Evaluation of Eye-Gaze Interaction Methods for Security Enhanced PIN-Entry

Alexander De Luca
Media Informatics Group
Amalienstr. 17, 80333 Munich, Germany
alexander.de.luca@ifi.lmu.de

Roman Weiss
Media Informatics Group
Amalienstr. 17, 80333 Munich, Germany
weissr@cip.ifi.lmu.de

Heiko Drewes
Media Informatics Group
Amalienstr. 17, 80333 Munich, Germany
heiko.drewes@ifi.lmu.de

ABSTRACT

Personal identification numbers (PINs) are one of the most common ways of electronic authentication these days and used in a wide variety of applications, especially in ATMs (cash machines). A non-marginal amount of tricks are used by criminals to spy on these numbers to gain access to the owners' valuables. Simply looking over the victims' shoulders to get in possession of their PINs is a common one. This effortless but effective trick is known as shoulder surfing. Thus, a less observable PIN entry method is desirable. In this work, we evaluate three different eye gaze interaction methods for PIN-entry, all resistant against these common attacks and thus providing enhanced security. Besides the classical eye input methods we also investigate a new approach of gaze gestures and compare it to the well known classical gaze-interactions. The evaluation considers both security and usability aspects. Finally we discuss possible enhancements for gaze gestures towards pattern based identification instead of number sequences.

Categories and Subject Descriptors

D.4.6 [**Operating Systems**]: Security and Protection – *access controls, authentication;* K.4.4 [**Computers and Society**]: Electronic Commerce – *security;* K.6.5 [**Management of Computing and Information Systems**]: Security and Protection – *authentication.*

General Terms

Experimentation, Security, Human Factors.

Keywords

Security, privacy, PIN entry, cash machine, eye gaze interaction, gaze gestures, look & shoot, dwell time.

1. INTRODUCTION

The personal identification numbers (PINs) are numeric codes used to authenticate persons to computer systems. Their best-known operational area is the field of ATMs (cash machines). But PIN authentication is also commonly used for mobile phones, door locks etc. Withdrawing money from an ATM mostly takes place in a public space, e.g. on the street outside a bank building or in a shopping mall. This makes the security provided by the combination of the bank card and the matching PIN vulnerable. Due to the fact that most PIN numbers only consist of 4-digits, a person observing the cash withdrawal can easily memorize the PIN. So a usual fraud scenario is that the PIN number is observed by a criminal, either by simply looking over a victims shoulder, known as *shoulder surfing*, or using technical devices like binoculars or miniature cameras to get the entered number sequence. The second step is to acquire the cash card from the victim. Here a wide range of methods is reported. Some scammers use technical devices, which they place on top of the card slot of the ATM that copy the card data, but often the cash cards are simply stolen afterwards. Various larcenies by trick are used to get in possession of the cards. [10]

In this work, we provide and discuss different gaze-based techniques for decreasing the security risks of public PIN entry since we consider observing somebody's finger movements on a number pad to be very simple and unobtrusive. That is, we evaluated different means to control an ATM with the users' eyes. We analyzed these methods on their appropriateness and on security issues for this specific purpose, both theoretically as well as practically in a user study. In contrast to recent work on this field [7], we did not only study the traditional approaches of gaze interaction but also evaluated the new concept of gaze gestures [2][3]. The gaze gestures imply a number of advantages both on implementation as on security aspects, which we will outline in this paper. To do so, we implemented a prototype of a gaze-enabled ATM and let volunteers perform previously defined tasks with the different interaction types. Additionally, we performed a security analysis on all interaction types to build a basis for our suggestion of preferring eye-gestures to the traditional approaches.

2. RELATED WORK

The first category of related work can be subsumed to the attempt of providing mechanisms increasing the security of finger-based PIN entry. Roth et al. provide a system that uses a randomized binary input method for numbers [11] that impedes shoulder-surfing attempts because for human attackers it is almost impossible to find out which number has been pressed since every number is built out of 4 binary choices. Tan et al. go a step further and provide the Spy-Resistant Keyboard [12], which was created for secure password entry on public touch screens like display walls. The advantage is that it provides a secure entry method for the whole alphabet plus numbers based on a randomly arranged keyboard in combination with spatial memory. Unfortunately, their approach noticeably increases the complexity of password entry. Finally it should be noted that there is also variety of commercial products dealing with this problem, which shows the need and also a market demand for this topic. One of them is the Scramble Pad by Hirsch

OzCHI 2007 Proceedings, ISBN 978-1-59593-872-5

Electronics [5] that randomly orders the numbers on the pad each time it is used to decrease the threat of shoulder surfing.

