Quantifying the Effect of Graphical Password Guidelines for Better Security

Mohd Jali[1,3], Steven Furnell[1,2], and Paul Dowland[1]

[1] Centre for Security, Communications and Network Research (CSCAN),
Room A304, Portland Square, University of Plymouth, Plymouth PL4 8AA, UK
[2] School of Computer & Security Science, Edith Cowan University,
Perth, Western Australia
[3] Faculty of Science & Technology, Universiti Sains Islam Malaysia,
Nilai, 71800, Negeri Sembilan, Malaysia
zalisham@usim.edu.my

Abstract. Authentication using images or graphical passwords is one of the possible alternatives for traditional authentication based upon passwords. This study aims to investigate the practicality of giving guidelines or advice to users before they start choosing their image passwords, the effectiveness of using a smaller tolerance (clickable areas) and the optimum combination of click and image passwords. An alternative graphical prototype known as the Enhanced Graphical Authentication Scheme (EGAS) was developed in order to achieve these aims which implemented two different types of data collection (internal and external). From the findings, both internal and external groups indicated that the implementation of guidelines alone cannot guarantee the security of image passwords created by participants; but, in combination with other usability measurements this study has shown positive outcomes.

Keywords: Graphical passwords, Authentication, Usability, Security, HCI.

1 Motivation

Using images to authenticate users is one possible alternative for password-based authentication. Previous work has divided image-based authentication into three categories; namely ‘click-based’, ‘choice-based’ and ‘draw-based’. The click-based approach refers to the action of clicking on the provided/chosen image(s) (i.e. selecting an element of the image), choice-based refers to the action of selecting a series of images (i.e. choosing images from a selection on screen) and draw-based refers to the action of drawing/sketching in order to be authenticated.

Regardless of the methodologies, previous studies have reported positive results, especially in the aspects of recall and memorability (i.e. participants were able to remember their secrets (i.e. image passwords) accurately after long periods of time) and usability (i.e. using images is user friendly) [1], [2] and [3]. Conversely, studies have also reported the disadvantages. Davis et al. [7] and Tullis and Tedesco [8] found that users chosen secrets were influenced by gender. Chiasson et al. [4] reported that the concept of clicking on images (e.g. Passpoint [5]) was not secure as

J. Camenisch et al. (Eds.): SEC 2011, IFIP AICT 354, pp. 80–91, 2011.

users tended to create hotspots (i.e. focussing upon one area in an image) and generating similar patterns (e.g. a straight line from top-bottom or left-right). Oorchot et al. [6] claimed that it was possible to crack users' secrets regardless of the background image, with the study by Everitt et al. [9] reporting that having multiple secrets resulted in more errors when compared with password-based authentication.

With respect to security, the main problem with the click-based method can be referred to as 'hotspot' while the problem with the choice-based method can be referred to as 'hot-image'. The problem of hot-image happens when a similar image is selected by many users. This problem could also be associated when users choose similar categories/themes or through gender preferences (e.g. males choose cars and females choose flowers). The hotspot problem could occur in two conditions. Firstly, the user clicks within the same or similar point on the given image or clicks on the same point or area when two or more images are given. Secondly the user produces predictable shapes such as straight lines and clicks on obvious/predictable objects within the image. Studies related to security in graphical passwords can be found in [10], [11], [12] and [13].

In an attempt to address or reduce the aforementioned problems and at the same time maintain users' memorability, many studies have been published with regards to the effect of using various types of images. Examples include using images of cartoon characters [3], images of geometric shapes [14] and using images that were later transformed into unclear or distorted forms during login [15] and [16]. With respect to the click-based method, a technique known as persuasion has been proposed [17] where the software recommends to the user possible 'safe' areas in which to create their secrets.

As far as the authors are aware, no study was found to have investigated or introduce user guidelines as part of the enrolment process. Therefore, the authors introduced a set of guidelines for graphical authentication, referred to as the Graphical Password Guidelines (GPG) which was presented to the user before they began choosing their secrets.

The authors also conjectured that GPGs on their own (Table 1) would not be a universal solution due to inherent human behaviour (i.e. certain users, although aware of the guidelines, sometimes violate the rules). To address this, restrictions were applied during registration. Two restrictions implemented in this study are as follows:

1. Users were only permitted to choose one image per category.
2. Users were not permitted to click on the same areas within an image. If they choose more images, they were also not permitted to click on the same area within the images.

