They are all after you: Investigating the Viability of a Threat Model that involves Multiple Shoulder Surfers

Mohamed Khamis[1], **Linda Bandelow**[1], **Stina Schick**[1], **Dario Casadevall**[1]
Andreas Bulling[2], **Florian Alt**[1]
[1] LMU Munich, Germany
[2]Max Planck Institute for Informatics, Saarland Informatics Campus, Germany
firstname.lastname@ifi.lmu.de, bulling@mpi-inf.mpg.de

ABSTRACT

Many of the authentication schemes for mobile devices that were proposed lately complicate shoulder surfing by splitting the attacker's attention into two or more entities. For example, multimodal authentication schemes such as GazeTouchPIN and GazeTouchPass require attackers to observe the user's gaze input and the touch input performed on the phone's screen. These schemes have always been evaluated against single observers, while multiple observers could potentially attack these schemes with greater ease, since each of them can focus exclusively on one part of the password. In this work, we study the effectiveness of a novel threat model against authentication schemes that split the attacker's attention. As a case study, we report on a security evaluation of two state of the art authentication schemes in the case of a team of two observers. Our results show that although multiple observers perform better against these schemes than single observers, multimodal schemes are significantly more secure against multiple observers compared to schemes that employ a single modality. We discuss how this threat model impacts the design of authentication schemes.

ACM Classification Keywords

H.5.2 Information Interfaces and Presentation: User Interfaces; K.6.5 Computing Milieux: Security and Protection: Authentication

Author Keywords

Multimodal Authentication; Gaze Gestures; Multiple Observers; Threat Model; Shoulder Surfing; Privacy

INTRODUCTION

Motivated by the need to secure access to personal mobile devices that hold a variety of sensitive and private data, research has brought forth a plethora of authentication schemes that are secure against different types of side channel attacks. In particular, several proposed schemes complicate shoulder surfing attacks by splitting the observer's attention. For example, in contrast to PINs and Android patterns where an attacker needs to observe the user's touch input only, multimodal schemes require attackers to keep track of multiple modalities.

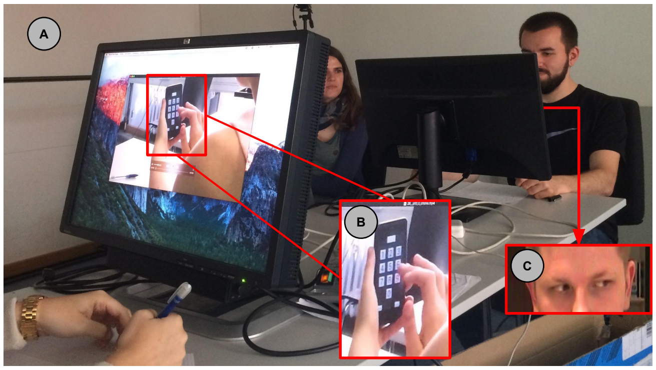

Figure 1. We investigate how well multiple observers perform against schemes that split the attacker's attention. The figure shows a team of two attackers attempting to observe a multimodal password that consists of gaze and touch input. An attacker watches a video simulating a shoulder surfer observing the user's phone (B), while the other one watches a video simulating an attacker watching the user's eyes (C).

To date, schemes were always evaluated assuming a single observer performing a one-time attack (e.g., side attack [7, 9, 10, 11]), multiple attacks (e.g., iterative [9, 11, 12], intersection [16] and insider attacks [15]), or video attacks [7, 14]. However, in theft, pick-pocketing and burglary situations, there are often multiple adversaries. For example, it is often the case that pickpockets work in teams where one distracts the victim while the other steals valuables. ATM users are sometimes subject to a team of shoulder surfers[1]. Furthermore, we know from previous research that there are situations in the real world where multiple people shoulder surf a user. For example, Eiband et al. reported a case where a user was shoulder surfed by two strangers while on the bus, the strangers then started talking about the user's destination and followed her after leaving the bus [8]. Hence, a practical and foreseeable threat that was never investigated before, is the case of multiple attackers observing the user simultaneously during authentication. Threat models have traditionally been developed with optimal conditions for the attackers in mind. This

