Analysis and Improvement of a PIN-Entry Method Resilient to Shoulder-Surfing and Recording Attacks

Taekyoung Kwon, *Member, IEEE,* and Jin Hong

***Abstract*—Devising a user authentication scheme based on personal identification numbers (PINs) that is both secure and practically usable is a challenging problem. The greatest difficulty lies with the susceptibility of the PIN entry process to direct observational attacks such as human shoulder-surfing and camera-based recording. This work starts with an examination of a previous attempt at solving the PIN entry problem, which was based on an elegant adaptive black and white coloring of the ten-digit keypad in the standard layout. Even though the method required uncomfortably many user inputs, it had the merit of being easy to understand and use. Our analysis that takes both experimental and theoretical approaches reveals multiple serious shortcomings of the previous method, including round redundancy, unbalanced key presses, highly frequent system errors, and insufficient resilience to recording attacks. The lessons learned through our analysis are then used to improve the black and white PIN entry scheme. The new scheme, which we name TictocPIN, has the remarkable property of resisting camera-based recording attacks over an unlimited number of authentication sessions without leaking any of the PIN digits.**

***Index Terms*—PIN, authentication, shoulder-surfing.**

I. Introduction

The most widespread user authentication method in use today is obviously the password-based authentication, where the user enters a pre-arranged textual, graphical, and/or numerical password directly through the user interface of the authentication system. However, the password submission process is prone to direct observational attacks, such as shoulder-surfing, and this is a source of security concerns. The entry of a password can easily be observed by nearby adversaries in crowded places, aided by vision enhancing and/or recording devices, and the information that should be kept secret is leaked in a relatively non-technical manner [13]. Even partial information leakage can be greatly harmful, since users tend to use similar or even identical passwords on multiple systems, some of which may be more important than others.

The personal identification number (PIN), typically consisting of four decimal digits, is especially susceptible to observational attacks, due to its short length and the simplicity of the ten-digit keypad. The whole secret PIN could be leaked through even a single authentication session. Since PINs are so popularly used in a variety of common devices, such as smartphones, automated teller machines (ATM), and point-of-sale (PoS) terminals, there is a great need for a secure PIN entry scheme that does not significantly sacrifice usability. Various security enforcement methods have been proposed to deal with this situation, but achieving both security and usability still remains a challenging goal [24].

Of the many previous attempts, this work focuses on a remarkably simple PIN entry method proposed by Roth et al. [18]. We will refer to the scheme as the *BW method* in this paper. The basic BW method presents the decimal digit keypad to the user, in the standard layout, with random half of the keys colored in black and the other half colored in white, and the user must indicate the color of his PIN by pressing a separate black or white button. A 4-round procedure identifies each PIN digit, so that the 4-digit PIN entry requires 16 rounds to complete. Each single round operation is quite simple and intuitive to the user, but the large number of rounds causes practical usability issues.

There are four versions of the BW method. Two of these are meant to resist shoulder-surfing attacks done by human adversaries that are limited in their observational capabilities. The other two versions attempt to be resilient to even the stronger camera-based recording attacks by having the amount of information transferred from the user to the system, and thus exposed to the adversary, insufficient for unique determination of the PIN, even at the cost of making naive guessing attacks slightly easier.

Although the justifications presented by [18] has brought about the wide acceptance [5], [17], [24] of the view that the BW scheme achieves its security objectives, our study uncovered issues that seriously contest its security and reliability. The first author of this paper recently showed [11] that the basic version of the BW method was actually vulnerable to a shoulder-surfing attack that employed sophisticated strategies and training. In this paper, we study the BW method further and obtain the following results, both experimentally and theoretically, concerning the scheme.

- The shoulder-surfing resilient versions hold severe round

Manuscript received June 01, 2014; revised September 04, 2014; accepted November 16, 2014. Date of publication ????? ??, ????; date of current version November 19, 2014. This work was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) with fundings from the Ministry of Education, Science and Technology (NRF-2012-R1A1B3000965) and the Ministry of Science, ICT & Future Planning (NRF-2012R1A1B4003379). This work was also supported by the ICT R&D program of MSIP/IITP [10039180, Intuitive, convenient and secure HCI-based usable security technologies for mobile authentication and security enhancement in mobile computing environments]. J. Hong is the contact author. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Nitesh Saxena.

T. Kwon is with the Graduate School of Information, Yonsei University, Seoul, 120-749, Korea. E-mail: taekyoung@yonsei.ac.kr.

J. Hong is with the Department of Mathematical Sciences and ISaC, Seoul National University, Seoul, 151-747, Korea. E-mail: jinhong@snu.ac.kr.

redundancy, and this can be exploited by adversaries.

- The black and white key presses during PIN entries are unbalanced, and this can be exploited by the adversaries.
- The frequency of system errors reported by one version (the delayed oracle choices version) is unacceptably high.
- The two recording resilient versions provide very little protection against recording attacks.

The insight obtained through the above findings has lead us to strengthen and improve the BW scheme into a new viable scheme, which we refer to as TictocPIN.

In Section II, we summarize the threat model and review the BW method. The BW method is then thoroughly analyzed in Section III. In Section IV, we introduce the improved TictocPIN scheme and evaluate its security and usability. The related works are briefly discussed in Section V and the paper is concluded in Section VI.

II. Preliminaries

A. Threat Model

The PIN-based authentication process may be straightforwardly abstracted as communication between two entities, human user and computing system, through a user interface. Although a PIN is usually linked to other settings, such as ID and/or token, this simple model is sufficient, considering the fact that the associated information and/or object can quite easily be stolen and/or copied in the real world. The user first makes a one-time registration of a PIN to the system through a secure channel. When the user later needs to be authenticated to the system, the system presents challenges to the user through the user interface, without referencing the stored PIN. For example, the BW method system presents a series of challenges chosen from a pool of $\binom{10}{5}$ possible patterns, and the regular PIN entry system may be interpreted as presenting an empty challenge. The user answers the challenges appropriately, based on his knowledge of the PIN. The system compares the information conveyed by the user with the stored PIN and either authorizes or denies the user of further access to the system.

The threat model focuses on a passive adversary who tries to observe a user-system interaction at a user interface in order to obtain the user's PIN. There are two types of passive adversaries. The *shoulder-surfing attacker* is a weaker adversary whose capabilities are confined to those of a human. She does not have any automatic recording device and relies only on manual tools, such as paper and pencil [18]. On the other hand, the camera-based *recording attacker* is a stronger adversary equipped with automatic recording devices, such as a concealed camera, to capture the complete interactions. The BW method is known to be resilient against shoulder surfers, and its probabilistic variants are meant to provide security up to a few camera-based recordings [18], [24].

B. Review of the BW Scheme

The BW PIN entry method [18] can be used with any finite set of PIN characters and with PINs of arbitrary lengths, but let us restrict it to the case of decimal digit PINs of length 4 in our description. There are two versions of the BW scheme and both versions can be modified in the same manner to produce two more variants.

Algorithm 1 Immediate Oracle Choices (System Procedure)

```
$Q = \{0, 1, \ldots, 9\}$; $\tilde{Q} = \emptyset$
for $i = 1, \cdots, 4$ do
    $(L, R) \leftarrow \gamma \circ \pi(Q)$; $(O, P) \leftarrow \gamma \circ \pi(\tilde{Q})$
    display $B = L \cup P$ and $W = R \cup O$ in black and white
    [User: submit PIN color by pressing black/white button]
    receive user input: choice $\in$ {black, white}
    if choice $=$ black then
        $Q \leftarrow L$; $\tilde{Q} \leftarrow \tilde{Q} \cup R$
    else
        $Q \leftarrow R$; $\tilde{Q} \leftarrow \tilde{Q} \cup L$
    end if
end for
return $Q$
```

The most basic version will be referred to in this work as the IOC (immediate oracle choices) BW method. Noting $\lceil \log_2 10 \rceil = 4$, the IOC BW scheme executes a certain 4-round procedure per PIN digit, so that the delivery of the full 4-digit PIN requires 16 rounds to complete. In each round, the numeric keypad in the standard layout is somewhat randomly colored in black and white, and the user presses a separate color button to indicates which of the two colors her key digit belongs to. The system combines the information obtained through the four color choices to single out the PIN digit the user intended to submit. Figure 1 illustrates this concept.

The speed of user reaction at each round of the IOC BW scheme is unpredictable, and the colored patterns of the keypad could be left exposed to the adversary for too long. The DOC (delayed oracle choices) version of the BW method deals with this problem by first displaying the colored numeric keypads for the four rounds sequentially for preset time periods and asking for the four color inputs only later.

