Shoulder Surfing Susceptibility of Bend Passwords

Sana Maqsood
School of Computer Science
Carleton University
Ottawa, Canada
sana.maqsood@carleton.ca

CHI 2014, April 26–May 1, 2014, Toronto, Ontario, Canada.
ACM 978-1-4503-2474-8/14/04.
http://dx.doi.org/10.1145/2559206.2579411

Abstract
The emergence of flexible displays provides us with an opportunity to explore new forms of user authentication on mobile devices. In prior work, we developed an authentication scheme utilizing bend interaction on flexible displays. A common concern among users was that this scheme may be susceptible to shoulder surfing attacks. In this paper, we evaluate the susceptibility of our scheme to such observation. We found that bend passwords are extremely difficult to observe and replicate, with only one participant correctly guessing a single password. This contradicts users' initial impressions and suggests that bend passwords are secure against shoulder-surfing.

Author Keywords
Usable Security, Authentication, Flexible Displays

ACM Classification Keywords
H.5.2 [Interfaces and Representation: User Interfaces]: Input devices and strategies

Introduction
With the emergence of flexible displays, researchers are increasingly exploring the use of bend gestures as an input technique [4, 5, 9]. As an interaction technique, bend gestures have been successfully used in the context of

Figure 1: A bend gesture performed on the flexible display.

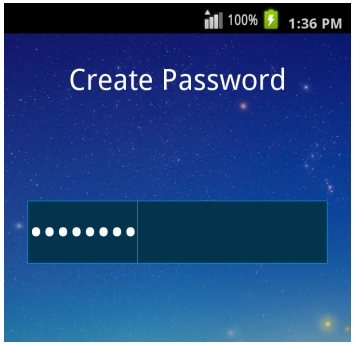

Figure 2: User Interface for both the Flexible and mobile phone prototypes.

smartphones, e-books and maps [4, 5]. When flexible displays become mainstream handheld wireless devices, they will require a means of authenticating legitimate users. Current mobile authentication schemes have weaknesses which cause users to create insecure passwords. Text passwords and PINs are difficult to remember, leading users to resort to insecure coping strategies [3]. Pattern Lock, a touch-based authentication scheme commonly used on Android devices creates low entropy passwords [1] susceptible to smudge attacks [2]. Hence, we have an opportunity to explore new forms of authentication on flexible displays by utilizing the bend interaction modality of these devices.

In prior work, we developed and user tested an authentication scheme for flexible displays [7]. Passwords in this scheme are created by performing a series of bend gestures on the flexible display. Results from our user study showed that participants liked bend passwords and could use them well, but expressed concerns that these passwords would be susceptible to shoulder surfing attacks. In this type of attack, attackers learn a password by observing users enter it in a public space, such as a bus stop. In the context of bend passwords, the visible physicality of bend gestures may make them susceptible to shoulder surfing attacks, which in turn might reduce user adoption. To address this issue, we designed a user study to determine whether bend passwords are susceptible to shoulder surfing attacks and compared the results with PINs. In this paper, we present the design, methodology and results of our user study.

Background

Bend Passwords

In our authentication scheme, passwords are created by performing a series of bend gestures on the flexible display [7]. A total of 20 gestures are available: each corner of the display could be bent up or down (8 gestures) and pairs of corners could be bent up or down simultaneously (12 gestures). We developed a flexible display prototype to test our authentication scheme because real flexible displays are not yet available. Our flexible display prototype is composed of a flexible PVC with the dimensions of 135 x 95 x 1.5 mm. Four 2" Flexpoint bidirectional bend sensors are placed in the top-left corner, top-right corner, bottom-left corner and bottom-right corner of the display. The display is connected to an Arduino Uno Microcontroller which is connected to a computer. A user interface is projected onto a screen (via a pico-projector) placed in front of the flexible display and when a gesture is entered an asterisk appears in the UI (Figure 2). Figure 1 shows how a bend gesture is performed on the flexible display. In prior work, we conducted a user study to test our authentication scheme and prototype [7]. Participants were trained on how to use the system to create a bend password. Next, they were asked to create a strong bend password on the flexible display and a PIN on a mobile phone. After creating each password, participants were asked to confirm it three times and rehearse it five times. We found that participants could successfully rehearse both bend passwords and PINs equally well. However, they took more time to rehearse their bend password ($M = 31.2s$, $SD = 18.08$) than their PIN ($M = 7.2s$, $SD = 3.42$). The results from this study show that participants could successfully use the system to create a bend password and remember it shortly after creation.

