PIN Scrambler: Assessing the Impact of Randomized Layouts on the Usability and Security of PINs

Daniel Kirkwood
Cagdas Tombul
Calum Firth
Finn MacDonald
Konstantinos Priftis
Florian Mathis
Mohamed Khamis
Florian.Mathis@glasgow.ac.uk
Mohamed.Khamis@glasgow.ac.uk
University of Glasgow
Glasgow, United Kingdom

Karola Marky
University of Glasgow
Glasgow, United Kingdom
Leibniz University Hannover
Hannover, Germany
karola.marky@itsec.uni-hannover.de

ABSTRACT

Randomizing the layout of the keypad has been proposed to improve the security of PIN entry. However, there has been no empirical quantification of its impact on usability and security. We present the first usability (N=17) and security (N=24) evaluations to compare PIN entry with the standard vs randomized layout. Our results show that randomizing the layout increases resistance to shoulder surfing and thermal attacks significantly, and has a very minor impact on entry accuracy, but it increases entry time (from ≈1.4 seconds to ≈2 seconds). We discuss how this simple approach can improve security with little impact on usability.

CCS CONCEPTS

• **Human-centered computing** → **Interaction techniques**; • **Security and privacy** → **Authentication**.

KEYWORDS

privacy, user-centered security, authentication

ACM Reference Format:
Daniel Kirkwood, Cagdas Tombul, Calum Firth, Finn MacDonald, Konstantinos Priftis, Florian Mathis, Mohamed Khamis, and Karola Marky. 2022. PIN Scrambler: Assessing the Impact of Randomized Layouts on the Usability and Security of PINs. In *MUM '22: Mobile and Ubiquitous Multimedia, November 27–30, 2022, Lisbon, Portugal.* ACM, New York, NY, USA, 6 pages. https://doi.org/10.1145/3568444.3568450

1 INTRODUCTION

Despite the increased adoption of biometric authentication, authentication schemes that rely on the knowledge factor, such as PINs, are still indispensable on smartphones. This is because they are often required as fallback schemes for biometric authentication, or because in some cases users prefer to use them to avoid sharing biometric data with third parties [21]. Driven by the need to improve the security of PINs, a simple approach is to randomize the layout of the keypad to make PINs more secure against common side channel attacks, such as shoulder surfing [9, 11] and thermal attacks [1–3, 5, 8].

Even though randomizing the layout is a simple approach, there is surprisingly no work that quantifies its impact on usability and security. In this paper, we present results of a usability study (N=17) and a security study (N=24) where we evaluated the use of scrambled and standard PIN layouts in terms of entry time, error rates, resistance to shoulder surfing attacks and resistance to thermal attacks in two different attack difficulties for each threat. Our results show that entry time is slower on scrambled layouts (≈ 2 seconds vs ≈1.4 seconds), but error rates are largely the same. We also found that while shoulder surfing from 34-inch and 68-inch distances succeeds 95.83% to 100% of the time when using the standard layout, success rates are significantly lower when using the scrambled layout (12.5% to 33.33%). Similarly, while visually inspecting thermal images taken two seconds and five seconds after authentication reveals 37.5% of PINs, they never succeed against scrambled PINs as the attacker would not know where the digits were assigned. We conclude by discussing the impact of layout randomization on the usability and security of authentication.

2 RELATED WORK

Research presented in this work draws from prior work on human-centered authentication, including novel knowledge-based authentication schemes and threat modeling.

2.1 Usable and Secure PIN Authentication Schemes

There is a large body of work aiming at protecting user authentication on everyday devices such as smartphones, virtual and augmented reality (VR/AR) headsets, and public displays. Von Zezschwitz et al. [23] presented a gesture-based authentication system where users perform simple touch gestures to provide PIN input. Their proposed system, SwiPIN, allows for fast authentications (3.7 s) and exhibits low error-rates (3%) and high shoulder surfing resistance. Mathis et al. [17] investigated the combination