Although both the IOC and DOC BW methods provide some level of protection against shoulder-surfing attacks, an adversary with a camera recording of a successful PIN entry session can easily identify the PIN. The RR (recording resilient) variants of the IOC and DOC BW methods attempt to solve this problem. The two RR variants are identical to the IOC and DOC BW methods except that a smaller number of rounds are executed for each digit. The amount of information made available to the system and the adversary is reduced, and both are forced to work with a pool of possible PIN candidates rather than a uniquely identified PIN. Details of the BW schemes are summarized below, following the original description [18] closely.

IOC BW Scheme. Algorithm 1 presents a formal description of the IOC BW scheme. The operator $\gamma \circ \pi$ should be understood as dividing the input set into two parts of similar sizes, and its exact definition can be inferred from the algorithm description given below. The symbols Q and $\tilde{Q}$ denote the current set of possible and eliminated key digits maintained by the system, and their sizes are written as $q = |Q|$ and $\tilde{q} = |\tilde{Q}|$, so that $q + \tilde{q} = 10$. Initially, we have $Q = \{0, 1, \ldots, 9\}$ and $\tilde{Q} = \emptyset$. At each round, the system divides Q randomly into

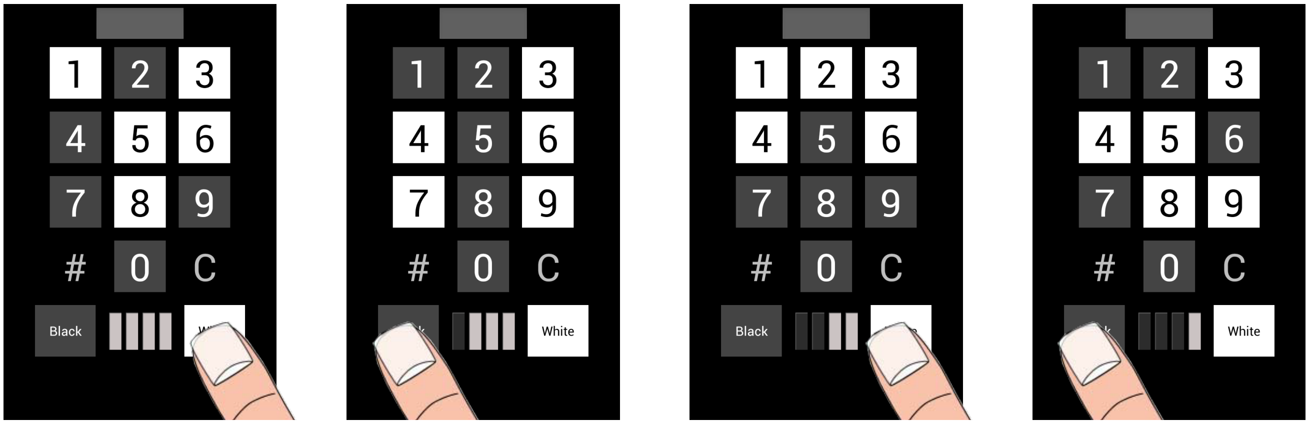


Fig. 1. The IOC BW PIN entry scheme. The digit 1 is being submitted by the user in this example of a 4-round procedure.

two parts, consisting of a $\lceil \frac{q}{2} \rceil$-sized L and a $\lfloor \frac{q}{2} \rfloor$-sized R. The system also divides $\tilde{Q}$ into O of size $\lceil \frac{\tilde{q}}{2} \rceil$ and P of size $\lfloor \frac{\tilde{q}}{2} \rfloor$. The four parts are then recombined as $B = L \cup P$ and $W = R \cup O$. The digits from the two 5-digit sets B and W are displayed to the user in colors black and white, respectively. Then, depending on the user input, the system sets one of L or R to be the set of possible digits Q for the next round, and joins the other part to the set of eliminated digits $\tilde{Q}$. After the 4-th round, only a single digit remains in Q and it is taken as the key digit submitted by the user.

Figure 1 illustrates the entry of a single PIN digit through a set of 4 rounds of the IOC BW scheme. The user input is received by the system at each round, immediately after the display of the colored numeric keypad.

DOC BW Scheme. Algorithm 2 presents a formal description of the DOC BW scheme. The system maintains a division of the digit space consisting of two 5-digit sets $P_{i,0}$ and $P_{i,1}$, where $0 \leq i \leq 4$. The initial division of $P_{0,0}$ and $P_{0,1}$ is chosen randomly. At the i-th round, $P_{i-1,0}$ is divided into two halves, L' of size $3 = \lceil \frac{10}{4} \rceil$ and L'' of size $2 = \lfloor \frac{10}{4} \rfloor$. The other half $P_{i-1,1}$ is likewise divided into two halves, R' of size 3 and R'' of size 2. The four parts are recombined as $P_{i,0} = L' \cup R''$ and $P_{i,1} = R' \cup L''$, so that each contains 5 digits. The system displays digits from $P_{i,0}$ and $P_{i,1}$ in black and white, respectively, for 500 milliseconds. This is done for $i = 1, \ldots, 4$, without any user interaction. During the display of each colored numeric keypad, the user memorizes the color of his PIN digit, but does not take any action. After the 4-th round, the user enters the four memorized colors in the correct order. Finally, the system derives Q as the intersection of the four $P_{i,j}$'s indicated by the user inputs. If Q contains more than one digit, the system reports error.

One could understand Figure 1 as a DOC procedure, if the bottom input buttons are removed and the delayed user inputs are illustrated after the 4-th box. However, the more careful reader may have noticed that the two pattern transitions from (a) to (b) and from (c) to (d) cannot occur through an execution of Algorithm 2.

RR Variants of the BW Scheme. The algorithms for the RR variants of the IOC and DOC BW methods are identical to those given by Algorithm 1 and Algorithm 2, except that the number of rounds, i.e., the range of i, is reduced, and that the $|Q| \neq 1$ error is not reported in the DOC case. The L and Q after the reduced final rounds could contain multiple digits, and the system must test all combinations of possible digits coming from each PIN position and verify if one of these candidates is the correct 4-digit PIN. In the example of Figure 1, the RR variant of the IOC BW method would stop with (c).

Algorithm 2 Delayed Oracle Choices (System Procedure)

```
$(P_{0,0}, P_{0,1}) = \gamma \circ \pi(\{0, 1, \ldots, 9\})$
for $i = 1, \cdots, 4$ do
    $(L', L'') \leftarrow \gamma \circ \pi(P_{i-1,0})$; $(R', R'') \leftarrow \gamma \circ \pi(P_{i-1,1})$
    $P_{i,0} \leftarrow L' \cup R''$; $P_{i,1} \leftarrow R' \cup L''$
    display $P_{i,0}$ and $P_{i,1}$ in black and white, respectively
    [User: note and memorize i-th color for PIN digit]
end for
for $i = 1, \cdots, 4$ do
    [User: recall and submit i-th color through black/white button]
    receive user input: $b_i \in \{\text{black}=0, \text{white}=1\}$
end for
$Q \leftarrow \bigcap_{i=1}^{4} P_{i,b_i}$
if $|Q| \neq 1$ then
    return error
end if
return $Q$
```

The article [18] did not specify the number of rounds to be used by the RR variant of the BW method, but our analysis will assume a 3-round RR variant and the reason for not treating a 2-round RR variant will be explained at the end.

III. Analysis of the BW Method

Although the BW scheme was evaluated to be resilient against practical attacks [18], [24], our investigation of the method will reveal various concerns about its security and reliability. In a previous work [11], we exploited the fact that perceiving the black and white keypad separations as visual patterns, in contrast to attending to the explicit digits, was sufficient in singling out the key digit and demonstrated that the IOC BW method could be defeated in practice. In this section, we further investigate the security of the BW method, both experimentally and theoretically, covering not just the

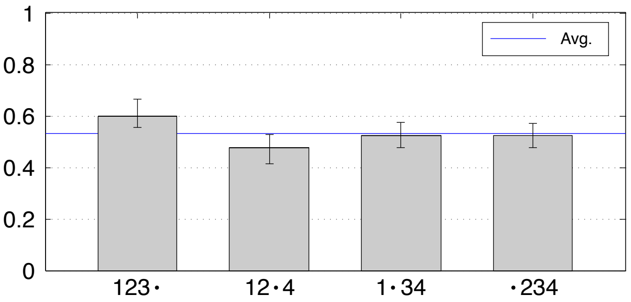


Fig. 2. Experimentally obtained round redundancies of the IOC BW method. The height of each bar gives the redundancy rate for the round marked with a dot.

