Shoulder Surfing Defence for Recall-based Graphical Passwords

Nur Haryani Zakaria[1], David Griffiths[1], Sacha Brostoff [3], Jeff Yan[*1, 2]
[1] School of Computing Science, Newcastle University, UK
[2] Dept. of Information Engineering, Chinese University of Hong Kong
[3] Dept. of Computer Science, University College London, UK
{n.h.zakaria, david.griffiths, jeff.yan}@ncl.ac.uk, s.brostoff@cs.ucl.ac.uk

ABSTRACT
Graphical passwords are often considered prone to shoulder-surfing attacks, where attackers can steal a user's password by peeking over his or her shoulder in the authentication process. In this paper, we explore shoulder surfing defence for recall-based graphical password systems such as Draw-A-Secret and Background Draw-A-Secret, where users doodle their passwords (i.e. secrets) on a drawing grid. We propose three innovative shoulder surfing defence techniques, and conduct two separate controlled laboratory experiments to evaluate both security and usability perspectives of the proposed techniques. One technique was expected to work to some extent theoretically, but it turned out to provide little protection. One technique provided the best overall shoulder surfing defence, but also caused some usability challenges. The other technique achieved reasonable shoulder surfing defence and good usability simultaneously, a good balance which the two other techniques did not achieve. Our results appear to be also relevant to other graphical password systems such as Pass-Go.

Categories and Subject Descriptors
D.4.6 [**Operating Systems**]: Security and Protection – *access controls, authentication*; K.6.5 [**Management of Computing and Information Systems**]: Security and Protection - *authentication*

General Terms
Experimentation, Security, Human Factors.

Keywords
Graphical Passwords, Shoulder-surfing defence, Usability.

1. INTRODUCTION
The alphanumeric password has been part of the authentication process for a very long time. However, this simple and ubiquitous technology has some well-known usability problems especially on the memorability aspect. The humans' ability to remember pictures better than text has been well documented in numerous cognitive and psychological studies (as reviewed in [2, 11, 24]). As a result, much research has been inspired in both the security and HCI communities in recent years to explore graphical authentication systems as an alternative or an enhancement to text passwords. As the name implies, graphical authentication uses graphics (pictures, icons, faces etc.) instead of the common used text strings [23].

Graphical passwords are still far from being perfect. For example, a password supplied for authentication by a user in a public place, if not properly protected, can be stolen by a bystander who observes over the user's shoulder. This is known as a *shoulder surfing attack* and commonly regarded as a drawback to various graphical password systems [22]. Alpha-numeric passwords are defended against this by substituting asterisks for the password characters in the display as the user logs in. To make graphical passwords reliable in the real world, it is essential to arm them with good shoulder surfing defence mechanisms.

In this paper, we study shoulder surfing defences for recall-based graphical password systems such as Draw-A-Secret (DAS) [11] and Background Draw-A-Secret (BDAS) [6]. DAS is a representative graphical password scheme and worthy of extensive study for the following reasons. First, its theoretical password space can be larger than that of text passwords. Second, unlike many other graphical password systems, DAS can be used for not only user authentication, but also for key generation. Although some research has revealed that the user choices of DAS passwords could render this theoretically sound scheme less secure in practice [26], it appears that many of the weaknesses could be improved by introducing a background image into the drawing grid [6], together with other countermeasures.

DAS and BDAS authenticate people by using a stylus input; provide an easy alternative to text passwords. They are particularly suitable for PDAs and mobile phones with a touch screen. As such mobile devices are highly portable it can be assumed that users venture into public places, and as a result will authenticate in areas where they may be left open to shoulder surfing attacks. Hence, a shoulder surfing defence is necessary to increase the security of the DAS/BDAS scheme which in turn could make the scheme a more appealing alternative to text passwords for mobile device users. To the best of our knowledge, currently there is little study of shoulder surfing defences for such graphical password systems except work in [17] and our previous work [15].

In this paper, we propose three innovative techniques to provide shoulder surfing protection for DAS and BDAS systems. A well-known lesson in computer security is that what security engineers expect to provide effective security, and what happens in reality,

Symposium On Usable Privacy and Security (SOUPS) 2011, July 20-22, 2011, Pittsburgh, PA, USA.

* To whom correspondence should be addressed. Email: jeff.yan@ncl.ac.uk

can differ greatly [30]. To evaluate our approach, we have implemented all three techniques, and conducted two separate controlled laboratory experiments to evaluate both security and usability perspectives of these techniques.

Our techniques do not aim to provide perfect shoulder surfing protection (e.g.: to make passwords invulnerable to attacks armed with a camera, video recorder or equivalent electronic devices). Rather, we aim to protect passwords from less dedicated attacks, those that can be carried out by human eyes alone. While attacks aided with a camera or the like can be a serious threat, casual human-based attacks may still pose real risks. It appears that the only feasible way of achieving a perfect shoulder surfing defence relies on zero-knowledge interaction as demonstrated in [20]. However, it is still an open problem to make such an approach usable for ordinary users.

The rest of this paper is organized as follows. In Section 2 we briefly discuss related work. Section 3 gives an overview of the DAS graphical password scheme. We describe our proposed defence techniques in Section 4, and report our experiments in Sections 5 – 6. We conclude with overall discussions and discuss future work in Section 7.

2. RELATED WORK

Graphical passwords have gained much attention as potential alternatives to text-based passwords. There are three categories of graphical passwords; those using either recognition-based, cued-recall based or recall based techniques [2]. In recognition based systems, a user is presented with a set of images and the user passes the authentication by recognising and identifying the images he/she had previously selected (e.g. DejaVu [5], Passfaces [19] and VIP [1]). In cued-recall systems, users are required to remember and target specific locations within an image (e.g. PassPoints [29]). In recall-based systems, however, a user is asked to reproduce something that he/she created or selected earlier during the registration stage (e.g. DAS [11] and Passdoodle [8]).

Shoulder surfing attacks have been identified as one of the main concerns against adopting graphical authentication in real use [23]. In order to protect from shoulder surfing attacks, various techniques have been proposed to overcome this problem. In recognition-based systems, Sobrado & Birget [21] developed the Convex Hull Click (CHC) scheme, using a huge number of pass-icons to confuse shoulder surfers trying to determine the correct pass-icon. However Man et al in [18] proved CHC to be unusable as so many objects had to be fitted on-screen at once that they were all too small, making it difficult for users to distinguish between pass-objects and non-pass-objects. Another possible technique to provide shoulder surfing resistance is by displaying degraded or distorted images as used in Use Your Illusion (UYI) scheme [10]. It is done with the intention to reduce the visibility of user's input in the hope to increase the protection.

In cued-recall based systems, Suo [22] creates a variation of PassPoints to protect the scheme by using blurring technique. The image is made obscured except for a small focus area where the authentication is achieved after ten rounds of click-on inputs on different focus area. Although the scheme looks promising, adversary can successfully recover the secret (password) by observing few rounds of login [2]. Researchers in [7] have introduced Cued Gaze-Points (CGP) scheme. Instead of using mouse click, users use eye-gazing technique to input their points and this has been claimed to increase shoulder surfing resistance. However their initial study had shown some clear trade-off between usability and security; obviously the larger tolerance size proved considerably more usable but would not enhance security.

As for recall-based graphical passwords, finger pressure technique has been introduce in [17] as another possible way to enter sensitive input that is resilient against shoulder surfing attack. By using haptic input device which measures pen pressure while users draw their password; an adversary would find it difficult to distinguish variances in pen pressure. However their user study revealed some usability challenges when users are found to apply very little pen pressure and hardly lifted the pen while drawing. The attempt to use haptic based input did not significantly increase the difficulty of guessing passwords [2]. In a previous effort [15], we revised the original DAS scheme by introducing qualitative spatial relations and dynamic grid transformation to enhance shoulder surfing resistance. Although this revised scheme provided good shoulder surfing defence, it has some usability issues. Due to space limitation, an intensive survey of graphical passwords together with some shoulder surfing defence techniques can be found in [2].

