Glass Unlock: Enhancing Security of Smartphone Unlocking through Leveraging a Private Near-eye Display

Christian Winkler[1] **Jan Gugenheimer**[1] **Alexander De Luca**[2,3]
Gabriel Haas[1] **Philipp Speidel**[1] **David Dobbelstein**[1] **Enrico Rukzio**[1]
[1]Institute of Media Informatics, Ulm University, Ulm, Germany
[2]University of Munich (LMU), Munich, Germany; [3]DFKI GmbH, Saarbrücken, Germany
[1]*<firstname>.<lastname>@uni-ulm.de,* [2,3]*alexander.de.luca@ifi.lmu.de*

ABSTRACT
This paper presents Glass Unlock, a novel concept using smart glasses for smartphone unlocking, which is theoretically secure against smudge attacks, shoulder-surfing, and camera attacks. By introducing an additional temporary secret like the layout of digits that is only shown on the private near-eye display, attackers cannot make sense of the observed input on the almost empty phone screen. We report a user study with three alternative input methods and compare them to current state-of-the-art systems. Our findings show that Glass Unlock only moderately increases authentication times and that users favor the input method yielding the slowest input times as it avoids focus switches between displays.

Author Keywords
Smartphone; Authentication; Near-eye display; User Study

INTRODUCTION
Recent findings suggest that about 43% of smartphone users rely on some form of lock-screen to protect their phone from unwanted usage [7]. However, currently deployed smartphone authentication mechanisms like PIN and the Android unlock pattern are susceptible to different real world attacks such as smudge attacks [1], shoulder-surfing [5], or camera attacks. Especially the latter is becoming more and more of a threat with the increasing prevalence of video surveillance.

One way of protecting authentication from these attacks is to use biometric properties like fingerprints or input behavior [3]. While these are highly usable alternatives, they suffer from trust issues and the fact that they make the devices hard or impossible to share [4]. Indirect input or other kinds of software distractions [8, 9] suffer from highly reduced authentication speed and thus, negatively influence usability. As opposed to this, hardware based approaches rely on additional, external devices to provide invisible channels to the user which affect the input [2] or relocate the input to a less observable position [5]. While increasing usability, they require additional devices to be carried around.

With the advent of smart wearable devices such as smart watches and smart glasses on the consumer market, such devices are not an additional burden anymore as they are carried around anyway as part of the users' daily lives. We already see that they can be used to enhance the usability of lock-screens. For instance, Google's Android now offers to automatically disable the lock screen whenever the user's smart watch is in the near vicinity. While this may be appropriate for less concerned users, it enables new types of attacks like stealing both devices together or leveraging moments when the phone is left unattended while still in range.

CHI 2015, April 18 - 23 2015, Seoul, Republic of Korea

http://dx.doi.org/10.1145/2702123.2702316

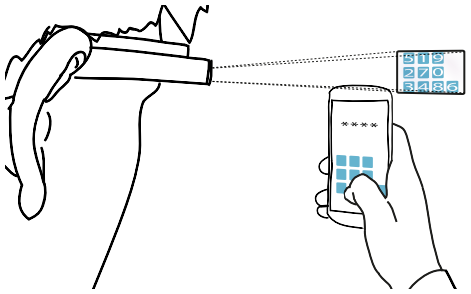


Figure 1. Glass Unlock concept: the scrambled PIN pad is only shown on the display of the user's smart glasses; input is performed on empty buttons on the smartphone, which does not give anything away to an attacker.

In this paper, we introduce a similar approach for phone unlocking, combining smart glasses with the phone's lock screen. The basic idea of Glass Unlock is to hide the lock information (e.g. PIN digits) on the phone and instead show it on the glasses' display. For instance, in a standard 10-digits PIN screen the phone would show *empty* buttons while the same layout including the digits would be visible on the glasses as shown in Figure 1. The random order of digits is required to achieve the desired security as explained later. By precluding any attackers of making sense of the users' input on the phone, Glass Unlock is secure against smudge attacks, shoulder surfing, and camera attacks.

