- Session 2: The second session was conducted two weeks after the first session and 26 participants took part in it. There were no practice sessions between the two sessions. The login success was collected and reported as recall 2. The authors did not report the number of participants in each group, who took part in session 2.

Results. The mean login success percentages for each session reported in the paper are discussed below:

- Recall 1: The mean login success percentage was 95% for MCP and 68% for MTP during the training session, when the participants logged in successfully in the first attempt. The mean login success for multiple attempts was 88% in case of MTP and 99% in the case of MCP. However, these are the mean success percentages in the training session, just after the passwords were created. Hence the results do not reveal much in context to the long term memorability of multiple graphical passwords.
- Recall 2: The mean login success percentage was 38% for MCP and 30% for MTP, when the participants logged in successfully in the first attempt. The mean login success for multiple attempts was 70% for MTP and 57% for MCP and this was found to be statistically insignificant.

Inferences. The results of recall 1 reveal that the short term memory for MCP is significantly better than MTP. However, the results for recall 2 revealed that users find it difficult to remember multiple click-based passwords over a longer term. Since the participation rate in the second session was low, the results may not be an actual reflection of the phenomenon of memory interference.

2.3 A Comprehensive Study of Frequency, Interference and Training of Multiple Graphical Passwords (CHI 2009) – Everitt et al. [7]

Everitt et al. [7] conducted a user study with 100 university students over a period of five weeks to examine the memorability of multiple facial passwords.

System. Each participant was assigned *x* number of passwords by the system. Each password comprised of five faces. During authentication, participants had to select the correct face from a sequence of 3x3 grids of decoy faces, at each step of a five step login process (figure 5).

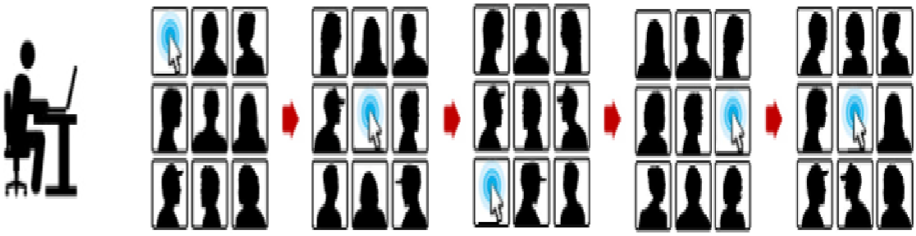


Fig. 5. Authentication screen reported in [7]

User Study. The user study was conducted using a between subject design where each participant was randomly assigned to one of the five conditions as given below:

- Condition 1: Participants used one facial graphical password (5 faces) once a week for a period of 5 weeks.
- Condition 2: Participants used one facial graphical password (5 faces) thrice a week for a period of 5 weeks.
- Condition 3: Participants used two facial graphical passwords (10 faces). One facial password was used thrice a week for a period of 5 weeks and the other was used once a week for a period of 5 weeks.
- Condition 4: Participants used four facial graphical passwords (20 faces). Each facial password was used once a week for a period of 5 weeks. Hence all the 4 different facial passwords were used at least once during the week.
- Condition 5: Participants used four facial graphical passwords (20 faces). In this condition only one password was used 4 times in a week. In the second week a different password was used. Thus distinct passwords were used during each week.

Results

- It was found that the participants using one facial password per week (condition 1) required more login attempts compared to the participants who used one facial password per day (condition 2). Hence the frequency of password usage would significantly affect the ease to login.
- The result indicated that participants accessing four passwords per week (condition 4), were ten times more likely to have an authentication failure compared to participants using a single password per week (condition 1). Thus interference occurring from the use of multiple graphical facial passwords would significantly affect user's memorability.
- The results also demonstrated that the participants who were trained using multiple facial passwords each week during a month (condition 4) were four times more likely to have an authentication failure, than the participants who were trained using one graphical password per week (condition 5). Hence the password training pattern would significantly affect the ease of access. The failure rate in the case of condition 4 was 15.23%.
- In context to long term recall, it was found that the participants using one graphical password throughout a month could remember their password correctly after four months. But, the participants using multiple passwords had problems remembering their passwords due to interference (failure rate 14.29%). Thus, the long term recall is significantly affected as the number of facial passwords increase.

Inferences. The performance of multiple facial passwords is better compared to the results reported in [3] and [8]. However, in [7] participants who were assigned to the condition 4 used each of their four passwords, at least once in a week. This may have helped to retrieve the passwords in subsequent use. The long term recall of the facial passwords also seems promising compared to the other password types. But [] does not report the number of participants in each condition, who took part in the long term recall study. It is also not known whether the participants recorded their passwords, or kept a copy of the same and used them during the long term recall study. This might have helped them to recall the multiple facial passwords after four months. Overall, the study demonstrates that memory interference and frequency of use significantly affected the memorability of multiple facial passwords.