MUM '22, November 27–30 202, Lisbon, Portugal

of AR glasses and keypad scrambling for secure and usable authentication on public displays. They found that scrambling the keypad layout significantly increases authentication times (5.29 s vs. 3.70 s for traditional 4-digit PIN input), but that there is no evidence of a significant impact on the number of PIN corrections and PIN entry errors [17]. Roth et al. [22] proposed an authentication scheme for public displays where the keypad's digits are randomly colored black or white. In their work, users were required to repetitively click one of the two colored buttons to enter their PIN. Other work by Khamis et al. [14] further showed how randomization (e.g., randomized visual cues) contribute towards highly secure and usable authentication on public displays. However, despite the promising results of prior works, Abdrabou et al. [4] argued that randomizing PIN layouts does not provide an acceptable usability-security trade-off. Their investigation of a gaze-based authentication system revealed limited improved observation resistance when scrambling the PIN layout, questioning the added security gained from scrambled PIN layouts w.r.t usability [4].

All in all, there is a large body of research on usable and secure authentication methods. While individual works commented on the impact of scrambling PIN/image layouts on usability and security (e.g., [4, 10, 17]), there is no formal evaluation of the impact of layout scrambling on the usability and security of traditional PINs when considering multiple different attacks (e.g., shoulder surfing attacks and thermal attacks). Section 2.2 will review and discuss different threats and threat modeling in more detail.

2.2 Threat modeling in Usable Security

Authentication schemes are commonly designed to protect users against one (or more) threat vectors. Threat modeling is often defined as the formal process of identifying, documenting, and mitigating security threats to a system [20]. Human-centered threats involve social engineering attacks [15], shoulder surfing attacks [9, 11], thermal attacks [1, 2, 5], smudge attacks [7, 24], just to name a few. Designing authentication schemes against multiple threats is challenging, mainly because protecting users against all security and privacy threats is close to impossible [18]. One of the most common threat models in the broader HCI field is *shoulder surfing* [11], where bystanders observe a user during their authentication. In contrast works that protect against observation attacks, Aviv et al. [7] and Abdelrahman et al. [1] respectively investigated how smudge attacks (i.e., oily residues) and thermal attacks (i.e., heat traces on a screen) impact the security of traditional user authentication on mobile devices. Abdelrahman et al. [1] found that thermal attacks are viable on mobile devices, with their ThermalAnalyzer prototype uncovering 72%– 100% of PINs in the first 30 seconds, and 100% of non-overlapping patterns. Aviv et al.'s smudge attacks resulted in extremely encouraging results: smudge attacks on mobile devices reduce the likely pattern space to two and there is no effect of smudge distortion caused by incidental contact with or wiping on clothing, making smudge attacks a crucial threat to Android password patterns [7]. Patterns were partially identifiable in 92% and fully identifiable in 68% of the tested settings [7].

In summary, threat modeling plays an important role in human-centered security. Novel technologies create new threats – "for example, attack models based on ubiquitously available high-resolution cameras or thermal imaging" [6], which need to be addressed when designing and implementing novel systems.

2.3 Summary Related Work

There are individual works that commented on the impact of randomized layouts on usability and security [13, 17, 23]. Previous work also carried out research on protection mechanisms against smudge attacks (e.g., [16]) and thermal attacks (e.g., different material [19]). However, there is a gap in research that studied scrambling PIN layouts and their impact on usability and security when designing against shoulder surfing attacks and more newfangled attacks such as thermal attacks. Revisiting the discussions around scrambling layouts and their impact on usability and security is important at times where attacks are no longer limited to observations but to more advanced techniques such as thermal imaging attacks [1]. As such, this paper presents the first study that assesses the impact of layout scrambling on the usability and security of PINs when designing against shoulder surfing attacks and thermal attacks.

3 CONCEPT AND IMPLEMENTATION

As shown in Figure 1-top, our implementation of the Standard PIN followed that of Android and iOS PIN keypads. For Scrambled PIN, the order of the digits was randomly generated at the beginning of the entry. The digits' positions are randomly changed when the user completes a successful or an unsuccessful authentication attempt.

4 USABILITY EVALUATION

The aim of the usability study was to evaluate the PIN scrambler's usability in terms of entry time and error rate. We followed a within-subjects design with one variable: the layout, which had two conditions: Standard vs Scrambled.

