Using Mobile Device Screens for Authentication

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

Ian Oakley
MITI – University of Madeira
Funchal, Portugal
ian@uma.pt

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

ABSTRACT
Authentication in public spaces, such as ATM PIN entry, is inherently susceptible to security attacks based on observation in person or via cameras. This paper addresses this problem with a system which allows users to enter a PIN on a standard mobile phone and then transmit it securely for authentication using modulated patterns of light shown on the screen and sensed by a cheap bespoke receiver unit. No pre-pairing is required as physical proximity guarantees security. The paper presents several hardware and software variations, evaluates the technical soundness of the system, and presents two user studies addressing usability and security against observation attacks.

Author Keywords
H5.2: Input devices and strategies.

ACM Classification Keywords
Authentication, mobile interface, PIN entry, light.

INTRODUCTION
Authentication to terminals in public spaces is commonplace. ATMs and credit card terminals, based on the entry of a numerical PIN, are the most prominent example of such systems. However, although clearly effective, these systems are prone to a range of security threats (e.g. observation, brute-force and tamper attacks) in part because terminals are fixed physical installations and therefore interactions are inherently observable [1].

A recent approach to address this problem takes advantage of users' personal devices, such as mobile phones, to which attackers have no direct physical access [4]. These devices act as private intermediaries for authentication interfaces that wirelessly communicate data to public terminals. Such techniques shift the problem from securing the interaction at the terminal to securing the communication between the device and terminal. Attacks on this channel are referred to as man-in-the-middle (MITM) attacks. To counteract these methods, a range of pairing schemes that rely on auxiliary out-of-band channels (OOB) have been proposed [6], such as shaking two devices simultaneously, or using infrared lights or barcodes to transmit unique IDs [8]. However these techniques typically add interaction steps to guarantee secure authentication, increasing the complexity of the process for users. Furthermore, their reliance on establishing secure paired connections reduces spontaneity, often cited as a desirable property for such systems [5].

Addressing these issues, this paper presents a novel method that allows users to authenticate to a public terminal using a mobile phone without requiring explicit pairing. The proposed system is a novel combination of two existing concepts: Stajano and Anderson's resurrecting duckling model [10] which suggests that a channel based on physical contact is resistant to MITM attacks, and Balfanz et al.'s notion that light can provide a secure OOB channel [2]. Our system works by displaying messages through modulated patterns of light on a mobile phone screen and by relying on close physical contact with a sensor terminal to ensure this channel is private. A key advantage of this approach is that it is based on standard component of a mobile device (the screen) making it economical and easy to deploy.

RELATED WORK
Researchers have proposed a wide range of OOB channels; readers seeking a comprehensive review are referred to Kobsa et al.'s [6] recent survey and study. Optical OOB channels include 2D barcodes shown on phone screens [8], communications over IR [2], via lasers [7] and through visible spectrum diodes and phone screens captured by cameras [9]. However optical communication is a challenging medium resulting in both technical and usability issues. For example, a recent study comparing 11 pairing techniques reported users perceived 2D barcodes as among the most difficult systems to use [6]. Issues of screen resolution and orientation whilst displaying markers and the inconvenience of setting up phone cameras to capture them are likely contributors to this rating. On the other hand, widely acknowledged disadvantages of lasers and IR communication stem from their reliance on non-standard hardware. However, the use of modulated illumination of mobile device screens to transmit information is a promising technique which avoids many of these problems. Although it has been explored in relatively slow camera based device to device communication [e.g. 9, or the Bloomberg B-Unit], this paper extends this concept for use as the sole and primary output channel in an rapid, spontaneous, public and secure one-way connection.

SYSTEM DESCRIPTION
The motivating scenario for this work is providing a secure and usable channel for authentication at public terminals. One of the key attacks in such scenarios is

OZCHI '11, Nov 28 – Dec 2, 2011, Canberra, Australia

Figure 1. Three receiver unit prototypes (left). The receiving unit and phone interface during user studies (center). The interaction diagram for our system (right).

observation, either in person or via video recording equipment. Our system eliminates the potential for this attack by shifting the act of PIN entry from the terminal to the user's mobile device as described below.

Users enter a PIN on their mobile device. This is locally retained for a period of time referred to as the PIN-entry-window and then deleted. In the current system, the PIN-entry-window is set to 10 seconds. Abrupt device movement, a accelerometer derived "shaking" gesture (e.g. stealing the phone from the users' hands), within the PIN-entry-window also results in immediate deletion of the PIN.

