Pobrane z czasopisma Annales AI- Informatica **http://ai.annales.umcs.pl**
Data: 24/09/2026 18:12:44

Annales UMCS Informatica AI XIV, 1 (2014) 75–84
DOI: 10.2478/umcsinfo-2014-0009

Annales UMCS
Informatica
Lublin-Polonia
Sectio AI
http://www.annales.umcs.lublin.pl/

Cheap and Easy PIN Entering Using Eye Gaze

Pawel Kasprowski[1*], Katarzyna Harężlak[1†]

[1] *Institute of Computer Science, Silesian University of Technology*
Gliwice, Poland

Abstract – PINs are one of the most popular methods to perform simple and fast user authentication. PIN stands for Personal Identification Number, which may have any number of digits or even letters. Nevertheless, 4-digit PIN is the most common and is used for instance in ATMs or cellular phones. The main advantage of the PIN is that it is easy to remember and fast to enter. There are, however, some drawbacks. One of them addressed in this paper is a possibility to steal PIN by a technique called shoulder surfing. To avoid such problems a novel method of the PIN entering was proposed. Instead of using a numerical keyboard, the PIN may be entered by eye gazes, which is a hands-free, easy and robust technique.

1 Introduction

Proper identification of a person getting access to some resources is an important and still challenging problem of nowadays systems. There are plethora of techniques used starting with simple passwords, through graphical passwords, tokens and ending with various biometric authentication methods. One of the simplest methods utilized in many access points is using a password that consists of 4-digits. It is commonly named PIN (for Personal Identification Number) and is used for instance for credit cards identification at ATMs. One of the main security problems while authenticating at ATM is possibility that somebody may see the PIN that was entered by the authenticating person. It is commonly known as shoulder surfing, and it may be for example done by using properly placed video camera or even e.g. by analyzing keyboard temperature directly after the PIN was entered. To avoid shoulder surfing many techniques have been proposed like adding some obfuscators (not important information entered

*pawel.kasprowski@polsl.pl
†katarzyna.harezlak@polsl.pl

Pobrane z czasopisma Annales AI- Informatica **http://ai.annales.umcs.pl**
Data: 24/09/2026 18:12:44

together with a password) [**8**] or using graphical passwords [**2**]. The solution presented in this paper gives opportunity to enter the PIN number without any keyboard. It uses gaze point information (e.g. information where the person is looking) and transforms gaze points into a sequence of digits. There were similar solutions proposed in the earlier studies but they used very expensive eye trackers and complicated experiment setups which made such solutions rather academic ones, not usable in practice. The solution presented here shows how to build a complete and robust setup that costs less than $100 and analyzes if it is possible to enter the PIN using eyes in time comparable to normal key-typed PIN entering.

2 Related Research

Eye contingent interfaces have been the subject of studies for many years [**5**]. However, the main problems of developing such interfaces is that human eyes are used by people as an input device and a human brain is not accustomed to use them to control something. In a poorly designed eye contingent interface a person automatically clicks everything she seeks and such interface becomes very annoying and not usable. The phenomenon is commonly named a Midas touch problem [**6**]. It is possible to use eyes as brain output but it must be done attentively and precisely. In most applications users issue commands by looking at a particular point (e.g. button on screen) for some time. It is called a dwell. The crucial parameter for such a system is how to choose a correct dwell time that triggers action [**7**] Of course, longer dwell times are expected to give more accurate results. On the other hand, shorter dwell times result in faster human-computer communication. The main idea of the work presented in this paper is to use eye gazes as input for an authentication application (like ATM). There is some research that utilizes the idea of using information about eye movements to enter a password. For instance Weaver et al. [**9**] created software that enables entering an alphanumeric password using eye gazes. It was tested for a specific complicated password and for different dwell times. The best results were obtained for static dwell times (80%) but the algorithm proposed for determining appropriate time adaptively did not work sufficiently well (45%). Similarly Kumar et al. [**7**] proposed the EyePassword software, which may be used for password entering. They have utilized two scenarios: a gaze-based, when the user just gazes at some particular point for some time and a trigger-based, when the user looks at some points and clicks the button. Surprisingly, the latter gave much worse results (15% of errors compared to 3% for the gaze-based). Another important contribution was the paper by DeLuca [**4**]. They performed both gaze-based and click-based scenarios and compared them to gaze gestures, which is yet another way to enter information using a gaze. Their work was continued in [**3**]. The main problem for gaze-based interfaces is its usability. Even if it is more secure, it will not be used if it is not convenient for users. An interesting study of usability of gaze-based interfaces may be found in [**1**]. Most studies mentioned above started with a priori defined dwell time threshold. The participants looking at the specified

Pobrane z czasopisma Annales AI- Informatica **http://ai.annales.umcs.pl**
Data: 24/09/2026 18:12:44

point were informed by a sound or visually that their choice has been registered. The most common such registration duration was longer than 10 seconds [**4**], [**1**]. In the research described in the paper a slightly different approach, in which a participant decides himself how long to focus at the point, has been applied.