MUM 2017, November 26–29, 2017, Stuttgart, Germany

ISBN 978-1-4503-5378-6/17/11...$15.00
DOI: https://doi.org/10.1145/3152832.3152851

[1] http://www.dailymail.co.uk/news/article-4443830/Footage-pensioner-chasing-gang-Roma-child-thieves.html

makes it crucial to understand whether this model would make the protective measures of advanced authentication schemes less effective, since each attacker can focus exclusively on one part of the password. For example, in case of multimodal authentication, each attacker could focus on a single modality.

In this work, we investigate how well GazeTouchPass and GazeTouchPIN perform against multiple observers. We found that although multiple observers perform better than single observers, multimodal schemes are still more shoulder surfing resilient compared to the unimodal baseline.

The contributions of this work are: (1) we propose a novel threat model that has the potential to render advanced authentication schemes less effective, and (2) we report on a security study in which we evaluate two state of the art schemes against this threat model and compare results to previous work.

BACKGROUND AND RELATED WORK

We build on two strands of related work: (1) Threat Models, and (2) Protection by Splitting the Attacker's Attention.

Threat Models

In the context of mobile devices, research in privacy and usable security discussed a variety of threat models. Perhaps the most discussed one is shoulder surfing, which is a type of side channel attacks where a malicious bystander tries to observe the user's interactions. A survey by Eiband et al. showed that shoulder surfing does indeed occur in daily situations, and that authentication is among the shoulder surfed activities [8].

In addition to casual shoulder surfing, more sophisticated forms of shoulder surfing were investigated in previous work. Wiese and Roth studied how the insider threat model can allow attackers to break secure schemes after multiple partial observations [15]. Several authentication schemes were evaluated against threat models in which the attacker has access to a video recording of the user during authentication (e.g., [6, 7, 13]). Other threat models include smudge attacks [2], and thermal attacks [1]. These attacks exploit the residues left on the touchscreen after the user has authenticated to find the entered password.

In contrast to previous work, in this work we explore a novel threat model in which multiple observers have the chance to simultaneously attack a user as a team.

Protection by Splitting the Attacker's Attention

Several authentication methods were introduced to overcome shoulder surfing. A widely adopted approach is to overwhelm the attacker with details that are difficult to keep track of. This is often done by splitting the attacker's attention, and requiring them to observe two or more entities. For example, XSide requires attackers to observe input on a two-sided touchscreen [7]. Similarly, being multimodal schemes, GazeTouchPass [9] and GazeTouchPIN [11] require attackers to observe the user's eyes to eavesdrop the gaze input, and the phone's screen to find the touch input. Schemes such as SwiPIN [13], PhoneLock [3], SpinLock [4], TimeLock [5] and Colorlock [5] rely on users responding to visual, haptic, or auditory cues. Hhence attackers would need to observe both the output cues in addition to the user's input in response to these cues.

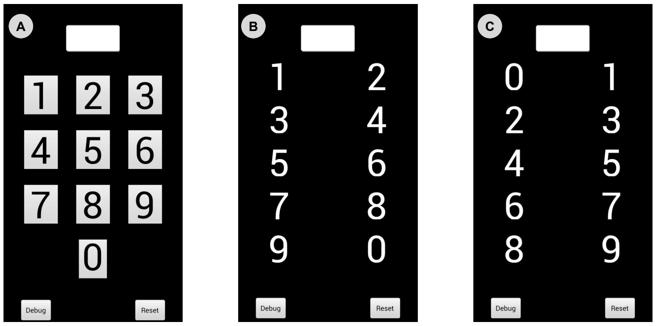

Figure 2. Layout (A) was used for GazeTouchPass. User enter passwords consisting of combinations of digits and gaze gestures to the left and right, while layouts (B) and (C) were used in GazeTouchPIN. Users select a digit by touching a row of numbers, then gazing left or right.