4.1 Participants and Procedure

We recruited 17 participants (8 males, 9 females; self-identified) aged between 23 and 37 (M=29.9, SD=3.7) for our usability study. The study was conducted remotely using an Android app that we developed and distributed. We recruited participants via mailing lists, social networks and word of mouth. Interested participants were emailed an info sheet and a consent form, and instructions on how to install the study app and take part in the study. Each participant had to enter 16 predefined PINs that were randomly generated and provided to the via the app, each 8 using one of the two layouts. Half of the participants started with the Standard layout while the other half started with the Scrambled layout. The participants concluded by answering questions about their demographics.

4.2 Usability Study Results

We analyzed the results by measuring the entry time and the input accuracy.

4.2.1 Entry time. We measured entry time from the moment the first digit is pressed until the moment the last one was entered. The difference scores in entry time between the standard layout and scrambled layouts were normally distributed, as assessed by Shapiro-Wilk's test ($p = .992$). A paired samples t-test showed that the scrambled layout elicited a statistically significant increase in

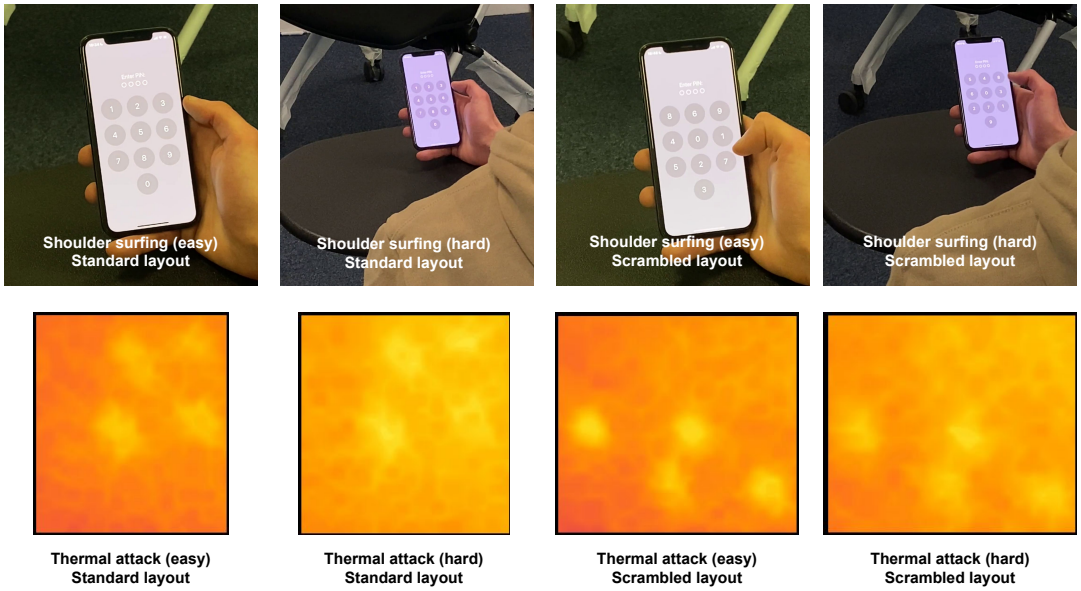

Figure 1: The figure illustrates the material shown to participants of the security study. In the shoulder surfing condition, participants watched videos of the PIN entry recorded from a 34-inch and 68-inch distance from the phone (for easy and hard attacks). In the thermal attacks condition, participants visually inspected thermal images taken of the interface taken 2-seconds and 5-seconds after PIN entry (for easy and hard attacks).

entry time compared to the standard layout M = 567.5 ms, 95% CI [182.4, 952.6], $t(16) = 3.124$, $p < .01$, $d = 0.758$. The overall mean time for entering PINs was 1,467 ms (SD = 866.6 ms) on the standard layout, and 2,034 ms (SD = 784.94 ms) on the scrambled layout.

This means entering PINs on a scrambled layout requires significantly more time compared to the standard one.

