Increasing the Security of Gaze-Based Cued-Recall Graphical Passwords Using Saliency Masks

Andreas Bulling
University of Cambridge
Lancaster University
andreas.bulling@acm.org

Florian Alt
University of Stuttgart
florian.alt@vis.uni-stuttgart.de

Albrecht Schmidt
University of Stuttgart
albrecht.schmidt@acm.org

ABSTRACT
With computers being used ever more ubiquitously in situations where privacy is important, secure user authentication is a central requirement. Gaze-based graphical passwords are a particularly promising means for shoulder-surfing-resistant authentication, but selecting secure passwords remains challenging. In this paper, we present a novel gaze-based authentication scheme that makes use of cued-recall graphical passwords on a single image. In order to increase password security, our approach uses a computational model of visual attention to mask those areas of the image that are most likely to attract visual attention. We create a realistic threat model for attacks that may occur in public settings, such as filming the user's interaction while drawing money from an ATM. Based on a 12-participant user study, we show that our approach is significantly more secure than a standard image-based authentication and gaze-based 4-digit PIN entry.

Author Keywords
User authentication; Eye tracking; Gaze-based; Cued-recall graphical passwords; Saliency masks

ACM Classification Keywords
H.5.2 Information Interfaces and Presentation: User Interfaces—*Input devices and strategies*; K.6.5 Computing Milieux: Security and Protection—*Authentication*

General Terms
Experimentation, Human Factors, Security

INTRODUCTION
Graphical passwords have long been investigated as a means of user authentication (see [2, 22] for reviews). Cued-recall graphical passwords (also known as locimetric passwords) have considerable advantages over traditional approaches, such as text passwords, as they leverage the vast capacity and capabilities of the human visual memory system [1, 8]. In addition to improved memorability and thus usability [21], graphical passwords promise increased resistance to guessing attacks, due to the potentially larger theoretical password space. Despite these advantages, graphical passwords that rely on physical interaction with the authentication system are still susceptible to shoulder-surfing attacks [9]. Particularly in public spaces, e.g. in front of an ATM, direct observation techniques such as video cameras or fake keypads can easily be used to eavesdrop and steal passwords or PINs.

CHI'12, May 5–10, 2012, Austin, Texas, USA.

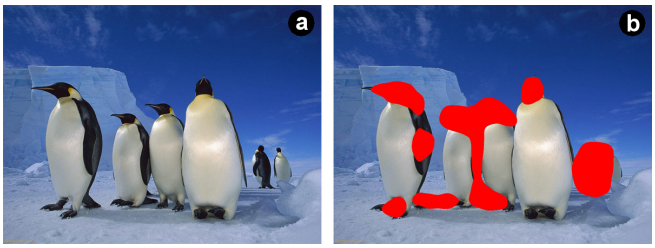

Figure 1: Sample images used in the study without (a) and with (b) saliency mask shown in red.

One solution to increase the security of graphical passwords is to use authentication schemes that do not require any physical contact with the system. Amongst the methods investigated in the past, the human gaze is particularly promising for implementing such schemes [14]. By its very nature, gazing does not require any physical contact and therefore, potentially works over greater distances. In addition, the human eye moves rapidly, which makes eavesdropping gaze-based passwords more difficult than touch-based input.

A key challenge in user authentication generally, and in graphical schemes in particular, is to define secure passwords. Previous research has shown that such schemes lead to hotspots, i.e. areas of the image that are more likely to be selected by users as password points. A password point is defined as a single fixation that is detected by the authentication system to be part of the chosen graphical password. These hotspots render such schemes more susceptible to dictionary attacks [24]. The only viable solution so far has been to select single password points across a sequence of several images [5].

In this present work, we present an alternative gaze-based authentication scheme that supports users in selecting secure gaze-based graphical passwords. To tackle the problem of hotspots, our scheme uses a computational model of visual attention – also known as saliency maps – to mask out those areas of the image most likely to attract visual attention (see Figure 1). We show that this approach significantly increases the security of gaze-based cued-recall graphical passwords.

The specific contributions of our research are 1) a shoulder-surfing resistant gaze-based authentication scheme that allows the user to select a sequence of arbitrary points in an image, 2) the introduction of computational models of visual attention to increase the security of gaze-based cued-recall graphical passwords, and 3) a security evaluation of three different gaze-based graphical passwords – PIN, picture without a saliency mask, and picture with a saliency mask – in a user study with 12 participants guessing passwords after watching close-up videos of the eye movements of other users.

RELATED WORK

Gaze-Based Authentication

Gaze has only recently started to be investigated as a means of user authentication. In one of the first works in this area, Kumar *et al.* presented an authentication system that mitigated the issues of shoulder-surfing by using gaze to enter a text-based password on an on-screen keyboard [18]. They found that gaze-based password entry required marginal additional time and that the error rates were similar to those of using a keyboard. Overall, the users in their study preferred gaze to traditional keyboard-based password entry.

One problem with text-based passwords is that they are more vulnerable to guessing attacks, due to the predictability of passwords, particularly for weak user-chosen ones. Several researchers have investigated gaze-based approaches that are less predictable and thus less vulnerable. For example, Maeder *et al.* demonstrated a graphical authentication scheme that requires users to look at pre-defined points in a single image [20]. Using a grid overlaid on the image, different sequences of these points were interpreted as different PINs. They evaluated different fixation thresholds and found that in all cases, users were able to successfully authenticate, using their own PIN sequence.

A key issue with gaze-based authentication schemes is the need for accurate point-of-regard gaze measurements. To address this issue, De Luca *et al.* proposed using sequences of relative eye movements, so-called eye gestures. They argued that gestures can be detected more robustly and are therefore better suited to real-world applications, such as public terminals or ATMs. De Luca *et al.* first demonstrated an authentication system in which each eye gesture corresponded to one digit of a standard pin [7]. In a subsequent work, they presented EyePassShapes – an authentication system that used a single eye gesture as the password [6]. The results of their usability studies indicated that EyePassShapes are more secure than common manual PIN entry and add less time to the login procedure than previous gaze-based authentication schemes. To assess the security of EyePassShapes, they performed a qualitative evaluation where a person familiar with their system tried to attack the password using video footage.