3 Experiment

The main objective of the presented studies was to check how fast the PIN can be entered using eyes and whether this time can be comparable to the keypad entering time. Therefore, no a priori dwell time was defined and no feedback was given to the users as it could influence the results of the experiments. Additionally, it is worth emphasizing that the registration of eye movements was done using the EyeTribe eye tracker (www.theeyetribe.com), which may be purchased for less than $100 which makes the solution accessible to ordinary users. Before each experiment the device was tuned with 7 points calibration that lasted approximately 7 seconds. This process was followed by two different types of trials (click-based and gaze-based) using a screen with 10 buttons marked by successive digits 0-9 as it is presented in Fig. 1. There were overall 370 trials performed with 23 participants including 185 trials for click-based and 185 trials for gaze-based trials (see explanation of types below). To achieve reliable results, the experiments were not conducted in a laboratory environment but in a crowded place with people trying to do shoulder surfing.

Fig. 1. Main application screen

The users task in the first type of trial (click-based type) was to click a key (trigger) while simultaneously looking at a subsequent digit of the PIN. The last click, after pointing with eyes all four digits, finished the trial. The participants' activities - eye movements and click moments - were recorded for further analysis. The second type of trial (gaze-based type) included only two clicks. The first one was done to start a

Pobrane z czasopisma Annales AI- Informatica **http://ai.annales.umcs.pl**
Data: 24/09/2026 18:12:44

trial, and the second one to finish it. Between clicks the users task was to look for some time at four subsequent digits of their PIN. As it was mentioned earlier, the users were instructed to look at the specified digit for some time without any feedback from the system that a time was sufficient to recognize the users intention. Similarly to the previous experiment, all eye movements were recorded together with the moments of initial and final clicks. Every single run consisted of three or four click-based trials and three or four gaze-based trials. Users were encouraged to try to enter the PIN number consisting of four digits as correct as possible but at the same time as fast as possible. Because every trial was recorded, the users had opportunity to examine the results directly after each run, which should supposedly improve an accuracy of their activities in subsequent trials. The results were presented as a list of scanpaths (see Fig. 2) in conjunction with a final score being a sum of Levenstein distances between the expected PIN and the PIN entered by a user, independently of each trial. Trials total time was calculated and provided as well.

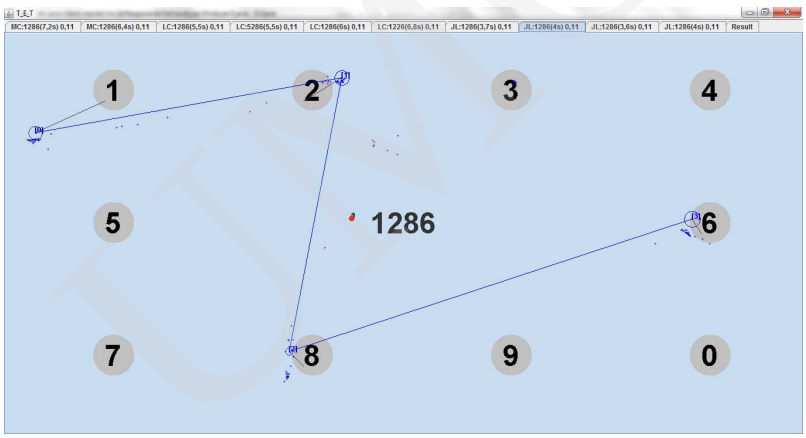

Fig. 2. Example of a recorded scanpath (the task was to enter PIN 1286)

The analysis of collected samples was performed using two algorithms developed for each of the experiments types. The algorithm processing clicks and searching for related gaze points worked as follows:

- Find three gaze points recorded directly before the click and three gaze points recorded directly after the click.
- For every found gaze point calculate distances to all digits displayed on the screen and choose the closest one as point value.
- Choose the most frequently repeated value for the points analyzed as the value of the click.