There are also several studies that have been carried out to improve shoulder surfing resistance of text-based authentication. For example, researchers in [12, 4] attempted to use similar gazing-based technique as in [7] to enter sensitive input from an on-screen keyboard. The evaluation of EyePIN in [4] showed promise, but it also seems to have revealed some usability problems when users needed to remember and understand the new alphabet gestures in order to use them. Also, additional hardware costs for eye tracking equipment are needed for this type of approaches.

Humans' cognitive ability has also been used as an approach to shoulder surfing defence. For example work in [20] requires users to answer a sequence of challenges posed by the system. Although this scheme provides an outstanding resistance to shoulder surfing attacks, it requires users to perform mentally demanding computation to pass the sequence of challenges and thus reduces the scheme's usability. Following similar thoughts, Weinshall [28] introduced a scheme to defend against eavesdropping attacks including shoulder surfing and spyware. However, Golle & Wagner [9] showed that a SAT solver can defeat Weinshall's design in a few seconds, after observing a small number of successful logins.

Finally, a type of screen filters specially made for mobile devices known as privacy screen protectors uses a polarization technique to enhance the privacy of its users [16]. The screen filter enables users to see from the front but is dark when viewed from the side at an angle of more than 30 degrees. However, this device burdens user with an additional hardware cost, compared to our proposed approach using software only.

3. DRAW A SECRET SCHEME

A DAS password is a free-form picture drawn on an N x N grid. The grid is denoted by discrete rectangular coordinates (x, y) which will be used to indicate the cells that are crossed by the user's drawn secret (password). Figure 1 illustrates an example of a DAS password (taken from [11]), which will be recorded by the system as a sequence of coordinate pairs: (2, 2); (3, 2); (3, 3); (2,

3); (2, 2); (2, 1); (5, 5), where (5, 5) is distinguished as a "pen-up" indicator.

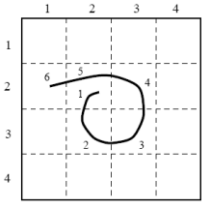

Figure 1. DAS passwords: an example (taken from [11])

In order for a drawn secret to be accepted in authentication, it needs to cross the same grid of cells while ensuring the breaks between the strokes occurring in the same place. The DAS password scheme however gives the user sufficient tolerance provided that the cell sequence follows the same encoding, even though the drawing produced is not exactly the same. A drawing-secret will be disallowed if it crosses through a cell corner or traces the grid lines. This is known as illegal crossings due to fuzzy boundaries [11] as it is difficult to ascertain which destination cell had been intended.

The following notions are important for understanding the DAS scheme, as quoted from our previous work [6]:

> "A *stroke* is made up of a sequence of cell crossings bounded at both ends by pen-up events, exclusive of the pen-ups themselves. For example: the sequence *Pen-up*, (1,2), (1,3), (1,4), *Pen-up* defines a stroke: (1,2), (1,3), (1,4). The *length of a stroke* is the number of coordinate pairs it contains. Therefore, the above stroke has a length of 3.
>
> Since a password in the DAS scheme is a sequence of strokes separated by pen-ups, the *length of a password* is the sum of the lengths of its components strokes, exclusive of pen-ups.
>
> The number of strokes (i.e. the *stroke counts*) and the password length are important security metrics measuring the strength of a DAS password. A high number of strokes or a high password length usually provides a high level of security as such secrets reside in a more secure part of the password space. "

4. THE THREE DEFENCE TECHNIQUES

In this section, we describe the following three techniques that we have designed for protecting DAS and BDAS systems from shoulder surfing attacks:

1) Decoy Strokes
2) Disappearing Strokes
3) Line Snaking

These techniques are not supposed to be used during password enrolment, where a user creates his or her new password. They are enabled **only** during a login procedure. The rationale is that we do not want users to be distracted when they are creating new passwords.

4.1 Decoy Strokes

The idea of using decoy strokes as a defence involves creating real time strokes alongside of a user's password, which are believable enough that they resemble strokes that could have been drawn by a user. The aim of this technique is to distract an onlooker's attention away from the actual password that is drawn by the user. The decoy strokes can make it harder for the user to enter a password. However this needs to be done without confusing the user so much that he cannot enter his password correctly, which would drastically reduce the usability of the system.

This defence is also expected to give the appearance of added complexity to a password by adding extra strokes, without actually modifying the user's password. This technique is expected to add security in particular to simple passwords that could be easily, memorised and replicated by an attacker. Consequently it was decided that the strokes had to be generated randomly and displayed as the user was drawing or otherwise, strokes that were from a library could become repetitive and would therefore be spotted by an attacker as the decoy (hence introducing weakness into the system). Also if the decoy strokes appeared randomly when a user had not started drawing, an attacker would be given further information which would help him/her to distinguish between the user stroke and the decoy stroke. Again, this would make the solution less effective at resisting shoulder surfing.

To keep the usability of DAS, two variables (the colour and thickness of decoy strokes) have been added as a user controlled feature in the prototype system. This controlled feature would enable the users to clearly distinguish between the two strokes, with the intention of keeping the attacker bemused by all the information on the screen. Figure 2 (a) and (b) were created as a story board to show how the output of the defence should appear.

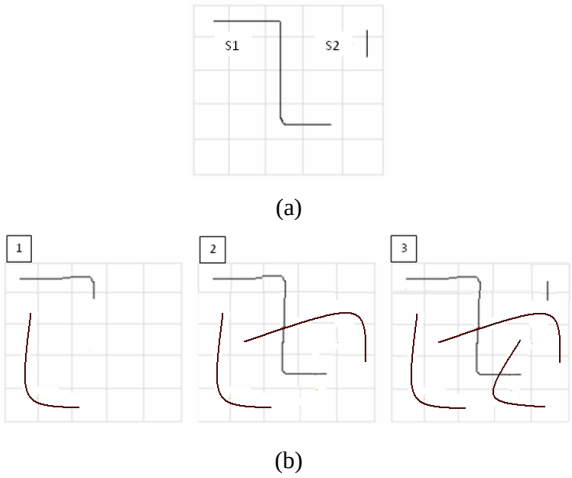

Figure 2. Decoy Stroke defence: (a) when the defence is not activated; (b) when activated (left to right proceeds with time)

Figure 2(a) shows a user password (in black colour) and Figure 2 (b) shows a decoy stroke (in dark brown colour), being drawn. We decided to use dark brown colour for the decoy stroke so as to ensure that it is not too similar to the password stroke (as using same colour might confuse the user and thus affect usability) but at the same time it must not be too different or otherwise the defence would be rendered useless. As the user draws his/her password the decoy stroke is drawn at a similar rate. In the interest of not confusing the user during the login phase, a relatively small number of decoy strokes were used (in this experiment the limit was four decoy strokes) but this could be easily changed or added as a user controlled feature.

In terms of implementation details, a decoy stroke begins at a randomly selected point within the DAS drawing grid that the

user is allowed to draw on with the stylus, and is limited to be of a 'believable' length. The algorithm then randomly generates new co-ordinate points which are to form part of the decoy stroke. However, in order to make the stroke realistic, the distance away from the previous stroke and the direction change had to be limited; although still random the distance away from the previous stroke has a maximum value and a random direction change is only allowed every four points. A typical decoy stroke could entail the following steps (where point refers to a co-ordinate):

1) Generate four random points that lie within grid confinements.
2) Add the points to the stroke data structure.
3) Display part of the stroke, slowly revealing more of the stroke.
4) Repeat while the user is still drawing until the maximum number of decoys has been created.

The decoy strokes are only generated whilst the user is still drawing. This is because if they continue to appear too long after the user had finished drawing a stroke (i.e. when the user had lifted the stylus away from the PDA) then it would be easier for an attacker to distinguish between the user's stroke and the decoy stroke.