Within the PIN-entry-window, users can identify themselves at a public terminal (for example using an ATM card) then transmit the authentication PIN from the mobile device to the terminal via a bespoke receiver unit. This is achieved by placing the device face down on the receiver. This action is sensed (by device accelerometers) and transmission of the PIN, via temporally varying patterns of light shown on the device screen, automatically commences. The light patterns are sensed using commodity electronic components integrated into the receiver unit. The transmission takes a few seconds and concludes with a denied or granted authentication. As suggested by the resurrecting duckling model [10] the physical proximity required by this communication channel implicitly assures confidentiality and protects against MITM attacks.

Hardware Design

The system is composed of two parts: a personal mobile device and a receiver unit. Two phones were tested: a recent Apple iPhone and an older Sony Ericsson M600i. Three versions of the receiver unit were constructed. The initial version of the system (Figure 1) used a single Texas Instruments TSL230 light-to-frequency converter. This small and inexpensive unit supports rapid, accurate measurement of light intensity. It was positioned facing upwards in a plastic box with a transparent top surface (but which was otherwise opaque) that was sized to comfortably accommodate a screen-down smartphone. The sensor was connected to an AVR micro-controller which communicated with a host PC via a serial link.

Two extensions to this system were developed. The first explored increasing the bandwidth by adding a second spatially separated sensor. This dual system required dividing the case into two equally sized, optically isolated portions, each of which could be simultaneously addressed by light originating in different segments of the mobile device screen. The second modification used a single optical sensor and addressed the potential susceptibility to observation attacks if light from the mobile device were to escape into the environment. This was achieved through isolating the sensor within the device housing and introducing two outward facing distracter LEDs which displayed random light patterns during system use. The mean hardware cost for each of the prototypes was $20.

Software Implementation

The mobile device software was developed on standard tools: Objective-C and iPhone OS for the iPhone and C++ and UIQ3 Symbian for the M600i. A simple proof of concept of the communication software was developed for the M600i, while the iPhone software also included a full user interface (a touchscreen keypad) and integration with the accelerometers (to activate the data communication and PIN-entry-window cancelation commands). The software for the receiver was implemented on Arduino, while the GUI hosted on the PC was developed in Java. The PC received, processed and logged raw data from the light-to-frequency converter via the serial port. Four simple data communication protocols, exploring trade-offs between efficiency, robustness and the constraints imposed by the hardware systems were developed and their performance evaluated.

Three shared a common approach: transmitting data one bit at time through successively illuminating (a 1) or blanking (a 0) the screen. Early feasibility testing of these flashing protocols led to the selection of a 40 ms presentation time for each bit. A number of overheads were also introduced to increase reliability: each data payload was preceded by a start bit (a 1) and followed by a parity and a stop bit (a 0). Two system versions used this protocol and varied the size of the payload from 4 bits (hexadecimal numbers) to 8 bits (ASCII codes). The third version was based on the hardware prototype with two light sensors and involved simultaneous transmission of two 4 bit payloads which were then combined into an 8 bit ASCII code. In each version of this system, the receiver operated in the same way. Luminance values were oversampled (at 1 KHz), quantized and processed with a median filter with a 3ms window to reduce noise.

Once a start bit was detected, the oversampled data and elapsed time were used to statistically infer the sequence of bit states.

The final communication protocol explored a different approach based on encoding numbers in light pulse duration. In this scheme, a duration of 50ms corresponded to a 0, and successive numbers were transmitted in increments of 50ms (until 500ms, which corresponded to a 9). Each transmitted digit was separated by a pause of 20ms. This protocol explored the use of transitions between high and low as key communication events. To detect these, the receiver oversampled the light to observe transitions and used a timer to determine pulse duration.

EVALUATION

In order to gauge the value and effectiveness of this system three different studies were conducted: 1) a technical evaluation of the hardware prototypes, transmission protocols and system performance in a range of lighting conditions; 2) a security user study which empirically challenges the security level of the system; and 3) a usability study to measure the user experience using a standard tool.

Technical Evaluations

Performance with the iPhone system was tested in three different environmental conditions, each representing a plausible deployment scenario: outdoors (in open space), indoors (in a well lit room) and in a dark room. Three different mobile device poses were also tested: in a normal position, firmly placed on the receiver surface, hovering approximately 1cm above the receiver surface and partially occluded (firmly placed on surface, but with half the device screen obscured by tape to simulate incorrect placement). These led to 36 test sessions: three environments by three device configurations by four communication protocols.