The above restrictions together with the GPG were integrated into a software prototype. The software applied these restrictions by displaying warning messages if the software identified the user attempting to breach the rules.

The study was conducted in order to examine the impact on usability as well as user perception towards the introduction of the GPGs and image selection restrictions. Each participant had two types of secret; namely click-secrets (based upon the action of clicking on an image) and image-secrets (based upon the action of choosing a sequence of images). In addition to this, the study aimed to find ideal (usable and secure) combinations of click and image-secrets. A third investigation was undertaken to evaluate the impact of reducing the tolerance of the click positions. Tolerance can be explained as the extent of the area surrounding the users' secret clicks which are still accepted as legitimate. Prior research has indicated that participants were quite good when entering their secrets, both during registration and login [19]. Thus, the authors believed that using a smaller tolerance is possible and for this reason, users' performance when using smaller tolerance was investigated.

Table 1. Graphical password guidelines

Task	Guideline	Explanation
Choosing images	Choose different themes and images	Users perceives image differently and previous studies have found gender bias in user image selections [7], [8] and [18]. As a result, the user is advised to choose different images, the image itself should not related to gender and more importantly, they are advised to choose images that they think could offer them memorable areas for placing their secret clicks.
	Try to avoid imagery that could be associated with your gender	
	Please choose images that offer you various memorable areas for placing your secret clicks	
Clicking on images	Try not to click within the same or adjacent areas	Oorchot et al., [6] showed that some users' secret were predictable. To reduce this, the user is advised to create their secret randomly. Specifically, they are not permitted to click on or within the same area (also applied to many images), advised not to create an easy to guess pattern (e.g. straight line) and encouraged not to click on obvious objects (e.g. edge, centre of each image).
	Try to click on various areas, not only on an obvious object	
	Please avoid predictable patterns (e.g. straight line, edges, central of images, etc)	

The next section of this paper highlights the methodology, followed by the results, discussion and conclusions.

2 Methodology

A graphical software prototype known as the Enhanced Graphical Authentication System (EGAS) was developed using Microsoft Visual Basic 2008. EGAS is an alternative graphical authentication employing a combination of both click and choice-based methods. In the EGAS software prototype, users are given the freedom to choose their preferred number of clicks (secret clicks), with the software assigning the number of images (secret images) they need to choose. Table 2 shows the combination of secret clicks and secret images.

Table 2. Click and Image details used in the software prototype

Secret click chosen	Secret image assigned	Image size/Tolerance
1	6	200x200 / 7x7
2	5	200x200 / 7x7
3	4	200x200 / 7x7
4	3	200x200 / 7x7
5	2	200x200 / 7x7

Two types of data collection were implemented; named as ‘internal’ and ‘external’. Internal means the experimenter observed participants during trial (similar with the one to one usability testing) and they had to complete current task before proceeding to the next (controlled by the software prototype). Participants within the external group had to install the software prototype into their personal computer and use it for three weeks, with all of their activities recorded into a database (no means of control was enforced by the software prototype).

Participants for both groups (internal and external) had to register their details (username and secrets) in the software prototype, were then required to log into the software using their chosen secrets and finally provide feedback via a questionnaire. All tasks were done within the software prototype.

During the secret registration (enrolment), the GPG were first displayed to them (by which they had to acknowledge the GPG) before they chose their secret. During image selection, participants were able to choose images from 10 different themes (buildings, abstract, food, animals, flowers, view, people, sport, transport and fruits), with each of them consisting of 9 distinct images (arranged in 3x3 grids).

Participants within the internal group were asked to login three times, while the external group needed to login on four different days in week 1, two different days in week 2 and finally login once in week 3. This aimed to examine their familiarity and competency (e.g. login time, clicking accuracy, total attempts).

The trial was conducted over two months with the participants of the internal group recruited via an open call for volunteers within the authors’ university. Participants of the external group were colleagues/friends of the author (external to the University) and invited via email, chat messengers and text messages.

The data were interpreted and reported into five main categories; namely number of attempt, timing, pattern, accuracy and finally users’ feedback. The number of attempt looks upon participants’ failure and success trials during both registration and login tasks while timing reports the time needed for these tasks. Pattern discusses the occurrences of ‘hotspot’ and ‘hot-image’, with accuracy mainly focuses upon the participants’ ability to click on their secret clicks and finally users’ feedback reports participants’ perception on the questionnaire.