MULTIPLE ATTACKERS THREAT MODEL

In the threat model we propose, two attackers are simultaneously observing the user. The attackers know how to authenticate, and decide upfront on an observation strategy. In this threat model, each of the two attackers has a single chance to observe the assigned part of the authentication process from an optimal angle (see Figure 1).

After combining their observations, the attackers then get hold of the phone (e.g., by stealing it or as the user leaves it unattended), and try to log in using the observed password.

GAZETOUCHPASS AND GAZETOUCHPIN

GazeTouchPass and GazeTouchPIN were introduced and evaluated in previous work [9, 11]. They are both multimodal authentication schemes that combine gaze and touch for password entry. Hence attacking them requires observing: (1) the touch input on the touchscreen (Figure 1B), and (2) the gaze input performed by the eyes (Figure 1C). Both schemes were evaluated previously against two threat models: (1) iterative attacks, where an attacker exclusively focuses on observing one modality at an occasion, and the other modality at a different occasion, and (2) side attacks, where a single attacker observes both modalities from the side. In the threat model we introduce in this work, each of the two attackers observes one of the two modalities, then they combine their observations to find the password.

In GazeTouchPass, the password consists of numerical digits entered through the touchscreen (0-9), and gaze gestures (left and right). It was found that passwords that contain more switches from one modality to another are more secure. For example, Touch(1), Gaze(left), Touch(2), Gaze(Right) has 3 switches from one modality to another, while Touch(1), Touch(2), Gaze(left), Gaze(Right) has only 1 switch. This means that the former password is more observation-resilient. This is because as the number of switches increases, the number of consecutive inputs using the same modality decreases, which results in a larger range of possible combinations of the inputs observed from the user's touch and gaze actions. When observing a view (e.g., Figure 1B), attackers noticed pauses in-between the inputs, and inferred that the user provided input using the other modality during these pauses.

Unlike GazeTouchPass, GazeTouchPIN uses traditional PINs that are entered using touch and gaze. Users first select a pair of digits from one of two predefined layouts (Figures 2B and 2C), and then gaze left or right to select the digit to be entered. The choice of which layout to show at each entry is determined randomly. This means that if an attacker observes the touch input, and then later observes the gaze input, the attacker would not know which layout the user is responding to. Hence there is only $\frac{1}{2^n}$ chance that an iterative attack would lead to observing matching gaze and touch input. Due to this random element, iterative attacks are very less likely to succeed against GazeTouchPIN.

In the case of multiple observers, attackers can observe both inputs in parallel, hence they neither need to rely on pauses and consecutive inputs to observe GazeTouchPass, nor is the random layout of GazeTouchPIN as effective.

OBSERVATION STUDY

The main goal of this study is to investigate how the multiple observers threat model influences the security of GazeTouchPass and GazeTouchPIN. We used videos of users entering passwords using both schemes. The videos were recorded from an optimal angle during the usability evaluations of the schemes. They are the same videos used previously to evaluate the security of GazeTouchPass and GazeTouchPIN [9, 11].

Design

The study was designed as a repeated measures experiment with a single independent variable: the password type. As explained previously, GazeTouchPass passwords could consist of multiple switches in input modality. Hence, we included four conditions: 3-switches, 2-switches, 1-switch, and 0-switches. The latter condition refers to having no switches in modalities when entering the password i.e., a unimodal password. This means that when two observers attack GazeTouchPass with 0-switches, they will be both observing the same modality. This was considered a baseline in our experiment. The fifth and last condition is GazeTouchPIN. Each team of attackers observed 3 passwords of each type. This means that each team attacked 15 passwords (3 passwords × 5 password types) The conditions were counter-balanced using a latin square.

