

TouchGazePath: Multimodal Interaction with Touch and Gaze Path for Secure Yet Efficient PIN Entry

Chandan Kumar
University of Koblenz-Landau
Institute WeST
Koblenz, Germany
kumar@uni-koblenz.de

Daniyal Akbari
University of Koblenz-Landau
Institute WeST
Koblenz, Germany
akbari@uni-koblenz.de

Raphael Menges
University of Koblenz-Landau
Institute WeST
Koblenz, Germany
raphaelmenges@uni-koblenz.de

I Scott MacKenzie
York University
Dept of EE and Computer Science
Toronto, Canada
mack@cse.yorku.ca

Steffen Staab*
University of Koblenz-Landau
Institute WeST
Koblenz, Germany
staab@uni-koblenz.de

ABSTRACT

We present TouchGazePath, a multimodal method for entering personal identification numbers (PINs). Using a touch-sensitive display showing a virtual keypad, the user initiates input with a touch at any location, glances with their eye gaze on the keys bearing the PIN numbers, then terminates input by lifting their finger. TouchGazePath is not susceptible to security attacks, such as shoulder surfing, thermal attacks, or smudge attacks. In a user study with 18 participants, TouchGazePath was compared with the traditional Touch-Only method and the multimodal Touch+Gaze method, the latter using eye gaze for targeting and touch for selection. The average time to enter a PIN with TouchGazePath was 3.3 s. This was not as fast as Touch-Only (as expected), but was about twice as fast the Touch+Gaze. TouchGazePath was also more accurate than Touch+Gaze. TouchGazePath had high user ratings as a secure PIN input method and was the preferred PIN input method for 11 of 18 participants.

CCS CONCEPTS

- **Human-centered computing** → **Interaction techniques**;
- **Security and privacy** → *Usability in security and privacy*.

* Also with University of Southampton, WAIS Research Group

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.
ICMI '19, October 14–18, 2019, Suzhou, China

© 2019 Copyright held by the owner/author(s). Publication rights licensed to ACM.

ACM ISBN 978-1-4503-6860-5/19/10...\$15.00

<https://doi.org/10.1145/3340555.3353734>

KEYWORDS

Authentication, eye tracking, gaze path, PIN entry, multimodal interaction, usable security

ACM Reference Format:

Chandan Kumar, Daniyal Akbari, Raphael Menges, I Scott MacKenzie, and Steffen Staab. 2019. TouchGazePath: Multimodal Interaction with Touch and Gaze Path for Secure Yet Efficient PIN Entry. In *2019 International Conference on Multimodal Interaction (ICMI '19)*, October 14–18, 2019, Suzhou, China. ACM, New York, NY, USA, 10 pages. <https://doi.org/10.1145/3340555.3353734>

1 INTRODUCTION

Entering a PIN (personal identification number) is a common and essential interaction for secure operation of ATMs (cash machines), smartphones, and other computing systems. Despite the advancement of biometric authentication through fingerprint scanners and face recognition, knowledge-based approaches are still the primary way of authentication for most users and serve as a fall-back solution when alternative approaches fail [21]. A PIN is a knowledge-based approach where numeric codes are used to authenticate users to computer systems. The use of PINs as passwords for authentication is ubiquitous nowadays. For example, entering a PIN is required to operate and access an ATM machine. Other public displays also support transactions using PINs for authentication. These include buying tickets for museums, trains, and buses. Furthermore, PIN authentication is commonly used for smartphones, door locks, etc.

End users typically input their PIN by touching a keypad that is physically placed beside the screen or virtually on the screen of a public display or personal device. Hence, PIN entry is susceptible to privacy invasion and possible fraud scenarios with different kinds of security attacks. The most common attack is shoulder surfing [11], especially because

observing somebody's finger movements on a keypad is simple and unobtrusive. Furthermore, most PINs consist only of four digits. Hence, the observer can easily memorize the PIN for later use. Additionally, devices like binoculars or miniature cameras can observe the entered digit sequence. Other methods like thermal attacks [1] exploit the heat traces from a user's interaction with the keypad; smudge attacks [2] exploit the oily residues left after authenticating on touchscreens. To prevent these attacks, it is important to conceal the touch pattern from adversaries during authentication. Adding eye gaze input can potentially hide the touch pattern from observers. In this regard, the conventional multimodal method of look and touch [18, 29] can obscure the spatial positioning of selection, but give away important information about the number of entered digits. Furthermore, these methods require hand-eye coordination for entering individual keys, instigating slow and erroneous entries.

We propose TouchGazePath, a secure PIN-entry input method that combines gaze and touch in a spatio-temporal manner. For PIN entry, touch input provides temporal bracketing: The act of placing the finger anywhere on the screen and then lifting it signals the start and end of PIN entry. This is combined with eye gaze for spatial information, whereby the user glances from the first through last digits of the PIN on the virtual keypad. Hence, to enter a PIN a user simply touches the screen at an arbitrary location, glances through the digits, then lifts their finger to confirm entry. The TouchGazePath approach is (i) secure (revealing neither the spatial position or count of the digits entered), (ii) usable (employing a conventional keypad layout), and (iii) fast (using a single touch down/up action concurrently with swift eye movement).

The chief focus in this paper is to investigate how this alliance of touch and gaze performs as a PIN entry method. To achieve this, we conducted a lab study with 18 participants (2160 PIN trials). The results reveal that TouchGazePath is fast ($M = 3.3$ s; $SD = 0.14$) with clear speed and accuracy advantages over the conventional method of combining gaze with touch (Touch+Gaze), often called look-and-shoot [6, 18]. In the subjective feedback, users cited TouchGazePath as a highly secure, easy to use, and viable as PIN entry method.