Each test session involved transmission of 1000 randomly selected data packets (composed of either 4 or 8 bits, depending on the communication protocol used). In each case, the phone was pre-configured with the data and transmission was instigated by inverting the phone over the scanner, an action sensed by the onboard accelerometer. Software on the receiver collected the transmitted data, which was then compared to the input file, yielding a percentage error rate. Total time to transmit the data was also recorded. These data are presented in Tables 1 and 2.

The tests were successful: with the exception of the hovering pose in the outdoor environment (with high ambient light) a high proportion of correct packets were successfully transferred in all conditions (Table 1). The lack of false positives in the hovering outdoors condition suggests the system is robust to environmental noise. The temporal data (Table 2) reveals that the pulse duration performed least efficiently, that the 4 and 8 bit flashing protocols performed roughly equivalently and that the two-channel, 2-sensor system provided an expected doubling in bandwidth. The error rate (including both omitted and erroneous packets) was generally low throughout, with a minimum in the 4-bit condition (at 1.3%). Consequently, all subsequent development and testing used this protocol. These figures demonstrates the viability of the system and are particularly promising given the relatively crude nature of the receiver unit – refinements to the physical design are likely to yield further improvements.

		Pulse Duration	4-bit	8-bit	2-sensors	Means
Indoor	Normal	1.3%	0.2%	3%	3.3%	*2%*
	Hovering	0.9%	0.6%	4%	3.8%	*2.3%*
	Occluded	3%	0.1%	4%	3.5%	*2.7%*
Dark	Normal	1.8%	0.3%	3.8%	4.5%	*2.6%*
	Hovering	0.9%	0%	2.2%	2.3%	*1.4%*
	Occluded	6.1%	1.2%	3.8%	6.1%	*4.3%*
Outdoor	Normal	1.1%	4.4%	9.4%	7.8%	*5.7%*
	Hovering	N/A 100%	N/A 100%	N/A 100%	N/A 100%	*N/A 100%*
	Occluded	11.7%	3.3%	6.1%	10.2%	*7.8%*
Means		*3.4%*	*1.3%*	*4.5%*	*5.2%*	*3.6%*

Table 1: Percentage errors in communication tests during transmission of 1000 data packets (size depends on protocols).

	Pulse Duration	4-bit	8-bit	2-sensors
Mean time to transmit 1000 packets (seconds)	305 (σ 0)	287 (σ 0.8)	557 (σ 0.5)	289 (σ 2.8)
Mean data rate (bits/ sec)	10.89	13.94	14.36	27.68

Table 2: Mean test times and derived data rates for communication protocols (packet size depends on protocols).

Performance with the Sony Ericsson M600i system was tested simply. 1000 random numerical characters were transmitted using the four bit flashing protocol in a well lit indoor environment using a slower system based on 160ms presentations of each bit - the M600i was incapable of accurate performance at 40ms. No errors were recorded suggesting that device dependent screen brightness is not an issue and that, subject to software adaption, this concept can be applied to a range of mobile devices.

Security Evaluations

To test the effectiveness of our prototype against observation attacks, a simulated attack was conducted in the lab. The threat model assumed an observer present in person and equipped with appropriate recording equipment. Tamper attacks, involving physical modifications to terminals, were not considered; they are beyond the scope of this work.

In total 10 volunteers were recruited (3 female, 7 male, mean age of 27 (SD 4.5)). They were university students and researchers, 70% of whom stated they were advanced computer users. After an introductory video and explanation of our system, participants were randomly grouped into 5 pairs. Initially one in each pair played the role of system-user while the other adopted the role of attacker. The system-user was provided with a randomly generated PIN, entered it into the system (mean time 2.4s, SD 0.4s) and authenticated three times. The attacker used detailed visual observation (both in person and via camera) of the receiver to attempt to determine the contents of the transmitted data. The two participants then switched roles and three further authentications and

observations were performed. Attackers were also offered a prize in case they would be able to crack the counterpart's PIN.

The observation attack centered on the receiver and data transmission process rather than on the standard numerical keypad on the mobile device. Attackers were supported in several ways. Firstly, they were briefed on the communication protocol structure (4-bit flashing) and provided with note taking material that reinforced this. Secondly, they were allowed to observe the data transmission in person freely and a video recording of the process was provided. This was captured from a digital video camera (60 FPS interlaced) pointed directly at the receiver unit from 10 cm. After the study, attackers examined the video frame by frame for a minimum of 10 minutes. Finally, the system used a reduced transmission rate of 160 ms per bit to ensure the camera was technically capable of recording the transmission (leading to quadrupling of total transmission time to 4.5s). In total, the study lasted 20-30 minutes and closed with an interview inquiring as to participants' perceptions of the system.