Security of Graphical Passwords

LeBlanc *et al.* identified eye tracking as a potential threat to the security of click-based graphical passwords [19]. They analysed participant gazes as they looked at images that had been used as the basis for passwords in an earlier study. They compared both datasets and analysed the likely success of guessing passwords. In their simulations, they found that the gaze data partly resembled the password click points. The authors argued that this might offer attackers an advantage over guessing click-based graphical passwords at random.

Others investigated means of making text-based and graphical passwords more secure. For example, Forget *et al.* developed a text-based password creation system that leveraged persuasive principles, in order to encourage users to select more secure passwords [10]. Their system randomly added characters to the passwords defined by the users. Afterwards, users were allowed to retrigger this process until they found a password they felt was easy to remember. The authors found that their approach significantly increased password security, but also that participants exhibited compensatory behaviour, by choosing weaker passwords in the first place.

Poorly chosen passwords in graphical password schemes like PassPoints [24] were found to lead to hotspots, i.e. parts of the image for which different users selected similar click-points as part of their passwords. These hotspots allowed attackers to mount more successful dictionary attacks [23]. Chiasson *et al.* addressed this issue by only using one click point on each of a sequence of different pictures; the next picture shown was determined by the location of the previous click point [5]. In subsequent research, they were able to successfully persuade users to select more secure passwords, by using a viewport positioned randomly on the picture that forced users to select a click point within its area [4].

Forget *et al.* conducted the only previous study on gaze-based graphical password entry via user-selected password points [9]. They described a shoulder-surfing-resistant gaze-based authentication scheme that relied on several points on a sequence of different images. They evaluated their scheme with respect to eye tracking accuracy, password entry time, as well as the number of successful logins by users with their own passwords. They also discussed some advantages, in particular the larger theoretical password space (TPS) and the cued-recall nature of their approach that helped users to more easily remember multiple passwords.

Summary

Previous research has investigated eye tracking for gaze-based authentication, as well as means of persuading users to select more secure graphical passwords. Only one study used a "guessing test" to qualitatively evaluate the security of gaze-based graphical passwords [6]. That study involved a single attacker and – like the study presented in [9] – required the user to press a key on a keyboard for several seconds to activate the password detection. In contrast, the current work presents a quantitative evaluation with a total of 240 password attacks from 12 users. In addition, we are the first to investigate computational models of visual attention in order to increase the security of gaze-based graphical passwords.

KEY IDEA

Visual attention is constantly attracted by different parts of the visual scene. Bottom-up computational models of visual

Figure 2: Sample images within the different image categories used in the studies: abstract image with low image complexity (a), abstract image with high image complexity (b), nature image (c), holiday image (d), and landscape image (e).

saliency aim at estimating the parts of a visual scene that are most likely to attract visual attention [15]. Given an input image or video, these models compute a so-called saliency map that topographically encodes for saliency at every location in the image. Visual saliency models were shown to predict human visual performance for a number of psychophysical tasks and have a large number of applications (see [11] for an extensive review). For example, in computer vision, saliency models were used for automated target detection in natural scenes, smart image compression, fast guidance of object recognition systems, or high-level scene analysis [16].

The key concept underlying this work is that by encouraging users to select password points that do not fall inside salient regions of the image, the security of gaze-based graphical passwords can be increased significantly. This is similar to the characteristics commonly required for text-based passwords, such as a minimum number of different alphanumeric or special characters.

Figure 1a shows one of the normal login images used in this work. Because the penguins' heads and feet are most likely to attract the user's visual attention – as predicted by the visual saliency model – these parts are masked out in Figure 1b. In a real-world authentication system, such masked images would be shown to the user in selecting the initial password. During operation, such as for authentication at a public terminal, the same image but without a mask would be used instead.

Method

To calculate saliency maps, we used a Graph-Based Visual Saliency (GBVS) model (see [13] for details on GBVS and [12] for the MATLAB toolbox we used). GBVS was shown to predict human fixations on natural images with superior performance to the original visual saliency algorithm presented in [16]. The saliency maps were calculated using the default parameters of the toolbox. The greyscale heatmaps returned by the GBVS algorithm were first normalised and a threshold applied at the 0.5 level so as to separate salient and non-salient areas. The salient areas were used as saliency masks that were overlaid in red onto the original images (see Figure 1b).

Hypotheses

In this work our aim was to compare gaze-based graphical passwords with and without saliency masks. As a baseline, we opted to use gaze-based PIN entry, which is a common approach for implementing gaze-based graphical passwords [18]. More specifically, we investigated the following hypotheses:

1. Gaze-based graphical passwords using several password points on an image are more secure, i.e. more difficult to attack, than gaze-based PIN entry.
2. Saliency masks increase the strength of gaze-based graphical passwords and thus the system's security.

APPARATUS

We designed and implemented a gaze-based authentication system consisting of a remote eye tracker and a custom login interface. The system allowed users to login using their gaze with two types of graphical passwords.

Hardware

For recording gaze data, we used a Tobii X120 remote eye tracker. The X120 has two integrated infrared cameras and was configured to track gaze at a sampling frequency of 60Hz. The login interface was shown to the participants on a 19 inch computer screen with a resolution of 1600x1200 pixels.

Login Interface

The login interface was implemented in Java and obtained gaze data from the eye tracker via TCP/IP. In addition, it recorded all events triggered by the user interacting with the system, such as the detection of password points or successful and failed login attempts. The interface included a standard 9-point calibration routine to adapt the tracking system to each user and a validation routine to assess the calibration quality. The testing routine involved the user looking at each of the calibration points in sequence until a fixation was detected by the system. Calibration quality was calculated for each participant as the mean Euclidean distance between all calibration points and the detected fixation points.