Shoulder Surfing

Authentication systems must protect users against several security threats, including password guessing attacks, interception, social engineering, and malware. With all

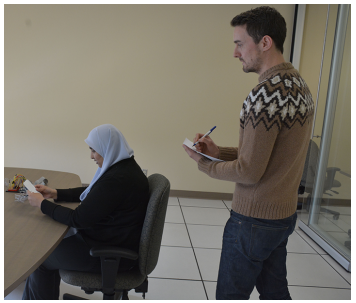

Figure 3: Shoulder Surfing study setup.

mobile devices, including flexible devices, one particularly relevant threat is shoulder surfing, an attack characterized by learning a password through observing password entry, since devices are frequently used in public spaces. Schaub et al. [8] completed a study comparing shoulder surfing susceptibility of five authentication schemes on mobile devices and found that those most resistant to shoulder surfing are also more difficult to use on mobile devices. Their study followed a similar methodology to Tari et al.'s [11] shoulder surfing study on desktop computers.

Shoulder Surfing User Study

Our study compares shoulder surfing of PINs and bend passwords. We chose PINs as a control because they are the most commonly used authentication mechanism on mobile devices. In the study, the experimenter played the victim and participants played the role of a malicious user, similar to Tari et. al's [11] and Schaub et al's [8] studies. The right-handed experimenter sat at a desk and entered passwords on a flexible display or mobile phone. The experimenter reviewed each password immediately before entering it to ensure consistency and reduce the risk of errors. Participants either stood behind the experimenter, to their right or to their left and observed them enter a password. Figure 3 shows the set-up of the experiment. During observation, participants were allowed to take notes on a piece of paper provided to them at the beginning of the study. After observing each password entry, participants were given three tries to correctly guess the observed password. After the shoulder-surfing task, participants completed an online questionnaire and a short interview. Prior to the shoulder-surfing task, participants were given an opportunity to familiarize themselves with our authentication scheme and prototype. We demonstrated each bend gesture and provided participants with the opportunity to practice and ask questions.

Passwords

Participants observed 8 passwords on each the flexible display and mobile phone. These passwords were selected using a factorial design, with password type (bend or PIN), hand position (moving or not) and password entropy (low or medium) as variables. The experimenter entered two passwords for each combination of factors, for a total of 16 passwords per participant (2 password types x 2 hand movements x 2 entropy levels X 2 trials). The presentation order of passwords was counterbalanced using a Latin square design. We selected bend passwords with a variety of gesture locations and directions, representative of the passwords created in our prior work [7]. For bend passwords, participants did not receive any visual feedback during observation but did when they guessed the password (i.e., asterisk appearing in the password field). This is a limitation of our study, but we believe it did not significantly effect our results because most participants focused on the placement of the experimenter's hands on the device rather than the UI. For PINs, we used the alphanumeric keyboard of the mobile device rather than the larger commonly used PIN keypad (Figure 4), because we believe the smaller keys of the alphanumeric keyboard would make it harder to shoulder-surf PINs. On mobile devices, the last character entered is normally displayed briefly before being obfuscated with a dot or star. We initially had this feature enabled for PINs and pilot tested our study with two participants. However, this made the PINs so easy to shoulder surf that participants correctly guessed all PINs. We decided that this would not provide a very effective comparison condition and disabled this feature before running our actual study. For both PINs and bend passwords, only a dot is displayed with each entered digit/gesture. In effect, we tried to devise the most difficult comparison condition possible to avoid overstating bend passwords' resistance to shoulder surfing.

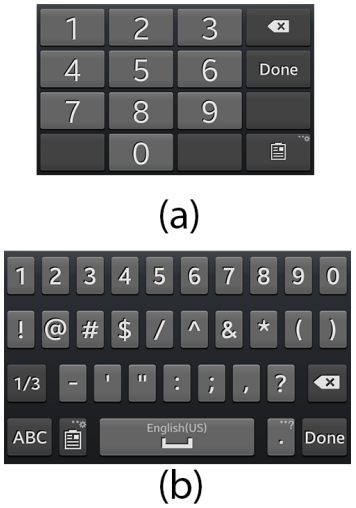

Figure 4: (a) Normal Android PIN keypad, (b) Full Android keyboard used in our study.

	Password Entropy	
Pwd	**Low**	**Medium**
PIN	6 digits	10 digits
Bend	5 gestures	8 gestures

Table 1: Length of passwords shoulder-surfed in the user study.

Hand Position: For hand position, *moving* means the experimenter's hands were moving across the device during password entry and *not-moving* means their hands were stationary. For bend passwords, the experimenter either positioned their hands on two corners of the device and performed gestures using only those corners (hand-not-moving) or performed gestures using all four corners of the device (hand moving). For PINs, the experimenter either held the phone in their right hand and entered the PIN using only their right thumb (hand-not-moving) or held the phone in their left hand and entered the PIN using their right hand, moving it across the screen (hand moving).