4.2.2 Accuracy. Participants performed very few errors in our experiment. In total, there were 8 errors when using the standard layout, and 9 errors when using the scrambled layout. We measured input accuracy using Levenshtein distance. The difference scores in Levenshtein distances between the standard layout and scrambled layouts were not normally distributed, as assessed by Shapiro-Wilk's test ($p < 0.001$). Thus, we ran a Wilcoxon signed-rank test but we could not find any significant differences between the Levenshtein distance when entering PINs on a standard layout compared to a scrambled one $Z = .905$, $p > 0.05$. The overall mean for Levenshtein distances was 0.09 for the standard layout and 0.1 for the scrambled layout.

This means that there is no evidence that either layout is more error prone than the other.

5 SECURITY EVALUATION

In our security evaluation of the PIN Scrambler, we considered two threat models:

- **Threat model S (Shoulder surfing)**: In this threat model, the attacker has an ideal view of the smartphone screen during PIN entry. To ensure optimal attack conditions, the attacker observes from over their victim's left shoulder, while the victim enters it with their right hand. To simulate this in our study, our participants watched videos (one time each) imitating the attack that are recorded at 34 inches away from the smartphone, and at 68 inches from the smartphone for the easy and hard conditions respectively (Fig. 1-top).
- **Threat model T (Thermal attack)**: In this threat model, the attacker records a thermal image of the smartphone screen after PIN entry. To ensure optimal attack conditions, the smartphone user does not interact with the device after PIN entry as this would typically distort the heat traces. To simulate this in our study, our participants visually inspected thermal images taken 2 seconds and 5 seconds after PIN entry for the easy and hard conditions respectively (Fig. 1-bottom).

In both threat models, a 4-digit PIN was taken to ensure comparability to prior work [1], and the attacker can make a maximum of three attempts to unlock the victim's phone using the PIN they retrieved, as normally phones lock users out after three incorrect attempts. No PIN included any duplicate digits.

The study followed a between-subjects design with two independent variables with two conditions each. The first one was the PIN layout which was either a standard layout or scrambled. The second variable is the attack difficulty. This was either easy or hard. In case of shoulder surfing, an easy attack is one where the observation is from a 34-inch distance from the screen. In case of thermal attack, it was against a thermal image taken two seconds after PIN entry. For the hard condition, we consider a 68-inch distance in case of shoulder surfing, and 4 seconds after PIN entry for thermal attacks.

The participants were divided into four groups. All participants in all groups performed a shoulder surfing and a thermal attack.

	Mean entry time	Total number of errors	Mean Levenshtein distance
Standard Layout	1467 ms	8	0.09
Scrambled Layout	2034 ms	9	0.1

Table 1: Participants were significantly faster in entering PINs on standard layouts. But we found no evidence that either layout is more error prone than the other.

Group 1 performed easy attacks against the standard layout (denoted as *EasyStandard*). Group 2 performed Hard attacks against the standard layout (*HardStandard*). Groups 3 and four performed easy and hard attacks respectively against a scrambled layout (*EasyScrambled* and *HardScrambled*).

5.1 Participants and Procedure

We invited 24 participants (14 males, 10 females; self-identified) aged between 19 and 35 (M=25.5, SD=4.1) to take part in our study. The participants were first asked to sign a consent form. The study complied with the ethics procedures of our institution. Then, the participants were explained the study which included informing them that they will have three attempts to guess each PIN, how the attack will take place, and that none of the PINs contains duplicated digits. Next, participants started with a trial run of each condition that was not included in the analysis. This was done to familiarize the participants with the task. Each participant then attacked eight different 4-digit PINs. The first four were attacked through shoulder surfing, while the other four through a thermal attack. No participant attacked the same PIN twice to avoid biased results in attacks that came later. Participants were then interviewed before concluding the study.

5.2 Security Study Results

We analyzed our results by calculating the Levenshtein distance to the correct PIN and the success rates for each threat model.

The average Levenshtein distance was calculated to understand how close the participants' guesses were to the correct PINs. To calculate the success rates, we first counted the number of correct guesses by summing up the cases where participants were able to correctly identify the PINs. The maximum number of possible guesses is 96 (24 participants×2 layouts×2 difficulties) for each threat model. Second, we calculated the success rate by dividing the number of correct guesses by the number of total guesses.