In the user study, depending on the type of graphical password, the interface showed two different screens:

- PIN: the login screen showed a grid of 10 tiles resembling a standard 10-digit keypad (see Figure 3). In order to perform a login attempt, the participant fixated at four of these tiles in sequence. Sequences of fixations to correct tiles in the correct order resulted in a successful login attempt; all other attempts were considered as failed.
- Image: the login screen showed a full-screen image. In contrast to PINs, the screen was not discretised by a grid, but arbitrary locations in the image could be selected.

We used a dwell-time-based method for selecting the password points of the graphical passwords, because this approach

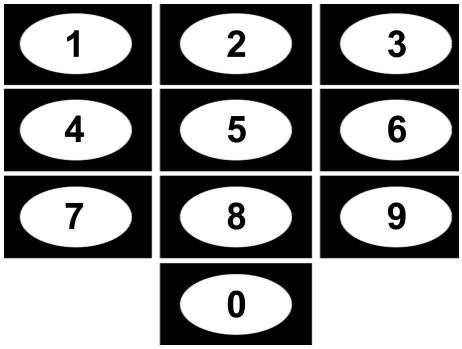

Figure 3: Standard 10-digit keypad used in the studies for gaze-based PIN entry.

is more natural and less error-prone than, for example, double blinks [17]. To select a password point, the participant had to fixate at a certain point in the image. The software continuously analysed the dispersion of the gaze points within a time window of 1.5 seconds. A fixation was detected by the system if at least 70% of the gaze points in this time window were inside a circular area with a radius of 1.7 degree of visual angle (about 75 pixels, the fixation detection threshold th_f).

EXPERIMENT

Study Design

The experiment consisted of a pre-study and a main study. In the pre-study, we asked one group of participants to define three different types of gaze-based graphical passwords (PIN, image with a saliency mask, image without a saliency mask). The goal of the main study was to assess the security of these passwords by asking a second group of participants to try to attack and guess these passwords by analysing close-up videos of the eyes recorded from the pre-study participants. The main study was designed as a randomised, controlled user study (repeated measures design).

For the main study, the binary output of all login attempts per image (successful: 1, unsuccessful: 0) was the dependent variable. The dependent variable was measured under three different conditions of the within-participants factor "graphical password": PIN-based, image-based without saliency mask, and image-based with saliency mask.

Image Selection

Selecting the images was an important step in the study preparation. Allowing participants to choose their own images would have made it difficult to control for similarity across images. In addition, pre-processing of the images would have been necessary to adjust for ratio and resolution. Finally, we expected side effects if people had chosen familiar images, for example containing family members or friends, for which certain areas in the image would have been more likely to be selected as password points.

Because they should also cover a realistic range of motifs and different visual complexities, we preselected all the images.

Figure 4: Experimental setup used in the pre-study. Participants selected image-based passwords, gazing at images with and without a saliency mask, as well as PINs on a standard 10-key keypad. Gazes were recorded using a remote eye tracker placed under, as well as a high-resolution video camera mounted above the computer screen.

We first downloaded a large number of images from the Internet with a resolution of 1600x1200 pixels and without any potentially offensive or sexual content. The images were then grouped into five different categories:

1. *Abstract images* showing *low-complexity* background images with several homogeneous areas (see Figure 2a).
2. *Abstract images* showing *high-complexity* scenes such as from computer games or star clouds (see Figure 2b).
3. *Nature images* showing mainly (multiple) animals in the wild, making it easy to select an individual or a feature of an individual, such as a leg (see Figure 2c).
4. *Holiday images* showing a beach or famous sites very similar to what we expected people to have among their own private holiday pictures (see Figure 2d).
5. *Landscape images* showing famous skylines or mountains (see Figure 2e).

We randomly chose four images per category and calculated a saliency mask version of each of these images, resulting in a total of 40 different images.

Pre-Study

The goal of the pre-study was to record a set of real gaze-based graphical passwords that would be attacked in the main study. For the pre-study, we recruited four participants (three male and one female), aged between 22 and 32 years (*mean* = 25, *sd* = 4.7), none of whom were familiar with the project. To have as wide a variety of faces as possible, we selected one participant wearing glasses, one with an Asian appearance, and one male as well as one female European.

The pre-study was performed in the lab using the setup shown in Figure 4. Participants were seated about 60 cm from the

computer screen, facing its centre. The eye tracker was placed under the computer screen. Above the screen, we mounted a video camera with a resolution of 720x576 pixels to record close-up videos of the participants' eyes (see Figure 5).

Experimental Procedure

After arriving at the lab we first explained the task to the participants and asked them to sign a consent form. Then, participants were instructed how to use the login interface, in particular how to enter and verify their passwords. We calibrated the eye tracker, validated the calibration and started the login interface. The interface automatically guided each participant through a sequence of screens, each involving their entering and verifying one image or PIN-based graphical password.

Each sequence consisted of 10 different PINs with four digits and a subset of 20 out of the 40 images – 10 images with a saliency mask and 10 without – shown to the participants in randomised order. The PINs were randomly generated to avoid people choosing easy or PINs that they were actually using. Participants were asked to look sequentially at the digits read out by the experimenter (cf. Figure 3).

For each image we asked participants to choose one graphical passwords consisting of four password points, but did not give them any indication as to how to choose them. We made sure that none of the participants was shown both the saliency and the non-saliency mask version of the same image and that, over all pre-study participants, each image was shown equally often. For images with a saliency mask, we asked participants to choose passwords points outside the red areas but did not explain how these areas had been defined.

Main Study

The goal of the main study was to evaluate the security of the three types of gaze-based graphical passwords recorded in the pre-study. We recruited 12 participants (one female and eleven males), aged between 23 and 29 years (*mean* = 25.9, *sd* = 1.7) via University mailing lists and bulletins in the neighbourhood surrounding the University building. Three of the participants wore glasses, one used contact lenses, and five had already participated in an eye tracking study before.