2 BACKGROUND AND RELATED WORK

Researchers have evaluated the eye gaze of users as an interaction method for several decades [13, 34]. The fact that gaze is a natural form of interaction between humans, and does not require physically visible actions, has received considerable attention in supporting secure authentication. Researchers have also integrated eye tracking in public displays, like ATM machines, in which the calibration for each user is stored in the system, making it a feasible technique for PIN entry [18]. The focus herein is to use eye gaze to improve

PIN entry. We begin with a review of related gaze-based input methods and discuss the limitations that we propose to overcome via TouchGazePath.

Dwell-based Methods

One of the major challenges of gaze-based interaction is distinguishing between the user intentions of inspection and selection, termed the Midas Touch problem [13]. In this regard, on-target dwell time for selection is a well-established gaze interaction technique, and has been used for PIN entry [4, 18, 24]. However, dwelling on each key slows the PIN entry process. Furthermore, to inform the user about a completed dwell time key selection, feedback is required and this compromises security. Best and Duchowski [3] experimented with numeric keypad design using dwell time to identify each key. To avoid any accidental entry, the user fixated on a * sign to initiate an entry and then fixated on each key with a digit of the four-digit PIN. To finalize entry, the user fixated on a # sign. Thus, for entering a four-digit PIN, six dwells are required, making the interaction slow. Seetharama et al. [33] replaced dwell with blink activation, whereby the user closed their eyes for a second to confirm digit selection. However, blink-based selection is slow and unnatural for end users [23].

Gesture-based Methods

Drawing passwords with eye gaze has been explored as an input method for authentication [31]. De Luca et al. [6, 8, 9] presented EyePIN, EyePass, and EyePassShapes, which relied on a gesture alphabet, i.e., they assigned specific eye gestures to each digit. Although these methods prohibit shoulder surfing, they were extremely slow (54 s per PIN [8]), rendering them impractical for real-world use. Moreover, users need to know and remember each gesture or have the gesture table nearby during PIN entry.

Rotary design [3] is a recently proposed gesture-based PIN entry method whereby the user enters each digit in a rotary circular interface using gaze transitions from the middle to each digit and then back to the middle. The approach achieves slightly better results than dwell-based methods; however, the use of an uncommon keypad design hampers the usability.

Multimodal Methods

Combining eye gaze with additional input is another approach to overcoming dwell-based selections with eye movements. In this regard, gaze+trigger [18] and look-and-shoot [8] used eye gaze to acquire a target and a physical button press to confirm the selection of each digit in the password, i.e., the interaction principle is to gaze at desired number and confirm the selection by pressing a key (or touch). In this paper, we refer this interaction method as Touch+Gaze. Although

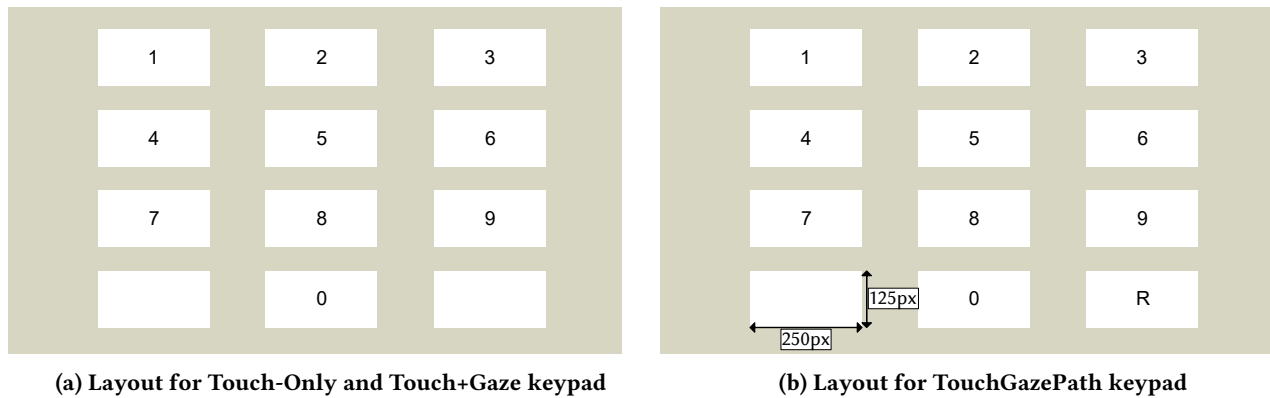


Figure 1: Keypad layouts of our study.

the approach avoids dwell time, substituting a button press for dwell time amounts to a corresponding keystroke-level delay. Furthermore, the number of button presses reveals information on the length of the password to the observer, which weakens security.

The interaction in the eye gesture-based methods of EyePass and EyePassShapes [6, 9] is also multimodal. It requires pressing a key to confirm the start and end of each gesture shape. To enter a PIN, the user draws multiple gestures, thus, multiple control key inputs are required. In contrast, we propose single-touch interaction for multi-digit PIN entry. Furthermore, the evaluation by De Luca et al. [6] compared standard touch-only interaction on a tablet PC with eye-gesture input on a desktop computer and used an external keyboard for key input. The interfaces (conventional PIN entry layout vs. drawing a shape) and interactions (touch vs. pressing a key) were notably different for each test condition. Hence, their evaluation does not aptly compare gaze-path PIN-entry with a traditional touch-only PIN method. In contrast, we keep the design and interface constant, with input method as the independent variable.

Khamis et al. [15] propose GazeTouchPIN, a multimodal authentication method for smartphones. Users first select a row of two digits via touch, then gaze left or right to select the digit. The approach avoids shoulder surfing, however, it is compromised if the attacker observes the screen as well as the user's eyes, because distinguishing between a left or right gaze is easy. GTmoPass [14] is an enhanced version of GazeTouchPIN for authentication on public displays. It is a multi-factor architecture, where the smartphone is used as a token to communicate the authentication securely to the public display. Although the method enhances security, users must hold their smartphones in their hand to gain access to their information in public displays.

3 CONCEPT OF TOUCH GAZE PATH

The TouchGazePath input method consolidates touch and gaze path interaction for text entry. Swiping one's finger through a touch-screen keyboard is now a recognized touch-path-based input method [35]. EyeSwipe uses the same concept for text entry by gaze path [19].