Participants

We invited 20 participants (9 females), each two were assigned to a team and took the role of a team of attackers. Each participant was awarded 10 Euro online shop voucher. An additional 20 Euro voucher was raffled among all teams such that the more successful guesses a team has, the more likely they win the additional prize. This was done to encourage participants to put an effort into trying to find the passwords.

Apparatus and Procedure

Each team watched two video clips on two different 17” displays (see Figure 1). Both videos started at the same time. Observers were free to examine the layouts (Figure 2) and to take notes at any time during the study. The pair were allowed to communicate at any time to, for example, discuss strategies. The pair were positioned at opposite sides of the table, to simulate an attacker observing the user’s face, and another one observing the user’s touchscreen. After each video, the participants had time to discuss their solution and could state up to three guesses for the password. We concluded with a semi-structured interview.

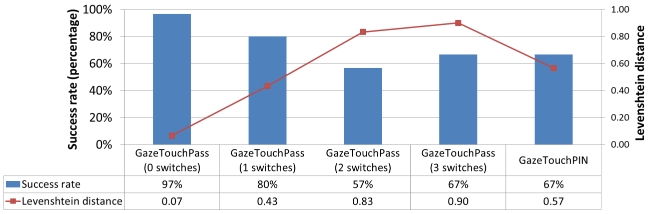

Figure 3. The figure shows that, similar to previous work, the Levenshtein distance is larger in case of 2- and 3-switches. This means that GazeTouchPass is more secure when more switches exist in the password. GazeTouchPIN is far less secure against our threat model compared to previously studied ones, since the random layout is no longer effective when two attackers observe simultaneously. Overall, while success rates are much higher in the multiple observers threat model compared to models studied in the past, both schemes still outperform the baseline.

RESULTS

To evaluate the success of the attacks, we measured (1) the Levenshtein distance between the guesses and the correct password, and (2) the success rate, which is the number of correct guesses out of all attacked passwords.

A repeated measures ANOVA with Greenhouse-Geisser correction showed a significant main effect for the password type ($F_{1.87,16.82} = 4.32$, $p < 0.05$). Post-hoc analysis with Bonferroni correction revealed a significant difference between GazeTouchPass with 0-switches and GazeTouchPass with 2-switches. Although the other pairs were not significantly different, we found a tendency for more successful guesses against GazeTouchPass with no switches, compared to GazeTouchPass with 1-, 2-, and 3-switches (see Figure 3). This matches results from previous work [9], which reported that the more switches in a GazeTouchPass password exist, the harder it is to observe. Furthermore, we found that GazeTouchPIN is less secure than many configurations of GazeTouchPass. This is expected since the random layout is no longer as effective when two attackers are observing the user at the same time.

Qualitative Feedback

The participants indicated their relationship to the other attacker in their team. In six teams, the attackers were friends, in three of them they were acquaintances, and the remaining pair were strangers. We did not find any effect of the relationship between the attackers on successful guesses. Participants reported that they devised strategies with their partners. For example, they would count in their heads to try to estimate the positions of the inputs from the other modality. The attacker who observed the touch input was able to see whether the successful login screen was shown after the last touch input, or if the last touch input was followed by a pause. This gave the attackers hints about the positions of the observed inputs.

DISCUSSION

The results indicate that a team of two attackers are better at guessing passwords compared to a single attacker. This can be seen in Table 1, which shows a comparison between the success rates in our study and success rates in previous work [9, 11]. The reason behind the higher success against the baseline condition (GazeTouchPass with 0-switches) is that both attackers saw the same video. Attackers were able to discuss their guesses afterwards, and this allowed them to fine-tune the three submitted guesses based on two observations instead of only one. Attackers performed better against the other conditions of GazeTouchPass as well due to the same reason: overall, the team had higher exposure to the password and was able to better identify the pauses between the inputs from different modalities. For example, observing Touch(1), Pause, Touch (2) in the phone view (Figure 1B) suggests that there is one or more gaze inputs in between those two touch inputs. These pauses in turn help the attackers identify how to order their observations. At the same time, the main reason behind incorrect guesses against GazeTouchPass is the ordering of inputs; in the vast majority of cases, the correct inputs were observed by the attackers, but the guessed order was incorrect (e.g., guessing Touch(1), Gaze(Left), Touch(2), Gaze(Right) instead of Gaze(Left), Touch(1), Touch(2), Gaze(Right)).

