Gaze-contingent passwords at the ATM

Paul Dunphy, Andrew Fitch, Patrick Olivier
Culture Lab
School of Computing Science
Newcastle University
{p.m.dunphy,p.l.olivier}@ncl.ac.uk

Keywords

Security, Graphical passwords, Passfaces, Shoulder surfing

Introduction

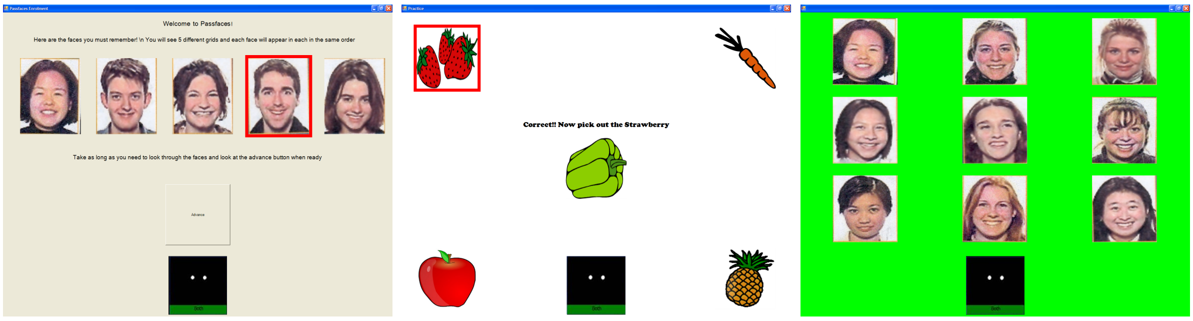
Knowledge-based authentication (e.g. passwords) has long been associated with a vulnerability to *shoulder surfing*; being stolen by attackers overlooking the interaction. In order to combat such threats, steps can be taken to either alter the form of the challenge made to the user, or make use of an interaction technique that is resistant to information leakage. We consider the latter, and empirically evaluate the usability of gaze-contingent interaction as a solution to shoulder surfing in an ATM scenario. We combine this technique with *Passfaces* graphical passwords; potentially more memorable than PINs and well suited to accept gaze-based input. To create a naturalistic setting for our study we utilise the *immersive video* technique originally deployed in the design of pervasive computing systems (Singh et al., 2006). We demonstrate the efficacy of the approach, a usable graphical password entry system that is impossible to attack by direct observation.

Gaze-Contingent Graphical Passwords

The use of an eye tracker to input alphanumeric passwords was presented in a system called *EyePassword* (Kumar et al., 2007), and the application to graphical password schemes has also been considered (Hoanca & Mock, 2006), however no implementations or evaluations of such a setup have been reported to date. Graphical passwords potentially offer improved memorability over passwords and PINs due to image recognition/recall being an easier memory task for humans over recall of words and numbers. Today, research continues into exploring the vast design space of graphical authentication. *Passfaces* is a simple image selection-based graphical scheme that also takes advantage of innate human ability to recognise previously seen faces. The authentication secret assigned to each user is a sequence of n (usually 5) face images which they are required to identify in a sequence of 3×3 grids of faces, each of which comprises an assigned face and eight decoy faces. A number of studies have reported Passfaces to be usable and memorable (Brostoff & Sasse, 2000) (Valentine, 1998a), even over a long period of time (Valentine, 1998b).

We developed software to loosely simulate ATM behaviour using a Tobii x50 eye tracker for input. Our system had 5 states: (1) *Idle*: the system displays various offers from banks on a loop in the fashion of a typical ATM, waiting for a customer; (2) *Calibration*: the user is led through a calibration process; (3) *Playtime*: the user is familiarised with the eye tracker by playing a short game involving selecting fruit on-screen; (4) *Enrolment*: the user is assigned the five face images to comprise their face password, and must spend time forming a memory association; (5) *Login*: the user must recognise and select the faces assigned at enrolment. The user is presented with a sequence of five 3×3 grids, containing eight decoy faces and one assigned face, upon selecting a face in a grid the next grid is displayed. A face selection was assumed after a 0.5 second dwell. For a successful login all 5 faces must be recognised and feedback to this effect was given only after the final face selection. A number of these states are illustrated in figure 3.

As users would be standing whilst interacting with the eye tracker (positioned at a fixed height), it was important to give feedback on the visibility of their eyes to the eye tracker, so they could re-create a good stance and debug any lack of responsiveness themselves. We included the Tobii *trackstatus* indicators of eye visibility; and dynamically changed the screen background colour depending on head distance from the eye tracker.



Welcome to Passfaces!