The result is a sequence of four digits one for each click. The algorithm was applied to all click-based trials. If the sequence consisted of four digits, exactly the same as those that were supposed to be entered, the trial was marked as correct. A different

Pobrane z czasopisma Annales AI- Informatica **http://ai.annales.umcs.pl**
Data: 24/09/2026 18:12:44

and more complicated algorithm was developed to retrieve the PIN number from the processed gaze points without any information about the clicks. It tries to build a sequence of digits based on found fixations moments when the eye is almost stable. It takes three parameters:

- *window* size of a window defining a number of points to be considered when evaluating the point as a part of fixation. Initially it is set to 3 points.
- *threshold* - the longest distance between points to be recognized as one fixation initially it is defined as 3 degrees.
- *sequence* sequence of currently recognized digits. Initially empty (length=0)

The main loop of the algorithm tries to find a sequence of length 4. If the run is not successful (sequence is shorter than 4), it decreases window by one, increases threshold by 0.4 and repeats the run until the sequence length is equal to 4 or window reaches 1 and threshold reaches 10 degrees. The run consists of the following steps:

- For every recorded gaze point classify it as part of fixation (F) if window previous points are closer than the threshold degrees of each other.
- Join neighboring F points into fixations.
- Calculate the fixation position as the average position of points belonging to the fixation
- For every fixation calculate distances to all digits displayed on the screen and choose the closest one as fixation value.
- Merge neighbouring fixations that have the same value (the same digit assignment) into one fixation.
- While a number of fixations is higher than 4 remove the shortest fixation.
- Build a sequence of digits from a sequence of fixations.

The end result is a sequence of 0 to 4 digits. Similarly to the previous algorithm if the sequence consists of four digits, exactly the same as those that were supposed to be entered, the trial is marked as correct. The algorithm described above was applied to both types of trials: click-based and gaze-based.

4 Results and Discussion

All conducted tests aimed at checking if not guided eye movement can be useful in the PIN delivering. The first parameter that provides such knowledge is the percent of correct trials (accuracy) - i.e. trials when the user entered the correct sequence of numbers - to the overall number of trials. Surprisingly, when considering data from the click-based type of trials, the algorithm that processed each trial taking users click moments into account (CBc) gave worse results than the algorithm considering the same signal but using only information about gaze points (CBg). Such outcome can result for two reasons: (1) imprecise users coordination of clicking and looking or (2) eye tracker delay. The best results were achieved for the gaze-based trials when the user did not have to worry about clicking during the trial. However, the differences in

Pobrane z czasopisma Annales AI- Informatica **http://ai.annales.umcs.pl**
Data: 24/09/2026 18:12:44

accuracy were not significant (p>0.05). Another interesting factor was the total time of each trial. As it can be seen in Table 1, the average time of runs was significantly shorter for the gaze-based (GB) trials.

TABLE 1. The accuracy and the time for different trial types and algorithms

Type	Average time	Accuracy	Rejection percent	Accuracy after rejection
CBc	6.5s (+/− 2.45s)	61.1%	-	-
CBg		66.5%	9%	73,2%
GB	4.36s (+/− 1.27s)	68.6%	15%	80.4%

For both CBg and GB types the algorithms returned from 0-4 digits. The number of digits was lower than four when, in spite of changing the thresholds (window and threshold), it was not possible to find four dominating fixations. Such situation occurred for 9% of click-based and 15% of gaze-based trials (see: Rejection percent in Table 1). Such error is easy to detect, contrary to an error when the algorithm returns wrong combination of four digits. If the number of digits is lower than 4 the trial can be automatically rejected and the user can be asked to do another attempt. Therefore, the results were analyzed once again after rejecting all too short combinations (see: Accuracy after rejection in Table 1). It obviously could not be done for the algorithm analyzing clicks because it always returns the four digits sequence (as there are always four clicks). Comparing the results for different users, it can be noticed that they vary significantly as shown in Fig. 3. There were two participants that were able to achieve 100% score all their attempts were successful (with 16 and 8 trials respectively). But there were also four participants with a lack of correct attempts. Three of them took part in 8 trials and one in 16 trials.

When analyzing the possible causes of errors, it seemed obvious that incorrect trails could result from too fast trial realization i.e. too short dwells duration on subsequent digits. It must be remembered that the users were not instructed to dwell at the digit for some specified time and there was no direct feedback that their dwell was accepted (as in similar works mentioned in section 2). They were told just to point the digit with their eyes. Surprisingly, it occurred that the average time for the correct trials is lower for both types of experiments and both algorithms used for determining sequences (Tab 2) and is significantly lower (p<0.05) for both algorithms with the click-based (CB) type .