5.2.1 *Threat Model S: Shoulder surfing.*

Levenshtein distance. Overall, the mean Levenshtein distance was low for the standard layout in the easy and hard attack settings (0.04; 0.00, see also Table 2). For the scrambled layouts the distance increased to 2.00 for the easy attack setting and 1.17 for the hard one. When considering the Levenshtein distance between the shoulder surfing attacks and the original PINs, we found a statistically significant interaction between the PIN layout and attack difficulty ($F(1, 92) = 5.381$, $p = .023$, partial $\eta^2 = .055$). Thus, we analyzed the simple main effects. We found a statistically significant difference in mean Levenshtein distances between attacks on standard layouts and scrambled layouts. The mean Levenshtein distance for easy attacks is 1.958 (95% CI, 1.48 to 2.44) higher when entered on scrambled than on standard layouts ($F(1, 92) = 65.855$, $p < .001$, partial $\eta^2 = .417$). For hard attacks, it is 1.167 (95% CI, 0.69 to 1.75) higher when entered on scrambled than on standard layouts ($F(1, 92) = 23.373$, $p < .001$, partial $\eta^2 = .203$).

This means that shoulder surfing attacks against PINs entered on the scrambled Layout are significantly farther away from the original PIN than those entered on a standard Layout.

Success Rate. When considering successful shoulder surfing attack rates, we could not find any statistically significant interaction between PIN Layout and Attack difficulty ($p > 0.05$), so we compared the main effects. We found a statistically significant main effect of layout on successful attack ($F(1, 92) = 139.29$, $p < .001$, partial $\eta^2 = .602$), but we found no statistically significant main effect of difficulty on successful attack ($p > .05$). Posthoc pairwise comparisons with Bonferroni corrected p-values to adjust for multiple comparisons revealed statistically significant differences between successful attack rates against scrambled (0.229, 95% CI, .14 to .32) and standard (0.98, 95%, .9+ to 1.1) layouts ($p < .001$).

This means that shoulder surfing attacks against PINs entered on the scrambled Layout are significantly less likely to be successful than those against PINs entered on a standard layout.

5.2.2 *Threat Model T: Thermal Attacks.*

Table 2 provides on overview of the thermal attack results. Considering the standard layout, the Levenshtein distance was 1.42 for easy attacks and 1.83 for hard ones. Using the scrambled layout, the distance increased to roughly 2.9 for both attacks. We did not find any statistically significant interaction between PIN layout and attack difficulty ($p > 0.05$), so we compared the main effects. We found a statistically significant main effect of layout on successful attack ($F(1, 92) = 32.802$, $p < .001$, partial $\eta^2 = .263$, but we found no statistically significant main effect of difficulty on Levenshtein distance ($p > .05$). Posthoc pairwise comparisons with Bonferroni corrected p-values to adjust for multiple comparisons revealed statistically significant differences between Levenshtein for scrambled (2.94, 95% CI, 2.62 to 3.26) and standard (1.63, 95%, 1.3 to 1.95) layouts ($p < .001$).

This means that thermal attacks against PINs entered on the scrambled layout are significantly farther away from the original PIN than those entered on a standard layout.

When considering successful thermal attack rates, we did not find any statistically significant interaction between PIN layout and attack difficulty ($p > 0.05$), so we compared the main effects. We found a statistically significant main effect of layout on successful attack ($F(1, 92) = 27.6$, $p < .001$, partial $\eta^2 = .231$), but we found no statistically significant main effect of difficulty on successful attack ($p > .05$). Posthoc pairwise comparisons with Bonferroni corrected p-values to adjust for multiple comparisons revealed statistically significant differences between successful attack rates against scrambled (0, 95% CI, -.1 to .1) and standard (0.375, 95%, .275 to .475) layouts ($p < .001$).

This means that thermal attacks against PINs entered on the scrambled layout are significantly less likely to be successful than those against PINs entered on a standard layout.