In order to make decoy strokes display as realistic smooth curves rather than straight interconnected (jagged) lines, we implemented the cubic Bezier curve fitting algorithm that allows four points to be fitted to a curve – making the stroke look more realistic and 'human-like' and hence more likely to improve shoulder surfing resistance.

4.2 Disappearing Strokes

The disappearing stroke solution entails the user stroke being removed from the screen after it has been drawn. The idea behind this is that the password information of an individual stroke is removed, which gives the attacker less time to store the image to memory. This solution is designed for both passwords that have multiple strokes, and passwords of one long stroke, although it might work better for the former type of passwords.

The stroke was designed to be wiped from the screen only after the user has finished drawing that particular stroke (i.e. when the stylus is removed from the screen). This was designed using a timer whose purpose was to remove the stroke after a certain period of time (after the pen up event).

Figure 3(a) shows an example of DAS password without any defence while 3(b) shows how the output of the defence technique should function as time proceeds. The variable factor in this defence technique is the amount of time that has to elapse before the stroke disappears from the screen should be kept within a restricted range. This is to ensure that the shortest time does not affect usability and the longest time does not have an adverse outcome on the effectiveness of the defence. A possible implementation for this defence could be starting a timer after a user pen up event, which would clear the stroke from the screen after a certain time period.

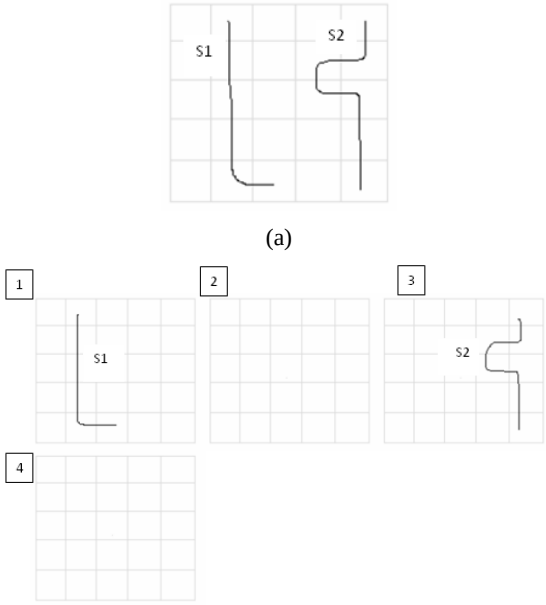

Figure 3. Disappearing Stroke defence: (a) when the defence is not activated; (b) when activated (left to right proceeds with time)

4.3 Line Snaking

This defence is based on the disappearing stroke solution but was intended to leave the vital password information onscreen for an even shorter period. An attacker is thereby not given a chance to see a complete user stroke onscreen. It would involve the start of the user stroke being removed from the screen as the user is still drawing, giving the appearance of the line snaking towards the user's stylus.

The line snake defence was designed to combat shoulder surfing for passwords containing long singular strokes. Hence, allowing stroke information to be removed from a long singular stroke, whilst the stroke is still being drawn. The variable factor for this solution was decided upon as being the speed at which the user stroke disappears (or snaking away) from the screen. This again was thought to be limited within a sensible range as to maintain the usability and effectiveness of the solution (i.e. otherwise the stroke may be removed too quickly or too slowly).

A possible implementation for this defence was decided as starting a timer from when the user begins drawing the stroke (which controls the line snake). A simple procedure removes points from the beginning of the stroke. This procedure is called every time the timer ticks. This would give the appearance of the line snaking towards the stylus's current position. Figure 4(a) shows an example of a DAS password without any defence while 4(b) shows how the output of the defence technique should function as time proceeds.

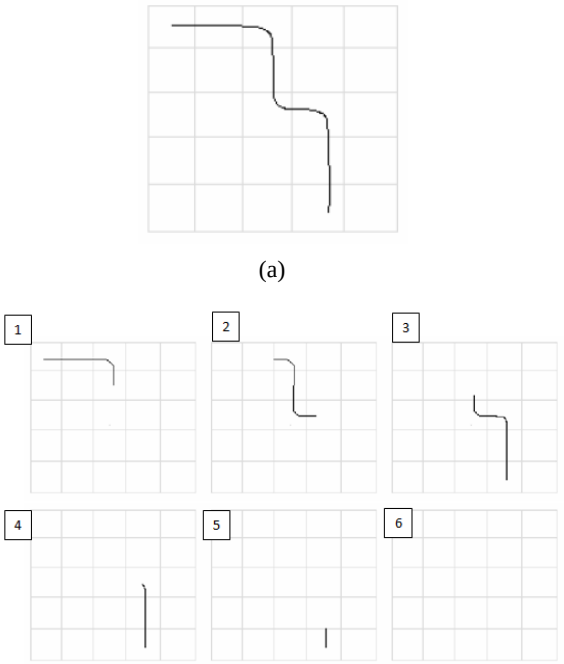

Figure 4. Line Snaking defence: (a) when the defence is not activated; (b) when activated (left to right proceeds with time)

5. USER STUDY 1

The aim of this user study is to determine the strongest defence technique among the three being proposed by conducting a controlled laboratory experiment. Before this user study was conducted, an approval was obtained from our University Ethics Committee (UEC). In contrast to previous work [25], we considered more than one password whereby our proposed defence techniques were tested across three levels of password difficulties (weak, medium and strong). In order to determine which technique offers the best shoulder surfing resistance, we decided to use a *between-subjects* design. Although this type of design requires more participants, it ensures that the exact same passwords were used in each experiment condition so that they would not be a confounding factor biasing the results. The main independent variable for this experiment is the defence technique, and we also explore the password strength as a secondary independent variable, whereas the participant's response is the dependent variable. Our hypotheses were as follows:

H_1 – *Disappearing Strokes* will not work as well as *Line Snaking*, as for the latter the strokes are snaking away while being drawn, leaving a very short time for the strokes to stay visible on the screen and be observed.

H_2 – The defence of *Decoy Strokes* is the weakest as all the strokes of the password still stay visible on the screen.

5.1 The Apparatus

A prototype of the DAS graphical password system was implemented together with all three shoulder-surfing resistance techniques, on a major-brand PDA. A 5x5 grid for the DAS scheme was chosen as a previous study [27] had shown that this size would provide a good balance between usability and security. The PDA has a 3.5 inch TFT active matrix display with dimensions of 2.9in x 0.7in x 4.7in (W x D x H) and a 240 x 320 display resolution. The proposed techniques were specifically tested on the DAS instead of the BDAS scheme for the following reasons:

1) To make sure that results produced reflect the effectiveness of the proposed techniques and are not due to other factors such as the background in the BDAS scheme.
2) To keep the instructions given to our participants as simple as possible so that the experiment is less burdensome and more fun to participants, avoiding biases in the results caused by participants who would be otherwise bored by the experiment and thus act less naturally – the DAS is less complicated to explain than the BDAS scheme.

5.2 Password Choices

It is interesting to see what effects each defence technique will have on passwords of different strength. Hence, three passwords of different security levels (weak, medium and strong) were tested for each experiment condition. The strength of a DAS password can be determined by its length and stroke count when the drawing grid size is fixed.

Table 1. Passwords choices and configurations [26]

Password category	Password length	Stroke count	Bit-size strength
Weak	7	3	~25.1
Medium	10	5	~39.1
Strong	13	7	~53

The configuration of the password length and stroke count for each security level, together with the bit-size strength of each level, is chosen as in Table 1. That is, our weak password has a search space of about 25 bits, the medium password about 39 bits. The strong password level is about 53 bits, equivalent to the strength of strong 8-character text passwords.

In order to maintain ecological validity of this experiment, the passwords tested must be memorable; otherwise they would be less likely to be chosen in the real world. As suggested by previous studies [11, 26] DAS users tended to employ centring and symmetry in their drawing as typical ways to help remember their secrets. Global symmetry and centring would lead to weaker DAS passwords. These are more likely to be vulnerable to dictionary attacks. In the case of BDAS, such weaknesses can be compensated for by the relatively larger bit size of BDAS passwords. Hence, it was speculated that in real life scenarios, people will still prefer passwords that exhibit symmetry and centring (at least to some extent). We therefore emulated these characteristics in our chosen passwords. On the other hand, this choice also arguably serves one of our experiment rationales: we aim to empower the shoulder surfer with an optimal attack setting. Figure 5 shows the three passwords we have chosen for the experiment.