Finally, the better performance against GazeTouchPIN is due to the parallel observations. Note that in its previous evaluations, GazeTouchPIN was very secure against iterative attacks because each time the user enters a digit, the layout could have been different. This made it unfeasible for attackers to identify which layout the user is responding to when observing their eye movements. This security advantage is no longer present in case of parallel multiple observers attacks; the attacker observing the screen could note down the touch input and the shown layout, while the other one observes the gaze input. Combining the observations in this case would be trivial.

Other Authentication Schemes

In this work we evaluated GazeTouchPass and GazeTouchPIN against the multiple observers threat model because they are state-of-the-art schemes that secure against shoulder surfing by splitting the attacker's attention. However, similar results are expected in case of other similar types of schemes. For example, XSide separates the input on a double-sided touchscreen [7]. Similarly, in our threat model each attacker would focus on only one side, before they combine their observations. Schemes that rely on the user's response to auditory, visual or haptic feedback can be vulnerable to multiple attackers if one of them observes the system's output, while the other one focuses on the user's input.

Splitting the Attacker's Attention is still the Way to go

Although multiple observers perform better than single ones when attacking GazeTouchPass and GazeTouchPIN compared to single observers, their success is significantly worse than when attacking the baseline (see Table 1). This means that while these schemes are not as effective against multiple attackers as they are against single observers, they are still more secure than the baseline and hence we would recommend using them.

	GazeTouchPass				GazeTouchPIN
Number of attackers	0-switches (baseline)	1-switch	2-switches	3-switches	
Single	63%	46%	37%	23%	4%
Multiple	97%	80%	57%	67%	67%

Table 1. Compared to previous evaluations of GazeTouchPass and GazeTouchPIN [9, 11], the multiple attackers threat model results in more successful attacks against the said schemes.

CONCLUSION AND FUTURE WORK

In this work we proposed a novel threat model and evaluated two state-of-the-art authentication schemes against it. We found that the multiple observers threat model is effective and renders some of the security features of authentication schemes less effective. However, multimodal schemes perform significantly better than single model schemes in resisting multiple observers attacks.

In future work, it would be interesting to evaluate other schemes such as XSide [7], SwiPIN [13], PhoneLock [3], SpinLock [4], TimeLock [5] and Colorlock [5] against this threat model. Another interesting direction for future work is to investigate combined threat models. For example, an attacker could observe a user's gaze input while authenticating using GazeTouchPIN or GazeTouchPass, and then perform a thermal attack [1] or a smudge attack [2] to infer touch input.

ACKNOWLEDGEMENTS

This work was partially funded by the Bavarian State Ministry of Education, Science and the Arts in the framework of the Centre Digitisation.Bavaria (ZD.B), and by the Cluster of Excellence on Multimodal Computing and Interaction (MMCI) at Saarland University, Germany.

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 (CHI '17)*. ACM, New York, NY, USA, 12.

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, USA, 1–7. http://dl.acm.org/citation.cfm?id=1925004.1925009

3. Andrea Bianchi, Ian Oakley, Vassilis Kostakos, and Dong Soo Kwon. 2011. The Phone Lock: Audio and Haptic Shoulder-surfing Resistant PIN Entry Methods for Mobile Devices. In *Proceedings of the Fifth International Conference on Tangible, Embedded, and Embodied Interaction (TEI '11)*. ACM, New York, NY, USA, 197–200. DOI: http://dx.doi.org/10.1145/1935701.1935740

