The Phone Lock: Audio and Haptic Shoulder-Surfing Resistant PIN Entry Methods for Mobile Devices

Andrea Bianchi
KAIST
Daejeon, Korea
andrea@kaist.ac.kr

Ian Oakley
Madeira-ITI
Funchal, Portugal
ian@uma.pt

Vassilis Kostakos
Madeira-ITI
Funchal, Portugal
vk@uma.pt

Dong Soo Kwon
KAIST
Daejeon, Korea
kwonds@kaist.ac.kr

ABSTRACT
Tangible user interfaces are portals to digital information. In the future, securing access to such material will be an important concern. This paper describes the design, implementation and evaluation of a PIN entry system based on audio or haptic cues that is suitable for integration into such physical systems. The current implementation links movements on a mobile phone touch screen with the display of non-visual cues; selection of a sequence of these cues composes a password. Studies reveal the validity of this approach in terms of task times and error rates that improve over prior art. In sum, this paper demonstrates the potential of non-visual PINs as a mechanism for securing access to a range of systems, ultimately incorporating mobile, ubiquitous or tangible interfaces.

Author Keywords
Security, Authentication, Mobile, PIN entry, Haptic, Audio.

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

General Terms
Design, Security.

INTRODUCTION

Security technologies underpin modern digital information systems. Clearly, the value of virtual data lies not only in the possibility to access it at anytime from anywhere, but also on the ability to securely restrict that access to particular individuals. This fact is illustrated by the ubiquity of personal passwords (or PINs), a security technique employed many times a day by millions of users (e.g. on bank ATMs and other public kiosks).

Although such techniques are simple and effective, they are susceptible to observation attacks in which viewing the PIN entry process results in uncovering the PIN contents. The difficulties that unintentional, non-malicious observations cause during secure data access in collaborative settings, such as group use of tabletop surfaces, has recently been highlighted [5]. Many other devices, such as mobile phones are subject to the same problems when used to perform secure transactions in public spaces. For example, a user initiating an online banking session via mobile phone in a crowded space will likely have the phone screen under the direct, but typically undesired or unintentional, visual observation of one or more other people.

This paper proposes a solution to this problem in the form of the design, implementation and evaluation of the Phone Lock, a PIN entry system based on audio or haptic cues. We argue that such non-visual PINs are less susceptible to observation attacks and that systems with this property are likely to grow in importance as authentication in tangible, ubiquitous, collaborative and public settings becomes more commonplace [4]. Furthermore, by considering and explicitly comparing the rich symbolic modality of audio with the less well understood haptic modality this paper aims to shed light on how haptic cues can be effectively designed to support authentication systems.

The remainder of this paper is structured as follows: previous mobile phone authentication techniques are introduced; the Phone Lock system and audio and tactile cues are presented; two short user studies and a security test are discussed; future avenues of research are highlighted.

RELATED WORK

Numerous researchers have explored techniques to make authentication on mobile devices less susceptible to observation attack. Typically these rely on the rich physical and multimodal cues that mobiles afford. For example, Vibrapass [3] uses tactile cues displayed on a user's mobile phone as a secondary channel to obfuscate a PIN entry that takes place on a public terminal. Mayrhofer [6] describes a system in which accelerometer data is used to pair two mobile devices - shaking them together results in similar sensor data and serves to secure the pairing. Awase-E [8] involves authenticating though a graphical password composed of a sequence of user-generated pictures. Although each of these methods uses interaction with a mobile phone to make authentication easier for users and safer against observation, none is entirely secure. The method proposed in this paper offers an improved resilience to observation and intersection (e.g. repeated observations) attack because relies on entirely non-visual cues (audio and haptic) as the means to represent PINs.

TEI'11, January 22–26, 2011, Funchal, Portugal.