IOC BW method, but also the DOC BW method and the RR variants. The list of concerns to be discussed include round redundancy, unbalanced key presses, frequent system errors from ambiguity, and recording non-resilience.

Let us introduce some terminology to be used in the proofs given below and in the appendix. We will refer to the $B \cup W$ and $P_{i,0} \cup P_{i,1}$ separations of the digit space appearing in Algorithm 1 and Algorithm 2 as *5+5 splits* and refer to the sets B, W, $P_{i,0}$, and $P_{i,1}$ as the *5-digit sets* for these splits. A natural *partition* of the digit space appears when appropriate intersections of the 5-digit sets from 2, 3, or 4 consecutive rounds of 5+5 splits are collected. Each component of the partition is a *cell* and a cell containing i-many digits is referred to as an i-*cell*. For example, the partition associated with the first two rounds of the DOC BW scheme always consists of four cells. The $P_{1,0} \cap P_{2,0}$ and $P_{1,1} \cap P_{2,1}$ are 3-cells, while $P_{1,0} \cap P_{2,1}$ and $P_{1,1} \cap P_{2,0}$ are 2-cells. Note that the first two rounds of the IOC BW method also creates a partition of these cell sizes.

A. Round Redundancy

The BW scheme specifies for $4 = \lceil \log_2 10 \rceil$ rounds to be executed for every PIN digit. However, since $\log_2 10 = 3.32$ is much closer to 3 than 4, one of the four rounds that are used to enter each PIN digit could quite often be redundant.

We first conducted a simulated experiment of entering 250 random 4-digit PINs (i.e., 1000 random digits) through the IOC BW method. As illustrated in Figure 2, one of the four rounds was redundant with mean probability 0.53, during our experiment. In particular, the fourth round was highly redundant with probability 0.596. The high probability of round redundancy implies that the identification of a PIN digit by the adversary can often be possible even when he has missed one of the four rounds. This observation also partially explains the successfulness of our previous human-based attack [11].

Our experimental figures are in full agreement with the theoretical analyses given below.

Lemma 1. *The 1-st round of the IOC BW method is redundant with probability* 0.524. *The 2-nd round is also redundant with the same probability.*

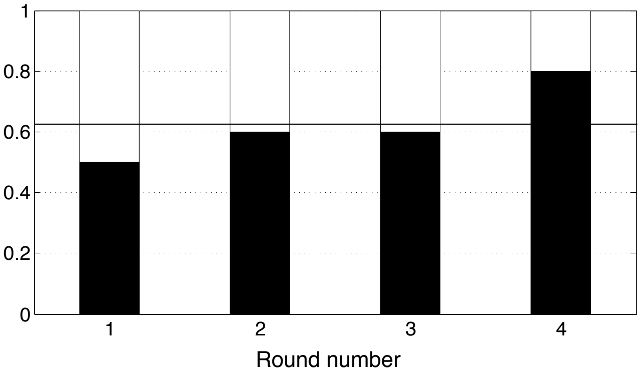


Fig. 3. Experimentally obtain ratios of black and white button presses for the IOC BW scheme. The horizontal solid line marks the mean probability of 0.625.

Lemma 2. *The 3-rd round of the IOC BW method is redundant with probability* 0.478.

Lemma 3. *The 4-th round of the IOC BW method is redundant with probability* 0.6.

The proofs of these lemmas are given in the Appendix. They essentially amount to a very careful listing and counting of all possible events. The following statement is a direct consequence of the above three lemmas.

Theorem 4. *A random round of the 4-round IOC BW PIN entry method is redundant in identifying a single key digit with probability* 0.531.

In other words, the attacker may miss a random round and still be able to recover the key digit in 0.531 of the cases.

B. Unbalanced Color Selection Frequencies

Note that the BW scheme specified the colors to be given to each of the two 5-digit sets, after each regrouping of the digit space into two new halves. We conducted an experiment to test whether the black (`B`) and white (`W`) inputs from the users would be equally likely. The simulation of entering 100,000 random 4-digit PINs to an IOC BW system resulted in the data presented by Figure 3. The two colors were pressed equally only in the 1-st round, and the `B` presses were more frequent in subsequent rounds. The bias is exceptionally large in the 4-th rounds with 80% of the inputs being `B`. In all, the `B` and `W` were pressed 1,000,449 and 599,551 times, respectively, during the 1,600,000 rounds, which translates to the ratio 0.625 of `B` presses. Furthermore, we noticed that the color sequences `BBWW`, `BWBW`, `BWWW`, `WBWW`, `WWBW`, and `WWWW` never occurred during the entry of any digit, so that only 10 of the 16 possible color combinations were ever used. This simply means that the amount of information conveyed by pressing the two color buttons are unequal, and does not directly imply weakness in the cryptographic sense. However, the property does make shoulder-surfing practically easier by allowing the observer to pay more attention to the black digits than the white digits, and this was actually done in our previous work [11].

Our corresponding theoretical analysis agrees exactly with the test results. The proof of the following statement is given in the Appendix.

Lemma 5. *The black/white button presses for the IOC BW method are expected to show the ratio* $5/3$*, assuming the user does not make mistakes. The black/white ratios expected from each of the four rounds are* $1/1$*,* $3/2$*,* $3/2$*, and* $4/1$*.*

C. Digit Identification Failures in the DOC BW Method

The DOC version of the BW method was devised to prevent the user from inadvertently exposing the black and white patterns for too long [18]. However, this version requires higher mental effort from the user, such as remembering a sequence of colors. It was estimated that the mean probability of user errors would be 0.2 for the DOC BW method, which is higher than the 0.09 expected of the IOC BW method [18].

Referring to Algorithm 2 of Section II-B, we note that a DOC BW system is suppose to "return error if $|Q| \neq 1$," i.e., if the key digit is not identified uniquely. This behavior needs to be investigated, since it would be undesirable to have the user experience frequent system errors during the authentication process.

We first conducted a simulation-based experiment of entering random PINs to a DOC BW system. A computer program simulated 100 users entering their 4-digit PINs, each for 100 sessions. Within the simulation, the system component forced the user component to re-execute the 4-round process, possibly multiple times, whenever it found any PIN digit to be ambiguous. The simulation results are summarized in Figure 4. The ratio of 4-digit PIN entry sessions that returned at least one error, averaged over all users, was 0.689. Of greater concern was the fact that three or more errors were experienced in 0.176 of the sessions. For instance, simulated-user #91 of Figure 4 experienced at least one error in 81 sessions, two or more errors in 46 sessions, and three or more errors in 18 sessions, among her 100 sessions. These repeated system errors are sure to frustrate many users, and the current form of DOC BW scheme does not seem fit for deployments that target the general public.

We also performed a theoretical analysis of the DOC BW method. The above experimental figures are in good agreement with Theorem 7 given below. The proof of Lemma 6 is given in the Appendix.

Lemma 6. *A 4-round execution of the DOC BW scheme will fail to uniquely identify the submitted key digit with probability* 0.25*.*

Theorem 7. *Consider a DOC BW PIN authentication system that is set to announce a system error and require the user to re-execute the 4-round process whenever it fails to identify a key digit uniquely. A user submitting a 4-digit PIN to this system has probability* 0.684 *of experiencing at least one error. The probabilities for the system to generate two or more and three or more errors are* 0.367 *and* 0.169*, respectively.*

Proof: It follows directly from Lemma 6 that the first claimed probability is $1 - \left(\frac{3}{4}\right)^4 = \frac{175}{256}$. Lemma 6 also allows us to state the probability for a single failure as $\binom{4}{1}\frac{1}{4}\left(\frac{3}{4}\right)^4$. Here, where the correct combination to be used is not $\binom{5}{1}$, since the last of the five attempts must be a successful one. Thus, we can state $\frac{175}{256} - \binom{4}{1}\frac{3^4}{4^5} = \frac{47}{128}$ and $\frac{47}{128} - \binom{5}{2}\frac{3^4}{4^6} = \frac{347}{2048}$ as the remaining two claimed probabilities. ■

We clarify that the error probabilities claimed by this theorem are for the system errors and are not related to the errors made by the user.

D. Inadequate Recording Resilience

The RR variant of the BW method attempts to provide security against adversaries that are equipped with camera-based recording devices [18]. The approach was to remove one round from the 4-round process required for each PIN digit entry. This creates ambiguity in the PIN digits to the observer (and to the PIN entry system), and the adversary is forced to guess the correct 4-digit PIN from a pool of possible PINs. However, the effectiveness of this approach can only be questioned after understanding Theorem 4.