Password Entropy: Password entropy was calculated using Shannon's formula for information entropy [10]. For password entropy, *low* was equal to approximately 20 bits and *medium* was 34 bits. We chose 20 bits as low entropy because it was the average entropy of passwords users created in our prior work [7], and we chose 34 bits as medium entropy to match the length of bend passwords to 8 character alphanumeric text passwords. Table 1 shows the length of our passwords.

Participants

We recruited 9 participants (7 male) with an average age of 28 years. All had participated in a bend gesture study within the last 6 months. We recruited participants with prior flexible display experience to ensure they had practice using bend gestures, making them moderate users. We believe this would make them more representative of real bend password shoulder surfers. All participants were aware of shoulder surfing attacks on mobile devices and were able to describe them. Participants were given a $10 compensation.

Results

We measured shoulder surfing success rates, degree of correctness of guessed passwords and user perceptions. Most participants stood behind and slightly to the right of the experimenter because this gave them the best viewing angle. Some changed their position to their improve viewing angle. None stood to the left of the experimenter.

Success Rates: We defined success rate as the number of passwords participants' successfully guessed within three attempts after shoulder-surfing the password. For both PINs and bend passwords, the success rates were very low. Out of 144 passwords (16 passwords x 9 participants), a total of 3 were guessed correctly: one bend and two PINs. This shows that users found it difficult to shoulder surf both bend passwords and PINs.

Degree of Correctness: Given the low success rates, we conducted post-hoc analysis to explore the composition of users' guesses using Levenshtein distance [6]. Levenshtein distance is commonly used to measure the dissimilarity of two strings. It computes the number of single character edits (inserts, deletes, substitutions) needed for one string to match another (e.g., car to cat = distance of 1). A distance of 0 indicates two identical strings. When two strings are completely different, the distance is equal to the length of the longest string. In our study, PINs could be compared directly and we represented each bend gesture as a single character to form a string for a bend password. For each participant, the Levenshtein distance was calculated to compare the original password with each of their three guessed passwords. We selected the guess with the lowest distance for each password per participant, giving us 16 best guesses per user (8 bend passwords, 8 PINs). Since we performed two trials for each type of password, we selected the trial with the

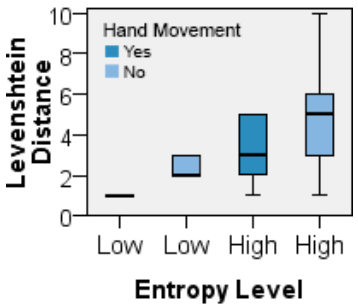

Figure 5: Levenshtein distance for PINs.

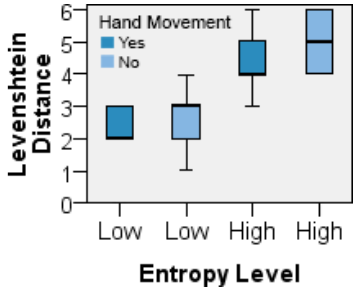

Figure 6: Levenshtein distance for Bend Passwords.

lowest Levenshtein distance, which gave us an end result of 8 passwords per participant (4 bend passwords and 4 PINs). Figures 5 and 6 show the Levenshtein distances for PINs and bend passwords respectively. Because this was post-hoc exploration with a small sample, we opted not to conduct any statistical analysis. However, the figures suggest that most users had several gestures/digits incorrect in their guesses; these were not simply one-off errors. It also suggests that passwords with hand movements or with more gestures/digits were more difficult, although this needs further testing.

Shoulder Surfing Strategies: Participants rated the difficulty of shoulder surfing passwords using a 10-point Likert scale question (1 = very difficult, 10 = very easy), and found it very difficult to observe and replicate both bend passwords and PINs. A Wilcoxon signed-rank test showed that participants found both types of passwords equally difficult to shoulder surf (Bend passwords: $Md = 2$, $SD = 3.07$; PINs: $Md = 3$, $SD = 2.55$) ($Z = -.212$, $p = 0.832$). We further asked participants to describe their strategies for shoulder surfing the passwords. For PINs, most observed the experimenter's hand movements and placement on the keys and made note of the keys pressed. A small majority (56%) wrote down the sequence of keys, and used them when entering their PIN. At the beginning of the study, many participants were confident that they would be able to easily shoulder surf PINs and were quite surprised when they could not. It appears that they initially overestimated their ability to shoulder surf PINs. Most (78%) did not change their shoulder surfing strategy throughout the study. For bend passwords, participants used a variety of strategies and changed them throughout the session (89% changed, usually more than once). The most common strategy was drawing a rectangle on paper, assigning numbers to each of the corners, and marking the observed gestures. This strategy was ineffective as participants had difficulty keeping track of the direction of each gesture. In general, taking notes did not prove to be an effective strategy, as it was difficult to observe and take notes simultaneously. When asked which type of password was most difficult to shoulder-surf, most participants answered that bend gestures were harder than PINs.

