NEW DIRECTIONS IN MULTISENSORY AUTHENTICATION

Madoka Hasegawa[1], Nicolas Christin[2], and Eiji Hayashi[3]

Abstract
This paper discusses and evaluates two novel multisensory user authentication mechanisms aimed at preventing observation attacks. These mechanisms improve the usability of our previous work by reducing authentication times, and are more suitable for portable and mobile devices.

1. Background and Motivation

The ability to authenticate users is crucial to most modern information systems. For instance, a mobile phone has to have some primitives in place to ensure that the person using the device is its legitimate owner, and not an unauthorized third-party; a banking terminal needs the ability to establish that the person using a given banking card is the holder of the corresponding account.

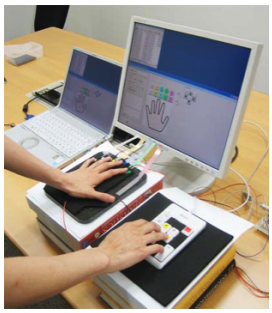
However, most user authentication systems rely on methods vulnerable to observation attacks. In an observation attack, an unauthorized party records the actions the legitimate user performs to authenticate and later reuses the recorded information to impersonate the user. A common instance of observation attacks is shoulder-surfing, where the impostor snoops over the shoulder of a legitimate user entering his/her password to successfully capture it. Besides shoulder-surfing, observation techniques that have proven successful include photo capture from a distance [1], audio recording of keyboard clicking sounds [2], miniaturized video recorders [3], and even biometric replication [4].

In previous work [3], we argued for a novel line of defense against observation attacks. Instead of trying to conceal user input, we hide the authentication challenges presented to the users. Think of an authentication session as a question-answer exchange ("What is the password? - 'Buddy'.") Proposals so far have focused on trying to hide the answer ('Buddy'). We postulate hiding (part of) the question is more usable and secure. To that effect, we break the challenge question into two parts. The first part of the challenge is conveyed through a visible, observable, channel, while the second part of the challenge is conveyed through a hidden channel. The user mentally reassembles both parts, and answers the reassembled challenge. The authentication system can verify the answer by combining its own knowledge of the correct response and of what was sent on both channels.

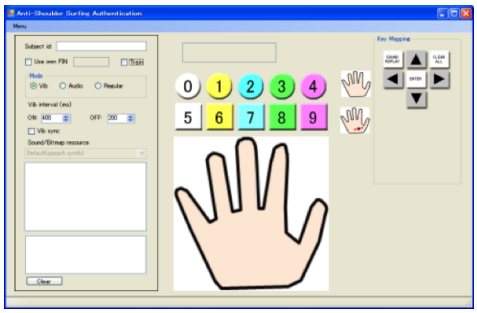
Conceptually, instead of asking "What is the password?", we ask aloud "Does your password contains a 'd'?" (visible challenge) while whispering to the user's ear "Tell a lie," or "Tell the truth" (hidden challenge). The user answers, e.g., "Yes." For an outsider unable to observe the whisper in the ear, the user's answer does not give her any clue whether there is a 'd' in the password or not. Either there is indeed a 'd' in the password, and s/he was told to tell the truth, or there is no 'd' in the password, but s/he was instructed to tell a lie. Both cases are indistinguishable to the observer.

An important research question lies in the implementation of the hidden channel. Our previous work

[1]Utsunomiya University, IS Dept., Utsunomiya, Japan and Carnegie Mellon University, CyLab Japan, Kobe, Japan
[2]Carnegie Mellon University, Information Networking Institute and CyLab, Pittsburgh, PA, USA
[3]Carnegie Mellon University, School of Computer Science, Pittsburgh, PA, USA



(a) System overview



(b) Visual channel

(c) Tactile channel

Figure 1. Tactile-based prototype.

[3] showed that using a trackball is successful in protecting against observation attacks, as the trackball movement is hidden by the hand of the user and very hard to observe by an attacker. However, 1) authentication times with such a device are quite long, in the order of tens of seconds to a minute, and 2) a bulky trackball is acceptable for fixed, large-size authentication terminals, but is unsuitable for portable devices such as mobile phones.

We set out to investigate alternate methods for delivering the hidden signal, and report on a pilot study we conducted of two novel multisensory authentication mechanisms. The first mechanism implements the hidden channel through a tactile device. The second mechanism uses an audio channel to convey hidden signals.