3 Results and Discussion

In total, there were 48 participants participated. Table 3 gives information for both groups highlighting the gender split and number of participants who had previously participated in graphical password studies [18].

Table 3. Participants' information

Demographic	Internal group	External group
Male participant	12	10
Female participant	18	8
Experienced using GA	10	2

3.1 Number of Attempt

3.1.1 Internal Group

Members of this group undertook 356 of authentication attempts. Of these, 94 logins were successful and 47 failed, 156 failed during registration and 66 were able to register successfully (note that software recorded two trials for each participant if they managed to register).

Participants who changed their click decided to choose the lowest click. Of the total seven participants who initially chosen three clicks on each image, five of them went to one click, while the remaining chosen two clicks. Moreover, all five participants who initially chosen two clicks and one participant who initially chosen four clicks also decided to choose one click.

During login, all participants within all click groups performed well where they managed to login, these results improved with experience. Only ten participants recorded a complete failure to login. There were six occurrences of failed attempts for login one, four occurrences for login two and only three occurrences for login three. The ability of participants to login with fewer failed attempts suggests participants performance improved with experience.

3.1.2 External Group

With eighteen participants within this group, the software recorded a total of 283 login attempts in week one, 61 trials for week two and finally 30 for week three. Of these, there were 92 successful logins for week one, 51 for week two and 20 for week three (note that there were participants who logged into the software more than was asked for).

Investigation of successful usernames who continued with the login tasks found mixed results. It was found only 12 participants followed the login interval task, with the remaining 6 participants using the software occasionally. For those who logged into the software according to specified tasks, 9 participants had chosen one click, 1 participant chose two clicks and 2 participants chose five clicks. Analysis has also found that 6 participants (who did not complete the login tasks) infrequently login during week one, with three of them logged twice for week two and finally all of them logged into the software in the third week. Five of them had chosen one click, while the remaining participant went for five clicks.

Only eight of the external group participants managed to register by using their first username. Of the remaining 10 participants who used a second username, six of them changed their secret click to the least click. Unless otherwise stated, most of the analysis for this group was based upon 18 participants who completed the specified tasks.

3.2 Timing

3.2.1 Internal Group

The time for participants to register and then log into the software prototype was recorded with the time during registration calculated from the point when they pressed the ‘register account’ button until to the result for registration is displayed. The time for login was calculated from when the participant started to enter their username until the last click for their secret images.

Table 4 shows participants’ time (average, shortest, longest and standard deviation) during registration and three logins, in minutes, (m) and seconds, (s).

Table 4. Timing for the internal group

Click	Participant	Time	Registration	Login One	Login Two	Login Three
1	18	Average	5m 23s	24	20	18
		Shortest	1m 43s	15	11	9
		Longest	21m 58s	42	42	4
		SD	4m 43s	8	8	6
2	5	Average	10m 29s	40	35	27
		Shortest	2m 23s	28	25	18
		Longest	23m 33s	71	69	40
		SD	7m 56s	17	18	8
3	3	Average	9m 56s	39	33	33
		Shortest	5m 47s	36	23	22
		Longest	16m 46s	43	39	42
		SD	5m 58s	4	9	10
5	2	Average	2m 56s	26	20	23
		Shortest	1m 12s	24	19	21
		Longest	4m 40s	28	21	24
		SD	2m 27s	3	1	2

For all click groups, the registration time can be considered long due to the action of selecting images and then clicking on the chosen images. It can be reported that for all click groups, the time to login during login attempts one to three are significantly shorter. The study also found that participants do not immediately select their click area, often taking several seconds before they start clicking on it. This action is believed to be due to the small tolerances used, which suggests it could directly affect login time and security if the users were observed.

3.2.2 External Group

Table 5 shows the time for 12 participants as they managed to login according to the specified login intervals. L1 to L5 refers to the login times (measured in seconds, (s)) for week one, L6 and L7 are login times for week two and finally L8 refers to the login time for the third week. It was found that the login time across the three weeks varied, although with one click, participants showed little change.