Our main goal of this paper is to assess the additional costs of this approach compared to the state-of-the-art of unlocking. According to Harbach et al. [7], this is PIN unlocking, which about a third of all smartphone users (78% of all lock screen users) rely upon. Besides the analogue 4 out of 10 digits implementation, we further evaluated alternative variations of Glass Unlock: one that proofed to decrease the visual search time by reducing the number of digits from 10 to 6; another that proofed to support eyes-free input on the phone by requiring swipes instead of touches, thus removing any need to switch focus between the phone and the display of the glasses.

The contributions of our paper are (1) a new PIN entry concept that separates visual output from the input by moving output to the near-eye display, thereby drastically increasing unlock security; (2) a thorough study of three alternative input methods for Glass Unlock (*with* and *without* glasses). The study reveals that introducing the smart glasses for increased security only moderately increases the authentication time (by about the factor two) and users prefer the slower eyes-free input over faster alternative input methods.

GLASS UNLOCK CONCEPT
As people owning smart glasses will likely wear them most of the time, it makes sense to combine them with the people's phones to increase their security. With Glass Unlock we only look at using

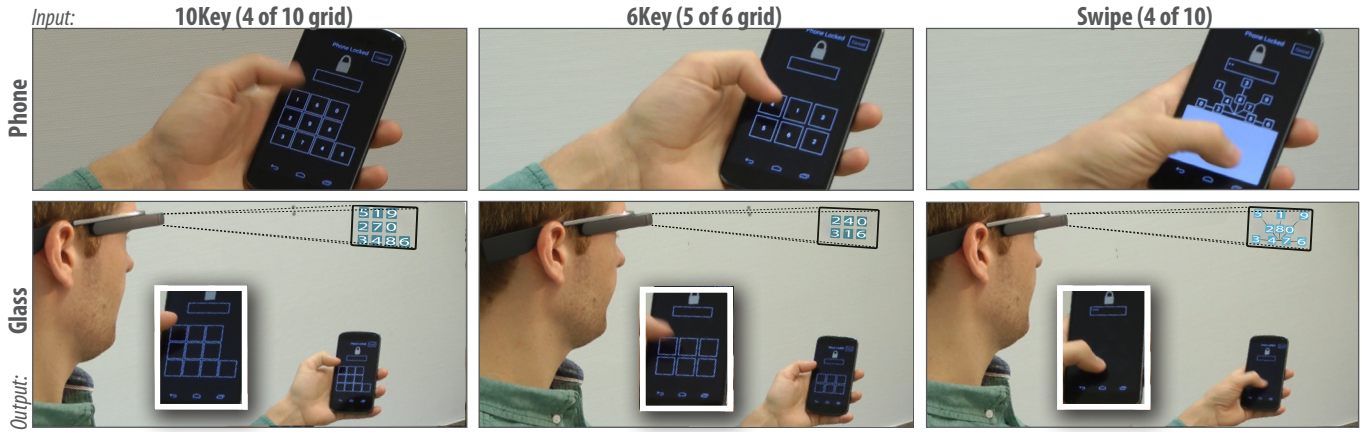


Figure 2. The 6 study systems: 3 input and 2 output methods (with and without Glass), additionally compared against standard PIN baseline (not shown).

the Glass for unlocking the phone as current smart glasses do not implement lock security so far. While the whole phone unlock could be performed on the glasses alone, users should not be forced to switch input to another device while they interact with their phone. Hence, for the Glass Unlock concept we decided to move the authentication challenge to the near-eye display while retaining authentication input on the phone. When the Glass is not available Glass Unlock gracefully degrades to scrambled PIN-entry with visible numbers on the phone.

Threat Model

Our threat model considers smudge attacks, shoulder-surfing, and camera attacks. By moving the authentication challenge from the (public) phone display to the (private) near-eye display, neither shoulder-surfing nor multiple synchronized camera observations give away the password simply because it is not shown on the phone. Small digits on the near-eye display are not visible to onlookers and cameras. In addition, Glass Unlock scrambles the order of digits after every successful unlock attempt, thus preventing attackers of merely repeating observed input on the phone, which also makes it resistant against smudge attacks.

As an attacker of Glass Unlock still has to acquire knowledge of the password, even stealing both devices together will not facilitate phone unlocking any more easily than without the glasses – this is a huge difference to previously mentioned automatic unlocking between multiple smart devices. Glass Unlock further assumes a secure Bluetooth connection between the devices, but even if the connection was compromised, an attacker of Glass Unlock would have to simultaneously record the digital transmission and observe the input on the phone. This is because no sensitive information is transmitted, only the randomized PIN layout.