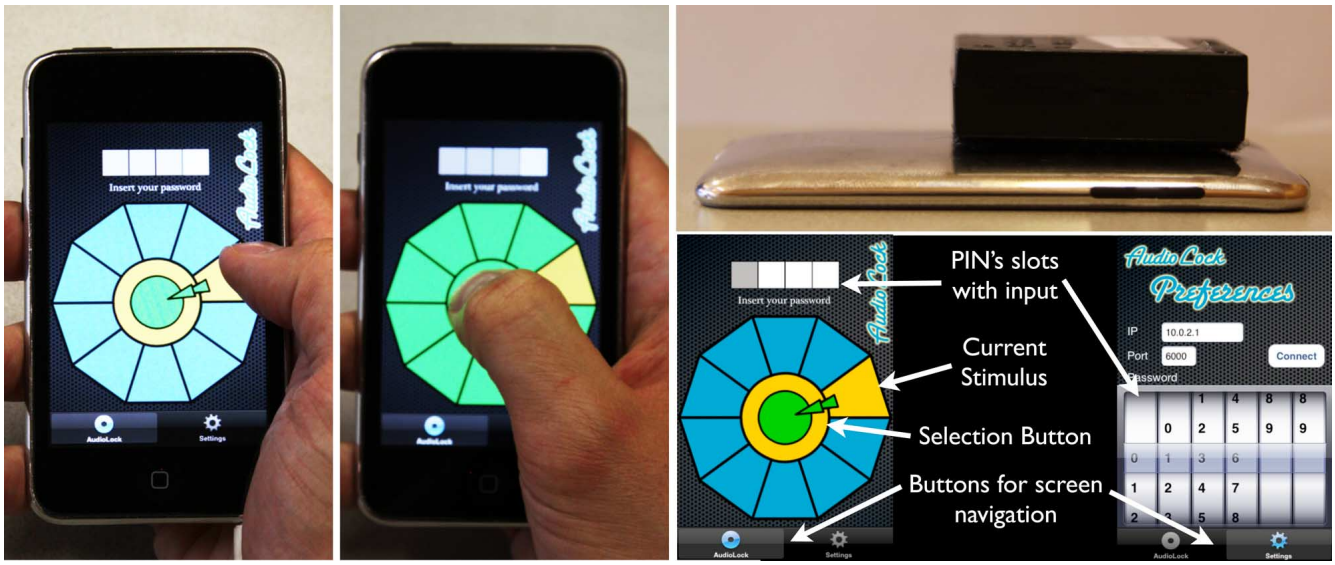

Figure 1. Virtual Wheel in use: navigation and selection (left). External motor attached to the device (top right) and screenshots of input and configurations screens with details about the GUI components (bottom right).

THE PHONE LOCK

System Description and Interaction

The Phone Lock is a PIN entry system for mobile phones that allows users to insert a PIN solely based on auditory or tactile stimuli. In the system, PIN items are drawn from a limited set of tactile or audio cues and PINs are composed of a sequence of these items. Both the size of the cue set and length of the PIN can be varied. Due to its lack of visual cues, the system is both an observation-resistant and an eye-free input technique.

The system displays a graphical wheel composed of a set of 10 equally sized, undifferentiated targets (Figure 1). Cues are mapped to particular targets and cue sets are required to possess a predetermined sequential order. This order is always maintained, although its orientation on the set of wheel targets is randomized for every digit of the PIN. In order to allow the sequences to wrap consistently around the wheel, cue sets must match or perfectly divide the number of targets in the wheel.

Cues are triggered by touching one of the on-screen targets. After experiencing a cue, users can move to another target via scrolling to an adjacent item or jumping to a more distant one. Alternatively, if they identify the item as the one they are seeking (e.g. the next PIN item), they simply select it by moving to the center of the wheel and releasing the screen. After an item has been selected in this way, it is recorded as a PIN item, the position of the cues is randomized and the system is ready to accept input relating to the following item. At any time during the interaction the user can erase the current PIN data by shaking the device.