Here are the faces you must remember! \n You will see 5 different grids and each face will appear in each in the same order

Take as long as you need to look through the faces and look at the advance button when ready

Correct!! Now pick out the Strawberry

Figure 1. Display designs for the different interaction stages: (1) enrolment; (2) playtime (experimenting with gaze-contingent interaction); and (3) *Passfaces* login challenge.

Study Environment

A key design concern of the study was the recreation of the sights, sounds and experiences that are typical of ATM machine usage in busy public settings. The configuration of our ATM study had 4 such elements: (1) *RFID* credit cards; (2) spatial and physical configuration of the ATM; (3) use of an array of large video screens in which to immerse the user at our ATM; and (4) the display of ambient video and sound of an actual ATM to simulate the sensory experience of real ATM usage. The simulated ATM was crafted to be of a familiar physical configuration, and identified users by their unique RFID credit cards assigned to them at the outset of the study, the unique identifier encoded on the card enabled the system to determine if the user required a calibration or could proceed directly to login. The recognition that ATM usage in a public setting can be very different to the desktop led us to re-create many of the common distractions. To capture ambient visuals and sound we recorded video footage at an actual ATM. We positioned our ATM in the centre of three screens and displayed ambient footage on a loop on the left and right screens with a still image of a wall on the centre screen (figure 2).

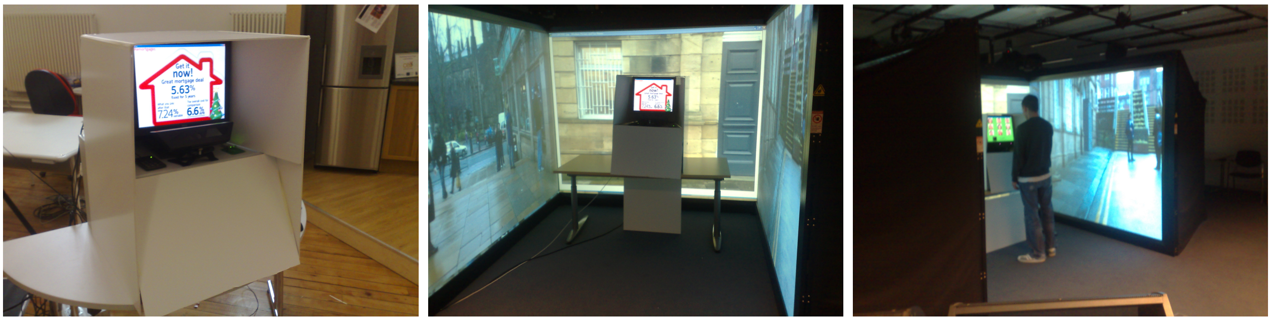


Figure 2. Configuration for the ATM study.

Procedure and results

Twenty student participants (16 male and 4 female) took part in the study; 11 of these were majors in Computing Science or Information Systems ($average\ age = 21$, $STD = 0.8$). Participants played the role of bank customers (users of the ATM) and would start the procedure by placing their RFID credit card on the RFID reader underneath the screen. After undergoing a satisfactory calibration (where a moderator could assist to determine quality) and passing the playtime and enrolment phases, the participants attempted to login with no number of incorrect attempts causing a lockout. Where a reminder was required, this option was included and access to it logged by our system. Participants were asked to login 3 times per session, with each login approximately 15 minutes apart. After a successful login the system thanked the participant for using the ATM and instructed them to remove their card. Participants returned 3 days later to the second session to perform further logins, but this time the calibration, playtime and enrolment phases were not required as details were stored in the database.

In the enrolment phase participants spent some time forming a memory association with the faces that comprised their graphical password. On average users spent 24.9 seconds on this screen but there was a wide variation ($STD = 12.88$) with one participant taking as long as 64 seconds. Figure 3 describes the number of attempts participants required to login at both the first and second sessions. In the first login of the first session performance was worst, as participants came to terms with their face password (a new experience for all). The first login saw participants just as likely to need one attempt as more than three attempts. In fact, in the first login alone, participants made 62 attempts; as many as in the following two logins combined. Performance improved for logins 2 and 3 as users increasingly logged in at the first attempt.

Three days after the first session participants returned and we were keen to observe any retention of faces in memory, and skill at using the eye tracker. At this session there was a marked improvement in performance over the first session with only 8 instances out of 60 (20 participants and each was asked to login 3 times in the session) where more than one attempt was needed, 127 attempts in total were required at the first session and just 73 at the second session ($t = 3.42; p < 0.01$). This superior performance was also reflected in memorability of the faces as reminders of the face passwords were requested significantly less in session 2. The system recorded 24 at session 1 and only 2 at session 2 ($t = 3.99; p < 0.01$)