A first test showed that attacking all 120 passwords recorded in the pre-study took too long and would have affected participants' motivation and ability to keep concentrated. Consequently, for the study to take at most one hour, we had to reduce the number of passwords. We deemed crucial to still use the same number of passwords from each pre-study participant in order to minimise potential effects from different password selection strategies. We therefore selected five passwords per pre-study participant. Counterbalancing these passwords for image category and study condition finally resulted in a set of seven PINs, as well as seven images with and six images without a saliency mask and the corresponding videos.

Each participant was asked to attack the passwords in this set in randomised order. We made sure that, overall, all passwords would be attacked the same amount of times. We did not reveal which passwords were originally defined with and without a saliency mask and ensured that the eyes could be well recognised in all videos. Participants were rewarded with 10 EUR for participating in the study. In addition, we offered them a 1 EUR bonus for each password they managed to guess successfully.

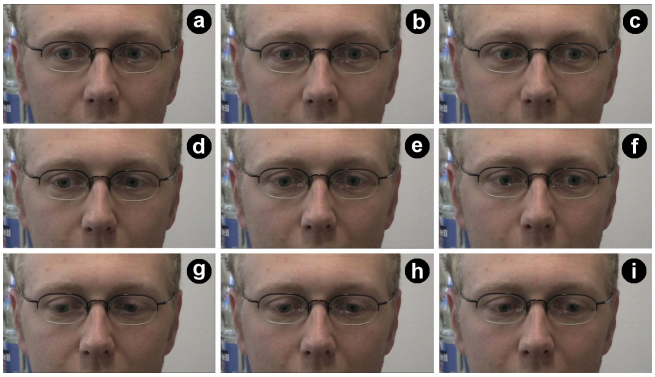

Figure 5: Sample images extracted from the video tutorial with the person looking at the top-left (a), top-centre (b), top-right (c), middle-left (d), middle-centre (e), middle-right (f), bottom-left (g), bottom-centre (h), and bottom-right (i) points of the calibration procedure.

Experimental Procedure

After the participants arrived at the lab, we introduced the study and asked them to sign a consent form. We asked them to complete the first part of the questionnaire on demographics and on their use of PINs and passwords. We guided them to the room where we had set up our authentication system.

All participants were first shown a video tutorial on how to guess gaze-based graphical passwords. In the first part of the tutorial, participants watched close-up videos of another person (who had not participated in the pre-study) looking at the nine points of the calibration procedure (see Figure 5). This was to give them a feeling of the maximum gaze range they could expect later on. In the second part of the tutorial, we showed them close-up videos of the person's gaze, while entering a PIN-based and an image-based password, and asked them to guess the password. In addition, we gave them some hints that we had discovered ourselves when trying to guess passwords using the same approach, e.g. to watch out for characteristic horizontal or vertical sequences of eye movements. All videos in the tutorial were mirrored horizontally, so that the gaze direction matched the on-screen gaze location.

After finishing the video tutorial, we ensured that participants were seated about 60 cm away from the computer screen and faced its centre. We calibrated the eye tracker, validated the calibration and started the login interface. The interface guided the participants through a sequence of login procedures, each involving their guessing one image or PIN-based password. For each password, the participants could watch the corresponding close-up video from the pre-study as often as they wished on a laptop placed next to the screen (see Figure 6). Participants could pause the video at any time so as to closely examine gaze direction. In addition, a printout of the image was provided for marking potential password points.

Figure 6: Experimental setup used in the main study. Participants tried to guess passwords by analysing close-up videos of other users' eye movements on the laptop. An additional printout of the image was provided to take notes and to mark potential password points.

Once participants thought they knew the password, they tried to log into the system using this password. For each image, participants were given a maximum of three login attempts before the interface automatically switched to the next image. If the password was guessed correctly, the system immediately switched to the next image. Participants were asked to guess as many passwords as possible while we recorded their gaze paths. We explicitly told them that guessing the passwords was more important than finishing the study quickly. However, when participants spent a lot of time watching the video, we applied a soft deadline of 2 minutes, after which the experimenter asked the participant to finish. Consequently, it was not possible to analyse post-hoc how quickly passwords could be guessed by the participants.

After the study, we asked participants to complete the second part of the questionnaire, with questions on the experiment and how they perceived the PIN-based and image-based graphical passwords with respect to robustness, security, and usability. In addition, we asked them how difficult it was to guess passwords, whether the tutorial had been useful, and whether they had noticed any differences in visual behaviour between the various people they had observed.

Data Analysis and Validation

In the main study, we collected quantitative and qualitative data. Quantitative data was recorded from participant interactions with the login interface and the questionnaires, using 5-point Likert scales. To verify that all login attempts were counted correctly in the post-hoc analysis, for each login attempt, we plotted the image with the password from the pre-study and the password guessed by the participants in the main study (see Figure 7 for an example). In total, we individually validated 541 images. Questionnaire responses were transcribed to a spreadsheet. Qualitative data was gathered from observations during the study, open questions in the questionnaire and from interviews following the study. This data was transcribed and important themes identified.

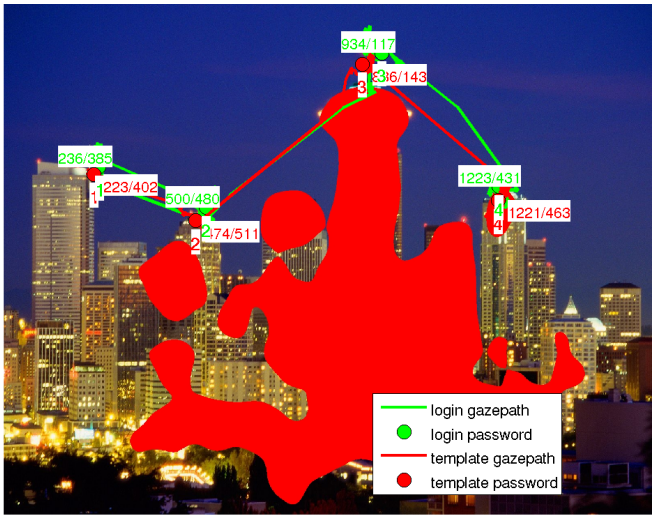

Figure 7: Image, password points as well as gaze paths of the password defined in the pre-study (template) and the login attempt (login) of the only successful login attempt on an image with a saliency mask.