Audio and Tactile Cues

Stimuli sets in the *Phone Lock* system must be clearly ordered. This was achieved for audio cues by the selection of generated computer speech expressing the digits 0-9 in English. The implementation of this feedback closely followed that described in Zhao et al.'s earPod [9]. The haptic stimuli were generated by variations of duration and beats to compose a scale with crescendo, a technique inspired by the work of Brown and Brewster [2]. They were composed of combinations of repetitions of beats (one, two or three) of different durations (40, 80 and 160 ms). The time between beat start points was always 160 ms, such that the final item in the sequence was a single sustained burst. A tenth tacton, featuring no vibration took the first place of this sequence.

Implementation

The *Phone Lock* is implemented for Apple's iPhone OS and intended for the physical form factor of the iPhone and iPod Touch devices. It relies on the touch screen for input, uses the devices' inbuilt capabilities for audio output (earphones are used to keep the audio authentication secure) and connects via Bluetooth to a SHAKE SK6 [7] device to generate tactile stimuli. The SHAKE is a matchbox-sized unit manually mounted on the back of the phone. Onboard processing allows it to natively support the generation of a rich set of tactile cues. Future versions of the system will provide haptic feedback directly on the mobile device via its internal pager motors.

The *Phone Lock*'s GUI has two screens, one to customize settings and the other to enter PINs. In the former, users can specify the current system PIN and length while in the latter they can insert their PINs and receive a visual confirmation of their progress and success or failure. This is shown as a series of initially white slots at the top of the device screen. Each represents a PIN item and is grayed out as items are entered. When the PIN is complete the set of slots turns green or red to indicate the correctness of the PIN as a

whole. The lower part of the screen displays the decagonal wheel on which the users make input. The inner and outer diameters of the wheel are 2cm and 4.5cm respectively.

USABILITY EVALUATION

Experiment Setup

Two studies were conducted. The first was a short pilot to ascertain raw recognition rates for the tactile cues. The second evaluated user performance during PIN entry and took approximately 1 hour to complete. Four participants were recruited for the pilot, aged between 25 and 28. Three were female and one male. 12 participants age between 19 and 43 years old completed the full experiment. Five were female and seven male, and each was compensated with 10 Euro. In both studies, participants were a mix of university students, staff and employees of companies.

Pilot

The pilot study used the SHAKE SK6 hardware (connected via Bluetooth to a PC) to render the tactile stimuli. It was composed of a series of trials in which participants were exposed to a random tacton that they had to identify using a simple GUI showing a graphical representation of each of the ten cues. Each participant completed a practice stage composed of 30 trials (three of each cue) followed by an experimental stage of twice that length. Task completion time and errors (in the form of incorrect identification of tactons) were measured. The results showed a mean task completion time of 2.25 seconds and a mean error rate of 14% indicating the task could be accomplished relatively quickly, but was challenging. An analysis of results for each tacton indicated that those featuring the mid-length 80ms element were the most challenging to recognize.

Phone Lock Experiment

This experiment had two objectives. Firstly, it compared performance between the haptic and audio interfaces. Beyond a simple modality comparison, this also acted a comparison between performance with a highly familiar stimulus set (spoken numbers) and a novel one composed of vibro-tactile cues. Secondly, it explored two configurations of the PIN length and cue set size. This allowed exploration of the optimal arrangement of these parameters for each modality and facilitated comparison with previous research that has taken this approach [e.g. 1].

The study used a partially balanced repeated measures design with two binary independent variables: modality (audio and haptic) and two variations of PIN length and cue size (*4 PIN-10 cue* and *6 PIN-5 cue*). Modality was balanced among participants, while PIN length was balanced among each modality sub-group. The 4 PIN-10 cue configuration was chosen because 4-PIN systems are a standard in applications such as bank ATMs; 4 tokens selected from a set of 10 options leads to 10000 possible PINs. The 6 PIN-5 cue variation offers an increased level of security - a total of 15625 possible PINs and was chosen to offer a direct comparison with a previous system, the Haptic Wheel [1]. PINs were randomly generated and assigned to participants on notes showing numerical digits and icons representing the different audio and haptic cues.