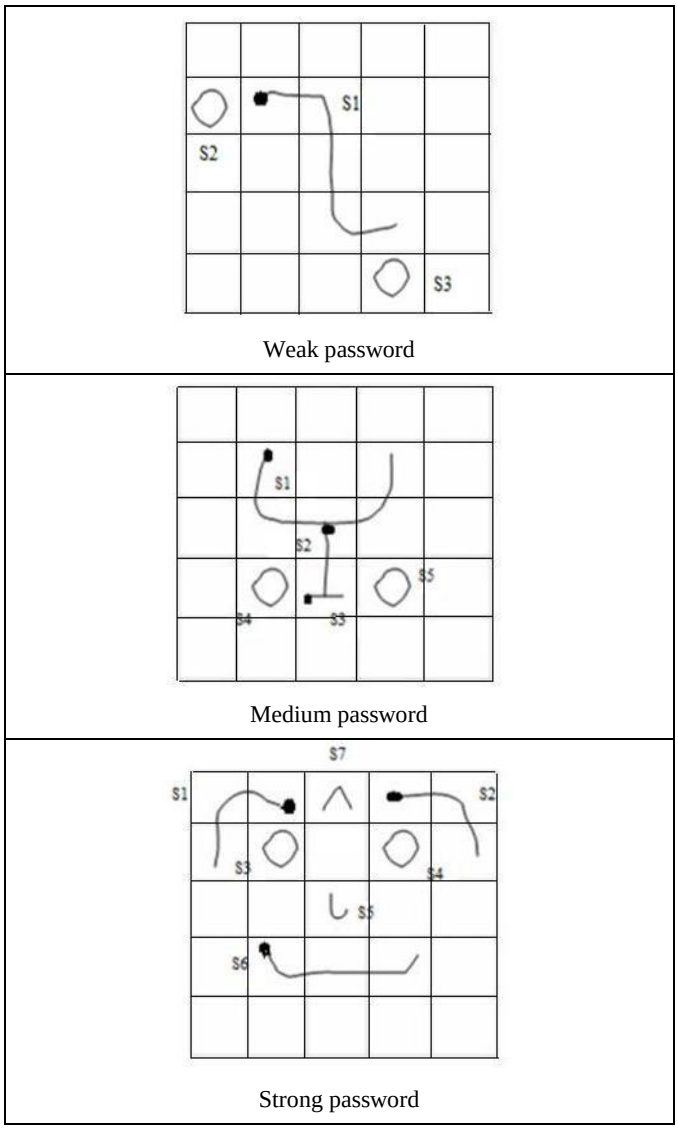

Figure 5. Three passwords of different strength used in the experiment. (Note: Si indicates the order of strokes and the black dots indicates the starting point of a stroke)

5.3 Procedures

The experiment was conducted in a controlled laboratory environment to avoid any distractions. Sixty-eight students were recruited for this experiment, 43 male and 25 female. Each participant was offered a printing credit of 50 A4 pages, which can be used on campus as a reward for completing the experiment.

At the beginning of the experimental session: Informed consent and demographic information was obtained from each participant. They were asked to provide their age, gender, educational background, and experiences using PDAs. The mean age for the participants was approximately 28 years with a standard deviation of 7.9. The majority of the participants (56) came from a technical background while the remaining came from a non-technical background. Participants from the technical category included university students from science and engineering while the non-technical category came from business and social sciences disciplines. Less than half (27) of the participants had used a PDA frequently. The participants were randomly assigned to one of the following four experimental groups:

- Decoy Stroke
- Disappearing Stroke
- Line Snaking
- Control Group (The undefended DAS scheme)

Each group had equal numbers (17) of participants who each spent approximately 15 minutes to complete the experiment. Reviewing the issues rose in previous work [15] regarding ecological validity, this experiment attempted to simulate shoulder surfing scenario. The participants were asked to play the role of shoulder surfers, trying to steal 3 passwords by observing them during individual login attempts. To ensure they acted as shoulder surfers (a true shoulder surfer would have the intention to steal the passwords), an additional incentive to increase their motivation was offered. The participants were told that there will be a competition between them. There will be only one prize given to the participant who performs the best from the best defence technique group.

During the experiment: Each participant was given a brief introduction to the DAS graphical password scheme. For participants who were assigned to the treatment groups, extra information about the assigned defence technique was also provided. This was done under the assumption that the shoulder surfers are aware of the defence technique being employed and are equipped with similar knowledge. Printed information was also supplied to support the briefings. Participants were highly encouraged to ask the experimenter any questions especially on how to construct the passwords as they need to reproduce them later in the experiment. Then, a quick demonstration on how the prototype system works was shown to the participants. The participants were then allowed to get a quick hands-on experience using the prototype system. A short training on shoulder surfing was then conducted using the same password as shown in the example provided in the printed information given to them earlier. It was decided that the experimenter acted as "the victim" to the shoulder surfing attack throughout this experiment. The reason for having just one person (the experimenter) being the victim is to reduce inconsistency-bias produced by two different person's login skills from effecting the results. The experimenter also underwent sufficient training to ensure constant speed in drawing the passwords. The training was proven sufficient as the experimenter managed to conduct the login procedure without any failures in all sessions undertaken during the experiment. The experimenter remained seated throughout. Also important to note is that the experimenter was not trying to cover up the PDA screen or applying any defence technique other than the one being tested. The purpose of having this scenario is to have a tight control so that the victim has no other protection mechanism (other than the one being tested) – although in real life situation, PDA users might tilt the screen to avoid from being seen.

Each participant played the role of shoulder surfer. They had free roam of the laboratory room and were asked to choose an optimal viewing position (on the left of "the victim" as the experimenter is a right-handed person). The participants were given only a single chance to observe each login session. The rationale behind this design is to emulate a casual shoulder surfer. The participants were allowed to take notes on their observations. The details of the experimental task that a participant carried out are the following:

- The experimenter attempted to login by drawing the passwords (one password at a time) on the PDA screen and then clicking the login button.
- After a password was drawn, the participant was asked to reproduce on a piece of paper containing (5x5) grid lines (similar to the interface of the DAS prototype) the password that he/she had captured by observation.
- Then, the participant was asked to play a mini jigsaw puzzle game for about one minute. This was to help clear up their recent memory on the password that they had attempted to shoulder surf, getting rid of potential interference that might cause to the next password.
- The same procedure was repeated for the second and third passwords.

At the end of the experiment: Before leaving the room, each participant handed over the papers (containing the three passwords as they have captured) to the experimenter.

5.4 Results & Analysis

In this experiment, all the participants successfully completed their given tasks. A password that is captured by a shoulder surfer is considered correct only if it contains all the following elements (an example is given in Figure 6):

- Strokes are in the correct order (2 strokes)
- Strokes are in the correct direction (as shown by the arrows)
- Strokes are the correct length (that is, the number of cells the stroke cuts through).

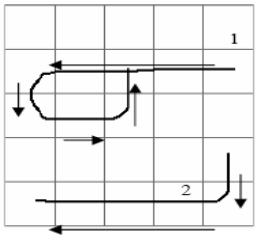

Figure 6. An example of how DAS password is correctly reproduced (hand-drawn)

A stroke by stroke analysis using the three elements stated above was performed on each password captured for each participant. For example, the weak password we tested has a stroke count of 3. Therefore the possible outcome for a shoulder surfing attack on this password can be 0 (an entire failure), 1/3, 2/3 or 3/3 (a complete success).

We will present results at two levels. First we examine the overall strength of the defences, by looking at the proportion of strokes that were captured by shoulder surfing.