RESULTS

Eye Tracking Accuracy

We first analysed the eye tracking accuracy (mean Euclidean distance between the calibration and detected fixation points) in the pre- and main study using the gaze data recorded during the validation routine after calibration. In the pre-study, the mean Euclidean distances ranged between 21 and 46 pixels (*mean* = 33, *sd* = 10); in the main study the distances were between 22 and 52 pixels (*mean* = 34, *sd* = 9). At a distance of 60 cm, these distances correspond to an accuracy of between 0.5 and 1.05 degree of visual angle – a range similar to the manufacturer-reported accuracy of the Tobii eye tracker.

Influence of Saliency Masks

Using the gaze data collected in the pre-study, we analysed whether the saliency masks actually influenced participant visual behaviour and thus the way they selected their password points. To this end, we compared the number of password points selected inside and outside the saliency regions for all images with and without a saliency mask. Figure 8 shows the results of this comparison for all pre-study images and participants. On average, for images without a saliency mask, 34.5% of the password points fell inside the saliency regions. For images with a saliency mask, this percentage dropped to 1.3%, a difference that was statistically significant ($p < .001$, one tailed Fisher's exact test).

Security of Graphical Passwords

We first calculated the theoretical password space, that is the total number of all possible distinct passwords in a system, for the different graphical password types. TPSs grow exponentially and are typically compared in log_2. The TPS for 4-digit PIN-based graphical passwords with a password length of four digits is $log_2(10^4) \approx 13.3$. An image size

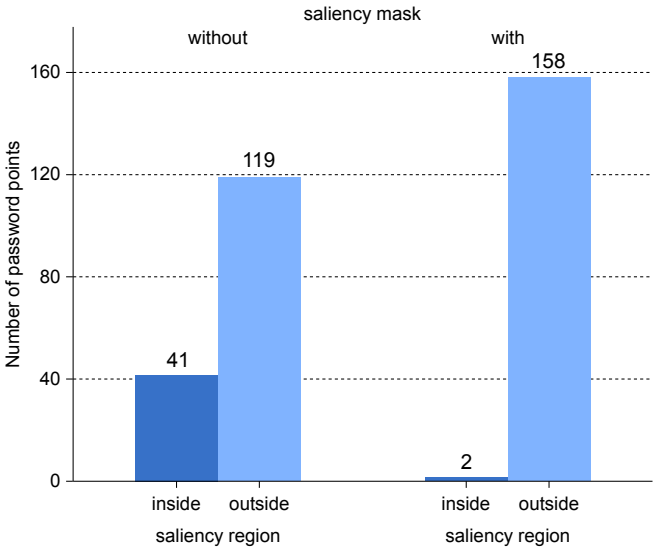

Figure 8: Number of password points inside and outside the saliency regions for images with and without a saliency mask. Values were calculated for all images shown to the four participants who took part in the pre-study.

Predictor	β	$SE\ \beta$	p	e^{β} (odds ratio)
intercept	-2.113	0.388	<0.001	N/A
saliency mask	-2.322	1.090	0.033	0.098
PIN	0.914	0.462	0.048	2.494

Table 1: Logistic regression analysis of 12 participants' performance in guessing three different types of graphical passwords. Both the "PIN" and image with a saliency mask ("saliency mask") conditions are significantly different to the "image without a saliency mask" reference condition.

of 1600x1200 pixels can contain 80 non-overlapping circles with a diameter of 150 pixels (fixation detection threshold th_f = 75 pixels). Thus, with the same number of password points per password, the TPS for image-based passwords is much larger, namely $log_2(80^4) \approx 25.3$.

We then evaluated the security of the different graphical passwords using the data recorded in the main study. Each successful login attempt was counted as a "1", whereas three failed login attempts on a single image were counted as a single "0". Overall, 19 out of 81 PINs were successfully guessed by the participants compared to 8 out of 72 images without a saliency mask and 1 out of 82 images with a saliency mask. We used a logistic regression to model the relationship between the password type and the relative probabilities of a successful login attempt. The regression was conducted using a generalized linear mixed effect model with "image without a saliency mask" as the reference condition. Participants was a random effect; the "PIN" vs. "image without a saliency mask", as well as "image with a saliency mask" vs. "image without a saliency mask" were fixed effects. As can be seen from Table 1, PINs are significantly less secure than

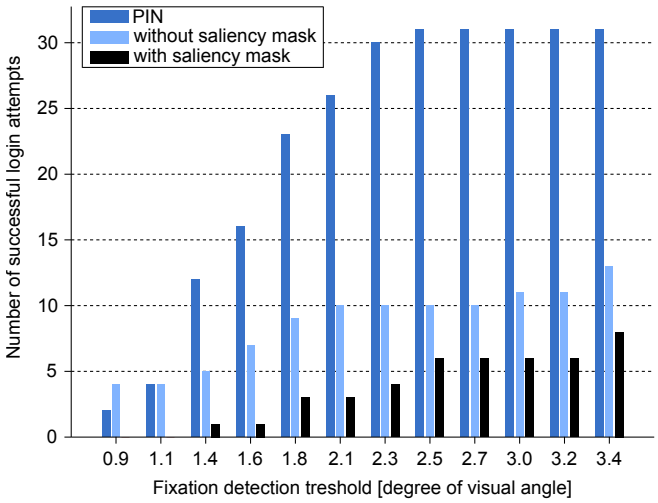

Figure 9: Number of successful login attempts for the different graphical password schemes for different values of the fixation detection threshold th_f.

image-based passwords without a saliency mask ($p = .048$) and saliency masks significantly increase password security ($p = .033$). A likelihood ratio test between the intercept-only (null) model and the logistic model was also significant (log-likelihood difference = 11.248, $\chi^2(2) = 22.497, p < .001$).