The challenge with swiping via eye gaze is correctly identifying the start and end point of interaction, which is critical to avoid the Midas Touch problem and to bracket the time sequence analysis of the gaze path. For this purpose, Kurauchi et al. [19] used target reverse-crossing in EyeSwipe: The user looks at the target, wherein a pop-up button appears above the target; the user looks at the pop-up button then back at the target to perform the selection. This interaction is externally observable, and, hence, is not secure. In TouchGazePath, we combine the gaze and touch modality for secure and efficient PIN entry: touch to specify the start and end point in time, and gaze to spatially glance through the digits.

Traditional Design

Novel approaches should support the entry of traditional passwords (e.g., PINs) to gain wide user acceptance [20]. Hence, we chose the conventional keypad layout (Figure 1a) for PIN entry used in ATMs or mobile devices. Figure 1b shows the layout of TouchGazePath.

One implementation issue is the entry of consecutive digits. A fixation filter might split a single fixation on a key into multiple consecutive fixations, e.g., if the precision of the eye tracking system is poor because of internal or external factors, or the user looks at different positions of the same digit rendered on one key. Thus, only the initial fixation on a key is considered as selection and consecutive fixations on the same key are ignored. Yet, some PINs might consist of repetitions of the same digit. We used an additional repeat

key (R) to allow for repetitions of entries. The repeat key is placed at the bottom-right of the layout.

In traditional keypads, the * or # keys at the bottom left and right are used to submit the PIN. However, in TouchGazePath there is no need for additional symbols, as the release of touch interaction explicitly confirms user input. During the design process, we considered various arrangements of keys to generate spatially distinct gaze paths for better accuracy. However, the focus in this paper is to assess the performance of TouchGazePath as an input method. Hence, we minimized the layout modifications in comparison to traditional PIN entry interfaces.

Gaze Path Analysis

The gaze path is analyzed between the start and end points of an interaction. The start and end points are provided by touch: The user touches the screen (start), then looks at the keys to enter the PIN, and finally releases the touch to confirm entry (end). The eye tracking system records the fixations between the start and end points of interaction. Figure 2 shows the gaze path for an example trial, plotted on the keypad stimulus. In the example, seven fixations are recorded by the eye tracking system. Fixations are first checked whether they are within the fixation-sensitive boundaries of a key. The fixation-sensitive area is larger than the key, covering the space between keys. This provides robust detection as the enlargement compensates for offsets in gaze estimation of eye tracking system [27].

In Figure 2, the first fixation f_1 is recorded over the key with the digit 2, thus 2 is the first digit, yielding PIN = “2”. The next fixation f_2 is withdrawn, as it is again registered over 2. Then fixation f_3 is registered on 8, thus, 8 is appended to the collected PIN = “28”. Fixation f_4 is not on any key and is not considered further. Fixation f_5 is on 8 and is also

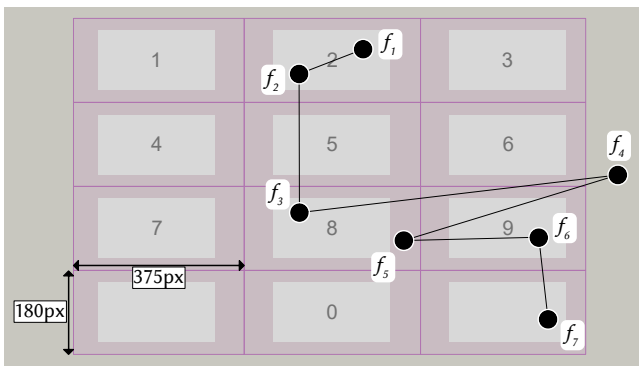


Figure 2: A gaze path with seven fixations f_1 to f_7 on the keypad. The fixation-sensitive area of each key is enlarged to cover the space between keys. The fixation-sensitive area is purple in the image. The generated PIN is “2899”.

not considered, because 8 was the last addition to the collected PIN. Fixation f_6 is on 9 and therefore, 9 is appended to the collected PIN = “289”. The last fixation within the start and end points of interaction is f_7 , which is on the R key, thus, R is appended to the collected PIN = “289R”. Finally, all occurrences of R in the collected PIN are replaced by the preceding digit, which transforms “289R” into the outputted PIN = “2899”. If a PIN has three consecutive digits, the user first fixates on the key with the desired digit, then the repeat key, and then again on the key with the desired digit.

Better Security

Similar to other gaze-based input approaches, TouchGazePath does not require the user to touch the digits on the keypad. No fingerprints or traces are left for thermal or smudge attacks. Thus, TouchGazePath is highly resistant to shoulder surfing and video analysis attacks. Conversely, dwell-based methods inherently reveal the interaction, as do touch-based methods, for which finger movements are easily observed. Unlike look-and-touch, TouchGazePath does not divulge the number of digits entered. Moreover, the number of digits can not be identified with other kinds of brute force attacks. For example, on public displays an attacker can enter random digits to understand how many digits the system accepts. In TouchGazePath, even if the system accepts an entry, an attacker would find it hard to determine the number of digits. Furthermore, most other gaze and touch techniques are vulnerable if an attacker records a video of hand and eye movements simultaneously. However, in TouchGazePath it is improbable to estimate the rapid eye movement on the intermediate digits of the PIN, especially because there are no explicit signals to correlate touch input with eye fixations.

4 METHOD

In the following, we discuss our methodology to assess the performance and usability of the TouchGazePath PIN-entry method against other state-of-the-art methods for PIN entry. We compare three input methods in our study:

- (1) Touch-Only – traditional method (baseline).
- (2) Touch+Gaze – multimodal; gaze for targeting and touch for selection [6, 18].
- (3) TouchGazePath – our method.