Each participant completed four identical conditions composed of 10 practice followed by 10 experimental PIN entries. The measures were authentication time, error rate and reset (or self-cancelation) of a PIN entry. Participants continued with each condition until 10 successful PIN entries were made and data for authentication time was drawn only from successful trials. All interaction data was streamed from the mobile phone to a host PC. Prior to the experiment participants were given instructions on the system operation and a chance to test it out. They were also presented with tactile and audio stimuli used by the system in order to gain experience with them prior to the start of the study. In the audio modality participants used earphones to listen to the audio cues; in the haptic modality the SHAKE was attached to the back of the mobile device.

Results

Experimental data are shown in Figure 2. All variables were tested using two-way ANOVAs. The authentication speed showed a significant main effect of modality ($F(1,44) = 65.1$ $p<0.01$), while the effect of PIN length did not attain

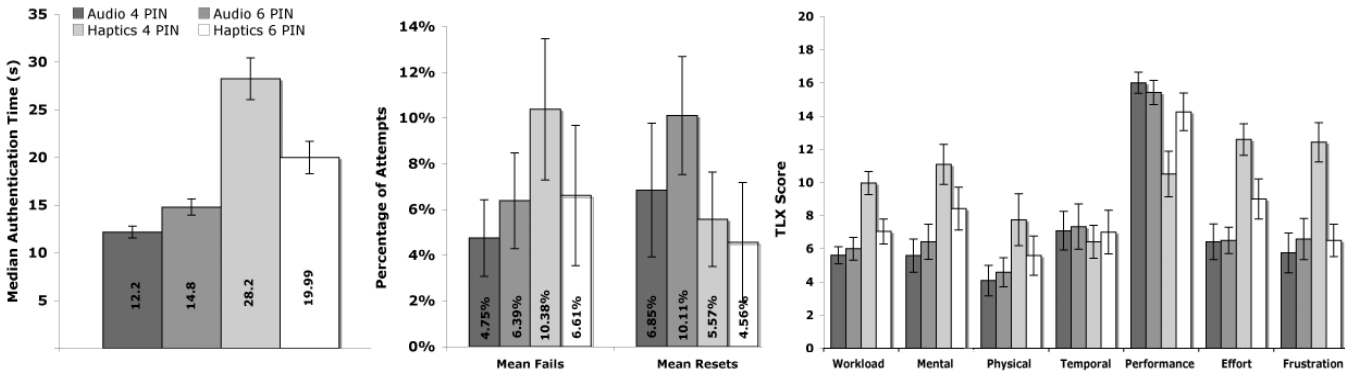

Figure 2. Mean authentication time (left); percentage of failed/reset authentications (center); NASA TLX (right). In each chart, columns from left to right show: Audio 4 PIN-10 cue; Audio 6 PIN-5 cue; Haptic 4 PIN-10 cue; Haptic 6 PIN-5 cue.

significance ($F(1,44) = 3.07$, $p=0.086$). The interaction between these two variables was significant ($F(1, 44) = 12.01$, $p<0.01$) indicating that users performed significantly faster with the smaller set of haptic cues and the larger set of audio cues. Authentication errors showed no significant differences in performance ($F(1, 44) < 1.7$, $p>0.18$ in all cases), while the workload measures revealed significant differences between the two modalities ($F(1,44) = 16.36$ $p<0.001$) but not between PIN lengths ($F(1,44) = 3.6$ $p=0.064$). A significant interaction mirroring the result from the temporal data was found ($F(1,44) = 6.16$ $p=0.016$).

Discussion

Perhaps unsurprisingly, the audio interface performed significantly faster than the haptic one: it used a familiar cue set composed of highly distinct items. This fact is clearly illustrated by the optimal performance seen in the audio 4 PIN-10 cue condition. However, in the haptic condition, users showed improved performance and workload with a reduced stimulus set of five tactons and a longer PIN entry. This highlights the fact that recognition among large sets of tactile cues is challenging.