2.4 A Comprehensive Study of the Usability of Multiple Graphical Passwords (INTERACT 2013) – Chowdhury et al. [4]

Chowdhury et al. [4] presented a study with 100 university students, who used multiple image passwords over a period of eight weeks. The study compared the usability of four distinct image types: Mikon, doodle, art and everyday object, when used as graphical passwords.

Study. The study used an independent measure style of experimental design with four conditions (equal number of participants in each condition) namely Mikon, doodle art and everyday objects. Each participant was randomly assigned to only one of the conditions. Each participant in each condition had to choose four passwords from four distinct image collections presented by the system. Each password comprised of four target images, chosen by the participants. Authentication was a four step process. At each step, a challenge set consisting of 15 decoy images and 1 target image was displayed as a 4X4 grid. The participants had to recognize and select the target image at each step. Upon completing the registration, participants had to login with each of their passwords over a period of eight weeks. The frequency of login differed for each week (high frequency-low frequency).

Results

- The findings showed that the memorability of the graphical passwords is significantly affected by the type of images used. In this context, the results revealed that the mean login success percentage over the period of eight weeks is highest for objects (77.31%), closely followed by Mikons (74.17%), then doodles (67.04%) and lowest for the art images (54.90%). The results of the study complement cognitive literature on the picture superiority effect [12], visual search process [16] and nameability of visually complex images [17].
- The results demonstrated that the mean login success percentage for each of the image type drops from week 2 to week 8, as the frequency of usage of the passwords decreases. The mean login success percentage dropped off by 11.44 % in case of Mikon, 12.55 % in case of doodle, 7.74 % in case of art and 14% in case of object, from week 2 to week 8.

Inferences. The performance of the multiple image passwords reported by Chowdhury et al. [4] is better than the results reported by [3] and [8], but inferior compared to [7]. The superior results can be attributed to the fact that all the participants used each of their passwords every week. This may have helped the participants to retain them in the memory through an elaborative encoding. The results reported in the paper clearly demonstrated that multiple image passwords are difficult to remember, even when they are created by the participants and used regularly. The study reported by [4], did not consider the scenario, where multiple image passwords are not used for a considerable period of time. This could have further degraded the performance of the participants, when they are required to remember multiple graphical passwords, without any practice.

3 Discussion

The reviews of the existing studies that have explored the cognitive demands of using multiple graphical passwords clearly demonstrated that users find it difficult to remember the passwords. Hence the memorability problem in case of authentication still exists, and GASs in their current state-of-the-art cannot be considered as a viable alternative to the traditional alphanumeric passwords. There is a need to develop authentication systems that would ease the burden of remembering many passwords. This would help to prevent the use of unsafe coping strategies to store or disclose the passwords, which inherently compromises the security of the system.

In the context of existing interest in image passwords, we propose a hint-based authentication system (PHAS), as a potential solution to address the problem of remembering multiple image passwords. In this system, the users have to choose four images and create hints for each one of them to form a password. During authentication, they have to recognize only the target images, which are displayed with their corresponding hints, among a collection of 15 decoy images, in a four step process. Our system would not rely just on recognition memory, but it would have an additional component, a 'hint', which will act as a cue to recognize the password. The hints can be in any language (we suggest a maximum of 5-6 words), but should be typed in English characters.

In the proposed approach, users give a hint for each target image and can use any strategy to do so. They do not need to create a story or use a mnemonic strategy, nor do they need to remember or reproduce the hints at any stage. This is because all hints are stored in the system and displayed with the challenge set to enhance memorability. We believe that the hints will act as cues while recognizing the images in future, which should enhance memorability [18].

In context to security, i.e. guessability of images using hints, we believe that an image can be guessed easily, if the hint given by a user denotatively describes the elements in it. But if the hint is connotative, where the user relates it to something personal (such as an episode in one's life), a sign or state (how it makes them feel), a context (an idea or event that only has relevance to them), then it might be very difficult for an attacker to guess, without being aware of the relation between the hint and the image [19 and 20]. This also creates a new avenue of research in the field of GASs, i.e. advice that should be given to the users, while they create the hints and select the target images.