TABLE 2. The time and the accuracy for the correct trials

Type	Correct	Incorrect
CBm	6.21s (+/− 2.44s)	6.95s (+/− 2.43s)
CBg	6.07s (+/− 1.56s)	7.35s (+/− 3.5s)
GB	4.33s (+/− 1.23s)	4.42s (+/− 1.36s)

Pobrane z czasopisma Annales AI- Informatica **http://ai.annales.umcs.pl**
Data: 24/09/2026 18:12:44

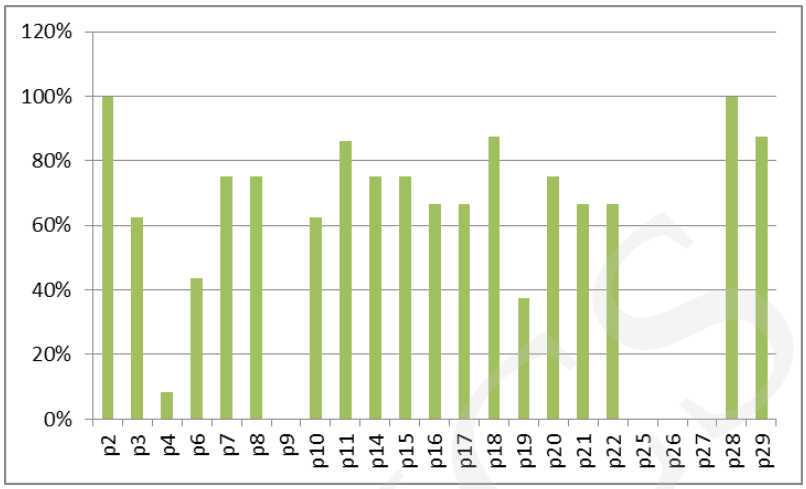

FIG. 3. Accuracies for different participants

The findings of the studies presented so far show that the algorithms developed in the research are able to find a correct sequence using a very low number of recordings. In fact it occurred that the shortest correctly entered trial was 1.97 sec. and 16% of correct trials were entered in less than 3 sec. Taking into account, that sequence of digits was not known to the participants before experiments started, it can be expected that in the case of well-known numbers arrangement, percentage of the correct results featuring by short time of its entering will be higher. Further analyses revealed that the distribution of accuracy is characterized by higher density near the boundary values.

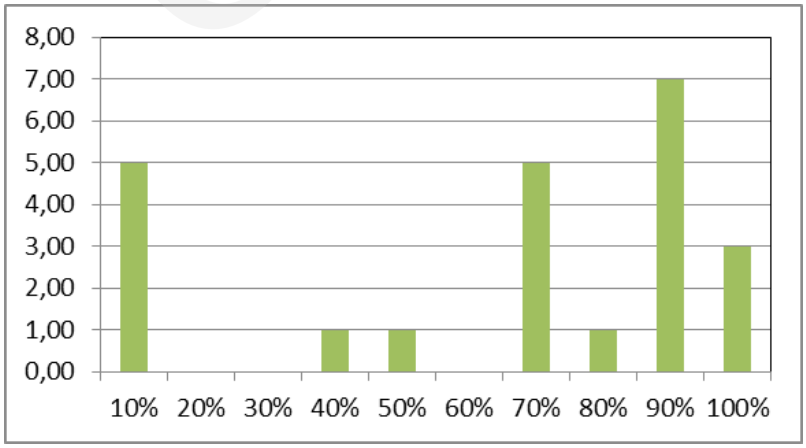

FIG. 4. The distribution of users accuracies

There were five participants with the result less than 10% (0% for 4 of them) but majority of participants were able to achieve accuracy higher than 75%. As it could be

Pobrane z czasopisma Annales AI- Informatica **http://ai.annales.umcs.pl**
Data: 24/09/2026 18:12:44

expected the users get used to the application and their later attempts were more successful than the first ones. Correlation between a number of attempts and a percentage of correct trials per user is 0.213, which indicates that users with more attempts tend to have better results.

5 Summary and Future Work

Using eye movement to control chosen areas of humans life is an interesting and challenging task. The experiment presented in the paper aimed at developing methods and tools making entering the PIN number using eyes possible. This basic goal of the research was extended with the analyses of time, which has to be spent to correctly point out appropriate sequence of digits. It was checked if it is possible to enter the PIN number using only eyes in time comparable with the usage of classic keyboard and without any direct feedback from the application. There were two types of experiments proposed. First of them assumed providing the PIN digits using eye movement signal confirmed by clicks. In the second solution the users were expected, for the same purpose, to utilize only their eyes. Analysis of the obtained results allowed drawing some interesting conclusions.