A different approach that eliminates the need to use fingers for PIN entry and thus providing additional security is explained in the recent work of Kumar et al. [7]. The authors conducted an experiment of using traditional gaze-based interaction for PIN and password entry and consider it a suitable interaction technique for this purpose. In our work, we describe a further approach based on gaze gestures [2][3]. We will show that gaze gestures have various advantages compared to the traditional methods. No calibration process is necessary and it is less error-prone for input without feedback. The lack of feedback is crucial for the intended security enhancements.

But authentication is not limited to PIN entry. Thus, different scientific work is performed on alternative authentication methods. One of these fields is Biometrics, which uses unique physical marks of a human (e.g. a fingerprint) for authentication also for ATMs as shown in [1]. Another alternative are graphical passwords as has been done in the work of Tullis and Tedesco [13] and Moncur and Leplâtre [8]. The idea is to find the password photos in a set of photos on the screen. While the Tullis and Tedesco were using personal pictures instead of passwords for authentication because they are easier to remember, Moncur and Leplâtre argue that the security of the password is noticeably increased when using randomly generated images as passwords. Of course, numerous other techniques exist, but the few outlined are meant to provide an overview of this field.

3. TECHNIQUES AND PROTOTYPE

Originally, eye-tracking has been developed for impaired, mostly paraplegic individuals. This way, they are enabled to interact with computing systems despite their handicap. In our work we adapt eye-tracking as a security enhanced interaction technique. As our goal is a more secure but still usable solution, we set aside the potential advantages for the handicapped and also use interaction techniques that require button pressing and holding. Using these preconditions, we developed and adapted three possible interaction techniques (the traditional dwell time and look & shoot as well as the novel gaze gestures method) that will be outlined in this section. Security aspects of each method will be discussed in section 4.1.

3.1 Dwell Time Method

The standard eye-gaze interaction technique is the dwell time method. Here the user stares for a certain time (dwell time) at an area on the display to trigger an action. This dwell time is typically below 1000ms. As an easy to use method, which has already been shown in [14], we decided to integrate it in the prototype. In addition to its simplicity it is an intuitive and easy to learn method. Much depends on the used dwell time. Short dwell times allow fluent working, but are more likely to produce errors by unwanted fixations also called the Midas-Touch effect [6] whereas long dwell times slow down the whole interaction process. Another advantage next to the ease of use is that there is no need for any extra button. The disadvantage of this technique is the need for a calibration process.

3.2 Look & Shoot

Another method taken from classic eye gaze interaction research is what we call the look & shoot method. Here the user fixates an interaction object on the screen and simultaneously hits a button to trigger the action. Ware et al. investigated that this method is very easy to use and is especially fast as no dwell times have to elapse before an action is triggered [14]. The need for a calibration process also applies to this method.

3.3 Gaze Gestures

The gaze gesture concept is rather new. It has recently been introduced in [2] and [3]. The idea is similar to the mouse gestures available for the Firefox browser [9]. Performing a specific eye movement pattern triggers an action.

As outlined in [2], the gaze gestures have several advantages compared to the previously mentioned approaches. First of all, the gaze gestures do not need a calibration because they use relative movements only. As the size of the gaze gestures is several degrees of visual angle the accuracy is not an issue. This means that gaze gestures require only low-tech eye-trackers like a normal webcam combined with an infrared LED light (the infrared light produces the red eye effect, known from photographs and simplifies the pupil recognition because of the higher contrast against the iris). Since many ATMs are already equipped with security cameras, deploying the gaze gesture method should be easy on current hardware. Another advantage of the method is that there is no need for interaction objects on the screen. The space on the display can be fully used for visualization of required information.

In this work, we used a modification of the gaze gesture algorithm developed in [2] and [3]. Instead of continuously analyzing the users' gazes, the modified algorithm only considers the eye-movements done while a specific button is pressed. Thus, in order to perform a gesture, the users have to press the button and then perform the gesture. When releasing the button, the eye movement is analyzed for an occurrence of a numeric gesture as depicted in Figure 1. The use of the gesture button gives more freedom in the design of the gesture alphabet used in our prototype, which is based on the well known EdgeWrite developed by Wobbrock et al. as a stylus based text entry method for handheld devices [15]. The gestures follow the outline of the characters to improve learning performance. Without the gesture button the digit 9 and 5 would not be distinguishable.