By using two separate displays, Glass Unlock may require users to constantly switch between such, possibly resulting in a negative impact on completion times and user satisfaction. The following alternative input methods try to provide solutions by simplifying the visual search task or supporting eyes-free input.

Reducing Visual Search Time with *6Key*

With the standard 4 out of 10 digits PIN variant (*10Key* in Figure 2 left), users have to shift their attention multiple times between the displays unless they are able to instantly remember 4 of 10 possible positions. One possibility to decrease visual search time is to reduce the number of displayed digits (e.g. 6 digits). To maintain a large enough password space, we can increase the password length to 5 (see Figure 2 middle), resulting in 7,776 possible combinations compared to the former 10,000).

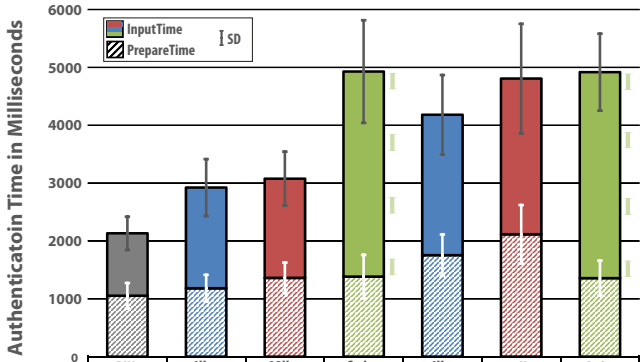
Reducing display switching time with *Swipe*

Currently available near-eye displays provide only a fixed focal length for their optics. For instance, Google Glass's virtual display is perceived in 244 cm distance while the handheld phone display is in closer proximity of only 45-60 cm to the user's eyes. Because both displays are in different focal planes, users are required to re-accommodate when switching focus between them. In our preliminary tests, this was less of an issue than initially anticipated. Yet, every focus switch adds up to the input time.

With the aforementioned 10 or 6-digit versions, focus switches are required for eye-hand coordination to hit the correct empty boxes on the phone. Relative movement such as swiping supports eyes-free input much better as touches can be performed anywhere on the screen. Swipes demand for a different mapping of the numbers, though, since these gestures are based on direction rather than position. The literature speaks against simply splitting the available space into even parts of 36 degrees because people struggle with performing eyes-free swipes of arbitrary angles [6]. We therefore designed the swipe area to consist of 5 directions (left, up/left, up, up/right, right). To increase the input space to 10 digits, all directions support two lengths, a short and a long swipe, whereby *short* is up to 1/6 of the display width and *long* anything beyond. In consequence, the input can be performed eyes-free on the phone, which in turn allows the user to focus on the glasses display all the time without having to switch attention between the displays (see Figure 2 right).

STUDY

As outlined before, we assume the presence of the Glass display and the scrambling of the PIN pad to have major negative effects on usability and input time. To test for these factors, we ran a user study with 18 participants ($M = 28$ years; $SD = 3$; 1 female). All participants were experienced smartphone users ($M = 4.39$ years; $SD = 2.05$) and 83.33% were active users of an unlock screen. Our study comprised three input techniques: *10Key*, *6Key*, and *Swipe*; and two output methods: *Phone* and *Glass* as independent variables, resulting in 6 systems overall (see Figure 2). We employed a repeated-measures design and the six systems were tested in counterbalanced order (6x6 Latin Square). In a separate entry-study, the same participants had previously provided their input to a system that mimicked a standard 4 of 10 digits PIN pad without any modification. However, participants were shortly distracted with a mental rotation task between login attempts to the 7th system to avoid input based solely on muscle memory. This study was conducted to record a baseline for each participant which we report but do not include in the statistical analysis.