(1) It was confirmed that utilizing eye movements as an output signal is possible even if a cheap eye-tracker is used.
(2) Such signal turned out to be valuable even for the participants that used eye tracking for the first time (as most of the participants during the experiment). It can be expected that more experienced participants, that have tried eye pointing multiple times, would achieve better results. It was partially confirmed during this research but more comprehensive conclusions require more extensive experiments.
(3) Time measured during experiments proved to be comparable with that, which is needed using keyboard.
(4) Allowing the users to decide how long to gaze at a given digit proved to be a good idea, shortening time of performed task. The user did not have to wait for a signal to continue a task. The findings show that shorter duration of experiment not necessarily has to give a worst result. On the contrary: correct trials were related to shorter task realization.
(5) Comparison of the results calculated based on clicking and without it indicated that necessity of correlating eyes and hand can lead to worse outcomes.

The studies presented in the paper will be continued and will concern more profound analyses of users dwell durations to find if there are significant differences among people as it was suggested in [**9**]. Furthermore, the possibility of defining one universal dwell threshold will be checked. Moreover, spatial errors of fixations should be analyzed to determine the minimal size of components, which may be pointed by gaze.

Another important problem is calibration of the device. Currently the calibration lasts 7 seconds it will be verified if the same results could be achieved using a template

Pobrane z czasopisma Annales AI- Informatica **http://ai.annales.umcs.pl**
Data: 24/09/2026 18:12:44

calibration as it was suggested in [**7**]. Entering PIN using eye movements seems to be an interesting alternative to the classic keyboard based methods. Firstly, it may be easier for people who for some reason have difficulties with keyboards (like disabled people). Secondly, it reduces a shoulder surfing problem. However, it must be emphasized, that it is still possible to steal PIN number entered using eye gazes. An impostor should place two cameras, one in front of the person (e.g. under a screen) and one pointing at the screen. Proper synchronization of images from both cameras, together with ensuring a high quality image from the camera located in front of the person, should give sufficient amount of information to resolve the PIN. It is also theoretically possible to use only one camera in front of the person to obtain some valuable information about PIN. However, these methods are more complicated than a classic shoulder surfing so, in general, the eye gaze based PIN entering may be treated as more secure than the keyboard based one.

References

[1] Brooks M., Aragon C. R., Komogortsev O. V., Perceptions of interfaces for eye movement biometrics, In: Biometrics (ICB), 2013 International Conference on, IEEE (2013): 1–8, Forget A., Chiasson S., Biddle R., Input precision for gaze-based graphical passwords, In: CHI'10 Extended Abstracts on Human Factors in Computing Systems, ACM (2010): 4279–4284.

[2] Bulling A., Alt F., Schmidt A., Increasing the security of gaze-based cued-recall graphical passwords using saliency masks, In:Proceedings of the 2012 ACM annual conference on Human Factors in Computing Systems, ACM (2012): 3011–3020.

[3] De Luca A., Denzel M., Hussman H., Look into my eyes!: Can you guess my password?, In: Proceedings of the 5th Symposium on Usable Privacy and Security, ACM (2009): 7.

[4] De Luca A., Weiss R., Drewes H., Evaluation of eye-gaze interaction methods for security enhanced PIN-entry, In: Proceedings of the 19th australasian conference on computer-human interaction: Entertaining user interfaces, ACM (2007): 199–202.

[5] Goldberg J. H.; Kotval X. P., Computer interface evaluation using eye movements: methods and constructs. International Journal of Industrial Ergonomics, 24(6) (1999): 631–645.

[6] Jacob R. J. K., The use of eye movements in human-computer interaction techniques: what you look at is what you get, ACM Transactions on Information Systems (TOIS), 9(2) (1991): 152–169.

[7] Kumar M., et al., Reducing shoulder-surfing by using gaze-based password entry, In: Proceedings of the 3rd symposium on Usable privacy and security, ACM (2007): 13–19.

[8] Tan D. S., Keyani P., Czerwinski M., Spy-resistant keyboard: more secure password entry on public touch screen displays, In:Proceedings of the 17th Australia conference on Computer-Human Interaction: Citizens Online: Considerations for Today and the Future, Computer-Human Interaction Special Interest Group (CHISIG) of Australia (2005): 1–10.

[9] Weaver J., Mock K., Hoanca B., Gaze-based password authentication through automatic clustering of gaze points, In: Systems, Man, and Cybernetics (SMC), 2011 IEEE International Conference on, IEEE (2011): 2749–2754.