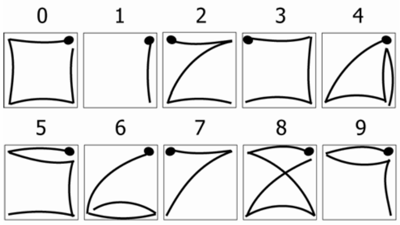
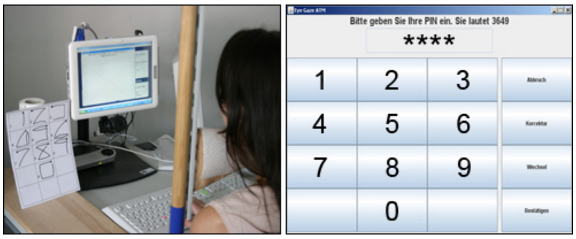


Figure 1: The numeric gaze gestures used for the prototype.

3.4 Prototype

For the user study we implemented a prototype with the three different interaction techniques mentioned above. The design followed the typical design of a German ATM consisting of a number pad with the digits zero to nine, two empty keys and four command keys for canceling the transaction, correcting the entered number, changing the language and committing the transaction.

The prototype application is written in Java and runs on a Windows PC together with the commercial Eye-Tracker ERICA [4]. For the dwell time method 800ms were used as the threshold. This value showed good results in a pre-test. For the look & shoot method the space bar of a standard keyboard was used as the button triggering the event. For the gesture recognition technique the space bar was used as the button the users had to hold down during making a gesture. In order to simplify the usage and limit the number of necessary gestures, the command functions of the ATM like “Cancel”, “Correct” and “OK” were chosen with the keyboard.



Bitte geben Sie Ihre PIN ein. Sie lautet 3649

Figure 2: The user study setting (left) and screenshot of the prototype application (right)

The prototype displays an ATM number pad of an overall size of 730 by 450 pixels. Each button has a size of 180 by 90 pixels, which is about 5° by 3° visual angle and hence clearly above the typical eye tracker accuracy of ± 0.5°.

A problem that occurred in the early design stage was the question how to provide feedback. As the security benefit would be nullified if the chosen numbers were highlighted on the display – an observer would only have to record the screen output – we decided to restrict feedback to only asterisks as shown in Figure 2 (right). As with conventional ATMs the asterisks indicate successful recording of digits by the system. However, the users had no direct feedback about what single digit they entered.

4. EVALUATION

4.1 Security Analysis

As stated before, shoulder surfing is a common method used by criminals to get in possession of another person’s PIN. For our security analysis we assume the attacker is able to observe the screen and the number pad of the ATM. Eye-gaze interaction as input method is resilient against such an attack, as observing the movement of the user’s pupil will surely arouse suspicion.

We note that a camera recording the eye movements could help the attacker, but first this camera must necessarily be placed in the field of vision of the user, which makes it harder to conceal it and second determining the actual PIN from the raw data will not be trivial (Especially filtering the eye-input sequences from the “normal” eye-movement caused by reading etc.). A successful attack against the look & shoot or the gesture recognition method would also require a synchronous recording of the button needed to trigger the event. Using a hardware button can be credited an extra security benefit. The raw eye movement data itself can only be used for extracting the PINs with an enormous effort by the attacker.

It is reasonable to assume that using eye gaze input for the PIN will prevent the vast majority of shoulder surfing attacks.

4.2 User Study

To evaluate the usability of the eye tracking enhanced ATM, we conducted a user study in our laboratory. 21 volunteers completed the three different PIN entry tasks and answered a questionnaire afterwards. Seven of the participants were female and all participants were aged between 22 and 37 years. Five of them had already used an eye tracker before.

The participants had to perform specific tasks with the prototype as shown in Figure 2 (right). At first the volunteers completed a training sequence for each technique in order to get used to the device. For this they had to enter the digits from zero to nine sequentially. After that, the participants had to enter three randomly generated 4-digit PINs with the specific interaction technique. This was repeated for each interaction technique with each user. To avoid ordering effects, the subjects applied the three different methods in random order. Each participant had to complete three sets of these tasks. Afterwards demographic data, information about the users’ habits using ATMs and experiences regarding user experience and usefulness of the system were collected with a questionnaire.