	PIN Baseline	6Key	10Key Phone	Swipe	6Key	10Key Glass	Swipe
InputTime	1078	1740	1715	3544	2426	2691	3560
ANOVA		$F_{Output}(1,17)=29.25$, $\eta^2=.632$		$F_{Input}(1.5,25.3)=97.96$, $\eta^2=.852$		$F_{Input*Output}(2,34)=9.66$, $\eta^2=.362$	
PrepareTime	1053	1182	1362	1384	1755	2115	1357
ANOVA		$F_{Output}(1,17)=41.28$, $\eta^2=.708$		$F_{Input}(1.5,25.7)=12.31$, $\eta^2=.420$		$F_{Input*Output}(2,34)=11.22$, $\eta^2=.398$	
TotalTime	2131	2122	3077	4928	4181	4806	4917
ANOVA		$F_{Output}(1,17)=87.14$, $\eta^2=.837$		$F_{Input}(2,34)=87.22$, $\eta^2=.837$		$F_{Input*Output}(2,34)=27.18$, $\eta^2=.615$	

Figure 3. Authentication times divided into preparation time (time until first touch) and input time (remaining time). Swipe input time also shows – next to it – input times per digit and preparation time in-between. An ANOVA revealed all effects to be highly significant at the $p<.001$ level.

Our prototype consists of an Android Nexus 4 and a Google Glass. On every unlock attempt, the phone sends the scrambled layout to the Glass; no transmission back to the phone is required.

Procedure

For all seven systems, participants started with a training phase. They were explained the technique and had time to train until they felt comfortable with the technique and had achieved at least two correct PIN entries in a row. Participants trained with the password they would use throughout the system *at turn*. When logging started, they were presented with 5 authentication trials, for which they each had 3 possible attempts. Passwords were chosen to consist only of unique numbers to homogenize the complexity of all possible combinations. That said, resulting times should be taken as worst case estimates. Before each attempt, the password was shown on the phone and Glass to avoid memorability issues with unfamiliar passwords during the study. Password entry started only upon touching the phone screen. When an attempt failed, the user was notified by flashing borders of phone and Glass display in red. If it was not the last failed attempt, the same scrambled layout as before was presented. When the trial was successful or had finally failed, the layout was randomized and the same password was shown again, for the next trial to start. There was no possibility to correct or undo any input in order to avoid confounding input times. Also, no instructions to approach PIN-entry were given as such would be unrealistic to enforce in a real deployment.

As any input, and scrambled PIN pads in particular, require a mental preparation time before the input can actually be performed, we logged the time between starting the attempt and performing the first touch as *preparation time* and the time between the first and last touch as *input time*, respectively. After each system, the user filled out a short questionnaire on the phone that asked for perceived speed, frustration, etc. and at the end, we asked participants to compare the systems against each other and give final feedback.

Results and Discussion

Authentication Speed

Figure 3 shows the preparation, input, and total times of the systems. First we see that the total time of the baseline was lower than total times of other phone methods. This is as expected since all other systems used a scrambled PIN pad that introduced a visual search task. An Anderson-Darling test revealed all three dependent variables to be normally distributed. Thus we averaged all 5 successful trials per user and condition and conducted a two-factorial repeated measures ANOVA on the data. Mauchly's test indicated that sphericity was violated for *input method* of preparation and input times, which were corrected using Greenhouse-Geisser estimates. The ANOVA revealed both independent variables (*input method* and *output method*) and their interaction to have a highly significant effect ($p<.001$) on all three dependent variables. Further post-hoc comparisons of means (Bonferroni corrected) indicate the success of our attempt to reduce the visual search effort with the 6Key input method as it significantly decreased ($\delta=-389.61$ *ms*, $p<.001$) the preparation time compared to 10Key. Furthermore, we see that the preparation times of 6Key and 10Key significantly increase ($\delta=-432.79$ *ms*, $p<.001$) when used with Glass, but Swipe remains almost the same. This can be attributed to Swipe's support for eyes-free input when used with the Glass. In contrast, 6Key and 10Key require users to perform a mapping to the phone once they switched their focus. This leads to a higher preparation time. Also, to minimize attention shifts, users may have tried to find and remember multiple positions from the very beginning. Swipe on the other hand allowed the input to start as soon as the first digit was discovered.