We conducted a simulated experiment to measure the ambiguity left to the recording observer of a 4-digit RR IOC BW PIN entry session. As before, the computer simulated 100 users entering random 4-digit PINs for 100 sessions, and transcripts were made of one randomly chosen session per simulated-user. Then, each transcript was studied to derive the PIN candidates, in exactly the same manner as would have been tried by an attacker. The results are summarized in Figure 5. It was highly probable that the number of PIN candidates was extremely small. The rate of unique PIN identifications was just 0.129, but that of finding three or less PIN candidates was 0.469 and that of at most five was 0.816. We also experimented with the recording of multiple sessions for the same user. As expected, the PIN could be identified uniquely with high probability.

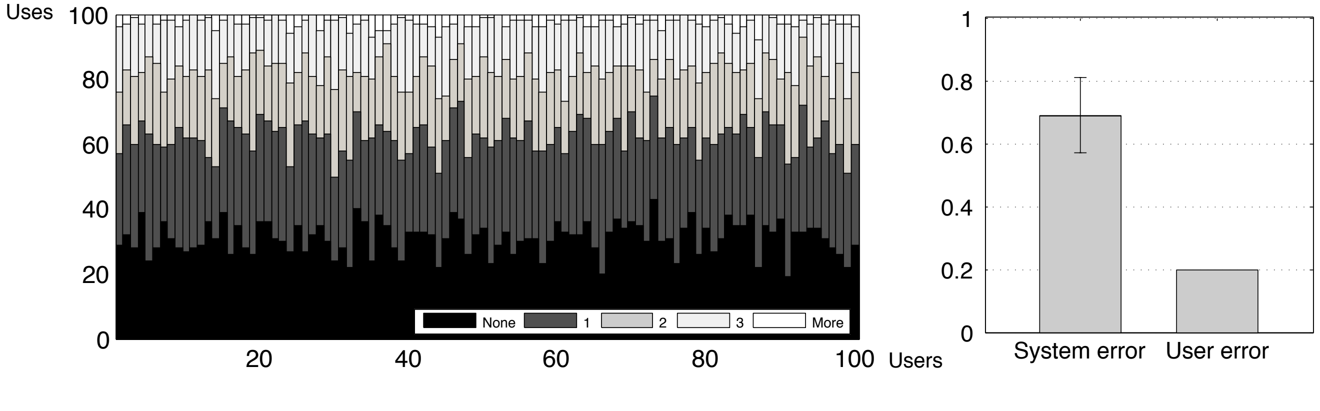
The experimentally obtained figures can be explained theoretically as well. First, note that since the 3-round IOC BW scheme always restrict each PIN digit to a set of size 1 or 2, the 4-digit PIN candidate sets can only be of sizes 1, 2, 4, 8, and 16. The following is a direct consequence of Lemma 3.

Lemma 8. *Consider the IOC BW scheme variant that executes only 3 rounds per digit. The probability for a 4-digit PIN entry session to reduce the number of 4-digit PIN candidates to a set of size* 2^i *is* $\binom{4}{i}\frac{2^i 3^{4-i}}{5^4}$*, for each* $i = 0, \ldots, 4$*.*

Proof: The probability for none of the four digits to be ambiguous is $\left(\frac{3}{5}\right)^4$. The probability for the PIN to be restricted to a set of two PINs is $\binom{4}{1}\frac{2\cdot 3^3}{5^4}$. The general situation should now be clear. ■

Thus, the 4-digit PIN candidate pool is of size at most 3 with probability $0.475 = \frac{81+216}{625}$ and at most 5 with probability $0.821 = \frac{81+216+216}{625}$, in agreement with Figure 5.

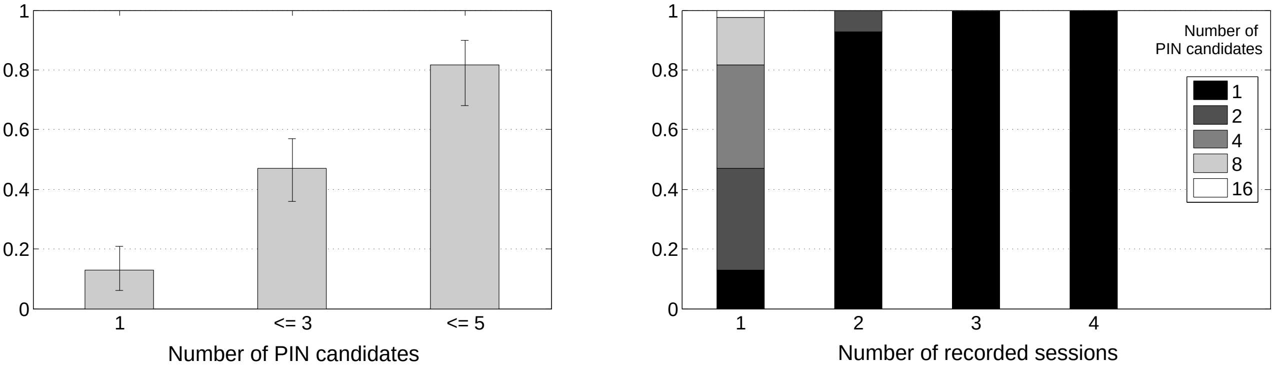
No user would feel adequately protected, knowing the existence of an adversary that could make three guesses from a pool of candidate PINs that is expected to be quite small. The following claim shows that the attacker will indeed success with a very high probability.



(a)

(b)

Fig. 4. System errors of the DOC BW scheme. (a) Error count distributions for 100 4-digit PIN entry sessions for 100 users. (b) Error rates. (The user error rate has been taken from [18].)



(a)

(b)

Fig. 5. Recording (non-)resilience of the 3-round IOC BW scheme. (a) 4-digit PIN candidates derived from a single recorded session. (b) PIN candidates derived from multiple recorded sessions.

Theorem 9. *Assume that the adversary is given a transcript of a 3-round IOC BW 4-digit PIN entry session. Then the possible PINs can be reduced to a set of expected size* 3.84*, and the adversary can guess the correct 4-digit PIN within three trials with probability* 0.797.

Proof: Lemma 8 states that the adversary's set of candidate PINs will be of size 2^i $(0 \le i \le 4)$ with probability $\binom{4}{i}\frac{2^i 3^{4-i}}{5^4}$ and we know that these sizes exhaust all possibilities. Thus, the expected candidate set size can be computed as $\sum_{i=0}^{4} 2^i \binom{4}{i}\frac{2^i 3^{4-i}}{5^4} = \frac{2401}{625}$.

When the number of possible PINs is either $1 = 2^0$ or $2 = 2^1$, the observer can certainly guess the correct PIN within three trials. When there are larger 2^i candidates, there is probability $\frac{3}{2^i}$ for the adversary to be successful in choosing the correct 4-digit PIN within three random trials. Gathering all cases together, we can claim $\sum_{i=0}^{4}\binom{4}{i}\frac{2^i 3^{4-i}}{5^4} \times \min\left\{1, \frac{3}{2^i}\right\} = \frac{498}{625}$ to be the adversary's success rate in correctly guessing the 4-digit PIN. ■

Note that the probability for the adversary to be successful in being authenticated by the PIN entry system is higher than what is stated by the above claim. The reduction in rounds brings ambiguity to the system also, and the system will occasionally grant access to even incorrect PIN entries, referred to as *shadows* in [18].

An analogous statement for the RR variant of the DOC BW scheme can also be obtained.

Theorem 10. *Assume that the adversary is given a transcript of a 3-round DOC BW 4-digit PIN entry session. Then the possible PINs can be reduced to a set of expected size* 7.23*, and the adversary can guess the correct 4-digit PIN within three trials with probability* 0.639.

Proof: Let us just provide a sketch of proof. By carefully following through the DOC BW algorithm, one can list all partitions of the digit space that could occur after the 3-rd round, together with their probabilities of appearances. This information directly implies that, after the 3-rd round, a random digit will find itself in a 1-cell, 2-cell, and 3-cell with probabilities $\frac{21}{50}$, $\frac{13}{25}$, and $\frac{3}{50}$, respectively. Then one can argue as in the proof of Theorem 9 to calculate the two claimed figures, which are $\frac{2825761}{390625}$ and $\frac{99813}{1562505}$. ■

As in the IOC case, the adversary can deceive the 3-round DOC BW PIN entry system with probability that is higher than the probability stated above.

The lower bounds 0.797 and 0.639 to the attacker's success rate in penetrating the 3-round IOC and DOC BW systems are clearly too high for any user to be comfortable with. One cannot claim the 3-round variant of either the IOC or DOC BW schemes to be providing adequate resilience against attacks that utilize recordings.