Participants

We recruited 19 participants for the study. For one participant (who wore thick glasses) the eye tracking calibration was not successful and a significant offset was observed. After few trials she left the experiment, and hence her data were not included. In the end, 18 participants (ten female) with mean age 27.2 years ($SD = 3.3$) completed the study, and the data from all these 18 participants are reported in the paper (we



Figure 3: Experimental setup: A participant performing the experiment using TouchGazePath on a laptop computer equipped with a touch-screen and an eye tracker.

did not artificially remove any data samples post experiment). All participants were university students. Vision was normal (uncorrected) for ten participants, while six wore glasses and two used contact lenses. Six participants had previously participated in studies with eye tracking, but these studies were not related to PIN entry. The other twelve participants had never used an eye tracker. All participants were familiar with the standard PIN entry layout. The participants were paid 10 euros for their effort after the experiment.

Apparatus

The PIN-entry interfaces were implemented as a graphical application using Python (Figure 1). We used a Lenovo 11.6" touch-screen laptop computer with a screen resolution of 1366×768 pixels. The key size is annotated in Figures 1 and 2. Eye movements were tracked using a SMI REDn scientific eye tracker with tracking frequency of 60 Hz. The eye tracker was placed at the lower edge of the screen (Figure 3). No chin rest was used. The eye tracker headbox as reported by the manufacturer is 50×30 (at 65 cm). The distance between the participant and monitor was within 45 cm, smaller than the recommended 65 cm (because the participants had to reach the touch screen), and gaze coordinates within a 2.4 degrees radius were registered as fixation. Fixation duration of 300 ms was experimentally decided, inspired from the literature discussions which vary between 250 and 600 ms [3, 32]. We utilized PyGaze [5] to perform velocity-based fixation filtering of the gaze data.

Procedure

The study was conducted in a university lab with artificial illumination and no direct sunlight. Figure 3 shows the experimental setup. Upon arrival, each participant was greeted

and given an information letter as part of the experimental protocol. The participant was then given a pre-experiment questionnaire soliciting demographic information. Testing was divided into three parts, one for each input method. Each part started with a calibration procedure performed with the eye tracker software. The software also provides the possibility to check the visual offset after a calibration. The experimenter verified this in-between blocks (entering 5 PINs is considered one block). A few times, when a participant took a longer break between two blocks (changing physical position), there was a notable visual offset and the calibration was performed again. Eye tracker calibration was not applicable to the Touch-Only method. Therefore, participants had more freedom for movement during the touch trial sequences.

A single PIN-entry trial involved a sequence of three screens in the interface. First, an information screen presented a randomly generated four-digit number. The user memorized the PIN and when ready, touched the screen to advance to the second screen to enter the PIN. A timer then started and continued to run until a four-digit PIN was entered through the virtual keypad with the current input method. After entering all four digits (and lifting the touch in TouchGazePath), timing stopped and the system checked the entered PIN. A third screen then appeared showing the result: green for a correct entry, or red for a wrong entry of the PIN.

Participants were tested on each input method, entering eight blocks of five PINs for each input method (entering five PINs is considered one block). The participants were allowed a short break between input methods and between blocks. After completing the trials, participants were asked to complete a questionnaire to provide subjective feedback on perceived familiarity, ease of use, accuracy, and security. The questions were formulated as "how would you rate the accuracy of the method", on a Likert scale of 1 to 5. We also asked participants their overall preference on which method they would like to use for PIN entry. There was no briefing on the advantage or disadvantages of any methods, nor did we provide information on different kinds of shoulder surfing or other security attacks. The experiment took about 40 minutes for each participant.

Design

The experiment was a 3×8 within-subjects design with the following independent variables and levels:

- Input Method (Touch-Only, Touch+Gaze, TouchGazePath)
- Block (1, 2, 3, 4, 5, 6, 7, 8)

A block was included to capture the participants' improvement with practice. The total number of trials was 2,160 ($=18$ participants $\times 3$ input methods $\times 8$ blocks $\times 5$ trials/block).

The dependent variables were entry time (s) and accuracy (%). Entry time was measured from the first user action to submitting the last digit of a PIN and was averaged over the five PINs in a block. Accuracy is the percentage of correct PIN entries, e.g., if 3 trials were correct in a block, the accuracy was reported as 60%. In addition, the questionnaire responses served as a qualitative dependent variable.

To offset learning effects, the three input methods were counterbalanced with participants divided into $3! = 6$ groups, thus using all six possible orders of the three input methods. A different order was used for each group; thus, “Group” was a between-subjects independent variable with six levels. Three participants were assigned to each group.

5 RESULTS

The results are provided in this section, organized by entry time and accuracy. For both these dependent variables, the group effect was not significant ($p > .05$), indicating that counterbalancing had the desired result of offsetting learning effects between the three input methods.

Entry Time

The grand mean for entry time was 4.1 s per PIN over all input methods. As expected, the conventional Touch-Only input method was fastest at 2.4 s. This was followed by our proposed TouchGazePath at 3.3 s. Touch+Gaze was comparably slow taking an average time of 6.5 s for PIN entry (Figure 4). The main effect of input method on entry time was statistically significant ($F_{2,24} = 53.97, p < .0001$).

The results for entry time by block are shown in Figure 5. The effect of block on entry time was statistically significant ($F_{7,84} = 31.62, p < .0001$), thus suggesting that participants were learning the methods over the eight blocks. However, the effect appears to be due mostly to the long entry time in block 1 for the Touch+Gaze input method. Lastly, the Input Method \times Block interaction effect on entry time was statistically significant ($F_{14,168} = 14.62, p < .0001$).

Accuracy

The grand mean for accuracy was 77.8%. Touch-Only at 96.3% was expectedly the most accurate method, followed by TouchGazePath at 72.2%, then Touch+Gaze at 65.0% (Figure 6). The differences by input method were statistically significant ($F_{2,24} = 69.97, p < .0001$).

The block effect for accuracy is shown in Figure 7. The effect of block on accuracy was statistically significant ($F_{7,84} = 6.72, p < .0001$). Although the effect was fairly flat for the Touch-Only input method, there was a clear improvement with practice for the Touch+Gaze and TouchGazePath input methods. Lastly, there was also a significant interaction effect ($F_{14,168} = 2.01, p < .05$).