Secondly, we examine the immediate threat of successfully collecting complete DAS passwords by shoulder surfing, with and without the protective techniques. We examine cases where shoulder surfing could not capture complete DAS passwords, and we investigate how much information in the password was still possible to recover by shoulder surfing. Following Lewis & Sauro [13], when reporting task completion (e.g. successfully stealing DAS passwords) we show 95% confidence intervals using the Adjusted Wald method. Moreover, we adjust for multiple comparisons using the Bonferroni method by setting the criterion value for significance to 0.0042 for the 12 inferential statistics reported in this section, to achieve an overall alpha of 0.05.

Table 2. Proportion of DAS password strokes stolen, reported according to defence used

Password Strength	Proportion of strokes shoulder surfed Mean (SD)			
	DAS only (Control)	Decoy Stroke	Disappearing Stroke	Line Snaking
Weak	1.00 (0)	0.86 (0.29)	0.69 (0.28)	0.59 (0.42)
Medium	0.8 (0.3)	0.79 (0.33)	0.45 (0.36)	0.41 (0.37)
Strong	0.52 (0.35)	0.7 (0.3)	0.33 (0.29)	0.2 (0.29)
Overall (n=51)	0.77 (0.33)	0.78 (0.31)	0.49 (0.34)	0.4 (0.39)

Table 2 shows the total proportion of strokes that were stolen for each defence type. They appear to fall into two groups – DAS only (the control group that had no defences) with Decoy Stroke, with approximately 77% of strokes captured, and then Disappearing Stroke and Line Snaking defences, that have between 40% and 50% of strokes stolen.

This characterization was confirmed by testing for differences between the defences using non-parametric Mann-Whitney U tests. U tests were chosen because the distribution of proportions was highly skewed and so not suitable for parametric tests. To be efficient with the total number of comparisons in this study, only three comparisons were made – testing for differences within each of the two groups (e.g. Control vs. Decoy Stroke, and Disappearing Stroke vs. Line Snaking), then between a member of each group (Decoy Stroke vs. Disappearing Stroke). No significant difference in proportions were detected between DAS only and Decoy stroke ($U=1285$, $z=-0.11$, $p=0.91$, $r=-0.01$), showing that Decoy Stroke did not offer any protection.

A statistically significant difference was found between Decoy Stroke and Disappearing Stroke ($U= 699.5$, $z= -4.1$, $p<0.0005$, $r=-.41$), showing that Disappearing Stroke offered improved strength compared to an undefended DAS. No statistically significant difference was found between Disappearing Stroke and Line Snaking ($U=1087.5$, $z= -1.4$, $p=0.149$, $r=-0.14$), showing that Line Snaking offered equivalently good strength.

We examine the effect of password strength on the proportions of strokes stolen, using Wilcoxon tests. These are non-parametric tests used for related data when there are only two groups. We use this in preference to an ANOVA because the data is highly skewed and would violate ANOVA's assumptions. Weak passwords were found to have a statistically significantly higher average proportion of strokes stolen (0.78 strokes, SD=0.33) than did Medium strength passwords (0.61 strokes, SD=0.38) ($z=-4.1$, $p<0.0005$, $r=-0.35$). Medium passwords had a statistically significantly higher proportion of password fragments stolen than Strong passwords (0.44 strokes, SD=0.36) ($z=-4.8$, $p<0.0005$, $r=-0.41$).

We now examine the immediate threat of successfully collecting complete DAS passwords by shoulder surfing.

Table 3 shows the numbers of passwords that were successfully stolen for each of the three password strengths. The pattern shown with proportions of strokes stolen is repeated here. Overall, 32 out of 51 DAS passwords were stolen (63%) in the undefended control group (which is compatible with a true rate in the population of between 25 and 38 passwords stolen). Decoy Stroke offered equivalent performance to the control group, with

29 out of 51 passwords stolen (57%) (compatible with a true rate in the population of between 22 and 35 passwords out of 51).

Table 3. Number of DAS passwords correctly shoulder surfed (out of 17) through each defence technique, with 95% adjusted Wald confidence intervals

Password Strength	DAS Passwords shoulder surfed Count (CI)			
	DAS only (Control)	Decoy Stroke	Disappearing Stroke	Line Snaking
Weak	**17** (14-17)	**13** (9-15)	**6** (3-10)	**7** (4-11)
Medium	**11** (7-14)	**11** (7-14)	**4** (2-8)	**3** (1-7)
Strong	**4** (2-8)	**5** (2-9)	**0** (0-3)	**0** (0-3)
Overall (n=51)	**32** (25-38)	**29** (22-35)	**10** (6-17)	**10** (6-17)

We now examine our two hypotheses about the numbers of passwords that would be stolen with each defence technique.

H_1 was that the Disappearing Stroke defence would not work as well as Line Snaking defence, as for the latter the strokes are snaking away while being drawn leaving a very short time for the strokes to stay visible on the screen. H_1 was not supported by the data – these defences had identical overall performance – allowing only 10 passwords out of 51 to be stolen each.

H_2 was that the Decoy Strokes defence would be the weakest of all defences, as it allowed strokes to remain on screen. This was tested using two Mann-Whitney U tests, which are appropriate tests where two unrelated groups of non-normally distributed data are being compared such as the password stolen/not stolen outcomes we are examining. The hypothesis was supported – a statistically significant difference was found between Decoy Stroke (with 29 out of 51 passwords stolen) and Disappearing Stroke (with 10 out of 51 passwords stolen) (U=699.5, z=-4.15, p<0.0005, r=-0.41). Disappearing Stroke has identical performance to Line Snaking, so Decoy Stroke is the weakest of all defences.

We further tested to see if the Decoy Stroke (with 29 out of 51 passwords stolen) performed better than having no defence at all (the control group, with 32 out of 51 passwords stolen). Using a Mann-Whitney U test again we found there was no statistically significant difference) (U=1285.5, z=-0.11, p=0.91, r=-0.01).

We also examined the effect of password strength. Stronger passwords should be stolen less frequently than weaker passwords. We tested this hypothesis using a Friedman test to check for a main effect of password strength. Friedman tests are appropriate for non-parametric data such as the password stolen/not stolen binary outcomes in this experiment, and are used when there are more than two groups and the data in each group are related. This is the case here, where each participant attempts to steal all three password types. A statistically significant effect of password strength was detected (Chi-square=51.5, N=68, df=2, p<0.0005).

Wilcoxon tests were used post-hoc to determine which password strengths were different from each other in numbers of passwords stolen. Wilcoxon tests are used for related non-parametric data when there are only two groups. Weak passwords were found to be stolen in statistically significantly larger amounts than Medium strength passwords (z=-3.7, p<0.0005, r=-0.32). Moreover, Medium strength passwords were found to be stolen in statistically significantly larger amounts than Strong passwords (z=-4.5, p<0.0005, r=-0.39).

We now examine differences between the defence techniques in their ability to prevent partial theft of DAS passwords. Table 4 gives the numbers of passwords of each type that were partially stolen by participants. We noted previously that DAS only and Decoy strokes had equivalent numbers of passwords completely stolen, but far more than both Disappearing Stroke and Line Snaking defences, which had approximately the same numbers of password completely stolen as each other. We restrict our comparisons only within those two groups of defences to avoid confounding. Mann-Whitney U tests were used, as they are suitable for comparing non-parametric data in two unrelated groups.