Since the original presentation of the RR variant to the BW method [18] did not specify how many rounds were to be removed from the normal BW method, let us briefly consider the possibility of using a 2-round BW scheme. With both the 2-round IOC and DOC BW methods, the system can only narrow down the 4-digit PIN to a candidate set of expected size $\left(2\frac{2}{5}+3\frac{3}{5}\right)^4 = 45.7$. In other words, a wild guess of the 4-digit PIN, without prior observation of any PIN entry session, will defeat the 2-round BW system with probability that is 45.7 times greater than the usual 4-digit PIN entry system. The break-in rate of 0.00457 per trial may seem small, but weakening the security against the most classical attacker this greatly cannot be acceptable. If it were acceptable, we would be using 2-digit or 3-digit PINs today. The use of a 2-round BW method does not seem reasonable. Furthermore, even with the 2-round variant, if two or more sessions are recorded, the PIN is likely to be identified.

The article [18] allocates much space to discuss how increasing the size of the PIN candidate set for their RR variant BW method makes recording attacks harder and raw guessing attacks easier. They claimed[1] that, for length-4 decimal digit PINs, the two opposite effects are balanced when the ambiguity is at the level of 100 PIN candidates. It is not clear if they are recommending systems to be designed to produce approximately 100 PIN candidates, but doing so does not seem advisable, since security against the naive guessing attacker would be greatly reduced.

E. *Comments on Other BW Configurations*

Recall that the original BW PIN entry method allowed for flexibility in the PIN character set to be used and in the length of the PINs. We had focused on just the 4-digit PIN configuration, because practical interest in this case greatly overwhelms those in all other cases. Furthermore, it is rather straightforward and easy to extend our results to the cases of decimal digit PINs of lengths other than 4. However, analyzing the BW method that utilizes a character set of size other than 10, which might still be of interest for certain applications, will require further work.

Let us briefly consider the BW method that is based on a character set of size 16 as an example. It is clear that, contrary to the 10-character case we had studied, none of the 4 rounds of the IOC BW method for the 16-character case will ever be redundant. It is also clear that the 16-character case will exhibit balanced black and white key presses. Hence, two of the undesirable characteristics of the 10-character case are completely absent in the 16-character case. In fact, the reader might recall that our discussion of the 10-character case began with the observation that $\log_2 10$ is closer to 3 than 4. Continuing the analysis of the 16-character case, one must expect the 4 rounds of the 16-character DOC BW method to be mostly insufficient in uniquely identifying an input character. In fact, one can argue that the probability for a single input character to be identified uniquely is only $\frac{\binom{4}{2}\binom{4}{2}}{\binom{8}{4}}\frac{\binom{2}{1}\binom{6}{3}}{\binom{8}{4}} = 0.294$, so that the entry of a length-4 PIN will fail with the huge probability $1 - 0.294^4 = 0.993$. Hence, the situation concerning system errors is even worse with the 16-character case than with the 10-character case. Finally, let us discuss the 3-round RR variants for the 16-character case. It is easy to see that the ambiguity faced by the attacker will be larger than the 10-character case, and this may initially seem to be a positive indication. However, since the system must also cope with such a larger ambiguity, one must consider arguments concerning the *shadows*, introduced in [18], before making the final judgement. This extra argument was not strictly necessary in the 10-character case to arrive at a negative conclusion, because the attacker ambiguity was small.

The 16-character example clearly shows that the characteristics of the BW method can vary greatly with ℓ, the size of the input character set. It also shows that even different approaches may be required, depending on how far $\lceil \log_2 \ell \rceil$ is from $\log_2 \ell$.

IV. TictocPIN: Colored PIN Entry, Strengthened through a Hidden Auxiliary Channel

The challenges displayed by the BW method through a colored ten-digit keypad allow for quick, intuitive, and simple user responses. However, the detailed analysis given in the previous section has taught us that the mismatch of entropies associated with a PIN digit and a set of color inputs incur side effects such as round redundancy, unbalanced key presses, and system errors. Furthermore, the RR variant of the BW method was shown to be non-resilient to camera-based recording attacks. In this section, we strengthen the BW method into a more viable PIN entry method, taking advantage of the lessons learned, and evaluate the security and usability of the reinforced scheme. We pursue minimization of the key entry count to remove the mentioned side effects and use a hidden auxiliary simplex channel to achieve recording resilience. Figure 6 illustrates an execution example of our new[2] scheme, which we refer to as *TictocPIN*.

A. *Description of TictocPIN*

The ingenuity of the BW method was in assigning colors to numeric keys and opting to receive each PIN digit through a *multi*-round challenge-response procedure. In particular, the colored challenges allowed for *intuitive* user responses. Our

[1] Their probability calculations seem inadequate. The cartesian product structure of the shadow set is disregarded, so that their arguments are meaningful only when the shadow sets are large. Furthermore, events A_k and B_k appearing in the arguments are erroneously taken to be independent, and $P[B_k] = \frac{1}{N-s+1}$ should be corrected to something closer to $\frac{s}{N}$.

[2] Our previous work [11] presented a novel human shoulder-surfing attack based on sophisticated cognitive strategies that defeated the IOC BW method and proposed a modified scheme which had better security. However, the improved method still required as many user inputs as the original method and was vulnerable to camera-based recording attacks.

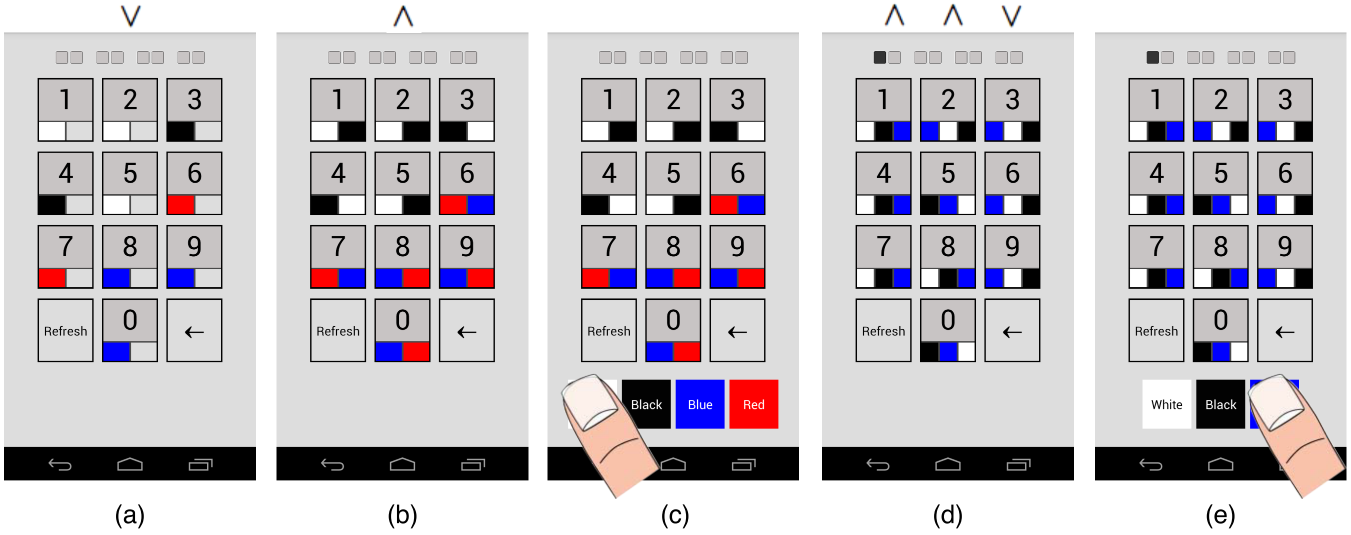


Fig. 6. TictocPIN scheme — A running example of submitting the digit 1 through a 2-round procedure. (The marking ∨ indicates a vibration and ∧ indicates its simulated sound.) (a) Phase 1-1 with short vibration; Colors appear in the left boxes. (b) Phase 1-2 with simulated vibration sound; Colors appear in the right boxes. (c) Phase 1-3; The four-color input keypad appears at the bottom. User presses white, the color under the 1-key that appeared with a vibration. (d) Phases 2-1 through 2-3 with two vibration sounds followed by a real vibration; Colors appear in the left, middle, and right boxes, incrementally. (e) Phase 2-4; The three-color input keypad appears at the bottom. User presses blue, the 1-key color associated with the vibration.

improved scheme retains this basic structure, but introduces tweaks that are explained below. In short, a smaller number of rounds is used and the user is informed through a vibrotactile[3] channel as to which of multiple displayed challenges is to be taken as valid.