2. Proposed Designs

We present the tactile-based prototype, before discussing differences with the audio-based mechanism. Fig. 1 shows an overview of the different components of the prototype. Fig. 1(a) highlights the separation between the visual channel (computer screen), the hidden channel (tactile interface covered by the user's left hand), and input device (cursor pad under the user's right hand). Fig. 1(b) shows the graphical user interface implementing the visual channel. The left part of the interface is a control box for the experimenter to configure the prototype. The right part shows both the visual challenge and graphical cues. The visual challenge consists of a 2×5 cell matrix corresponding to the digits from 0 to 9. Pressing any of the input cursor keys shifts the position of the digits (row, column) on the screen. The hand drawings on the screen guide the user to select the appropriate column and row to authenticate, as described next. Fig. 1(c) shows the hidden, tactile channel. Vibrating devices roughly the size of a button battery are fitted under the user's five left fingers and palm.

Both the tactile-based and audio-based prototype use a 4-digit personal identification number (PIN) as authentication token, but are easily extensible to other types of tokens, such as graphical passwords. An authentication sequence consists of four challenges, corresponding to each digit of the PIN. A challenge starts with exactly one of the finger-based devices vibrating. In addition, the device located in the palm may also vibrate.[2] Using the cursor pad, the user moves digits on the screen to align the column of his/her PIN digit with the vibrating finger. The signal or absence of signal in the palm determines if the PIN digit should be placed in the top or bottom row. The user then validates the entry, and the authentication sequence moves to the next PIN digit.

For instance, assume that the user's PIN starts with a "2," and that both the ring finger and the palm actuators vibrate. With the starting position of in Fig. 1(b), the user presses the left cursor once, aligning number 2 with his ring finger, then the down cursor once to place number 2 in the bottom

[2]Both actuators vibrate alternatively with a low frequency, to ensure that the user can accurately perceive both signals.

row, and validates. The process is repeated, with different actuators vibrating, until all four digits in the user's PIN have been entered.

While our prototype remains relatively bulky, such a set of actuators could be fitted in a portable "USB glove" that could be plugged into the terminal to be authenticated to. However, considering mobile devices, e.g., cell phones, led us to investigate more compact designs, resulting in our audio-based prototype.

The audio-based prototype works using the same principle – aligning the digits on the screen based on information conveyed through the hidden channel. Here the hidden channel is implemented by audio instructions, delivered through a headset. We compare three types of audio instructions. In the "color" type, all cells on the screen are of a different color, and the audio signal indicates the color of the "destination" cell in which the user's PIN digit should be placed. The "color-shape" instructions give the color and shape of the destination cell, such as "blue circle." Last, the "column-row" instructions indicate the location of the destination cell, such as "first column, second row."

3. Experimental Pilot Study

We conducted a pilot study with 33 participants (ages ranging from early 20s to late 50s), recruited by advertisement, to evaluate the usability and security of our prototypes. All participants knew how to use a numeric/cursor pad, and tested all conditions (within-subject design).

We used a quiet test room as shown in Fig. 2. The experimenter controlled a PC piloting the prototype. Participants wore a lapel microphone to record any sounds made during the experiment, and a headset to receive the audio signals in the audio prototype test. Although each audio signal was conveyed once for each digit, participants could use a replay button if they needed to listen to it again. Unbeknownst to the participant, a video camera, the lapel microphone and an assistant imitate malicious observers, checking 1) whether the participant says anything revealing, 2) hand movements (e.g., removing the left hand to see the actuators and/or pointing at the screen with the right hand), 3) possible leakages, e.g. of the hidden audio signal.

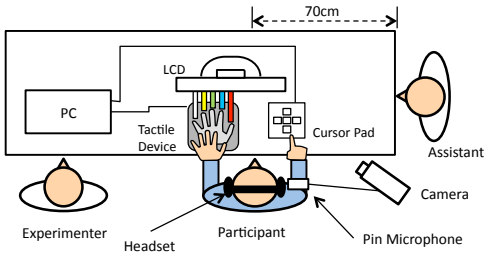


Figure 2. User study environment.