		Shoulder Surfing Attack		Thermal Attack	
		Levenshtein Distance	Success Rate	Levenshtein Distance	Success Rate
Standard Layout	Easy	0.04	95.83%	1.42	37.50%
	Hard	0.00	100.00%	1.83	37.50%
Scrambled Layout	Easy	2.00	12.50%	2.92	0.00%
	Hard	1.17	33.33%	2.96	0.00%

Table 2: Participants were significantly more successful in attacking Easy and Hard PINs entered on the Standard Layout compared to the Scrambled Layout. Guesses against PINs entered using a Scrambled Layout were significantly farther away from the correct PIN.

6 DISCUSSION

In this section, we discuss the results of both studies showing that PIN scrambling is a promising method to mitigate shoulder surfing and thermal attacks while not impacting usability to severely.

6.1 Security versus Usability

Our security study shows that scrambled layouts impact thermal and shoulder surfing attacks. The correct guess ratios for shoulder surfing on standard layouts are in the bands of 60 and 80 percent, while the ratios for scrambled keyboards decrease to 4 and 12 percent. In addition, the difference can also be seen in the average Levenshtein distances; attackers can understand pressed numbers more than three times when regular keyboards are used rather than scrambled ones.

As for the thermal attack, even though the participants were able to identify the PINs better in the shoulder surfing phase, we can still see that scrambled layouts improve security. No participant was able to correctly guess the PINs in the scrambled layout in the hard difficulty.

Our results show that scrambling PINs might be an effective and viable solution at mitigating shoulder surfing and thermal attacks. For the case of shoulder surfing, the PIN scrambler decreased the chance of participants making a correct guess by 59%, when compared against the regular method. This is a result of attackers finding it harder to follow the order of the PIN along with the digits themselves as they were not used to the foreign layout of the keyboard.

For thermal attacks, our PIN scrambler was successful at preventing all guesses except for one. This is due to the fact that for a thermal image it is impossible to see the ordering of numbers which is essential in making a correct guess. One participant managed to successfully guess a PIN however it can be noted that this participant stated he was simply guessing, thus the result was based on blind luck. Therefore, despite that outlier, the PIN scrambler was successful at preventing all guesses and is a viable solution against thermal attacks.

The usability study reveals that entry time is statistically significantly longer on scrambled layouts compared to standard ones: 1.467 seconds vs 2.034 seconds. We did not collect data on perceived usability – this is important to investigate in future work.

Based on the two studies, we can conclude that PIN scrambling is a promising solution to safeguard PIN credentials. It mitigates shoulder surfing and thermal attacks and likely impacts further attacks not investigated by us, such as smudge attacks. Finally, it is easy to implement and distribute to users.

6.2 Limitations and Future Work

In our study, we investigated a small sample, hence our results might not be representative and should be validated with a larger more diverse sample. Since our experiment was done in the lab, the attackers could not freely move around and use further strategies, such as multiple shoulder surfers [12], to carry out the attack. Since such strategies might have am impact on attack success, this should be investigated by future work. Lastly, there are other attacks that might be prevented by PIN scrambling, such as smudge attacks. Future work should consider these attacks as well.

7 CONCLUSION

This paper presented two investigations of PIN scrambling considering security and usability. Based on our study results, we can conclude that PIN scrambling is a promising safeguard mechanism for improving the security of PINs, with little impact on usability.

ACKNOWLEDGMENTS

This work was supported by a jointly funded PhD studentship from the University of Edinburgh and the University of Glasgow, the EPSRC (EP/V008870/1), and the PETRAS National Centre of Excellence for IoT Systems Cybersecurity, which is also funded by the EPSRC (EP/S035362/1).