Preparation time (search time) does not only happen before the first touch, but also between touches/swipes during the input time. This explains the significant rise ($\delta=-559.53$ *ms*, $p<.001$) in input times between phone and Glass methods. Again, like with preparation time, the times of 6Key and 10Key increase significantly when used with Glass as display switches occur during the input as well. Very interesting for us are the high input times of Swipe. They remain almost exactly the same between output methods, which gives strong evidence that Swipe supported eyes-free input, thus was not confounded by the separation of displays. Originally we thought the high input time of Swipe must mainly stem from performing the swipes that naturally take much longer than simple touches. Further investigation of the inter-digit preparation and input times revealed, though, that performing swipes made only for about 1s of the total time (depicted to the right of Swipe bars in Figure 3). The remaining preparation time nested in the input time of Swipe was even larger than the whole input time of 6Key or 10Key, at least on the phone. A general shift of preparation time to input time for Swipe seems plausible, as there is no need to remember two positions at once. The high difference however, can possibly be explained with the unusual layout of digits during Swipe that may have hampered the visual search task even more – but this requires further research.

Authentication Errors

During the study, we collected 681 authentications. Errors were very low across all key input systems (overall 7 errors) and thus only the *Swipe* errors with and without Glass will be discussed. Most errors occurred in the length (29 errors) of the swipe – too short or too long – or the angle (14 errors) – left or right slip. Using the Glass, participants produced more errors (19) in the length of the input than without (10). This can be attributed to the eyes-free input as the works of De Luca et. al. [6] already revealed that users struggle with swipe input more when performed eyes-free. To our very surprise, introducing the Glass did not lead to any more errors with the key input methods, despite the required switching and the possible out-of-focus touching.

Qualitative System Feedback

After each system we asked participants about their perceived speed and success during the interaction. With Glass input methods, we also asked whether the participants constantly switched

between displays or rather stayed on either of them (all using a positive and a negative formulation on Likert scales from 1 – strongly disagree to 5 – strongly agree which were averaged to a continuous interval scale). For speed, the output method ($F(1,16) = 10.39$, $p < .001$), input method ($F(2,32) = 12.16$, $p < .01$) and their interaction ($F(2,32) = 8.59$, $p < .001$) all turned out significant in a two-factorial ANOVA of the means of answers. Overall, 6Key ($M = 4.7$) was almost significantly ($\delta = 0.5$, $p = .059$) perceived faster than 10Key ($M = 4.4$) and significantly faster ($\delta = 0.84$, $p < .001$) perceived than Swipe ($M = 3.5$). Using Glass, 10Key ($M = 3.4$) was inferior to 6Key ($M = 4.0$), but not better than Swipe ($M = 3.6$) which scored between the others. Thus, although objectively Swipe was the slowest when used with the Glass, it was perceived better than 10Key. In contrast, participants estimated their own success in accordance with the actually measured errors. Finally, users were very sure about whether they switched between displays with almost total *dis*agreement that they switched with Swipe ($m = 1.5$) but significantly more often with 6Key ($M = 3.4$, $\delta = 1.97$, $p < .001$) and 10Key ($M = 4.3$, $\delta = 2.8$, $p < .001$). We also found a significant effect between 6Key and 10Key in favor of 6Key ($\delta = -0.83$, $p < .01$), which did not require as much display switches. This is interesting, as the number of digits to enter was even higher with 6Key. Yet, in the post-study questionnaire 5 participants explained to us that they could handle 6Key almost eyes-free which explains the results. 7 users attributed their increased success with 6Key to the larger keys (that would also facilitate eyes-free input) while only 3 mentioned the simplified search task.

Openly asked for comments about 10Key, 13 participants criticized the annoying display switches while regarding Swipe, 12 users explicitly mentioned to cherish canceling out of display switches. On the negative side, 4 reported problems with distinguishing between short and long swipes, 3 found short swipes harder to perform than long swipes, and 3 found Swipe too slow in general. Interestingly, in the final ranking of the three input methods by output method (Figure 4) participants shifted their sympathy nonetheless even more towards the Swipe technique when used with Glass, followed by 6Key gaining only half the sympathy on rank 1. Thus, display switches seem to be a very annoying factor in this new type of multi-display system and users would rather choose a slower input technique but which is less demanding on the eye. Finally, ≈65% of participants stated they would entirely replace their current lock screen with their favorite Glass Unlock variant if they owned compatible glasses and additional ≈18% would do so only for security critical apps.