- *Structured Partitions*: Let us consider the partition of the digit space $Q = \{1, \ldots, 9, 0\}$ into the four subsets $L_1 = \{1, 2, 5\}$, $L_2 = \{3, 4\}$, $R_1 = \{6, 7\}$, and $R_2 = \{8, 9, 0\}$. The partition of Q into the three subsets $Q_1 = \{1, 4, 7, 8\}$, $Q_2 = \{2, 3, 6, 9\}$, and $Q_3 = \{5, 0\}$ will also be used. These fixed structured partitions of the digit space are to be used in all sessions, unlike the BW method, which utilized randomly generated partitions.
- *Two-Round Key Entry and More Colors*: The number of rounds required to enter each PIN digit is halved from that of the BW method to two. In the first round, the four sets (L_1, L_2, R_1, R_2) are assigned distinct colors (black, white, blue, red) in some mixed order. The subsequent second round similarly assigns colors (black, white, blue) to the three sets (Q_1, Q_2, Q_3). Since each Q_i contains at most one element from each of the sets L_1, L_2, R_1, and R_2, each pair of colors collected over the two rounds corresponds to at most one digit, completely removing the possibility of system errors.
- *Multiple Assignments of Colors and the Haptic Signal*: Further tweaks are required to achieve recording resilience. Two different assignments of the four colors are displayed by the first round. A separate vibrotactile channel is used to inform the user as to which of the two assignments is to be considered as the true challenge. The second round consists of three different color assignments with one of them covertly signaled as the true challenge. The shoulder surfer with access to only the visual channel will not be able to distinguish the real challenge from the fake.
- *Audio Leakage Obfuscation — Tic and Toc*: Another tweak is used to protect against an attacker that has access to audio leakage of the PIN entry process. The PIN system either (a) simulates the noise of vibration when displaying the non-vibrated challenges, or (b) generates auditory noise that hides or scrambles the sound of vibrations throughout the whole period of multiple color assignments. When properly implemented, even the advanced attacker, equipped with a directional microphone, should not be able to distinguish the sound of a real vibration (tic) from a simulated vibration sound (toc) that accompanies a fake challenge.
- *Multi-Phased Color Assignments*: The first round contains two challenge phases. Each key of the numeric keypad is made to contain two small boxes. Let C_1 and C_2 be the colors black and white in some order, and let C_3 and C_4 be blue and red in some order. In the first challenge phase of the first round, the left boxes of L_1, L_2, R_1, R_2 are filled with colors C_1, C_2, C_3, C_4, respectively. The second challenge phase of the first round begins after a 500 msec delay. With the left boxes retaining their colors, the right boxes of L_1, L_2, R_1, R_2 are filled with colors C_2, C_1, C_4, C_3, respectively. Then, after a 500 msec delay, a keypad consisting of the four colors, in random order, is displayed to receive user input. One of the two challenge phases, randomly chosen at the time of execution, is accompanied by a short 30 msec vibrotactile signal. The display of the accumulated challenges is maintained until the user supplies a color input.

The second round basically consists of three 500 msec

[3] To the best of our knowledge, haptic channels were first leveraged to provide resilience to shoulder-surfing by the seminal Undercover [19] system.

challenge phases, with one of them, chosen at random, accompanied by a 30 msec vibrotactile signal. Each key of the numeric keypad contains three small boxes that are progressively colored from left to right through the three phases. Let C_5, C_6, and C_7 be the colors black, white, and blue, in any order. The first phase of the second round assigns colors C_5, C_6, C_7 to the left boxes of Q_1, Q_2, Q_3, respectively. The second phase further fills the center boxes of Q_1, Q_2, Q_3 with the colors C_6, C_7, C_5. The third phase adds colors C_7, C_5, C_6 to the right boxes. A keypad consisting of the three colors, in random order, appears after the three 500 msec challenge phases. Display of all three separate color assignments are maintained until the user input is received.

A refresh key can be used before color entry to restart a round, in which case, the phase to be vibrated is newly selected at random. A backspace key allows the latest key entry to be deleted.

The association of the actual colors to the C_is may either be done randomly at each round or hard-coded into the system and fixed for all rounds and sessions. The online randomization of this part may add a small amount of confusion to the human observer, but has no effect on the recording attacker. Since the added confusion also applies to the user, our later user studies were carried out with colors hard-coded to certain default values. Note that, in contrast, the randomization of the user input color pad is required, as it is designed to prevent a user from preemptively positioning her finger before all challenges have been displayed.

B. PIN Entry Example

Let us explain the execution example given by Figure 6, which illustrates the entry of the single PIN digit 1 through the TictocPIN system.

Figure 6-(a) is first displayed with a short 30 msec vibration. The left box under the 1-key is colored in white and the right box is left empty. Note that there is no color keypad at the bottom. After a 500 msec pause, which includes the 30 msec vibration, the right boxes under the digits are additionally colored as in Figure 6-(b), and a simulated sound of vibration is produced. This display is maintained for 500 msec, after which the display changes to Figure 6-(c). The color pad has appeared below the numeric keypad to serve as the user input interface. The first round of the key entry ends with the user pressing white.

The displays for the subsequent three 500ms intervals, which form the challenge presented by the second round, is compressed into Figures 6-(d). The left, middle, and right boxes under each of the ten digits are incrementally filled with colors, with each display lasting 500 msec. The first two phases are accompanied by simulated vibration sounds and the third phase arrives with a real short vibration. Finally, the input color pad appears as in Figure 6-(e), and the user presses the blue key, the color under the 1-key that appeared with the vibration, to completes the second round of the key entry.

The PIN digit 1 was successfully entered to the system through these two rounds. Further PIN digits may be entered through similar processes.

C. Security Analysis

1) Direct Assessment: It is easy to see that the information of whether each PIN digit belongs to $L = \{1, \dots, 5\}$ or $R = \{6, \dots, 9, 0\}$ can be gathered by a shoulder surfer. Such information reduces the search space for a 4-digit PIN to a set of size 625, but this is still large enough to deter brute force attacks. We claim that no additional information is leaked to even the strongest adversary that may analyze any number of PIN entry session recordings.

Consider a single PIN digit for a certain user, and suppose that the attacker already knows whether it belongs to L or R. At the next PIN entry session the attacker will observe the user submitting one of six possible color pairs for this digit. However, the color patterns laid out by the TictocPIN system for a single digit input is such that each of the six color pairs can be associated with any one of the five digits. In fact, for someone without access to the haptic channel information, all five possibilities are *equally* likely to have been meant by the user input color pair. No information beyond the L-R classification is revealed by any number of PIN entry sessions. Furthermore, the argument remains true even if the user occasionally produces failed PIN entry sessions.

2) TictocPIN as a One-time Pad: If one treats the L-R classification of each PIN digit as public information, the TictocPIN system can be understood as a one-time pad. A one-time pad that utilizes an alphabet size of six, rather than the usual two, would work as follows. For each alphabet p from the plaintext space $P = \{0, 1, \dots, 5\}$ to be sent, the sender and receiver share a key k from the key space $K = \{0, 1, \dots, 5\}$, through a separate secure channel. The ciphertext $c = p + k \pmod 6$ belonging to the ciphertext space $C = \{0, 1, \dots, 5\}$ is sent, and the receiver decrypts by computing $p = c - k \pmod 6$. We wish to interpret each key k, not as an element of K, but as the bijection $x \mapsto x + k \pmod 6$ from P to C. Note that, as long as the key is chosen at random for each plaintext alphabet to be sent, no information is revealed by the ciphertext characters, even if the plaintext is fixed.

In the TictocPIN situation, the plaintext space is either $P = L \cup \{\#\}$ or $P = R \cup \{*\}$, where $\#$ and $*$ represent two characters that will never be sent, and the ciphertext space is $C = \{\text{color pairs}\}$. Both P and C contain six elements. The key space K consists of the six possible bijections between P and C displayed through the multi-phase two-round challenges. The same set of six bijections are used in every session, and the only difference between sessions is in which of the six is signaled via the hidden channel to be used. Just as with the one-time pad encryption, as long as the hidden channel signals (one-time pad) are produced randomly and delivered securely, no information concerning each PIN digit (plaintext) is leaked through the corresponding color pair input (ciphertext).

3) General Strategies for Protecting PIN Entry: Let us briefly digress and present a more higher level discussion concerning the security of PIN entry systems. When under the assumption that the exchange of information between the user and the system is completely exposed to the adversary through the visual channel, the basic strategy for providing protection against the human shoulder surfer is to involve the requirement for information processing. The BW method hides the digit being submitted in a group of five digits that are colored the same, and the true intention of the user is revealed only through the combination of multiple submissions. This multi-round approach places heavy cognitive burden on the observer, while the user is only required to respond to easy challenges. Our TictocPIN method inherits this strategy from the BW method and hides each user response in a group of digits.