We also propose the following features, which could enhance the security of the proposed PHAS:

- PHAS could also offer secure authentication, if an additional lock out policy is implemented not only on a definite number of failed login attempts, but a threshold value of login time. For example, once a user has used the system for certain number of times, then a timer could be set for all the subsequent login sessions. If the user is unable to complete the login session within the set timer, then this will be recorded. After a definite number of failed attempts due to timer expiration, the account could be locked. But, different aspects such as how to customize the timer, the number of attempts before the account is locked have to be considered before this feature could be implemented in practice. The usability of the proposed security component would need to be examined.
- The research on challenge sets in the field of GASs is sparse. Hence we also propose a novel challenge set configuration, which could increase the security of PHAS. Let a user U select four images and give one hint for each one of them in PHAS (x_1- x_4). The system chooses 15 decoy images for each of the target images (x_1- x_4), generating four challenge sets (T_1- T_4). The system would now choose four random images with their corresponding hints (y_1- y_4), which do not belong to user U. Four false challenge sets (F_1- F_4) are generated. The system displays *m* number of true sets selected from (T_1- T_4) and *n* number of false sets selected from (F_1- F_4). The value of *m* and *n* can either vary for each login attempt or remain constant for all login attempts for the user. Each challenge screen will have 16 images, a hint and a button named "Ignore". This Ignore button could be used by the legitimate user, when a false challenge set is displayed.

If cognitive attacks are carried out to break a PHAS password, we believe that the false challenge sets would make it difficult for an attacker to follow a lead for breaking into the system and the lock out policy based on the login time will put further pressure on making it hard to succeed.

References

1. Adams, A., Sasse, M.A.: Users are not the enemy. Communications of the ACM, 40–46 (CACM December 1999)
2. Florencio, D., Herley, C.: A large-scale study of web password habits. In: Proceedings of International Conference on World Wide Web (WWW 2007), pp. 657–666 (2007)
3. Chiasson, S., Forget, A., Stobert, E., Oorschot, P.C.: Van, and Biddle, R. Multiple Password Interference in Text and Click-Based Graphical Passwords. In: Proc. of CCS, pp. 500–511 (2009)
4. Chowdhury, S., Poet, R., Mackenzie, L.: A comprehensive study of the usability of multiple graphical passwords. In: Kotzé, P., Marsden, G., Lindgaard, G., Wesson, J., Winckler, M. (eds.) INTERACT 2013, Part III. LNCS, vol. 8119, pp. 424–441. Springer, Heidelberg (2013)

5. Davis, D., Monrose, F., Reiter, M.: On user choice in graphical password schemes. In: Proc. of the 13th conference on USENIX Security Symposium, vol. 13, USENIX Association Berkeley, CA (2004)
6. Dhamija, R., Perrig, A.: Deja vu: A user study using images for authentication. In: Proc. USENIX Security Symposium, pp. 45–48 (2000)
7. Everitt, K.M., Bragin, T., Fogarty, J., Kohno, T.: A comprehensive study of frequency, interference, and training of multiple graphical passwords. In: Proc. of CHI, pp. 889–898. ACM, New York (2009)
8. Moncur, W., LePlâtre, G.: Pictures at the ATM - Exploring the usability of multiple graphical passwords. In: Proc. of CHI, pp. 887–894 (2007)
9. Renaud, K.: Web authentication using Mikon images. In: World Congress on Privacy, Security, Trust and the Management of E-Business, pp. 1-10
10. Renaud, K.: On user involvement in production of images used in visual authentication. Journal of Visual Languages and Computing 92, 1–15 (2009)
11. Tao, H.: Pass-Go, a new graphical password scheme. M.S. thesis, School of Information Technology and Engineering, University of Ottawa (2006)
12. Madigan, S.: Picture Memory. In: Yuille, J. (ed.) Imagery, Memory, and Cognition: Essays in Honor of Allan Paivio. Lawrence Erlbaum Associates, Hillsdale (1983)
13. Tulving, E., Pearlstone, Z.: Availaibility Versus Accessibility of Information in Memory for Words. Journal of Verbal Learning and Verbal Behaviour 5, 381–391 (1966)
14. Baddeley, A.: Human Memory:Theory and Practice. Psychology Press, Hove (1997)
15. Jermyn, I., Mayer, A., Monrose, F., Reiter, M., Rubin, A.: The Design and Analysis of Graphical Passwords. In: Proceedings of 8th USENIX Security Symposium (1999)
16. Wolfe, M.: Guided Search 2.0 A Revised Model of Visual Search. Psychonomic Bulletin & Review 1(2), 202–238 (1994)
17. Szekely, A., Bates, E.: Objective Visual Complexity as a Variable in Picture Naming. In: CRL Newsletter Center for Research in Language, University of California, pp. 3–33 (2000)
18. Mantyla, T.: Optimising cue effectiveness. Journal of Experimental Psychology: Learning Memory and Cognition 12, 66–71 (1986)
19. Mathur, P.N.: Barriers to effective visual communication, 3rd edn. Media Asia (1978)
20. Sturken, M., Cartwright, L.: Practices of Looking: An introduction to visual culture. Oxford Press (2012)