The results of the user study show that eye gaze interaction is a suited method for PIN entry. We can partially approve the results found by Kumar et al. regarding the dwell time and look & shoot techniques. In contrast to their results, in our study the look & shoot method emerged as the faster and less erroneous one compared to dwell time. It was also the method favored by the majority of the subjects. They considered it the most easy to use, the fastest, the less error-prone and the most secure technique.

The statistically evaluated data using analysis of variance (ANOVA) showed no significant advantage regarding execution times for the look & shoot method. A four digit PIN entry took the subjects 12 seconds in average, whereas a PIN entered using dwell time took 13 seconds. The probability for errors also showed no significant difference. Using dwell time 15 of the entered 63 PINs were faulty (23.8%), using look & shoot 13 contained errors (20.6%).

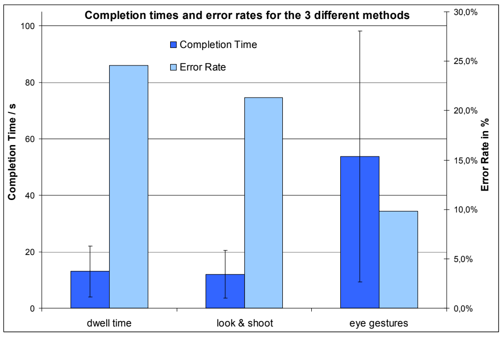


Figure 3: Performance and error rates for 3 techniques. On the right: Gestures, the slowest but the least error-prone.

Evaluating the gaze gesture method, we found that entering PIN numbers took much longer than using the “classic” methods (an average of 54 seconds per PIN entry was measured) but was also much more robust against errors than the methods described above. Only 6 of the entered PINs using gestures were erroneous (9.5%). We consider this as a great improvement, especially in regard of the fact that most PIN-authentication systems block the access after some incorrect attempts due to security reasons. Using a binomial test also shows a significant enhancement of the error rate ($p<0.008$).

The gaze gesture method is less intuitive than the classic methods as many subjects initially had problems to produce recognizable gestures. Furthermore the gesture alphabet was unknown to all participants. This explains the big difference in time for completing the PIN entering task. Much time was used for looking at the sheet with the gestures for the single digits. Theoretically a stroke within a gaze gesture needs about 500ms (200ms for the saccade and 300ms for the fixation) and a digit with four strokes takes about 2 seconds to perform. A four-digit PIN with one second break between the inputs of the digits will last a little bit more than 10 seconds. Indeed we had a participant in the study, who entered the PIN correctly within 14 seconds. We expect this time for all users once they are trained for gaze gesture input.

In addition to the absence of a calibration process, the big advantage of the gaze gesture method is its robustness against input errors. Due to the abandonment of feedback for enhanced security each wrong gaze leads to an incorrect PIN entry if the dwell time or look & shoot method is used and thus to the high error rates as shown in Figure 3. Using the gestures a wrong gaze leads most probably to an unrecognizable gesture, thus recognized errors even without immediate feedback.

In general our doubts that the lack of feedback will make the prototype very difficult to use were resolved in the end. None of the users even mentioned that using an ATM without knowing what was entered was confusing them. The classic methods benefit from the large target sizes that made interaction easy and the gesture recognition method appeared very error-resistant itself.

4.3 Gaze Gesture Enhancements

There is a lot of open space to improve the usability of eye gaze recognition. Due to the rather new approach using gaze gestures we used a quite simple recognition algorithm in our prototype. We observed that the main reason why a gesture performed by a user is not recognized by the system is a lack of exactness in the hand eye coordination. As a button has to be pressed and hold while performing the gesture, often an additional stroke was detected directly before or after the proper gesture. These unintended upstrokes or tails could be filtered out by the algorithm and improve the recognition rate significantly. Another approach would be an adapted design of the gestures. If the “distance” of any gestures to any other would be at least three strokes, the algorithm could auto-correct gestures with a single false stroke.

5. A NOVEL PASSWORD ENTRY METHOD BASED ON EYE-GESTURES

As a side effect of our evaluation, we found that many participants stated to remember their PINs as a shape on the number pad instead of the actual number sequence. Thus, we conducted an online survey with 86 participants. The results show that more than 40% of the participants stated that they would remember their PINs by a shape or at least use a shape to support their memory, which supports our assumption.