Table 4. Number of DAS passwords that were partially stolen, with 95% adjusted Wald confidence intervals

Password Strength	DAS Passwords shoulder surfed Count (CI)			
	DAS only (Control)	Decoy Stroke	Disappearing Stroke	Line Snaking
Weak	**0** (0-4)	**3** (1-7)	**11** (7-14)	**6** (3-10)
Medium	**6** (3-10)	**5** (2-9)	**11** (7-14)	**9** (5-13)
Strong	**11** (7-14)	**11** (7-14)	**12** (8-15)	**8** (4-12)
Total (out of 51)	**17** (11-24)	**19** (13-26)	**34** (27-40)	**23** (16-30)

There was no statistically significant difference between the number of passwords partially stolen through the Decoy Strokes defence (19, from 51 passwords) than from the undefended control group (17, from 51 passwords) (U=1249.5, z=-0.41, p=0.68, r=-0.04). No statistically significant difference was found between the numbers of passwords partially stolen through Disappearing Stroke (34, from 51 passwords) and through Line Snaking (23, from 51 passwords) (U=1020, z=-2.2, p=0.029, r=-0.22) (although the difference had been statistically significant before our Bonferroni adjustment to the criterion value for significance). Previously, we found that Line Snaking and Disappearing Stroke allowed equal numbers of passwords to be fully stolen. When partial thefts are taken into account, it appears that both defences are still equal.

Table 5. Number of DAS passwords that were completely resistant to shoulder surfing, with 95% adjusted Wald confidence intervals

Password Strength	DAS Passwords shoulder surfed Count (CI)			
	DAS only (Control)	Decoy Stroke	Disappearing Stroke	Line Snaking
Weak	**0** (0-4)	**1** (0-5)	**0** (0-4)	**4** (2-8)
Medium	**0** (0-4)	**1** (0-5)	**2** (0-6)	**5** (2-9)
Strong	**2** (0-6)	**1** (0-5)	**5** (2-9)	**9** (5-13)
Total (out of 51)	**2** (0-7)	**3** (1-8)	**7** (3-13)	**18** (12-25)

We investigated this further by comparing the defences against the numbers of passwords that could not be stolen through them (Table 5), focussing our inferential tests only on Line Snaking and Disappearing Stroke. We used a Mann-Whitney U test again, and for the same reasons. Although the Line Snaking defence

appeared to defend more passwords completely (18 of 51) than Disappearing Stroke (with 7 passwords fully defended out of 51) again, after Bonferroni adjustment no statistically significant difference was found – (U=1020, z=-2.5, p=0.012, r=-0.25). However, an effect of this size (r = -0.25) is substantial – approaching Cohen's criterion [2] value of 0.3 for a "medium" sized effect. Line Snaking and Disappearing Stroke should be subject to more focused study, in order to achieve greater statistical power.

As a separate experiment, we conducted a more focused study involving only the Line Snaking and Disappearing Stroke group. We recruited 34 participants, and each group had 17 participants. It is important to note that all the procedures and apparatus used in the previous experiment were maintained exactly the same, but all the participants of this focused study did not participate in our previous experiment. In general, slightly more than half of the participants in this focussed study came from the technical background and were female. The average age of the participants was approximately 29 years old with less than half of them having had some experience using PDAs previously.

The distribution was found to be highly skewed; hence the Mann Whitney U test was used to compare the two groups. A statistically significant difference was detected (U= 72.5, z=-3.57, p=0.001, r=-0.31) which indicates that differences exist between the Disappearing Stroke and Line Snaking technique.

Next, we examine the number of passwords that were successfully shoulder surfed. The Disappearing Stroke group managed to successfully shoulder surf 14% of the passwords compared to the Line Snaking group with none. This was found to be statistically significant (U= 102, z=-2.38, p=0.017, r=-.41). In terms of the number of passwords that was completely resistant to shoulder surfing, Line Snaking managed to have about 12% out of the total passwords completely un-captured compared to Disappearing Stroke with just 4%. However, we did not find this difference to be statistically significant (U= 102, z=-1.71, p=0.087, r=-.29). As such, we find that Line Snaking outperformed Disappearing Stroke, and hence provided better protection.

5.5 Discussion

The Decoy Stroke defence technique performed the worst in our experiment, and this can be explained that it allows all the strokes to remain on screen visibly during the whole login process. We expected that this defence technique would work to some extent. However, it turns out that this technique provided little protection, as Tables 2-4 all show similar performance between the Decoy Stroke group and the undefended DAS group, and applicable statistical tests did not yield any significant difference between these two groups. In some circumstances, the Decoy Stroke group even performed worse than the undefended DAS group. This discrepancy can be likely explained as follows. The decoy strokes helped the attackers to locate and remember the legitimate strokes that were intended to be obfuscated by the decoys. Why so? Relative positions between decoys and legitimate strokes can be exploited by the attackers to aid their locating starting and ending cells of legitimate strokes in the drawing grid, for example.

6. USER STUDY 2

We conducted a separate experiment to evaluate the usability of our defence techniques across three levels of password difficulty (weak, medium and strong). It was decided that decoy stroke defence technique should be excluded in this study as the results obtained from Study 1 indicated that this technique provided little protection – our research aims to find a technique that is good in terms of both security and usability. This user study used a *within subject* design, which means that each participant is assigned to all of the following experimental conditions:

- Disappearing Stroke
- Line Snaking
- Undefended DAS scheme (control group)

Our hypotheses were as follows:

H_1 – It takes more time for users to login when *Line Snaking* is enabled, than when *Disappearing Stroke* is enabled.

H_2 – *Line Snaking* technique will cause a higher login error rate than *Disappearing Stroke* does.

H_3 – *Disappearing Stroke* technique is more preferred by the users compared to the *Line Snaking* technique (as in the latter the line starts snaking away while users are still drawing).

Similar to User study 1, this study was approved, before carried out, by our University Ethics Committee (UEC) as a study with minimal risk.

6.1 Procedures

The experiment was conducted in a controlled laboratory environment to avoid any distractions. The same apparatus used in the first user study were retained (refer to *Apparatus* in section 5.1 of the User Study 1). Thirty participants were recruited for this experiment, 20 male and 10 female. For their participation, each participant was offered £10. Informed consent and demographic information were obtained from each participant. The participants were asked to provide their age, gender, educational background, and experiences using PDA. The mean age for the participants was 31.1 with a standard deviation of 9.1. More than half of participants (18) came from technical backgrounds while the remaining came from non-technical backgrounds. Participants from the technical category included university students from science and engineering while the non-technical category came from business and social sciences disciplines. More than half (19) of the participants reported having used PDAs frequently.

The experiment began with an introductory session where participants were given a brief explanation of the DAS system and all the defence techniques. Printed information was also supplied to support the briefings. This was followed by a short demo to show how the system works and a quick hands-on was allowed to ensure the participants had some experience using the system prototype.

As mentioned, this study aims to evaluate the two defence techniques across three levels of password difficulty (weak, medium and strong). Throughout this experiment, the same three passwords were used. The three passwords used in User Study 1 were retained in this study (refer to Passwords Choices in section 5.2-User Study 1). The reason for controlling the password choice rather than allowing the participants to create their own ones is to avoid bias caused by different password choices. All the three passwords were drawn on a separate piece of paper with (5x5) gridlines similar to the prototype system. Participants were shown one password at a time beginning with weak followed by medium

and then the strong password. With each password, the participants were instructed to perform the following tasks:

- Treat the shown password as theirs.
- Get familiar with the password by using it to login to the system several times. In order to ensure a consistent amount of training, each participant was allowed approximately 10 minutes (this amount of time was found to be adequate during a pilot study we conducted).
- Once the training period was over, the participants returned the paper containing the password to the experimenter.
- Participants were then instructed to login using the passwords they had used in the training session, for each of the experimental conditions. To minimise the training effect caused by the same password used, the experimental conditions were arranged in a random order. Time taken to login and login error rate were recorded.
- Then, the participant was asked to play a mini jigsaw puzzle game for about one minute. This was to help clear up their recent memory on the password they used, before moving on to the next password.
- The above procedure was repeated for the second and third passwords.

After participants had completed all the tasks with all the three passwords, they were asked about which defence technique they preferred and why. All answers and comments given were noted.

6.2 Results & Analysis

In this study, all the participants successfully completed their given tasks. Outcome metrics we measures were the following:

- *Login time*: time taken to complete a successful login. If a participant had to make several attempts, his/her login time was measured as the sum of the time taken by each attempt.
- *Login error rate*: measured by the number of attempts taken to complete a successful login.