Discussion

We found that bend passwords and PINs were equally difficult to shoulder surf and success rates were extremely low for both. However, it is important to note that for PINs we used a set-up (smaller keyboard and no feedback) that is more resistant to shoulder-surfing attacks than the set-up (larger keyboard and feedback) most commonly used on mobile devices. Given this, we believe that bend passwords are more resistant to shoulder-surfing attacks than PINs. However, further research is required to confirm this. For bend passwords, users found it difficult to observe gesture movements and identify exactly which gesture had been performed. Furthermore, participants had no easy way to write down the observed gestures whereas they were able to write down the PINs without looking at their notes. Shoulder-surfers may eventually develop better note-taking strategies, but we see it as a positive result that our participants found it very difficult to shoulder-surf bend passwords despite people's initial assumptions that it would be easy. Our results also suggest that participants found it more difficult to shoulder-surf long passwords than short passwords. Our success rates do not show any difference since nearly no passwords were guessed correctly, however the Levenshtein distances were higher for longer passwords. In passwords where each gesture/digit is independent, it is reasonable to expect that longer passwords will be more difficult to shoulder-surf because there are more components to

observe. User feedback revealed that participants found it more difficult to shoulder-surf passwords when hand movement was involved. Specifically, they had a hard time shoulder-surfing passwords when the experimenter's hand was moving across the device, because they were unable to keep track of which keys or gestures were pressed.

Conclusion and Future Work

In this paper, we designed a study to look at shoulder-surfing passwords on flexible display devices and compared the results with PINs. We found that bend passwords were extremely difficult to shoulder-surf, with only one participant correctly guessing one password in the entire study. Similarly, PINs were also very difficult to shoulder-surf. Compared to PINs, we found that bend passwords are more difficult to shoulder-surf. Although, all passwords were difficult to shoulder-surf, longer passwords and passwords without hand movement were more diffcult to shoulder-surf than shorter passwords and passwords requiring hand movement. In our previous work [7], participants liked bend passwords and could use them well, but believed these passwords would be easy to shoulder-surf and thus would not be secure. The findings from our shoulder-surfing study contradict participants initial impressions and show that bend passwords are extremely difficult to shoulder-surf. These findings combined with our previous work, show that bend passwords are a secure and usable authentication mechanism on flexible displays that requires further exploration. Participants in our study had very little knowledge of computer security which could affect their ability to shoulder-surf passwords. In future work, we will look at whether computer security knowledge can affect peoples' ability to shoulder-surf passwords. We will also use the results from our studies to develop guidelines for creating passwords on flexible display devices.

Acknowledgements

I would like to thank my supervisors Sonia Chiasson and Audrey Girouard for all their guidance and support.

References

[1] http://beust.com/weblog2/archives/000497.html.

[2] http://techcrunch.com/2008/10/12/androids-login-is-cool-but-is-it-secure.

[3] Adams, A., and Sasse, M. A. Users are not the enemy. *Commun. ACM 42*, 12 (1999), 40–46.

[4] Kildal, J., Paasovaara, S., and Aaltonen, V. Kinetic Device: Designing Interactions with a Deformable Mobile Interface. In *Proc. CHI EA* (2012), 1871–1876.

[5] Lahey, B., Girouard, A., Burleson, W., and Vertegaal, R. PaperPhone: Understanding the Use of Bend Gestures in Mobile Devices with Flexible Electronic Paper Displays. In *Proc. CHI* (2011), 1303–1312.

[6] Levenshtein, V. I. Binary codes capable of correcting deletions, insertions, and reversals. *Soviet Physics Doklady 10*, 8 (1966), 707–710.

[7] Maqsood, S., Chiasson, S., and Girouard, A. Poster: Passwords on flexible display devices. In *Proc. CCS* (2013), 1469–1472.

[8] Schaub, F., Walch, M., Könings, B., and Weber, M. Exploring the design space of graphical passwords on smartphones. In *Proc. SOUPS* (2013).

[9] Schwesig, C., Poupyrev, I., and Mori, E. Gummi: a bendable computer. In *Proc. CHI* (2004), 263 – 270.

[10] Shannon, C. E. A mathematical theory of communication. *SIGMOBILE Mob. Comput. Commun. Rev. 5*, 1 (Jan. 2001), 3–55.

[11] Tari, F., Ozok, A. A., and Holden, S. H. A comparison of perceived and real shoulder-surfing risks between alphanumeric and graphical passwords. In *Proc. SOUPS* (2006), 56–66.