Based on these results, we are discussing an enhancement of the gaze gesture input methodology for secure authentication. The idea is to improve the algorithm to enable the input of arbitrary shapes or combination of shapes instead of inputting numbers. We claim that shapes are by far easier to remember even if very complex shapes are chosen (that is, much harder to copy by an attacker). Thus, in future work, we will be creating a complete new authentication mechanism and evaluate it, as almost 50% of the survey’s participants are interested in using such a password entry method.

6. CONCLUSION AND OUTLOOK

Talking about PIN entry, we are often facing insecure situations, even more since locations of input are mostly public or semi public places like ATM cash machines on the street outside bank buildings. In this paper, we discussed and evaluated three different gaze-based input methods for PINs: the two classical approaches dwell time and look & shoot as well as a new interaction type of gaze gestures.

In a security analysis, we showed that these approaches provide much higher security against common attacks compared to the commonly used input methods. Out of the chosen methods, the gaze gestures provide a couple of further advantages. They are easy to deploy (also on already existing machines), do not need a calibration for each individual, offer the same amount of security and are less error-prone for input without feedback As providing real feedback is not possible since it would nullify the security advantages of those methods, this seems to be a main advantage for eye-gestures.

In future work, we plan to focus on the previously discussed complex gaze gestures and evaluate shape-based passwords on their appropriateness for this specific task.

7. ACKNOWLEDGMENTS

This work is partially supported by the European Union, in the framework of the FP6 – IST Project DISCREET.

8. REFERENCES

[1] Coventry, L., De Angeli, A., and Johnson, G. Usability and biometric verification at the ATM interface. In: Proceedings of CHI '03, Fort Lauderdale, Florida, USA, April 5 - 10, 2003.

[2] Drewes, H., Schmidt, A. Interacting with the Computer using Gaze Gestures. In: Proceedings of Interact’07. Rio De Janeiro, Brasil. September 10 – 14, 2007.

[3] Drewes, H., De Luca, A., Schmidt, A. 2007. Eye-Gaze Interaction for Mobile Phones. In: Proceedings of Mobility’07. Singapore, September 10 - 12, 2007.

[4] http://www.eyeresponse.com, September 2007.

[5] http://www.hirschelectronics.com/Products_ScramblePads.asp, August 2007.

[6] Jacob, R. J. What you look at is what you get: eye movement-based interaction techniques. In: Proceedings of CHI ’90, Seattle, Washington, USA, April 01 - 05, 1990.

[7] Kumar, M., Garfinkel, T., Boneh, D., Winograd, T. Reducing Shoulder-surfing by Using Gaze-based Password Entry. In: Proceedings of SOUPS '07, Pittsburgh, USA, July 18 - 20, 2007.

[8] Moncur, W. and Leplâtre, G. Pictures at the ATM: exploring the usability of multiple graphical passwords. In: Proceedings of CHI ’07, San Jose, California, USA, April 28 - May 03, 2007.

[9] http://optimoz.mozdev.org/gestures/, August 2007.

[10] Rogers, J. “Please enter your 4-digit PIN”. In Financial Services Technology,U.S. Edition, Issue 4, March 2007.

[11] Roth, V., Richter, K., and Freidinger, R. A PIN-entry method resilient against shoulder surfing. In: Proceedings of CCS '04, Washington DC, USA, October 25 - 29, 2004.

[12] Tan, D. S., Keyani, P., and Czerwinski, M. Spy-resistant keyboard: more secure password entry on public touch screen displays. In Proceedings of OZCHI ’05, Canberra, Australia, November 21 - 25, 2005.

[13] Tullis, T. S. and Tedesco, D. P. Using personal photos as pictorial passwords. In: CHI '05 Extended Abstracts, Portland, OR, USA, April 02 - 07, 2005.

[14] Ware, C. and Mikaelian, H. H. An evaluation of an eye tracker as a device for computer input. In: Proceedings of CHI '87, Toronto, Ontario, Canada, April 05 - 09, 1987.

[15] Wobbrock, J. O., Myers, B. A., and Kembel, J. A. EdgeWrite: a stylus-based text entry method designed for high accuracy and stability of motion. In: Proceedings of UIST '03, Vancouver, Canada, November 02 - 05, 2003.