CONCLUSION & FUTURE WORK

Users spend much of their time on unlocking their phones. With mobile devices becoming more and more a medium for highly sensitive data, secure unlock methods are researched that yield acceptable input times without requiring additional hardware to carry. In this paper we presented such an authentication system, Glass Unlock, that is inherently secure against the most common visual attacks against mobile phone locks while increasing the unlock time only moderately. Glass Unlock achieves this by outsourcing the security critical output to the near-eye display, which is believed to become a regular companion of many smartphone users, thus adding no additional hardware requirements. If a form of randomization is used to obfuscate the link between the displays, our quantitative results speak for the reduction of visual search tasks, for instance by performing a trade-off between the number of PIN digits and PIN length as in 6Key. But even more important seems the support of eyes-free input, either through large buttons on the phone (6Key) or through alternative input methods such as swiping, as it greatly improves the user experience.

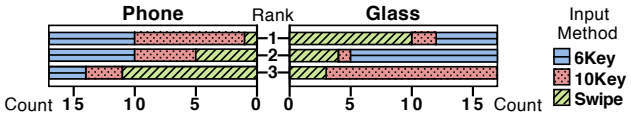


Figure 4. Final ranking of participants' preferred input method.

We also see room for future work. By comparing baseline and 10Key we have seen that scrambling of the PIN pad alone accounted for a rise of ≈44% in unlock time. Because smart glasses like the Google Glass have inbuilt proximity and motion sensors, they recognize when they are put on or off. We could take advantage of this and, instead of randomizing the PIN layout after every successful unlock attempt, only randomize it when the Glass was put off or the devices get out of range. Depending on how long the Glass is continuously worn, users may be able to memorize and adjust to the current instance of the scrambled PIN pad, thereby reducing the visual search time further. In addition, the feedback we received about improving the Swipe technique and maybe reducing it to 6 digits may lead to lower input times and fewer errors. Besides, it is important to note that the general Glass Unlock concept does not only relate to smartphone unlocking. People are required to enter secrets all the time, at the ATM, when paying with debit cards, etc. We can envision a general framework that would automatically transfer the challenge to the user's smart glasses using input methods similar to those presented in this paper.

ACKNOWLEDGMENTS

This work was conducted within the projects "Mobile Interaction with Pervasive User Interfaces" and "CompanionTechnology for Cognitive Technical Systems SFB/TRR 62" both funded by the German Research Foundation (DFG).

REFERENCES

1. Aviv, A. J., Gibson, K., Mossop, E., Blaze, M., and Smith, J. M. Smudge attacks on smartphone touch screens. In *Proc. WOOT '10*, USENIX Association (2010), 1–7.
2. Bianchi, A., Oakley, I., Kostakos, V., and Kwon, D. S. The phone lock: Audio and haptic shoulder-surfing resistant pin entry methods for mobile devices. In *Proc. TEI '11*, ACM (2011), 197–200.
3. Burgbacher, U., and Hinrichs, K. An implicit author verification system for text messages based on gesture typing biometrics. In *Proc. CHI '14*, ACM (2014), 2951–2954.
4. Coventry, L., De Angeli, A., and Johnson, G. Usability and biometric verification at the atm interface. In *Proc. CHI '03*, ACM (2003), 153–160.
5. De Luca, A., Harbach, M., von Zezschwitz, E., Maurer, M.-E., Slawik, B. E., Hussmann, H., and Smith, M. Now you see me, now you don't: Protecting smartphone authentication from shoulder surfers. In *Proc. CHI '14*, ACM (2014), 2937–2946.
6. De Luca, A., von Zezschwitz, E., Nguyen, N. D. H., Maurer, M.-E., Rubegni, E., Scipioni, M. P., and Langheinrich, M. Back-of-device authentication on smartphones. In *Proc. CHI '13*, ACM (2013), 2389–2398.
7. Harbach, M., von Zezschwitz, E., Fichtner, A., De Luca, A., and Smith, M. It's a hard lock life: A field study of smartphone (un)locking behavior and risk perception. In *Proc. SOUPS '14*, USENIX Association (July 2014), 213–230.
8. Khot, R. A., Kumaraguru, P., and Srinathan, K. WYSWYE: Shoulder Surfing Defense for Recognition Based Graphical Passwords. In *Proc. OzCHI '12*, ACM (2012), 285–294.
9. Kim, S.-H., Kim, J.-W., Kim, S.-Y., and Cho, H.-G. A new shoulder-surfing resistant password for mobile environments. In *Proc. ICUIMC '11*, ACM (2011), 27:1–27:8.