The results were encouraging. None of the attackers successfully retrieved an entire PIN. One correctly identified two digits, most likely due to misaligned placement of the phone on the receiver unit, leaving the screen open to camera observation. Other attackers were unable to retrieve meaningful information and typically reported that the attack task was "impossible" in interview.

User Evaluations

The final evaluation collected usability data for our system using the System Usability Scale (SUS) [3], a tool recently employed to contrast among a wide range of pairing systems [6]. This enables a comparison of user perceptions of our prototype with those of systems in the literature.

30 volunteers were recruited (14 female, 16 male with a mean age of 27 (SD 5.4)). They were a mix of university students, staff and company employees. Each viewed an introductory video of the system, which explained the user scenario, functionality and features. The participants were then free to use the system for as long as they wished, with the requirement that they successfully authenticate, using a supplied sample PIN, at least 3 times. They then completed an SUS questionnaire and were given an opportunity to make subjective comments. The experiment took between 5 and 10 minutes for each user.

The mean SUS score of 0.81 (SD 0.11) suggests our work is highly usable and contrasts well with the ratings reported for the 11 authentication techniques studied by Kobsa et al. [6]. Furthermore, ratings did not vary significantly with genders (female 0.8, SD 0.13; male 0.81, SD 0.09) or reported computer literacy level (advanced 0.8, SD 0.08; medium 0.82, SD 0.1; low 0.8, SD 0.1), suggesting that the system is broadly appealing. Furthermore, informal comments were generally positive, including statements such "is a good idea" and "seems safe and easy to use".

CONCLUSIONS AND FUTURE WORK

This paper presented a novel authentication method which uses modulated pattern of light displayed on a standard mobile phone screen to transmit PINs to public terminals via physical contact. The system was introduced and evaluations of its technical performance, security against observation attacks and usability were presented. These indicate that our work is an immediately feasible and affordable way to use standard mobile devices to counter observation attacks without decreasing usability.

The opportunities for future work are broad. Improvements to the receiver hardware are clearly possible, both in terms of physical design (e.g. better obscuring ambient light) and in terms of the basic sensing (e.g. increasing bandwidth via color sensing). Other applications of the communication, such as securely transmitting encrypted data (e.g. salting and hashing PIN), or acting as an OOB channel to establish shared keys need also be considered. In fact, early tests of the 4-bit protocol using MD5 128-bit encryption show unchanged error rates but lengthier authentication times (of 5.3 seconds). Further work will be devoted to improving system bandwidth so as to incorporate encryption.

REFERENCES

1. Anderson, R., Why cryptosystems fail. In Proc. ACM CCS'93, pp. 215–227, 1993
2. Balfanz, D., Smetters, D. K. , Stewart, P., Wong., H. C., Talking to strangers. In Proc. NDSS 2002, pp.23-35.
3. Brooke, J., SUS: a "quick and dirty" usability scale. In Usability Evaluation in Ind. Taylor & Francis, 1996.
4. Kainda, R., Flechais, I., and Roscoe, A. W. Usability and security of out-of-band channels in secure device pairing protocols. In Proc. of SOUPS '09, pp. 1-12.
5. Kindberg, T., Zhang, K., Secure spontaneous devices association. In Proc. UbiComp 2003, pp 124–131.
6. Kobsa, A., Sonawalla, R., Tsudik, G., Uzun, E., Wang, Y., Serial hook-ups: a comparative usability study of secure device pairing methods. In Proc. of SOUPS '09.
7. Mayrhofer, R., Welch, M.: A human-verifiable authentication protocol using visible laser light. In: Proc. Conf. on Availability, Reliability & Security, 2007
8. Mccune, J.M., Perrig, A., Reiter, M.K., Seeing-is-believing. In Proc. of Symposium on Security and Privacy, IEEE, pp. 110–124, 2005.
9. Saxena, N., Ekberg, J. E., Kostiainen, K., Asokan, N., Secure device pairing based on a visual channel, IEEE Symposium on Security and Privacy, 2006.
10. Stajano, F., Anderson, R. J., The resurrecting duckling: Security issues for ad-hoc wireless networks. 7th Security Protocols Workshop, LNCS 1796, pp.172–194.