4. Andrea Bianchi, Ian Oakley, and DongSoo Kwon. 2011. Spinlock: A Single-Cue Haptic and Audio PIN Input Technique for Authentication. In *Haptic and Audio Interaction Design*, EricW. Cooper, VictorV. Kryssanov, Hitoshi Ogawa, and Stephen Brewster (Eds.). Lecture

Notes in Computer Science, Vol. 6851. Springer Berlin Heidelberg, 81–90. DOI: http://dx.doi.org/10.1007/978-3-642-22950-3_9

5. Andrea Bianchi, Ian Oakley, and Dong Soo Kwon. 2012. Counting clicks and beeps: Exploring numerosity based haptic and audio {PIN} entry. *Interacting with Computers* 24, 5 (2012), 409 – 422. DOI: http://dx.doi.org/10.1016/j.intcom.2012.06.005

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, NY, USA, Article 7, 12 pages. DOI: http://dx.doi.org/10.1145/1572532.1572542

7. Alexander De Luca, Marian Harbach, Emanuel von Zezschwitz, Max-Emanuel Maurer, Bernhard Ewald Slawik, Heinrich Hussmann, and Matthew Smith. 2014. Now You See Me, Now You Don't: Protecting Smartphone Authentication from Shoulder Surfers. In *Proceedings of the 32Nd Annual ACM Conference on Human Factors in Computing Systems (CHI '14)*. ACM, New York, NY, USA, 2937–2946. DOI: http://dx.doi.org/10.1145/2556288.2557097

8. 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 (CHI '17)*. ACM, New York, NY, USA, 11.

9. Mohamed Khamis, Florian Alt, Mariam Hassib, Emanuel von Zezschwitz, Regina Hasholzner, and Andreas Bulling. 2016. GazeTouchPass: Multimodal Authentication Using Gaze and Touch on Mobile Devices. In *Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems (CHI EA '16)*. ACM, New York, NY, USA, 2156–2164. DOI: http://dx.doi.org/10.1145/2851581.2892314

10. Mohamed Khamis, Regina Hasholzner, Andreas Bulling, and Florian Alt. 2017a. 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, NY, USA, Article 8, 9 pages. DOI: http://dx.doi.org/10.1145/3078810.3078815

11. Mohamed Khamis, Mariam Hassib, Emanuel von Zezschwitz, Andreas Bulling, and Florian Alt. 2017b. 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, NY, USA, 5. DOI: http://dx.doi.org/10.1145/3136755.3136809

12. Vijay Rajanna, Seth Polsley, Paul Taele, and Tracy Hammond. 2017. A Gaze Gesture-Based User Authentication System to Counter Shoulder-Surfing Attacks. In *Proceedings of the 2017 CHI Conference Extended Abstracts on Human Factors in Computing Systems (CHI EA '17)*. ACM, New York, NY, USA, 1978–1986. DOI: http://dx.doi.org/10.1145/3027063.3053070

13. 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 (CHI '15)*. ACM, New York, NY, USA, 1403–1406. DOI: http://dx.doi.org/10.1145/2702123.2702212

14. Emanuel von Zezschwitz, Paul Dunphy, and Alexander De Luca. 2013. Patterns in the Wild: A Field Study of the Usability of Pattern and Pin-based Authentication on Mobile Devices. In *Proceedings of the 15th International Conference on Human-computer Interaction with Mobile Devices and Services (MobileHCI '13)*. ACM, New York, NY, USA, 261–270. DOI: http://dx.doi.org/10.1145/2493190.2493231

15. Oliver Wiese and Roth Volker. 2016. See you next time: A model for modern shoulder surfers. In *Proceedings of the 18th International Conference on Human-computer Interaction with Mobile Devices and Services (MobileHCI '16)*.

16. David Isaac Wolinsky, Ewa Syta, and Bryan Ford. 2013. Hang with your buddies to resist intersection attacks. In *Proceedings of the 2013 ACM SIGSAC conference on Computer & communications security (CCS '13)*. ACM, New York, NY, USA, 1153–1166. DOI: http://dx.doi.org/10.1145/2508859.2516740