However, this strategy is completely useless in view of the camera-based recording attacker, who may perform the information processing at his leisure. To overcome this problem one must focus on the initial assumption that all information exchanged between the user and the system is made accessible to the adversary through the visual channel. Covering the PIN entry interface with one's belongings is an effective way of blocking the visual channel itself. For example, [10] takes this approach of obfuscating the visual channel.

A more sophisticated approach is to make the information submitted during the PIN entry session insufficient for the recovery of the complete secret, accepting the side-effect that even the system becomes unable to verify the complete secret. The RR variant of the BW scheme is clearly an example of this approach, and the ColorPIN [5], which assumes a pre-shared secret that is larger than the PIN, also belongs to this category.

The approach taken by TictocPIN, as with Undercover [19], is to employ a separate secret channel. Of course, a fast protected duplex channel, which is impractical to assume, would solve all problems, but we are utilizing only a very limited simplex channel. This approach is fundamentally different from the approaches mentioned before in that one is utilizing temporary secrets that are generated in real time, as opposed to a pre-shared fixed secret, which would gradually be exposed over multiple sessions. One big difference between TictocPIN and the previous schemes [23], [16], and [25], that relied on hidden auxiliary channels, is the use of a haptic channel as opposed to the visual or phonic channels. We believe that the haptic channel, despite its deficiency of being low bandwidth, could be easier to protect from the adversary than the visual or phonic channels due to its inherent privateness.

4) Switching the Information Leakage Channel: In view of our previous interpretation of TictocPIN as a one-time pad, it can now be said that TictocPIN is a clever realization involving a separate protected channel of the following simple observation: If $\log_2 6$ bits of secret can be shared (received), one should be able to transmit (submit) a PIN digit in *perfect* secrecy, assuming the digit holds just $\log_2 5$ bits of secrecy. Thus, the security of TictocPIN no longer relies on the obfuscation of the visual channel. On the other hand, it is clear that the security of TictocPIN is directly affected by how well the haptic channel is protected.

5) Impeding Audio Leakage: The adversary has no means of receiving the actual vibrotactile signals, but one must consider the possibility of information leaking through the audio channel [4]. If the adversary employs a directional microphone and the victim is situated in a very quiet environment, it may be possible for him to gain access to both the visual and the haptic channels. The role of the simulated vibration sound is to mitigate this danger in our method.

The basic concept is to reduce the possibility of the vibrotactile channel leaking information through the audio channel by blanketing it in additional noise. However, note that, because each vibrotactile signal lasts only 30 msec, perceiving the real vibrations (tics) through the audio channel is already quite difficult in public places. With the addition of properly designed fake simulated vibration sounds (tocs), adversaries will not be able to distinguish between the real and fake challenges, even when using directional microphones.

6) Preventing User Misbehavior: Note that we have specified for the user input color keypad to be presented at the end of each round and for the colors to be positioned in random order. Since the user cannot predict the location of the color to be pressed, she cannot correctly position her finger until all challenges have been displayed. This deters the user from inadvertently disclosing her moment of decision, which would be correlated to the moment of real vibration.

The design choices of delayed and randomized input keypad were made to counter the timing attacks of [17], and we believe that these measures provide a high level of protection. However, we acknowledge that these measures are not perfect in preventing all detrimental user behaviors. For example, even though the separate placement of the input keypad works as a deterrent, we cannot prevent a user from absentmindedly pointing at his secret PIN digit with his finger. In fact, the eye movement of the user can be a valid subject of study for many PIN entry systems. Of higher concern is that there is a small possibility that the user response could be slightly slower when a later phase of a round is given the vibration, thus leaking the haptic channel information. Another point is that if a user develops the habit of double checking his secret PIN digit and the correct input color after the input keypad has appeared, then it might even be possible that the distance between the secret digit and the input color button could be reflected in his response time. These are interesting questions that call for a separate extensive user study.

7) Security Summary: TictocPIN obfuscates the visual information leakage channel by leveraging the inherent privateness of vibrotactile signals and further protects the haptic channel by obfuscating the audio informational leakage channel with scrambling noise, so that even the advanced adversaries would face serious difficulties in perceiving the haptic signals. We acknowledge that it should be possible to visually detect the 30 msec vibrotactile signals generated by a smartphone, using sophisticated equipments, such as a high speed video camera. However, at this point, we do not know how much effort or cost would be involved with such an attempt aimed at a hand-held device. This could be an interesting subject of future study, and one could also consider the use of a properly protected audio channel as an alternative to the haptic channel.

D. User Study — Usability Evaluation

In our user study, we evaluated the usability of TictocPIN in comparison with the standard PIN entry method. We implemented TictocPIN and the standard PIN entry systems in software, on the Google Android platform running on Galaxy Nexus smartphones (4.65″, 1280 × 720 pixels, 316 ppi). The whole interaction was logged for later analysis with regard to PIN entry time and authentication results.

1) Design: We performed a 2 × 2 *within-subject* design study to evaluate the usability of TictocPIN. The independent variables were PIN type (system-chosen, user-chosen) and PIN entry system (standard, TictocPIN). The standard PIN entry system was taken as the control condition. The system-chosen PINs were selected at random, while avoiding simple configurations, and the user-chosen PINs were selected by the test participants after being instructed to select easy-to-remember PINs. Each participant was tasked to authenticate oneself through the PIN entry system within three attempts, under every combination of the independent variables (PIN type × PIN entry system). We counterbalanced the order of the conditions to reduce learning effects. For the user-chosen PIN and TictocPIN combination, we tested whether the initial training done during the Day-1 experiment had affected the results by having the participants enter the PINs they had used on Day-1 again on Day-5.

2) Participants: We recruited 24 participants (17 males, 7 females) from the local university after gaining the official approval of the Ethical Review Board organized by the university's Student Affairs Section and Research Support Division. We tried to balance their majors and ages to represent the general population of users. The participants had (corrected-to-)normal eyesight and 22 of them were right-handed. Their average age was 28, ranging from 21 to 42, and their average experience of using smartphones (cell phones) was 3 (11) years. All the participants had prior experiences with the standard PIN system. We gave them a small gratuity for taking part in our user study.

3) Procedure: At the instructional meeting, we explained our user study and demonstrated both standard PIN and TictocPIN systems with the simple PIN, 1234. We repeatedly demonstrated the TictocPIN system with both fast and slow key entries during the tutorial session. We also collected demographic information about the participants. After the tutorial session, we explained the test procedure again and asked each participant to submit two (easy-to-remember) user-chosen PINs, one for each PIN entry system. We also generated two random PINs with the properties that the digits are distinct and not in monotone order, for each user.

On Day-1, each participant was asked to go through a training phase followed by an evaluation phase, for each of the four PIN type and PIN entry system combinations. Each training phase required the participant to first perform three training trials that resulted in successful authentications and to show us one successful practice entry, i.e., four successful trials in all. We did not count the failed trials during the training phase. The participants failing the practice entry that was shown to us were supposed to repeat the practice trials, but this did not happen. After the training phase, the participants were asked to enter their PINs for evaluation. Each participant was asked to authenticate him/herself to our system with the same PIN that was used during practice within a maximum of three trials. The PIN entry time and authentication results were logged for later analysis. After completing their tasks for all four combinations, the participants were asked to fill out post-test questionnaires. The Likert-type scale was used for ratings from 1 (strongly disagree) to 5 (strongly agree).

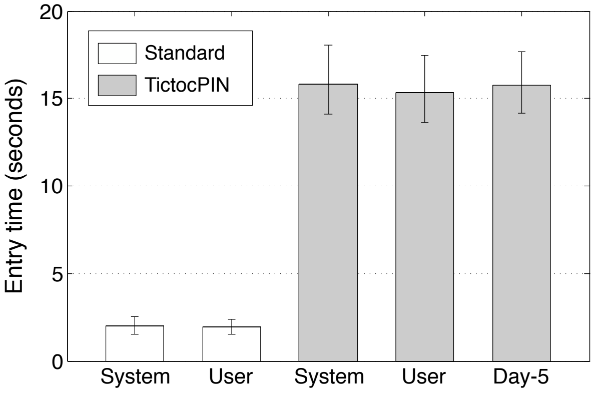


Fig. 7. PIN entry time of a successful authentication session.

The participants were summoned again on Day-5 and asked to re-enter their user-chosen PINs through the TictocPIN system. This was done without any formal instructions or practice trials. Each participant entered the same user-chosen PIN that was used on Day-1. The PIN entry time and authentication results were logged again for later analysis. Since our interest was in whether each participant could remember the PIN entry method, as opposed to the PIN itself, we provided any participant who was unsure of the PIN used on Day-1 with the information.