Influence of Fixation Detection Threshold

In the previous evaluation, the fixation detection threshold was fixed to th_f = 1.7 degree of visual angle. To analyse the influence of th_f on the number of successful login attempts, we simulated different thresholds by sweeping $th_f = 0.9 \ldots 3.4$ (in 12 steps). The lower bound of the sweep was motivated by the maximum eye tracking accuracy achieved after calibration. For login attempts in which one of these thresholds resulted in an earlier successful attempt than in the original study, all subsequent attempts on that image were excluded from the analysis. As can be seen from Figure 9, the fixation detection threshold exerts a considerable influence on the number of successful login attempts. Most importantly, the analysis shows that using saliency masks consistently, results in fewer successful login attempts and thus more secure passwords than without saliency masks or using PINs.

Questionnaires

From the questionnaires completed by the participants of the main study, we received valuable feedback on the perceived security and usability of the PIN-based and image-based graphical passwords, experiences with guessing them, preferred situations and devices, as well as general feedback on the study.

Security and Usability

We first analysed the perceived security and usability of the PIN-based and image-based passwords using a Friedman analysis of variance by ranks on the participant responses in both conditions (5-point Likert scale, 1: very low, 5: very high;

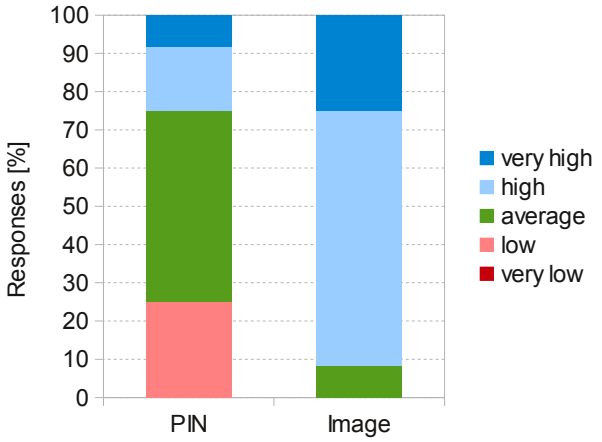

Figure 10: Distribution of responses to question "How do you rate the security of the different graphical passwords?".

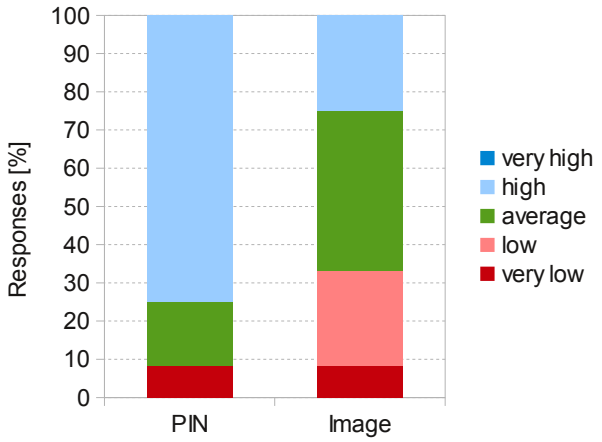

Figure 11: Distribution of responses to question "How do you rate the usability of the different graphical passwords?".

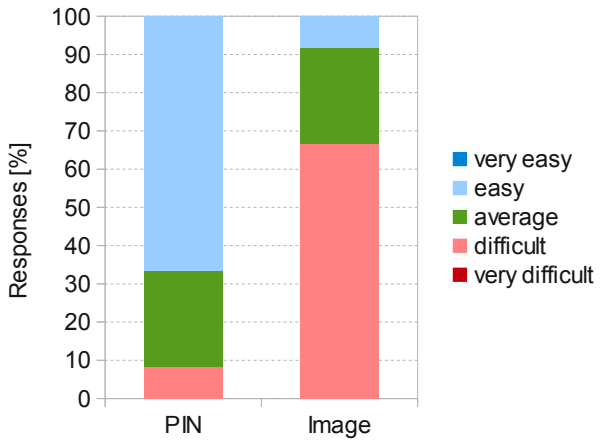

Figure 12: Distribution of responses to question "How do you rate the ease of monitoring eye movements in the videos?".

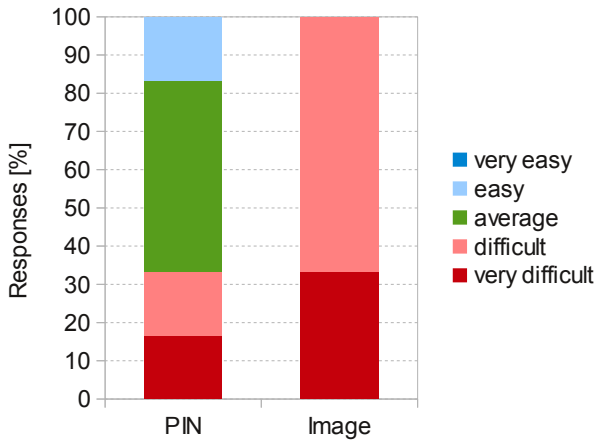

Figure 13: Distribution of responses to question "How do you rate the ease of guessing passwords from the videos?".

see Figures 10 and 11). Participants found image-based passwords to be significantly more secure than PIN-based passwords, $\chi^2(1) = 10.0, p < .05$. Yet, the password usability was rated significantly higher for PIN-based than for image-based passwords, $\chi^2(1) = 6.0, p < .05$.