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 2017 CHI Conference on Human Factors in Computing Systems* (Denver, Colorado, USA) *(CHI '17)*. Association for Computing Machinery, New York, NY, USA, 3751–3763. https://doi.org/10.1145/3025453.3025461

[2] Yasmeen Abdrabou, Yomna Abdelrahman, Ahmed Ayman, Amr Elmougy, and Mohamed Khamis. 2020. Are Thermal Attacks Ubiquitous? When Non-Expert Attackers Use Off the Shelf Thermal Cameras. In *Proceedings of the International Conference on Advanced Visual Interfaces* (Salerno, Italy) *(AVI '20)*. Association for Computing Machinery, New York, NY, USA, Article 47, 5 pages. https://doi.org/10.1145/3399715.3399819

[3] Yasmeen Abdrabou, Reem Hatem, Yomna Abdelrahman, Amr Elmougy, and Mohamed Khamis. 2021. Passphrases Beat Thermal Attacks: Evaluating Text Input Characteristics Against Thermal Attacks on Laptops and Smartphones. In *Human-Computer Interaction – INTERACT 2021*, Carmelo Ardito, Rosa Lanzilotti, Alessio Malizia, Helen Petrie, Antonio Piccinno, Giuseppe Desolda, and Kori Inkpen (Eds.). Springer International Publishing, Cham, 712–721.

[4] Yasmeen Abdrabou, Mohamed Khamis, Rana Mohamed Eisa, Sherif Ismail, and Amrl Elmougy. 2019. Just Gaze and Wave: Exploring the Use of Gaze and Gestures for Shoulder-Surfing Resilient Authentication. In *Proceedings of the 11th ACM Symposium on Eye Tracking Research & Applications* (Denver, Colorado) *(ETRA '19)*. Association for Computing Machinery, New York, NY, USA, Article 29, 10 pages. https://doi.org/10.1145/3314111.3319837

[5] Norah Alotaibi, John Williamson, and Mohamed Khamis. 2022. ThermoSecure: Investigating the Effectiveness of AI-Driven Thermal Attacks on Commonly Used Computer Keyboards. *ACM Trans. Priv. Secur.* (sep 2022). https://doi.org/10.1145/3563693 Just Accepted.

[6] Florian Alt and Emanuel von Zezschwitz. 2019. Emerging trends in usable security and privacy. *i-com* 18, 3 (2019), 189–195.

[7] Adam J Aviv, Katherine Gibson, Evan Mossop, Matt Blaze, and Jonathan M Smith. 2010. Smudge attacks on smartphone touch screens. In *4th USENIX Workshop on Offensive Technologies (WOOT 10)*.

[8] Paul Bekaert, Norah Alotaibi, Florian Mathis, Nina Gerber, Aidan Christopher Rafferty, Mohamed Khamis, and Karola Marky. 2022. Are Thermal Attacks a Realistic Threat? Investigating the Preconditions of Thermal Attacks in Users' Daily Lives. In *Nordic Human-Computer Interaction Conference* (Aarhus, Denmark) *(NordiCHI '22)*. Association for Computing Machinery, New York, NY, USA, Article 76, 9 pages. https://doi.org/10.1145/3546155.3546706

[9] Leon Bošnjak and Boštjan Brumen. 2020. Shoulder surfing experiments: A systematic literature review. *Computers & Security* 99 (2020), 102023. https://doi.org/10.1016/j.cose.2020.102023

[10] Antonella De Angeli, Mike Coutts, Lynne Coventry, Graham I. Johnson, David Cameron, and Martin H. Fischer. 2002. VIP: A Visual Approach to User Authentication. In *Proceedings of the Working Conference on Advanced Visual Interfaces* (Trento, Italy) *(AVI '02)*. Association for Computing Machinery, New York, NY, USA, 316–323. https://doi.org/10.1145/1556262.1556312

[11] Malin Eiband, Mohamed Khamis, Emanuel von Zezschwitz, Heinrich Hussmann, and Florian Alt. 2017. Understanding Shoulder Surfing in the Wild: Stories from Users and Observers. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems* (Denver, Colorado, USA) *(CHI '17)*. Association for Computing Machinery, New York, NY, USA, 4254–4265. https://doi.org/10.1145/3025453.3025636

[12] Mohamed Khamis, Linda Bandelow, Stina Schick, Dario Casadevall, Andreas Bulling, and Florian Alt. 2017. They are all after you: Investigating the Viability of a Threat Model that involves Multiple Shoulder Surfers. In *Proceedings of the 16th International Conference on Mobile and Ubiquitous Multimedia* (Stuttgart, Germany) *(MUM '17)*. ACM, New York, NY, USA, 5 pages. https://doi.org/10.1145/3152832.3152851