Table 6 shows the mean and standard deviation of login time (in seconds) for all techniques across three levels of password difficulty.

Table 6. Mean and standard deviation for login time (in seconds) for all techniques across three levels of password difficulties (N=30)

Password strength / Techniques	Weak	Medium	Strong
	Mean (SD)	Mean (SD)	Mean (SD)
DAS only	4.5 (0.52)	5.4 (0.73)	7.5 (0.54)
Disappearing Stroke	5.3(1.06)	6.5 (1.15)	9.6 (2.38)
Line Snaking	5.9 (1.35)	7.9 (2.62)	12.4 (4.42)

From defence techniques perspective, Line Snaking has the highest mean login time, followed by Disappearing Stroke and DAS only (undefended group). This suggests that Line Snaking technique poses some challenges for participants to complete the login task. On the other hand, from password level perspective, strong password requires more time to login compared to medium and weak passwords. This indicates that strong passwords require more time to login compared to the other two password levels. A two-way *within* subjects (repeated measures) ANOVA test was performed to compare the interaction effect of password levels on the defence techniques applied. This test was chosen since the same participant was exposed to all the conditions, and the data are ratio data and approximately normally distributed, while the test is robust against violations of homogeneity when the sample sizes are equal. The result (with a Greenhouse-Geisser correction) shows a statistically significant interaction effect ($F(2.03, 58.75) = 12.84$, $p=0.001$). This effect tells us that time taken to login using the techniques applied are different across the three levels of passwords. Paired sample t-tests were then used to make post hoc comparisons between conditions. All the nine t-tests show significant differences with ($p \leq 0.005$) indicating that password levels really do have an effect on the techniques used, where increasing password strength produces a larger increase in login time for Disappearing Stroke than for unprotected DAS, and a larger increase still for Line Snaking than for Disappearing Stroke. The results above show that our Hypothesis 1 is supported.

Table 7 summarises the number of attempts taken to complete a successful login for all techniques across the three levels of password difficulty. A Friedman test was used to assess the numbers of login attempts required across techniques, because the distributions of numbers of login attempts were not normal, and the observations were related. We found a statistically significant overall difference between the different techniques (Chi square=32.2, df=2, $p<.0001$). Wilcoxon signed ranks tests were used to test for difference in the numbers of login attempts required between each pair of techniques, because the data was not normal, though was related. The Line Snaking technique required statistically significantly more attempts to login than both undefended DAS ($z=4.7$, $p < .0001$) and Disappearing Stroke ($z=3.4$, $p = .001$). Disappearing Stroke also required more login attempts than the undefended DAS ($z=2.8$, $p=.005$). These results supported our Hypothesis 2.

Table 7. Mean for number of attempts taken to complete a successful login for all techniques across three levels of password difficulties (N=30)

Password strength / Techniques	Weak	Medium	Strong
DAS only	1.0	1.0	1.0
Disappearing Stroke	1.1	1.1	1.1
Line Snaking	1.1	1.3	1.4

Qualitative data we collected such as participants' preferences and additional comments on the techniques used also revealed interesting results. Most of the participants (77%, i.e. 23 out of 33) preferred the Disappearing Stroke technique, 10% (3 out of 30) of the participants preferred the Line Snaking technique, and the remaining (13%, 4 out of 30) did not have any preference. These results indicated that Hypothesis 3 is supported.

6.3 Discussion

User study 2 has shown that the Disappearing Stroke defence is generally more usable compared to the Line Snaking defence. As shown in Table 6, participants require more time (8.7 sec) on average to login and also use more attempts to login (1.3 sec) on average using Line Snaking compared to the other techniques. These create some usability challenges for the Line Snaking technique. The challenge is more obvious with strong passwords compared to medium and weak passwords. A possible explanation for this is the following. The participants were not given the freedom to choose their own passwords. Should they have this flexibility, different results might have been yielded. However, such a hypothesis requires a different experimental design, and is

a next step in the investigation of these defence techniques, building upon the first proof their effectiveness as reported here.

The results have also shown an interaction effect was detected between password levels and the defence techniques upon time taken to login, such that a combination of password and technique choice will determine login time. An implication of this is the possibility that users might be discouraged to use stronger passwords especially when using the Line Snaking technique.

The participants who preferred Disappearing Stroke did so for different reasons:

- More than 65% of them (15 out of 23 participants) stated that they preferred this technique because they felt more comfortable and confident while drawing the passwords, as the stroke only starts to disappear the moment the stylus is pulled up. This aspect is important as the previous stroke will become a vital reference point to draw the adjacent strokes.

  10 out of these 15 participants (67%) instantaneously spotted the security advantage of the Line Snaking technique. They all mentioned that the snaking effect quickly removed the strokes from the screen, which was obviously good for security defence. They realised the security advantage of LS, but because feeling more comfortable/usable with DS is more important to them, they chose DS as their favourite.

- The remaining 8 participants (35%) pointed out that the "*snaking stroke*" (as an effect of Line Snaking technique) was actually annoying and distracting, making them prefer the technique less.

The participants who did not have a preference between the techniques commented that as long as they knew their password, the techniques applied do not affect them. To confirm this, we cross-checked their performance on each technique and found that average difference of performance between the techniques was indeed small for these participants. This result possibly supports the earlier assumption that if users are highly familiar with their passwords the defence techniques might not affect their performance significantly.

7. CONCLUSION & FUTURE WORK

We have presented three new shoulder surfing defence techniques designed for recall-based graphical passwords, as well as two experimental evaluations of these techniques. The *Line Snaking* defence technique has the best overall performance in terms of defence, since the strokes of the 'password' are snaking away while they are being drawn leaving a very short time for the strokes to stay visible on the screen. The *Disappearing Stroke* technique is the second best defence. However, in many circumstances, both techniques worked equally well (see Tables 2-4).

We expected the *Decoy Stroke* technique to provide some defence against shoulder-surfing, but it turns out to achieve little protection. The reason is likely that all the password strokes stay visible on screen, and the decoys do not work well to distract the attackers. It is possible that the Decoy Stroke defence technique could be improved by introducing extra decoys, or the way the decoys are introduced, but this should be done with careful consideration and evaluation as it is important not to confuse the user. However, further work is required to establish this.

We conducted a second user study to compare the usability of *Line Snaking* and *Disappearing Stroke* techniques. In general, our results suggest that Disappearing Stroke is preferred by users, compared to the Line Snaking technique. Both average login time and the login error rate for Line Snaking were also higher than for Disappearing Stroke, indicating that the former imposes greater usability challenges for the users. However, our results also reveal that for some users, there is a possibility that usability will not be affected by the defence techniques applied especially when the users are highly familiar with their passwords. Further research should be conducted to further investigate this issue.

Reviewing the results from both user studies, we conclude that although Line Snaking has better defence performance, the Disappearing Stroke technique is more appropriate for general deployment, since it offers reasonable protection and good usability, and it is also preferred by the users. However, with regards to technique choice, it is possible that users themselves should be able to decide on which defence technique to apply depending on their situation, as our results indirectly show that although users reported Disappearing Stroke as more comfortable to use, they immediately spotted the security advantage of Line Snaking.

Our techniques and experimental results are directly relevant to other graphical password schemes such as Background Draw a Secret (BDAS) [6] and Pass-Go [24]. However, it is useful future work to empirically evaluate the effectiveness of these defence techniques on each of the schemes. Finally, it is interesting to see how the defence techniques can be combined together to provide better defence and at the same time maintain good usability.

Although our work provides a practical, low-cost and deployable shoulder surfing defence for recall-based graphical password systems, it is vulnerable to shoulder-surfing attacks equipped with a video camera. It is import future work to investigate other shoulder-surfing defence mechanisms that are invulnerable to camera attacks. An apparent direction is to combine our approaches with haptic input devices. Another worthwhile direction is to further investigate truly usable zero-knowledge interaction based techniques.