Guessing Passwords

We then analysed the perceived ease of monitoring eye movements and guessing passwords from the videos for PIN-based and image-based passwords (1: very difficult, 5: very easy; see Figures 12 and 13). We again used a Friedman analysis of variance by ranks on participant responses in both conditions. Participants found it significantly easier to monitor eye movements for PIN-based passwords than for image-based passwords, $\chi^2(1) = 10.0, p < .05$. Consequently, participants found it significantly easier to guess PIN-based passwords ($\chi^2(1) = 8.0, p < .05$) and felt that they had to concentrate significantly more (1: had to concentrate intensively, 5: did not have to concentrate at all) when trying to guess image-based password, $\chi^2(1) = 4.5, p < .05$.

Although we did not elaborate on this, our results indicate that difficulties in guessing passwords may be related to the gender of pre-study participants. Participants consistently found the passwords of our female pre-study participant significantly easier to guess than those of the Asian male, $\chi^2(1) = 6.4, p < .05$, as well as the European male not wearing, $\chi^2(1) = 8.0, p < .05$, and wearing glasses, $\chi^2(1) = 5.4, p < .05$ (1: very difficult, 5: very easy).

Situations and Devices

We further analysed in which situations and with respect to which devices participants could imagine using the different graphical passwords. Although none of the following findings was statistically significant, they nonetheless raise some interesting questions for future work.

For situations (1: would definitely use, 5: would not use at all), we found that the image-based passwords were favoured in non-private situations, such as in a bank (*mean* = 3.58, *sd* = 0.79) or in public spaces (*mean* = 3.33, *sd* = 1.27). Image-based passwords were less favoured at home (*mean* = 3, *sd*

= 1.45) and at work (*mean* = 3.25, *sd* = 1.17). The favoured devices were ATMs (*mean* = 3.33, *sd* = 1.21) and access systems (*mean* = 3.64, *sd* = 1.17), while laptops (*mean* = 3, *sd* = 1.55), mobile phones (*mean* = 2.92, *sd* = 1.67), and desktop computers (*mean* = 3, *sd* = 1.55) were less popular.

By contrast, PIN-based passwords were most popular in private situations such as at home (*mean* = 3.42, *sd* = 1.57) and in the office (*mean* = 3.33, *sd* = 1.12), and less popular in public spaces (*mean* = 3, *sd* = 1.22) and a bank (*mean* = 3.08, *sd* = 1.1). PIN-based passwords were also preferred when used on laptops (*mean* = 3.42, *sd* = 1.13) and desktop computers (*mean* = 3.33, *sd* = 1.12) than on mobile phones (*mean* = 3.08, *sd* = 1.22) and ATMs (*mean* = 3.08, *sd* = 1.25).

User Study

Finally, participants reported that they found the user study moderately exhausting (*mean* = 3.08, *sd* = 1.16; 1: very exhausting, 5: not exhausting at all). The video tutorial was perceived as being helpful by all participants (*mean* = 4, *sd* = 1.18; 1: not helpful at all, 5: very helpful).

Qualitative Feedback

The qualitative feedback from the pre-study revealed that some participants developed strategies for selecting passwords. One participant constructed stories around the password points to make them more memorable. A second strategy reported by another participant was to select password points along vertical or horizontal lines, such as by choosing points at the intersection of foreground objects.

Participants from the main study found PINs easier to guess, due to the grid and thus the eye movements were more easily recognisable. P5 and P8 reported that having a reference point (e.g. by initially looking at the centre of the screen or using the glasses as a reference) or knowing the starting gaze position would make guessing easier. P2, P7, P8 found vertical eye movements more difficult to detect than horizontal ones. P2 reported looking for visual strategies, while P8 and P9 noticed that one pre-study participant had chosen password points in the direction of reading, i.e. from left to right.

DISCUSSION

The results of our study demonstrate that image-based graphical passwords are significantly more secure than PIN-based passwords, both in an actual attack and in terms of participant perception, hence verifying Hypothesis 1 (see Table 1 and Figures 10 and 13). Using computational models of visual attention to mask the most salient areas of the images does significantly increase security, compared to the standard image-based approach, hence verifying Hypothesis 2. In combination with the much larger theoretical password space, these results make saliency masks a promising means of increasing the security of gaze-based graphical passwords.

While image-based graphical passwords were perceived as significantly more secure than PIN-based passwords, the usability was rated lower by the participants in our study (see Figure 11). Participants preferred image-based passwords for public terminals, while PIN-based passwords were preferred for mobile devices such as laptops and mobile phones. These responses may have been caused by the fact that participants could not imagine mobile devices being equipped with robust eye trackers in the near future. While currently, application domains are indeed mostly limited to ATMs or similar stationary systems, the advent of mobile eye trackers will pave the way for gaze-based authentication on smaller and thus more mobile devices [3].

A proper analysis of password memorability requires a long-term study and was therefore beyond the scope of this work. However, when we asked two pre-study participants to log in with their image-based passwords two days later (whom we did not ask to remember their passwords in the first place) they correctly remembered 14 out of 40 passwords (five images with and nine without a saliency mask). 13 of these image-based passwords were remembered by the pre-study participant who had selected the password points in the direction of reading, i.e. from left to right. While using such strategies seems to improve password memorability this may come at the cost of reduced security. We plan to investigate this trade-off between memorability and security in more detail in future work and particularly how password memorability can be improved without compromising security. In terms of security, it will also be interesting to see how saliency masks compare to other approaches, such as selecting password points on a sequence of images [5].

Finally, the study also reveals some of the issues researchers may face in the real-world implementation of gaze-based graphical passwords. Participants in our pre-study reported having used visual strategies for selecting their passwords points in the images. Two main-study participants noticed and exploited this behaviour by specifically looking for characteristic eye movement sequences such as in a vertical or horizontal direction. This suggests that, in addition to the saliency masks presented here, measures need to be taken to prevent users from choosing closely related password points (similar to preventing PINs like “1111” or “1234”). Additional user studies will be required to investigate whether users should be allowed to choose their own graphical passwords (and potentially the images as well), or whether both should be provided by the authentication system during registration. In the latter case, it would be useful to identify what defines such “good” passwords and images.