We assessed the authentication times and failure rates of the tactile- and audio-based mechanisms in two different treatments, where participants use 1) a self-chosen PIN, and 2) an assigned random PIN. We further ran a control experiment where the user is asked to enter both self-chosen and assigned PINs on a numerical pad. Thus, we tested six conditions in total. We also compared our tactile-based mechanism with the trackball mechanism of [3]. To that effect, we replaced the trackball used in the machinery described in [3] by our tactile device, while keeping the same graphical password-based visual channel. (An assigned set of graphical passwords is used.) In that experiment, the palm vibrator remains inactive, so that the five finger actuators are used as a replacement for the five possible trackball movements.

All participants practiced until comfortable with the prototype. They then started with the self-chosen PIN condition, took a 30-minute break, moved to the assigned PIN condition, and finally went through the graphical password condition. Each condition consisted of five authentication trials.

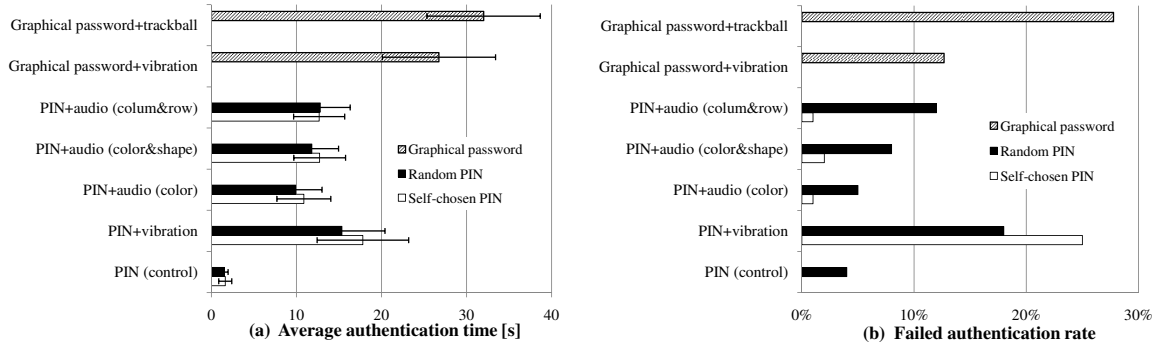


Figure 3. Experimental results. The "Graphical password + track ball" results are cited from [3].

Results. Authentication times were measured as the time elapsed between the first hidden signal being sent and the input of the last authentication token being completed. Fig. 3 (a) shows the average authentication times over all trials (error bars denote the standard deviation), and indicates that our mechanisms outperform the older trackball system; and that PINs allow for faster authentication that graphical passwords. Fig. 3 (b) gives the error rates for the different authentication systems, and evidences that the audio-based prototype performs very well, especially with self-chosen PINs. In contrast, the tactile-based prototypes have rather high error rates. Possible reasons for the apparent superiority of the audio mechanism would lie in 1) users lacking familiarity with our tactile device, and 2) cognitive difficulties when combining vibrations in a finger and the palm.

From a security standpoint, compared to the trackball mechanism, users of our novel mechanisms inadvertently revealed authentication secrets (e.g., by pointing at the screen) less frequently. However, the longer audio signals (e.g., "color-shape") could be partially heard despite the use of a headset. Audio signals thus need to be kept as brief as possible, and further secured.

4. Conclusion

We reported on a pilot study of two novel multisensory authentication mechanisms relying on different hidden channels. In our design, using vibrating devices is slightly faster than a trackball-based mechanism, at the expense of possibly higher error rates; and using audio as a hidden channel allows for faster authentication than other methods, while keeping authentication error rates lower. Combined with portability objectives, these results suggest that future efforts should consider using a combination of audio and visual channels to provide usable and secure observation-resilient authentication primitives to ubiquitous devices.

References

[1] B. Laxton, K. Wang, and S. Savage. Reconsidering physical key secrecy: Teleduplication via optical decoding. In *Proc. ACM CCS*, Arlington, VA, October 2008.

[2] L. Zhuang, F. Zhou, and J. D. Tygar. Keyboard acoustic emanations revisited. In *Proc. ACM CCS*, pages 373–382, November 2005.

[3] H. Sasamoto, N. Christin, and E. Hayashi. Undercover: Authentication usable in front of prying eyes. In *Proc. ACM CHI'08*, pages 183–192, Florence, Italy, April 2008.

[4] T. Matsumoto, H. Matsumoto, K. Yamada, and S. Hoshino. Impact of artificial gummy fingers on fingerprint systems. In *Proc. SPIE*, vol. 4677, pages 275–289, January 2002.