Table 5. Timing for the external group

Click	Participant	Time	Register	L1	L2	L3	L4	L5	L6	L7	L8
1	9	Average	11m 17s	17	20	17	20	14	17	17	14
		Shortest	2m 15s	14	14	13	11	12	10	10	9
		Longest	47m 23s	23	39	28	37	22	31	43	21
		SD	14m 2s	4	8	7	10	3	6	11	3
2	1	Average	3m 22s	19	26	21	31	16	15	18	24
		Shortest	3m 22s	19	26	21	31	16	15	18	24
		Longest	3m 22s	19	26	21	31	16	15	18	24
		SD	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
5	2	Average	10m 2s	33	29	23	35	25	28	20	27
		Shortest	2m 39s	29	24	22	22	23	23	19	26
		Longest	17m 25s	37	34	23	48	27	32	20	28
		SD	10m 26s	6	7	1	18	3	6	1	1

3.3 Accuracy

As reported earlier, the numbers of failed attempts during registration were high. As a result, participants had to use other usernames and changed their preference secret click or image. The authors discovered two main errors associated with such scenario, as indicates below.

a) Participant was unable to click within the allowable tolerance.

b) Participant did not click in sequence order, as the result of forgetting their secret order or areas.

From the data for both internal and external groups, it can be revealed that errors during both registration and login were correlated with participants who selected more clicks. In specific, there were slightly more participants who made tolerance errors than order errors. This is probably due to the software prototype using a small click tolerance.

Particularly within the external group, participants were unable to click accurately when they first started using the prototype. However, they managed to click within the clickable areas as they became familiar with using the software and understood what they needed to accomplish.

3.4 Pattern

Patterns are created during image selection when participants chose the same images (in the case of changing username of click), gender skew selection (e.g. men choosing sports car while women chose flowers) and following image order (e.g. participants choosing the first image in each theme). Moreover, patterns during the click selection are created when participants clicked on the same area across all images, producing obvious shapes or clicking their secrets in a straight line (e.g. top, bottom and left side of image area), and clicked on the image that appeared to be offering a pattern. Results for both groups are reported together within this section as they used similar software prototype.

With the internal group, the study found the majority of participants who changed their username or secrets (click or image) used their previously chosen images. One participant from the five clicks group used both of his previous images while two participants from the two clicks group used three and one of their previous images respectively. Of all the participants from the one click group who changed their username or clicks, only one did not used their previous image. Specifically for the one click group, two participants used four of their previous images while the others were ranging from one to three. In addition, it was also found one of these participants selected the first image for each theme as their secret images.

The external group also used their previous secret images with one participant using all of their previous images, with six other participants using between one to two of their previous chosen images .It was also found that two participants of the one click group chose their images in sequence (choose the first six themes); however their chosen images were different with each other.

Table 6. Image popular with their associated number of male and female

Theme	Number of participants choosing popular image	Male	Female
Buildings	11	3	8
Abstract	12	6	6
Food	8	4	4
Animals	7	3	4
Flower	10	4	6
View	14	7	7
People	5	2	3
Sport	13	8	5
Transport	4	1	3
Fruits	7	2	5

Table 6 presents the number of participants who chose popular images for each theme. It was found that the view and sport themes are the most popular, with the transport and people themes as the least popular selection.

Although it was found that a number of participants clicked within similar areas when creating their secret clicks, such action was eliminated due to the software prototype preventing participants from clicking on the same area within multiple images. Analysis was carried out to examine the area of clicking for participants who chose more clicks and although it can be reported that participants with two or three click groups create less obvious pattern, participants of the five clicks group clearly create patterns. The authors deduced that such scenarios are related to the images themselves, which clearly offer a pattern to be created.

Analysis was also done to examine the click areas in popular images for each theme. Analysis on the one-click group who chose the most popular image revealed that ten of the twelve participants who chose popular images in the sports theme clicked on the three most popular areas (see left side of the fig. 1), with seven out of twelve participants who chose the most popular image for the 'view' theme clicked on the same area (see right side of the fig. 1). Equally, all other popular images have shown a pattern where participants clicked on similar spots.