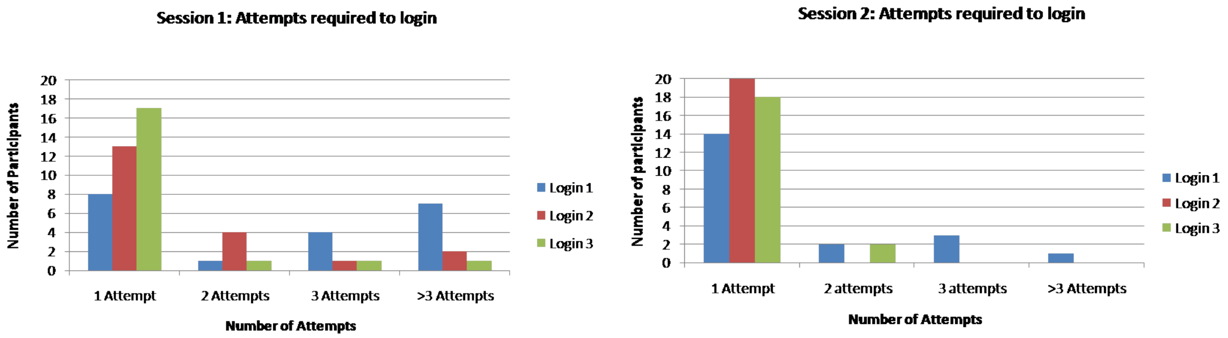


Figure 3. Attempts required for a successful login

At the end of the study participants were interviewed in small groups and asked for their perceptions and opinions on the proposed system and encouraged to relate it to their daily lives. Some of the taller participants criticised the height of the eye tracker as they had to adopt a slightly crouched position (a vulnerable position at an ATM). This problem has been experienced before in trials of iris biometrics at the ATM interface (Coventry et al., 2003). The calibration process frustrated some participants for whom it did not work first time, exacerbated by the sense that they were doing something wrong. One participant commented that the system felt more hygienic as no physical contact with a keypad on the system was required. Another found the experience of eye tracking to be confusing, trying to touch buttons on-screen as though interacting with a touchscreen. Also the ambient sights and sounds did not seem to cause undue distraction for the participants, perhaps indicating a *cocktail party effect* where focused attention on the task at hand drowns out the extra stimuli.

Concluding remarks

We have reported initial investigations into the potential of eye tracking to be applied as a usable solution to shoulder surfing in an authentication context. Despite initial uncertainty about an unfamiliar technology, user performance in the study was good and we witnessed a significant improvement in skill with the eye tracker technique across two short sessions. Also the eye tracking input seems to eliminate the threat of shoulder surfing, something we are interested to validate in future work. One concern regards our inclusion of the *trackstatus* indicator displaying visibility of the eyes to the camera within the eye tracker, but it seems unlikely this can leak accurate information of gaze to an adversary.

There are obstacles to commercial adoption aside from cost alone. The average time to login was 20 seconds, with the fastest time recorded being 8 seconds. Compared to a typical PIN entry of just a few seconds there is still a significant deficit. Eye trackers themselves have well known limitations in terms of *failure to enrol* errors as some users can consistently achieve a less than perfect calibration. Future plans include exploring the design space of authentication techniques suited to gaze-contingent interaction and further increasing the ecological validity of our ATM environment for future lab-based studies.

References

Brostoff, S., & Sasse, A. (2000). Are Passfaces more usable than passwords? A field trial investigation. In *Hci 2000: Proceedings of people and computers xiv - usability or else* (p. 405-424). Springer.

Coventry, L., Angeli, A. D., & Johnson, G. (2003). Usability and biometric verification at the atm interface. In *Chi '03: Proceedings of the sigchi conference on human factors in computing systems* (pp. 153–160). New York, NY, USA: ACM.

Hoanca, B., & Mock, K. (2006). Secure graphical password system for high traffic public areas. In *Etra '06: Proceedings of the 2006 symposium on eye tracking research & applications* (pp. 35–35). New York, NY, USA: ACM.

Kumar, M., Garfinkel, T., Boneh, D., & Winograd, T. (2007). Reducing shoulder-surfing by using gaze-based password entry. In *Soups '07: Proceedings of the 3rd symposium on usable privacy and security* (pp. 13–19). New York, NY, USA: ACM.

Singh, P., Ha, H. N., Kuang, Z., Olivier, P., Kray, C., Blythe, P., et al. (2006). Immersive video as a rapid prototyping and evaluation tool for mobile and ambient applications. In *Mobilehci '06: Proceedings of the 8th conference on human-computer interaction with mobile devices and services* (pp. 264–264). New York, NY, USA: ACM.

Valentine, T. (1998a). *An Evaluation of the Passface Personal Authentication System .* (Technical Report. London: Goldmsiths College University of London.)

Valentine, T. (1998b). *Memory for Passfaces after a long delay.* (Technical Report. London: Goldmsiths College University of London.)