It is interesting to note the drop in accuracy during block 5 for the Touch+Gaze method. Detailed analyses revealed that one participant had very low accuracy for the Touch+Gaze method (perhaps due to hand-eye coordination). For block 5, she entered all PINs wrongly with zero accuracy. At the same time her entry speed was very high, so she was focusing on faster input while making errors. After several consecutive wrong entries during block 4 and 5 she changed her strategy.

Subjective Feedback

We solicited participants’ feedback on their familiarity, ease of use, accuracy, and security of the three input methods. Questionnaire responses were on a 5-point Likert scale. Figure 8 shows the results. Higher scores are better. Not surprisingly, participants judged Touch-Only as the most familiar, as it is widely used. Understandably, Touch-Only was also considered the most easy to use due to its familiarity.

Touch+Gaze and TouchGazePath were novel input methods for participants. However, they still scored high on ease

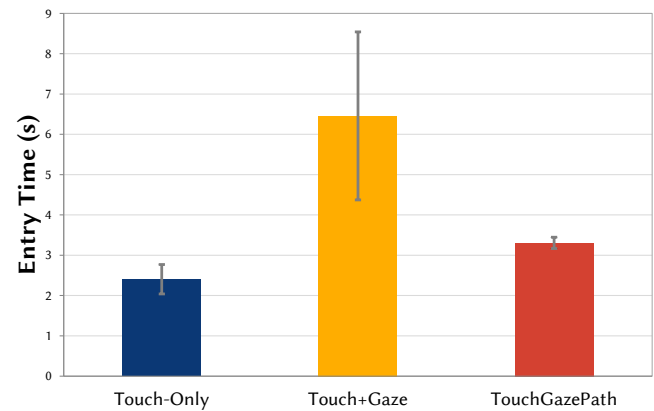


Figure 4: Entry time by input method in seconds. Error bars indicate ± 1 SD.

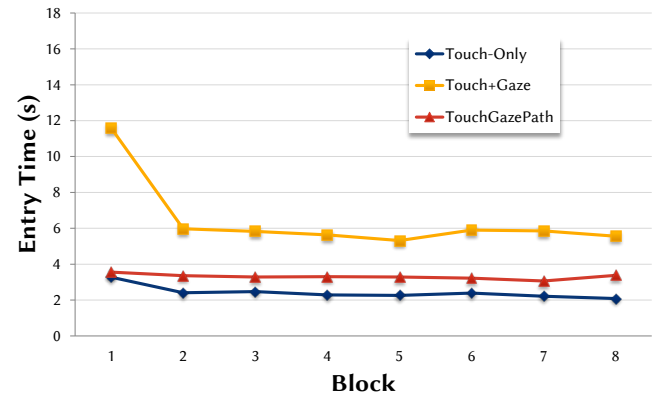


Figure 5: Entry time by input method and block in seconds.

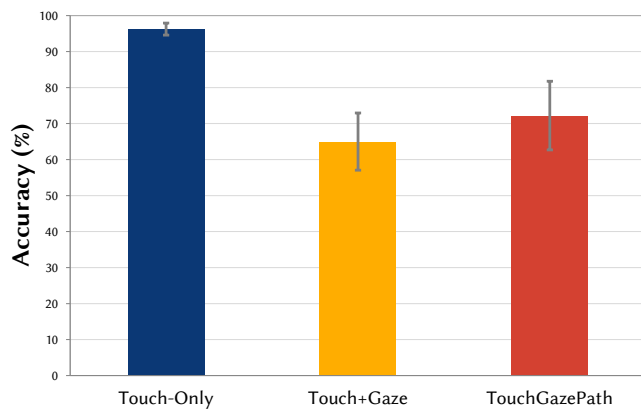


Figure 6: Accuracy (%) by input method. Error bars indicate ± 1 SD.

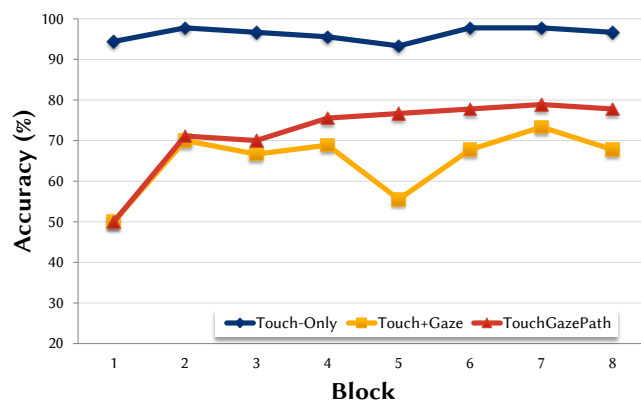


Figure 7: Accuracy (%) by input method and block.

of use. Subjective impressions of accuracy aligned with the quantitative outcomes discussed above. Most importantly, participants noted that Touch+Gaze and TouchGazePath are more secure than Touch-Only, which corresponds to our hypothesis.

Furthermore, we asked participants their overall preference on which method they would like to use for PIN entry. Eleven of the participants selected TouchGazePath as their preferred method, six participants opted for Touch+Gaze and one chose to stay with the Touch-Only method. The preference of TouchGazePath was primarily based on its novelty, enhanced security, and a kind of magical experience to touch – looking over the keys – and lift.

6 DISCUSSION

From the experimental results, it is evident that Touch-Only is the best performing PIN entry method. This is understandable because Touch-Only is in wide use today. Users also gave it a high average score of 4.6 on the familiarity scale.

The proposed TouchGazePath method, despite its novelty among users (average score of 2.3 on familiarity), showed commendable performance with an entry time of 3.3 s per PIN, slower by less than a second compared to Touch-Only at 2.4 s. Figure 4 shows the consistency in entry time of TouchGazePath and Touch-Only with very low standard deviations, i.e., all participants consistently took about two or three seconds with Touch-Only or TouchGazePath methods respectively.