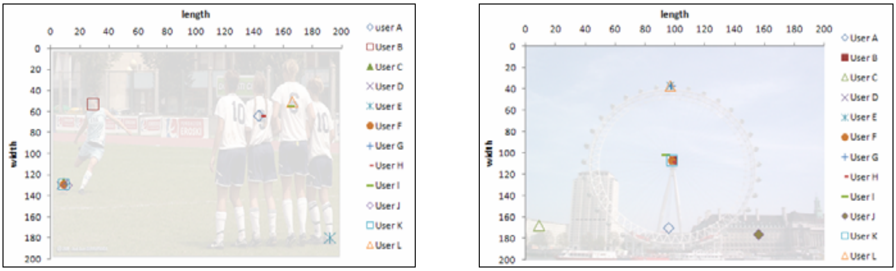

Fig. 1. Participants click areas for the popular image of the 'sport' (left) and 'view' (right) themes

From the collected data, the authors summarised that participants who chose more clicks tended to create patterns during their clicking task, while the existence of pattern during image selection was unidentified. Meanwhile, participants who chose more images (fewer clicks) tended to create patterns during both image and click selection. Patterns where users chose the first or last image for each theme was also reported. Although the authors' approach of not implementing restrictions for image selections and depending solely upon the guidelines is less effective then the introduction of guidelines. The GPG itself has resulted in the reduction of gender bias image selection and image order patterns.

It can, however, also be said that the restrictions together with the guideline during secret click selection played a minor role during the click selection task. Although not representative, participants with a higher number of clicks created more patterns (possibly as it is easier for them to remember), with analysis towards one click participants revealing the existence of hotspot.

3.5 Users' Feedback

A Likert five points scale rating was used to obtain participant feedback with the lowest score indicating participants' agreement with the statements while the highest score indicating disagreement. Table 7 reports the mean score of feedbacks for the first three questions within the internal group.

Table 7. Questionnaire results

Question	Mean score
Perception towards graphical password guidelines (GPG)	1.6
Perception towards restrictions	1.8
Perception towards combining GPG with the restrictions	1.8

When asked about participants average login time (with the software prototype displaying their average login time), it can be revealed that twenty participants found their login time were acceptable, with eight unacceptable. Seventeen participants agreed that their total registration time was acceptable while the remaining eleven disagreed. With the statement on the optimum combination of image and click, twenty two of the participants felt that having more images was more memorable than having more clicks, while five participants felt that the balance between both click and images were still memorable.

Participants who were new to the graphical method felt the method could be very useful and provided excellent protection. However, the majority of the participants who were involved in the previous trial felt that having larger clickable areas was more usable. In addition, they felt that having more clicks could be troublesome as they had to memorise too many spots and finally all participants agreed that in order for them to perform better, they needed to become more familiar with the method.

4 Conclusions and Future Work

This paper presented an investigation of the practicability of giving guidelines to a user before they chose their secrets for a graphical authentication system as well as evaluating user attitudes and opinions to the enhanced techniques.

During the registration task, participants struggled to click accurately within the allowable click tolerance and those who chose more clicks often failed to click in the correct order. As the result, they had to change to create new accounts or change to fewer clicks. The login task had shown improvement as they managed to login with fewer failed attempts, and the time to login to the software prototype was reduced marginally across login interval. The above findings reflect participants' familiarity with the software prototype as they used the software regularly.

Introducing guidelines to the participants before they start selecting their secrets had obtained positive perception from the majority of participants. However, this study has shown that guidelines on their own cannot guarantee the security and safety of the method itself. This is because participants used their previous images and created secret clicks using easy to remember spots, which resulted in predictable click-areas. By combining the introduction of guidelines with restrictions, user behaviour can be controlled to safeguard the method. This was proven where cases such as clicking on similar areas within the same or multiple images and where creating predictable pattern were reduced.

Finally, this paper has shown that the click patterns created by users who chose more clicks had a direct relationship with the nature of the image itself. It could be said that the introduction of guidelines gave no effect on participants' usability performance, but might give positive or negative effects on the security. The study also suggests that using one click per image is an ideal combination. This is because using one click per image requires less memorisation (i.e. it is more suitable for users with multiple accounts), less time to authenticate, convenience and significantly safer from predictability. The study also suggests that using a small tolerance without giving sufficient opportunity for familiarity to the user could result in a lack of usability of the proposed method.

It is suggested that future work could include a larger and more varied participant based for conducting significance testing to validate the collected data, testing different restrictions with the GPG, further evaluation of the claim that 'one click per image is better' and evaluating the technique known as the 'graphical-passwords-strength-meter', for safer secret creation based upon feedback from the system itself.