CONCLUSION

In this paper we have proposed computational models of visual attention to increase the security of gaze-based cued-recall graphical passwords. We introduced saliency masks as a promising method for supporting the user in selecting more secure passwords and thus reducing the risk of hotspots in the authentication images. In a study with a realistic threat model, we showed that saliency masks significantly increase password security on a single image, compared to a standard image-based method and gaze-based 4-digit PIN entry. This result is promising, as saliency masks can easily be computed. It also raises the issue of the wider applicability of this approach, such as in the development of quantitative measures of the security of cued-recall graphical passwords.

ACKNOWLEDGEMENTS

This work was supported by the European Union 7th framework programme under grant agreements no. 215893 and 244011. We would like to thank Yordan Terziev, Andreas Kaiser, and Ken Pfeuffer for their help with implementing the authentication software, Brian Bloch for editing a draft of this paper, as well as Simon Byrne and Julian Mennenöh.

REFERENCES

1. Angeli, A. D., Coventry, L., Johnson, G., and Renaud, K. Is a picture really worth a thousand words? exploring the feasibility of graphical authentication systems. *International Journal of Human-Computer Studies 63*, 1-2 (2005), 128 – 152.
2. Biddle, R., Chiasson, S., and van Oorschot, P. C. Graphical passwords: Learning from the first twelve years. *ACM Computing Surveys* (2011). to appear.
3. Bulling, A., and Gellersen, H. Toward Mobile Eye-Based Human-Computer Interaction. *IEEE Pervasive Computing 9*, 4 (2010), 8–12.
4. Chiasson, S., Forget, A., Biddle, R., and van Oorschot, P. C. Influencing users towards better passwords: persuasive cued click-points. In *Proceedings of the 22nd British HCI Group Conference on People and Computers* (2008), 121–130.
5. Chiasson, S., van Oorschot, P. C., and Biddle, R. Graphical password authentication using cued click points. In *Proceedings of the 12th European Symposium on Research in Computer Security* (2007), 359–374.
6. De Luca, A., Denzel, M., and Hussmann, H. Look into my eyes!: can you guess my password? In *Proceedings of the 5th Symposium on Usable Privacy and Security* (2009), 1–12.
7. De Luca, A., Weiss, R., and Drewes, H. Evaluation of eye-gaze interaction methods for security enhanced pin-entry. In *Proceedings of the 19th Australasian Conference on Computer-Human Interaction* (2007), 199–202.
8. Everitt, K. M., Bragin, T., Fogarty, J., and Kohno, T. A comprehensive study of frequency, interference, and training of multiple graphical passwords. In *Proceedings of the 27th SIGCHI International Conference on Human Factors in Computing Systems* (2009), 889–898.
9. Forget, A., Chiasson, S., and Biddle, R. Shoulder-surfing resistance with eye-gaze entry in cued-recall graphical passwords. In *Proceedings of the 28th SIGCHI International Conference on Human Factors in Computing Systems* (2010), 1107–1110.
10. Forget, A., Chiasson, S., van Oorschot, P. C., and Biddle, R. Improving text passwords through persuasion. In *Proceedings of the 4th Symposium on Usable Privacy and Security* (2008), 1–12.
11. Frintrop, S., Rome, E., and Christensen, H. I. Computational visual attention systems and their cognitive foundations: A survey. *ACM Transactions on Applied Perception 7*, 1 (2010), 1–39.
12. Harel, J. Graph-Based Visual Saliency Toolbox for MATLAB, http://www.klab.caltech.edu/ harel/share/ gbvs.php, 2006.
13. Harel, J., Koch, C., and Perona, P. Graph-based visual saliency. In *Proceedings of the 20th International Conference on Neural Information Processing Systems* (2006), 545–552.
14. Hoanca, B., and Mock, K. Secure graphical password system for high traffic public areas. In *Proceedings of the 2006 Symposium on Eye Tracking Research & Applications* (2006), 35–35.
15. Itti, L., and Koch, C. Computational modelling of visual attention. *Nature Reviews Neuroscience 2*, 3 (2001), 194–203.
16. Itti, L., Koch, C., and Niebur, E. A model of saliency-based visual attention for rapid scene analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence 20*, 11 (1998), 1254–1259.
17. Jacob, R. J. K. What you look at is what you get: eye movement-based interaction techniques. In *Proceedings of the 8th SIGCHI International Conference on Human Factors in Computing Systems* (1990), 11–18.
18. Kumar, M., Garfinkel, T., Boneh, D., and Winograd, T. Reducing shoulder-surfing by using gaze-based password entry. In *Proceedings of the 3rd Symposium on Usable Privacy and Security* (2007), 13–19.
19. LeBlanc, D., Forget, A., and Biddle, R. Guessing click-based graphical passwords by eye tracking. In *Proceedings of the 8th International Conference on Privacy Security and Trust* (2010), 197 –204.
20. Maeder, A. J., Fookes, C. B., and Sridharan, S. Gaze based user authentication for personal computer applications. In *Proceedings of International Symposium on Intelligent Multimedia, Video and Speech Processing* (2004), 727–730.
21. Moncur, W., and Leplâtre, G. Pictures at the atm: exploring the usability of multiple graphical passwords. In *Proceedings of the 25th SIGCHI International Conference on Human Factors in Computing Systems* (2007), 887–894.
22. Suo, X., Zhu, Y., and Owen, G. S. Graphical passwords: A survey. In *Proceedings of the 21st Computer Security Applications Conference* (2005), 463–472.
23. Thorpe, J., and van Oorschot, P. C. Human-seeded attacks and exploiting hot-spots in graphical passwords. In *Proceedings of the 16th USENIX Security Symposium* (2007), 8:1–8:16.
24. Wiedenbeck, S., Waters, J., Birget, J.-C., Brodskiy, A., and Memon, N. Passpoints: design and longitudinal evaluation of a graphical password system. *International Journal of Human-Computer Studies 63*, 1-2 (2005), 102–127.