Given the level of security with TouchGazePath, users might accept the small delay for enhanced privacy in authentication. This is supported by post-experiment feedback, as 11 participants expressed a preference for TouchGazePath for PIN entry. In comparison to Touch+Gaze (mean entry time 6.4 s), TouchGazePath performed significantly better, reducing the required time of PIN entry by almost half. One reason for the lower speed with Touch+Gaze could be the number of keystrokes, because multiple touch interactions are required for a four digit PIN entry. Furthermore, it could be difficult for users to coordinate touch interactions with gaze fixations for individual PIN digits.

Except for the experimental setup variations in different papers, TouchGazePath is the fastest in comparison to state-of-the-art gaze-based PIN entry authentication approaches (6 - 54 s) [6–8, 10, 15, 31]. To the best of our knowledge, Best and Duchowski [3] reported the lowest gaze-based PIN entry time with 4.6 s.

The average accuracy of TouchGazePath and Touch+Gaze over eight blocks was low at 72% and 65%, respectively. However, it is important to note that eye tracking is a novel interaction medium and requires training [13, 22]. This is also evident in the results, as accuracy was only 50% in the first block

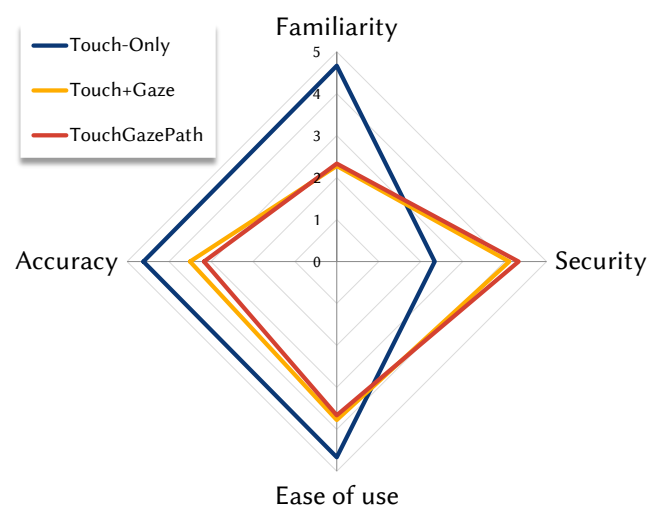


Figure 8: Subjective response (1 to 5) by input method and questionnaire item. Higher scores are better.

for both TouchGazePath and Touch+Gaze (Figure 7). Nevertheless, the reported accuracy of both gaze-based methods presents a significant gap compared to the baseline Touch-Only method. This could be attributed to various human factors such as hand-eye coordination and eye movement pattern, and most importantly the limitations of eye tracking technology itself (precision and accuracy issues), which is experienced in most gaze interactive applications [12, 17, 25]. Considering this, the accuracy of TouchGazePath is on par with other gaze-based PIN entry experiments in the literature [7, 8, 10, 15, 31]. For example, a recent gesture-based rotary design [3] exhibited 71% accuracy in a PIN-entry experiment. De Luca et al. [6] reported higher accuracy, however, they only considered critical errors – no authentication within three tries. This does not reflect the practical accuracy of gaze-based PIN entry, as presented herein.

Another interesting point is that participants entered randomly generated PINs during the experiment. However, a user in a real-world scenario has a single PIN to access their account. Hence, it is reasonable to expect better performance if users entered the same PIN repeatedly. This suggests an improvement in time and accuracy with further use of TouchGazePath.

In summary, the quantitative results showcase the effectiveness of TouchGazePath. Via qualitative feedback, users acknowledged the improved security and expressed a willingness to accept TouchGazePath as an authentication method in real world applications. In this regard, TouchGazePath provides additional opportunity for practical scenarios (instead of lifting the finger to confirm the PIN entry, the user could swipe the finger to repeat or abort an entry). Beyond its applicability as a general authentication approach, gaze-based methods are also valuable for people lacking fine motor skills [13, 16]. There are wide variety of user groups operating touch-screen tablets with eye gaze control [28, 30].¹ Many users with motor impairment can perform coarse interaction with the hands (tap), but cannot perform precise finger movements for targeting. TouchGazePath could suffice as an easy, yet, secure authentication approach for such assistive applications.

The real word deployment of gaze-based PIN entry methods is primarily dependent on reliable and robust eye gaze estimation. However in practical scenario, several factors influence eye gaze estimation of remote eye tracking systems [26]. Users may move their head, change their angle of gaze in relation to the tracking device or wear visual aids that distort the recorded geometry of the eyes. Further parameters could influence the capture of eye gaze, such as ambient lighting, the sensor resolution of the utilized camera, and

the calibration algorithm. Eye tracking technology has been continuously evolving to improve precision, accuracy, and the calibration procedure. The technological advancements would enhance the proposed work and its practical applicability in real world scenarios. In real PIN-entry scenarios where at maximum three trials are allowed, one strategy could be to provide a practice screen for user to rehearse a practice number via TouchGazePath, before entering their actual PIN.

7 CONCLUSIONS AND FUTURE WORK

It is often argued that biometric systems will solve the problems of user authentication with public displays and personal devices. However, biometric systems are still error-prone. For instance, fingerprint scanners are sensitive to humidity in the air. Furthermore, biometric features are unchangeable. Once recorded or given away they can not be changed, even though their security might be compromised. Thus, it is still worthwhile to evaluate and improve traditional knowledge-based approaches like PIN authentication.

In this paper, we presented TouchGazePath, a multimodal input method that replaces the PIN-entry touch pattern using a gaze path to enhance security. The evaluation results demonstrated that TouchGazePath is a fast input method, and significantly better than the conventional method of combining gaze with touch. The speed of TouchGazePath was fastest in comparison to other state-of-the-art approaches of gaze-based PIN entry, while maintaining similar or better accuracy.