References

1. De Angeli, A., Coventry, L., Johnson, G., Renaud, K.: Is a picture really worth a thousand words? Reflecting on the usability of graphical authentication systems. International Journal of Human Computer Studies 63(2), 128–152 (2005)
2. Chiasson, S., Oorschot, P.C.V., Biddle, R.: Graphical password authentication using cued click points. In: Biskup, J., López, J. (eds.) ESORICS 2007. LNCS, vol. 4734, pp. 359–374. Springer, Heidelberg (2007)
3. Hinds, C., Ekwueme, C.: Increasing security and usability of computer systems with graphical password. In: ACM Southeast Regional Conference, Winston-Salem, North Carolina, USA, pp. 529–530. ACM, New York (2007)
4. Chiasson, S., Forget, A., Biddle, R., Oorschot, P.C.V.: User interface design affects security: Patterns in click-based graphical passwords. International Journal of Information Security 8(6), 387–398 (2009)
5. 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)
6. Oorschot, P.C.V., Salehi-Abari, A., Thorpe, J.: Purely automated attacks on Passpoints-style graphical passwords. Transactions on Information Forensics and Security 5(3), 393–405 (2010)
7. Davis, D., Monrose, F., Reiter, M.K.: On user choice in graphical password schemes. In: Proceedings of the 13th USENIX Security Symposium, California, USA, August 9-13, pp. 1–11. USENIX Association (2004)
8. Tullis, T.S., Tedesco, D.P.: Using personal photos as pictorial passwords. In: CHI 2005 Extended Abstracts on Human Factors in Computing Systems, Portland, Oregon, USA, pp. 1841–1844. ACM, New York (2005)
9. Everitt, K.M., Bragin, T., Fogarty, J., Kohno, T.: A comprehensive study of frequency, interference, and training of multiple graphical passwords. In: Proceedings of the 27th International Conference on Human Factors in Computing Systems, Boston, MA, USA, pp. 889–898. ACM, New York (2009)
10. Dirik, A.E., Memon, N., Birget, J.-C.: Modelling user choice in the Passpoints graphical password scheme. Paper presented at the Symposium on Usable Privacy and Security, Pittsburgh, PA, USA, July 18-20 (2007)
11. Gołofit, K.: Click passwords under investigation. In: Biskup, J., López, J. (eds.) ESORICS 2007. LNCS, vol. 4734, pp. 343–358. Springer, Heidelberg (2007)
12. Golofit, K.: Picture passwords superiority and picture passwords dictionary attacks. Journal of Information Assurance and Security 2, 179–183 (2007)
13. Peach, S., Voster, J., Heerden, R.V.: Heuristic Attacks against graphical password generators. In: Clarke, N., Furnell, S., Solms, R.V. (eds.) Proceedings of the South African Information Security Multi-Conference (SAISMC 2010), Port Elizabeth, South Africa, pp. 272–284. University of Plymouth (2010)
14. Lin, P.L., Weng, L.T., Huang, P.W.: Graphical password using images with random tracks of geometric shapes. In: Proceedings of the 2008 Congress on Image and Signal Processing, pp. 27–31. IEEE Computer Society, Los Alamitos (2008)
15. Harada, A., Isarida, T., Mizuno, T., Nishigaki, M.: A User Authentication System Using Schema of Visual Memory. In: Ijspeert, A.J., Masuzawa, T., Kusumoto, S. (eds.) BioADIT 2006. LNCS, vol. 3853, pp. 338–345. Springer, Heidelberg (2006)

16. Hayashi, E., Dhamija, R., Christin, N., Perrig, A.: Use Your Illusion: secure authentication usable anywhere. In: Proceedings of the 4th Symposium on Usable Privacy and Security, Pittsburgh, Pennsylvania, pp. 35–45. ACM, New York (2008)
17. Chiasson, S., Forget, A., Biddle, R., Oorschot, P.C.V.: Influencing users towards better passwords: persuasive cued click-points. In: Proceedings of the 22nd British HCI Group Annual Conference on HCI 2008: People and Computers XXII: Culture, Creativity, Interaction, Liverpool, United Kingdom, vol. 1, pp. 121–130. British Computer Society (2008)
18. Jali, M.Z., Furnell, S.M., Dowland, P.S.: Assessing image-based authentication techniques in a web-based environment. Information Management & Computer Security 18(1), 43–53 (2010)
19. Chiasson, S., Biddle, R., Oorschot, P.C.V.: A second look at the usability of click-based graphical passwords. In: Proceedings of the 3rd Symposium on Usable Privacy and Security, Pittsburgh, Pennsylvania, pp. 1–12. ACM, New York (2007)