[13] Mohamed Khamis, Mariam Hassib, Emanuel von Zezschwitz, 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* (Glasgow, UK) *(ICMI '17)*. Association for Computing Machinery, New York, NY, USA, 446–450. https://doi.org/10.1145/3136755.3136809

[14] Mohamed Khamis, Ludwig Trotter, Ville Mäkelä, Emanuel von Zezschwitz, Jens Le, Andreas Bulling, and Florian Alt. 2018. CueAuth: Comparing Touch, Mid-Air Gestures, and Gaze for Cue-Based Authentication on Situated Displays. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 2, 4, Article 174 (dec 2018), 22 pages. https://doi.org/10.1145/3287052

[15] Katharina Krombholz, Heidelinde Hobel, Markus Huber, and Edgar Weippl. 2015. Advanced social engineering attacks. *Journal of Information Security and Applications* 22 (2015), 113–122. https://doi.org/10.1016/j.jisa.2014.09.005 Special Issue on Security of Information and Networks.

[16] Taekyoung Kwon and Sarang Na. 2014. TinyLock: Affordable defense against smudge attacks on smartphone pattern lock systems. *computers & security* 42 (2014), 137–150.

[17] Florian Mathis, Joseph O'Hagan, Kami Vaniea, and Mohamed Khamis. 2022. Stay Home! Conducting Remote Usability Evaluations of Novel Real-World Authentication Systems Using Virtual Reality. In *Proceedings of the 2022 International Conference on Advanced Visual Interfaces* (Frascati, Rome, Italy) *(AVI 2022)*. Association for Computing Machinery, New York, NY, USA, Article 14, 9 pages. https://doi.org/10.1145/3531073.3531087

[18] Florian Mathis, Kami Vaniea, and Mohamed Khamis. 2022. Prototyping Usable Privacy and Security Systems: Insights from Experts. *International Journal of Human–Computer Interaction* 38, 5 (2022), 468–490. https://doi.org/10.1080/10447318.2021.1949134 arXiv:https://doi.org/10.1080/10447318.2021.1949134

[19] Keaton Mowery, Sarah Meiklejohn, and Stefan Savage. 2011. Heat of the moment: Characterizing the efficacy of thermal camera-based attacks. In *Proceedings of the 5th USENIX conference on Offensive technologies*. 6–6.

[20] Ebenezer A Oladimeji, Sam Supakkul, and Lawrence Chung. 2006. Security threat modeling and analysis: A goal-oriented approach. In *Proc. of the 10th IASTED International Conference on Software Engineering and Applications (SEA 2006)*. Citeseer, 13–15.

[21] Alexander P. Pons and Peter Polak. 2008. Understanding User Perspectives on Biometric Technology. *Commun. ACM* 51, 9 (sep 2008), 115–118. https://doi.org/10.1145/1378727.1389971

[22] Volker Roth, Kai Richter, and Rene Freidinger. 2004. A PIN-Entry Method Resilient against Shoulder Surfing. In *Proceedings of the 11th ACM Conference on Computer and Communications Security* (Washington DC, USA) *(CCS '04)*. Association for Computing Machinery, New York, NY, USA, 236–245. https://doi.org/10.1145/1030083.1030116

[23] Emanuel Von Zezschwitz, Alexander De Luca, Bruno Brunkow, and Heinrich Hussmann. 2015. Swipin: Fast and secure pin-entry on smartphones. In *Proceedings of the 33rd annual acm conference on human factors in computing systems*. 1403–1406.

[24] Emanuel von Zezschwitz, Anton Koslow, Alexander De Luca, and Heinrich Hussmann. 2013. Making Graphic-Based Authentication Secure against Smudge Attacks. In *Proceedings of the 2013 International Conference on Intelligent User Interfaces* (Santa Monica, California, USA) *(IUI '13)*. Association for Computing Machinery, New York, NY, USA, 277–286. https://doi.org/10.1145/2449396.2449432