8. ACKNOWLEDGMENT

We thank Ahmad El Ahmad, Su-Yang Yu, Wayne Smith, David Greathead, Zoe Andrews and Ju Helen Wong for many useful discussions and constructive feedbacks. We thank Ayad Keshlaf and Su-Yang Yu for their valuable experiment support, and thank all the participants of our experiments. Comments from Robert Biddle and anonymous reviewers helped improve this paper.

9. REFERENCES

[1] D. Angeli, M. Coutts, L. Coventry, G. I. Johnson, D. Cameron, and M. H. Fischer. 2002. VIP: A Visual Approach to User Authentication. In Proc. of the Working Conference on Advanced Visual Interfaces (Trento, Italy, May 22-24, 2002). ACM Press, New York, NY, 316-323.

[2] R. Biddle, S. Chiasson, and P. C. V. Oorschot. 2011. Graphical Passwords: Learning from the First Twelve Years. *ACM Computing Surveys*, to appear.

[3] Cohen, J. 1992. A Power Primer. *Psychological Bulletin*, 112, 1 (1992), 155-159.

[4] A. De Luca, R. Weiss and H. Drewes. 2007. Evaluation of Eye-Gaze Interaction Methods for Security Enhanced PIN Entry. In Proceedings of the Computer Human Interaction Special Interest Group (OzCHI) (Adelaide, Australia,November 28-30, 2079), ACM Press, 199-202.

[5] R. Dhamija and A. Perrig. 2000. Deja Vu: A User Study Using Images for Authentication. In Proc. of the 9th USENIX Security Symposium (Denver, Colorado, USA, August 14-17, 2000). pp 4-19.

[6] P. Dunphy and J. Yan. 2007. Do Background Images Improve "Draw a Secret" Graphical Passwords? In Proc. of the 14th ACM Conference on Computer and Communications Security (Alexandria, Virginia, USA, October, 28-31, 2007). ACM Press, New York, 36-47.

[7] A. Forget, S. Chiasson, and R. Biddle. 2010. Shoulder-Surfing Resistance with Eye-Gaze Entry in Cued-Recall Graphical Passwords. In Proceedings of the 28th International Conference on Human Factors in Computing Systems (Atlanta,GA, USA, April 10-15, 2010). ACM Press, New York, 1107-1110).

[8] J. Goldberg, J. Hagman, and V. Sazawal. 2002. Doodling Our Way to Better Authentication. In Proc. of Human Factors in Computing Systems (Minneapolis, MN, USA April 20 - 25, 2002). ACM Press, New York, 868-869.

[9] P. Golle and D. Wagner. 2007. Cryptanalysis of a Cognitive Authentication Scheme. In Proc. of the IEEE Symposium on Security and Privacy (Oakland, California, USA, May 20-23, 2007) IEEE Computer Society, 66-70.

[10] E. Hayashi and N. Christin. 2008. Use Your Illusion: Secure Authentication Usable Anywhere. In Proceedings of the 4th Symposium on Usable Privacy and Security, (Pittsburgh, PA, USA, July 23-25, 2008), ACM Press, New York, 35-45.

[11] I. Jermyn, A. Mayer, F. Monrose, M. K. Reiter, and A. D. Rubin. 1999. The Design and Analysis of Graphical Passwords. In Proceedings of The 8th USENIX Security Symposium (Washington, DC, August 23-26, 1999) IEEE Computer Society, 1-14.

[12] M. Kumar, T. Garfinkel, D. Boneh, and T. Winograd. 2007. Reducing Shoulder Surfing by Using Gaze-based Password Entry. In Proceedings of the 3rd symposium on Usable Privacy and Security (Pittsburgh, PA, USA, July 18-20, 2007) ACM Press, New York,13-19.

[13] Lewis, J. R., & Sauro, J. 2006. When 100% Really Isn't 100%: Improving the Accuracy of Small-Sample Estimates of Completion Rates. *Journal of Usability Studies*, 3,1 (2006), 136-150.

[14] Z. Li, Q. Sun, Y. Lian, and D. D. Guisto. 2005. An Association-Based Graphical Password Design Resistant to Shoulder-Surfing Attack. In Proceedings of the IEEE International Conference on Multimedia and Expo (Amsterdam, The Netherlands, July 6-8, 2005) IEEE Computer Society, 245- 248.

[15] D. Lin, P. Dunhpy, P. Oliver amd J. Yan. Graphical Passwords & Qualitative Spatial Relations. In Proceedings of the Symposium on Usable Privacy and Security (Pittsburgh, PA, USA, July 18-20, 2007) ACM Press, New York, 161-162.

[16] Amzer Privacy Protector Shield for HTC Magic. http://www.cellsavers.co.uk/acatalog/Privacy_Screen-Protectors.html. Accessed on February 14, 2011.

[17] B. Malek, M. Orozco, and A. El Saddik. 2006. Novel Shoulder Surfing Resistant Haptic-based Graphical Password. In Proceeding EuroHaptics (Paris, France, July 3-6, 2006).

[18] S. Man, D. Hong, and M. Matthews. 2003. A Shoulder-Surfing Resistant Graphical Password Scheme – WIW. In Proceedings of the International Conference on Security and Management (Las Vegas, Nevada, US, June 23-26, 2003). CSREA Press,105-111.

[19] Real User Corporation, Passfaces: Two Factor Authentication for the Enterprise, 2005.

[20] V. Roth, K. Richter, and R. Freidinger. 2004. A PIN-Entry Method Resilient Against Shoulder Surfing. In Proceedings of the Computer and Communication Security (Washington DC, USA, October 25-29, 2004), ACM Press, New York 236-245.

[21] L. Sobrado and J. C. Birget. 2002. Graphical passwords. vol. 4: The Rutgers Scholar, 2002.

[22] X. Suo. 2006. A Design and Analysis of Graphical Password.M.S. thesis, College of Arts and Sciences, Geogia State University.

[23] X. Suo, Y. Zhu, and G. S. Owen. 2005. Graphical Passwords: A Survey. In Proceedings of the 21st Annual Computer Security Applications Conference (Tucson, Arizona, USA, December, 5-92005), IEEE Computer Society 463-472.

[24] H. Tao and C. Adams. 2008. Pass-Go: A Proposal to Improve the Usability of Graphical Passwords. *Int'l Journal of Network Security*, 7, 2008, 273-292.

[25] F. Tari, A. A. Ozok, and S. H. Holden. 2006. A Comparison of Perceived and Real Shoulder-surfing Resistant Risks between Alphanumeric and Graphical Passwords. In Proceedings of the Second Symposium on Usable Privacy and Security (Pittsburgh, PA, USA, July 12-14, 2006), ACM Press, New York, 56-66.

[26] J. Thorpe and P. C. V. Oorschot. 2004. Towards Secure Design Choices for Implementing Graphical Passwords. In Proceedings of the 20th Annual Computer Security Applications Conference (Washington, DC, USA, December 6-10, 2004), IEEE Computer Society, 50–60.

[27] J. Thorpe and P. C. V. Oorschot. 2004. Graphical Dictionaries and the Memorable Space of Graphical Passwords. In Proceedings of the 13th conference on USENIX Security Symposium (San Deigo, USA, August 9-13, 2004) 135-150.

[28] D. Weinshall. 2006. Cognitive Authentication Schemes Safe Against Spyware. In Proc. of the IEEE Symposium on Security and Privacy (Oakland, California, USA, May 21-24, 2006), IEEE Computer Society, 295-300.

[29] S. Wiedenbeck, J. Waters, J.-C. Birget, A. Brodskiy, and N. Memon. 2005. PassPoints: Design and Longitudinal Evaluation of a Graphical Password System. *International Journal of Human-Computer Studies.* 63,102-127.

[30] J. Yan, A. Blackwell, R. Anderson, and A. Grant. 2004. Password Memorability and Security: Empirical Results. *IEEE Privacy & Security*, 2, 5, 25-31.