In future, we intend to investigate interface design variations to generate unique gaze paths for distinct PINs, and evaluate their impact on performance. We are also keen in performing field studies (ATMs or public displays) to investigate the practical feasibility of the approach. Furthermore, we are interested in assessing the performance of TouchGazePath for text entry and compare performance with other eye typing and multimodal approaches.

ACKNOWLEDGMENTS

The work was partly supported by project MAMEM that has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement number 644780. We also acknowledge the financial support by the Federal Ministry of Education and Research of Germany under the project number 01IS17095B. We would also like to thank all the participants for their effort, time, and feedback during the experiment.

¹<https://www.tobiidynavox.com/products>

REFERENCES

- [1] Yomna Abdelrahman, Mohamed Khamis, Stefan Schneegass, and Florian Alt. 2017. Stay cool! Understanding thermal attacks on mobile-based user authentication. In *Proceedings of the ACM SIGCHI Conference on Human Factors in Computing Systems (CHI '17)*. ACM, New York, 3751–3763. <https://doi.org/10.1145/3025453.3025461>
- [2] Adam J. Aviv, Katherine Gibson, Evan Mossop, Matt Blaze, and Jonathan M. Smith. 2010. Smudge attacks on smartphone touch screens. In *Proceedings of the 4th USENIX Conference on Offensive Technologies (WOOT'10)*. USENIX Association, Berkeley, CA, 1–7. <http://dl.acm.org/citation.cfm?id=1925004.1925009>
- [3] Darrell S. Best and Andrew T. Duchowski. 2016. A rotary dial for gaze-based PIN entry. In *Proceedings of the ACM Symposium on Eye Tracking Research & Applications (ETRA '16)*. ACM, New York, 69–76. <https://doi.org/10.1145/2857491.2857527>
- [4] Virginio Cantoni, Tomas Lacovara, Marco Porta, and Haochen Wang. 2018. A Study on Gaze-Controlled PIN Input with Biometric Data Analysis. In *Proceedings of the 19th International Conference on Computer Systems and Technologies*. ACM, 99–103.
- [5] Edwin S Dalmaijer, Sebastiaan Mathôt, and Stefan Van der Stigchel. 2013. PyGaze: an open-source, cross-platform toolbox for minimal-effort programming of eye-tracking experiments. *Behavior Research Methods* 46, 4 (2013), 1–16.
- [6] Alexander De Luca, Martin Denzel, and Heinrich Hussmann. 2009. Look into my eyes! Can you guess my password?. In *Proceedings of the 5th Symposium on Usable Privacy and Security (SOUPS '09)*. ACM, New York, Article 7, 12 pages. <https://doi.org/10.1145/1572532.1572542>
- [7] Alexander De Luca, Katja Hertzschuch, and Heinrich Hussmann. 2010. ColorPIN: Securing PIN entry through indirect input. In *Proceedings of the ACM SIGCHI Conference on Human Factors in Computing Systems (CHI '10)*. ACM, New York, 1103–1106. <https://doi.org/10.1145/1753326.1753490>
- [8] Alexander De Luca, Roman Weiss, and Heiko Drewes. 2007. Evaluation of eye-gaze interaction methods for security enhanced PIN-entry. In *Proceedings of the 19th Australasian Conference on Computer-Human Interaction (OZCHI '07)*. ACM, New York, 199–202. <https://doi.org/10.1145/1324892.1324932>
- [9] Alexander De Luca, Roman Weiss, Heinrich Hussmann, and Xueli An. 2008. Eyepass - Eye-stroke Authentication for Public Terminals. In *CHI '08 Extended Abstracts on Human Factors in Computing Systems (CHI EA '08)*. ACM, New York, USA, 3003–3008. <https://doi.org/10.1145/1358628.1358798>
- [10] Paul Dunphy, Andrew Fitch, and Patrick Olivier. 2008. Gaze-contingent passwords at the ATM. In *The 4th Conference on Communication by Gaze Interaction (COGAIN '08)*. COGAIN, Prague, Czech Republic, 2–5.
- [11] Malin Eiband, Mohamed Khamis, Emanuel von Zetzschwitz, Heinrich Hussmann, and Florian Alt. 2017. Understanding shoulder surfing in the wild: Stories from users and observers. In *Proceedings of the ACM SIGCHI Conference on Human Factors in Computing Systems (CHI '17)*. ACM, New York, 4254–4265. <https://doi.org/10.1145/3025453.3025636>
- [12] Anna Maria Feit, Shane Williams, Arturo Toledo, Ann Paradiso, Harish Kulkarni, Shaun Kane, and Meredith Ringel Morris. 2017. Toward everyday gaze input: Accuracy and precision of eye tracking and implications for design. In *Proceedings of the ACM SIGCHI Conference on Human Factors in Computing Systems (CHI '17)*. ACM, New York, 1118–1130. <https://doi.org/10.1145/3025453.3025599>
- [13] Robert J. K. Jacob. 1990. What you look at is what you get: Eye movement-based interaction techniques. In *Proceedings of the ACM SIGCHI Conference on Human Factors in Computing Systems (CHI '90)*. ACM, New York, 11–18. <https://doi.org/10.1145/97243.97246>
- [14] Mohamed Khamis, Regina Hasholzner, Andreas Bulling, and Florian Alt. 2017. GTmoPass: Two-factor authentication on public displays using gaze-touch passwords and personal mobile devices. In *Proceedings of the 6th ACM International Symposium on Pervasive Displays (PerDis '17)*. ACM, New York, Article 8, 9 pages. <https://doi.org/10.1145/3078810.3078815>
- [15] Mohamed Khamis, Mariam Hassib, Emanuel von Zetzschwitz, Andreas Bulling, and Florian Alt. 2017. GazeTouchPIN: Protecting sensitive data on mobile devices using secure multimodal authentication. In *Proceedings of the 19th ACM International Conference on Multimodal Interaction (ICMI 2017)*. ACM, New York, 446–450. <https://doi.org/10.1145/3136755.3136809>
- [16] Chandan Kumar, Raphael Menges, Daniel Müller, and Steffen Staab. 2017. Chromium based framework to include gaze interaction in web browser. In *Proceedings of the 26th international conference on world wide web companion*. International World Wide Web Conferences Steering Committee, 219–223.
- [17] Chandan Kumar, Raphael Menges, and Steffen Staab. 2016. Eye-controlled interfaces for multimedia interaction. *IEEE MultiMedia* 23, 4 (Oct 2016), 6–13. <https://doi.org/10.1109/MMUL.2016.52>
- [18] Manu Kumar, Terry Winograd, Andreas Paepcke, and Jeff Klingner. 2007. *Gaze-enhanced user interface design*. Technical Report 806. Stanford InfoLab, Stanford, CA.
- [19] Andrew Kurauchi, Wenxin Feng, Ajjen Joshi, Carlos Morimoto, and Margrit Betke. 2016. EyeSwipe: Dwell-free text entry using gaze paths. In *Proceedings of the ACM SIGCHI Conference on Human Factors in Computing Systems (CHI '16)*. ACM, New York, 1952–1956. <https://doi.org/10.1145/2858036.2858335>
- [20] Mun-Kyu Lee. 2014. Security notions and advanced method for human shoulder-surfing resistant PIN-entry. *IEEE Transactions on Information Forensics and Security* 9, 4 (April 2014), 695–708. <https://doi.org/10.1109/TIFS.2014.2307671>
- [21] Mun-Kyu Lee, Jin Yoo, and Hyeonjin Nam. 2017. Analysis and improvement on a unimodal haptic PIN-entry method. *Mobile Information Systems Volume 2017*, Article ID 6047312 (2017), 17 pages. <https://doi.org/10.1155/2017/6047312>
- [22] I. Scott MacKenzie. 2012. Evaluating eye tracking systems for computer input. In *Gaze interaction and applications of eye tracking: Advances in assistive technologies*, P. Majaranta, H. Aoki, M. Donegan, D. W. Hansen, J. P. Hansen, A. Hyrskykari, and Riihã K.-J. (Eds.). IGI Global, Hershey, PA, 205–225. <https://doi.org/10.4018/978-1-61350-098-9>
- [23] Päivi Majaranta. 2011. *Gaze Interaction and Applications of Eye Tracking: Advances in Assistive Technologies: Advances in Assistive Technologies*. IGI Global.
- [24] Mehrube Mehrubeoglu and Vuong Nguyen. 2018. Real-time eye tracking for password authentication. In *2018 IEEE International Conference on Consumer Electronics (ICCE)*. IEEE, 1–4.
- [25] Raphael Menges, Chandan Kumar, Daniel Müller, and Korok Sengupta. 2017. GazeTheWeb: A Gaze-Controlled Web Browser. In *Proceedings of the 14th Web for All Conference on The Future of Accessible Work (W4A '17)*. ACM, New York, Article 25, 2 pages. <https://doi.org/10.1145/3058555.3058582>
- [26] Raphael Menges, Chandan Kumar, and Steffen Staab. 2019. Improving User Experience of Eye Tracking-based Interaction: Introspecting and Adapting Interfaces. *ACM Trans. Comput.-Hum. Interact.* (2019). Accepted May 2019.
- [27] Darius Miniotas, Oleg Špakov, and I. Scott MacKenzie. 2004. Eye gaze interaction with expanding targets. In *Extended Abstracts of the ACM SIGCHI Conferences on Human Factors in Computing Systems (CHI EA '04)*. ACM, New York, 1255–1258. <https://doi.org/10.1145/985921.986037>