However, beyond these broad trends, the results are encouraging. No significant differences were detected in the error rate among the conditions and the overall mean of 7% is low for an experimental setting. This suggests that, overall, users found the Phone Lock PIN entry process to be relatively easy to learn and use. This is held up by a comparison with the performance on PIN entry supported by tactile feedback in the literature [1], which reports authentication times up to 23.5 seconds and errors rates of up to 18%. Furthermore, the concept appears secure against prolonged observation attack – one way ANOVAs on the time spent recognizing each cue within each set showed no significant differences ($F(9,110) <1.27$, $p<0.29$ in all cases) except in the 4 PIN-10 cue haptic condition ($F(9,110) = 2.09$, $p=0.035$). These results suggest that, this one haptic condition aside, an attacker would not be able to reverse engineer a password from observing a user's dwell times on individual targets in the Phone Lock system. Security implications are discussed further in the following section.

THREAT MODEL AND SECURITY EVALUATION

The threat model in this work supposes an attacker capable of repeatedly observing a user authenticating at short range (i.e., from over the shoulder). Furthermore, it assumes they can also capture a close-up audiovisual recording of the interaction for analysis. To validate the security of the systems to observation, a short study was run on the 6 PIN-5 cue interfaces in both modalities. This took place in a quiet lab and involved a single user entering the same six digit PIN three times using both audio and haptic cues. This was observed from a position directly behind the shoulder (60 cm from the device and affording a clear view of its screen) both in person and with a camcorder with built-in microphone. The video and audio were analyzed in an attempt to deduce the PIN. It could not be recovered from either modality suggesting that both interfaces are secure; more rigorous testing is required to confirm this finding.

CONCLUSIONS AND FUTURE WORK

This paper introduced a novel observation-resistant authentication system that uses either audio or haptic cues to obfuscate a PIN entry process. Studies showed the feasibility of the concept in terms of task completion time, error rate and user acceptance. Future work should explore other combinations of PIN length and stimuli set size. We also believe the tactile PIN entry method has considerable potential in the context of securing tangible interfaces; in such systems, users may have to insert PINs into physical elements that have no graphical display. A haptic system to achieve this seems an elegant solution worthy of further investigation. In summary, the work presented in this paper makes steps towards the production of a rapid, easy to learn, reliable and secure non-visual authentication system; further work is required to fully realize this vision.

ACKNOWLEDGMENTS

This work was partially supported by the Portuguese Foundation for Science and Technology (FCT) grant CMU-PT/SE/0028/2008 (Web Security and Privacy).

REFERENCES

1. Bianchi, A., Oakley, I., Lee, J., Kwon, D. The haptic wheel: design & evaluation of a tactile password system. In CHI'10 EA, pp. 3625-3630.
2. Brewster, S. A. and Brown, L. M. 2004. Non-visual information display using tactons. In Procs of CHI '04 Extended Abstracts, pp.787-788.
3. De Luca, A., von Zezschwitz, E., and Hußmann, H. 2009. Vibrapass: secure authentication based on shared lies. In Procs. of CHI '09. ACM, NY, pp. 913-916.
4. Kainda, R., Flechais, I., and Roscoe, A. W. Two heads are better than one: security and usability of device associations in group scenarios. In Proceedings of SOUPS '10, vol. 485, pp. 1-13.
5. Kim, D., Dunphy, P., Briggs, P., Hook, J., Nicholson, J., Nicholson, J., Olivier, P. Multi-touch authentication on tabletops. In Proceedings of CHI 2010, pp. 1093-1102.
6. Mayrhofer, R., Gellersen., H., Shake well before use. In Proc. Pervasive 2007: 5th International Conference on Pervasive Computing. Springer, 2007.
7. SHAKE SK: http://code.google.com/p/shake-drivers
8. Takada, T., Koike, H., Awase-E: Image-based authentication for mobile phones using user's favorite images. In MHCI, LNCS 2795, pp. 347–351, 2003.
9. Zhao, S., Dragicevic, P., Chignell, M., Balakrishnan, R., and Baudisch, P. Earpod: eyes-free menu selection using touch input and reactive audio feedback. In Proceedings of CHI '07, pp. 1395-1404.