- [28] Thies Pfeiffer. 2018. Gaze-based assistive technologies. In *Smart Technologies: Breakthroughs in Research and Practice*. IGI Global, 44–66.
- [29] Ken Pfeuffer, Jason Alexander, and Hans Gellersen. 2015. Gaze + touch vs. touch: What's the trade-off when using gaze to extend touch to remote displays?. In *Proceedings of the IFIP Conference on Human-Computer Interaction (INTERACT '15)*. Springer, Berlin, 349–367. https://doi.org/10.1007/978-3-319-22668-2_27
- [30] Ken Pfeuffer and Hans Gellersen. 2016. Gaze and touch interaction on tablets. In *Proceedings of the 29th Annual Symposium on User Interface Software and Technology (UIST '16)*. ACM, New York, 301–311. <https://doi.org/10.1145/2984511.2984514>
- [31] Vijay Rajanna, Seth Polsley, Paul Taele, and Tracy Hammond. 2017. A gaze gesture-based user authentication system to counter shoulder-surfing attacks. In *Extended Abstracts of the ACM SIGCHI Conference on Human Factors in Computing Systems (CHI EA '17)*. ACM, New York, 1978–1986. <https://doi.org/10.1145/3027063.3053070>
- [32] Simon Schenk, Marc Dreiser, Gerhard Rigoll, and Michael Dorr. 2017. GazeEverywhere: enabling gaze-only user interaction on an unmodified desktop PC in everyday scenarios. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*. ACM, 3034–3044.
- [33] Mythreya Seetharama, Volker Paelke, and Carsten Röcker. 2015. SafetyPIN: Secure PIN Entry Through Eye Tracking. In *International Conference on Human Aspects of Information Security, Privacy, and Trust*. Springer, 426–435.
- [34] Colin Ware and Harutune H. Mikaelian. 1987. An evaluation of an eye tracker as a device for computer input. In *Proceedings of the ACM SIGCHI/GI Conference on Human Factors in Computing Systems and Graphics Interface (CHI+GI '87)*. ACM, New York, 183–188. <https://doi.org/10.1145/29933.275627>
- [35] Shumin Zhai and Per Ola Kristensson. 2012. The word-gesture keyboard: Reimagining keyboard interaction. *Commun. ACM* 55, 9 (2012), 91–101. <https://doi.org/10.1145/2330667.2330689>