4) Hypotheses: Our hypotheses for the user study regarding the usability were as follows:

(H1) TictocPIN is slower than the standard PIN system.
(H2) TictocPIN is more error-prone than the standard PIN system.
(H3) TictocPIN is slower on Day-5 than on Day-1.
(H4) TictocPIN is more error-prone on Day-5 than on Day-1.

5) PIN Entry Time: For both PIN entry systems, we defined the PIN entry time to be the time span beginning with the first display of the numeric keypad and ending with the final user input through the touchscreen. The participants were allowed to use refresh and backspace keys. The execution time was measured during the evaluation trials for only the successful sessions. Figure 7 graphically illustrates the resulting successful PIN entry times for the two PIN entry systems in combination with how the PINs were chosen. The same content is summarized numerically in Table I.

The fastest combination was the standard PIN system with user-chosen PINs. This was followed by the standard PIN system with system-chosen PINs, the TictocPIN system with user-chosen PINs, and the TictocPIN system with system-chosen PINs.

A 2×2 (PIN type × PIN entry system) Repeated Measures-ANOVA test showed that there was a significant main effect for PIN entry system ($F(1, 23) = 7782.713$, $p < 0.001$). However, there was no significant main effect for PIN type

TABLE I
PIN ENTRY TIME (SECONDS).

PIN entry methods	Mean	Min	Max	Sd
Standard PIN, system-chosen	2.017	1.515	2.512	0.218
Standard PIN, user-chosen	1.934	1.546	2.397	0.186
TictocPIN, system-chosen	15.806	14.107	18.061	1.051
TictocPIN, user-chosen	15.314	13.639	17.493	0.967
TictocPIN, user-chosen, Day-5	15.750	14.155	17.695	0.919

($F(1, 23) = 4.001$, n.s.). The interaction effect between PIN type and PIN entry system was also not significant ($F(1, 23) = 1.937$, n.s.). Based on these results, we can accept hypothesis H1.

6) Error Rate: During the evaluation phases, we observed whether each participant could succeed in authenticating oneself within three trials. With the standard PIN system, all participants were successful at their first trials, regardless of how the PINs were chosen. With TictocPIN, two participants submitting system-chosen PINs succeeded at their second trial (4.167% overall), while the remaining participants succeeded at their first trial. With both PIN entry systems, no participant experienced three consecutive authentication failures during the evaluation phase. There were no significant differences between the different PIN types ($F(1, 23) = 2.091$, n.s.) and between the different PIN entry systems ($F(1, 23) = 2.091$, n.s.). It can be stated that we have failed to find support for hypothesis H2.

7) Effect of Intermittent Use: The Day-5 experiment was conducted to check whether the results of Day-1 were not positively affected by the training that had immediately preceded the evaluation phases and to verify whether the results would be valid for applications where the PIN entry method is used only occasionally. The results are given in Figure 7 and Table I. A paired-samples t-test suggested that there was no significant difference in the PIN entry time between Day-1 and Day-5 experiments ($t(23) = -1.549$, n.s.). One participant succeeded at the second trial (4.167%) while all other participants succeeded at their first trial. A paired-samples t-test suggested that there was no significant difference in the error rate between Day-1 and Day-5 experiments ($t(23) = -1.000$, n.s.). These results fail to support hypotheses H3 and H4.

8) User Evaluation Report: The participants of our user study provided the following evaluations through the questionnaires: TictocPIN (mean: 4.71, sd: 0.550) was more secure than the standard PIN entry system (mean: 1.25, sd: 0.442). However, the standard PIN entry system (mean: 4.92, sd: 0.282) was more usable than TictocPIN (mean: 3.04, sd: 0.955). For rarely-used but security-sensitive applications, such as on-line banking, TictocPIN (mean: 3.21, sd: 1.179) was likely to be chosen for PIN entries, when in a public place. The short vibrotactile signals were not difficult to perceive (mean: 3.96, sd: 0.751). Figure 8 illustrates the main results of the questionnaires.

9) Discussion: The user experiment has shown that the PIN entry time of TictocPIN is significantly slower than that of the standard PIN system, which is to be expected, but that the error rates of the two are not significantly different. Considering the fixed base time spent on displaying the challenges, we can see that the time required by TictocPIN purely for the user input process is in the range of $5.314 \sim 5.806$ seconds, or $0.664 \sim 0.726$ seconds for each of the eight color key presses that follow two or three 500 msec displays. This is slightly longer than the $0.483 \sim 0.504$ seconds per single digit entry required by the standard PIN system. The difference in key pressing times may be attributed to various factors, which would include, among others, the randomized ordering of the input pad colors and the intermittent nature of the entry method itself. We believe that these multitude of reasons can create an averaging effect and, to a certain degree, contribute to the resistance of TictocPIN against timing attacks [17]. However, we have already discussed in Section IV-C6 that it would be difficult to prevent all detrimental user behaviors. On Day-5, we were able to observe that the participants could remember how to use TictocPIN and, under the assumption that they still remembered their PINs, perform without significant differences with their Day-1 performances. The user evaluation report also supports the experimental results: Although the scores were significantly lower than those for the standard PIN system, the participants still evaluated TictocPIN to be higher than score 3 regarding usability, usage, and haptic experience.

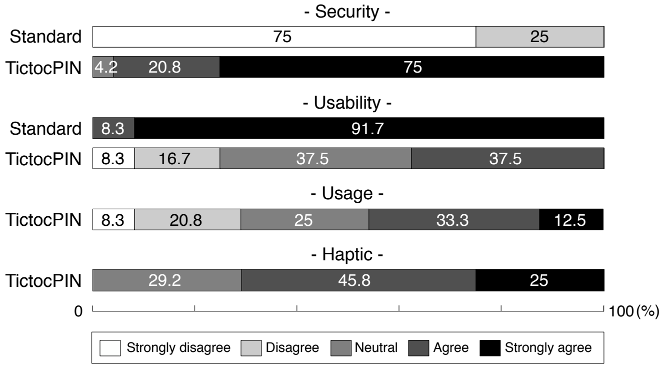


Fig. 8. Security and usability. The Likert-type scale was used with ratings from 1 (strongly disagree) to 5 (strongly agree). Questions: Security - I believe this method is secure against camera-based recording attacks; Usability - I was able to use this method without difficulty; Usage - I am likely to choose this method for seldom-used but security-sensitive applications, such as on-line banking, when in a public place; Haptic - It was not difficult to perceive the short vibrotactile signals.

V. RELATED WORK

In this section, we review the related works and discuss several issues closely related to our work.

The fact that shoulder-surfing attacks are directed at the human user makes their prevention through cryptographic techniques quite infeasible. Since the seminal work of Matsumoto and Imai [14] and the work by Wang et al. [20] that discussed its security, a large number of studies have considered alternatives that are within the limitations of humans. The central theme has been to incorporate an *indirect* method for secret transfer, that is, to separate the visible key entry procedure from the secret itself. Some of the attempts have focused on textual passwords [3], [26], graphical passwords [21], [22], and PINs [5], [16], [18], [23]. Some have leveraged the use of haptic channels [4], [6], [19], with the work [19] even taking into account the possibility of the haptic channel leaking information visually. Even the approach of physically occluding the leakage from at least some component of the visual channel [10], [25] can be found. The very existence of these diverse schemes testifies as to how challenging it is to design an authentication scheme that is both secure and usable [9], [14]. Measures that strengthen security are likely to result in highly complex, error-prone, and tedious user procedures, while putting more emphasis on usability can lead to insecure schemes [2], [24].

The difficulty of the task is also evident from the following non-exhaustive list of related works: Golle and Wagner revealed the insecurity of the cognitive authentication scheme [8], [21]. Li et al. represented a brute force attack against the so-called PAS scheme [3], [12]. Dunphy et al. showed a replay-based shoulder-surfing attack against the recognition-based graphical password system [7]. Asghar et al. revealed the insecurity of the Convex Hull Click (CHC) and its related methods [1], [22], [26]. Yan et al. reported on general attacks against leakage-resilient password systems and discussed the security-usability tradeoff [24]. Asghar et al. revisited Yan et al.'s work and showed how to theoretically estimate a lower bound on the number of authentication sessions that are safe against passive observers [2]. Kwon et al. showed that the basic IOC version of the BW method was vulnerable to human shoulder-surfing attacks if those attackers were trained and prepared [11], [18]. Bianchi et al. discussed how a directional microphone or similar device was a realistic threat to vibrotactile signaling schemes [4], [6]. Perkovic et al. disclosed the insecurity of Undercover by exploiting the user's behavioral (timing) characteristics or the systematic intersection of multiple random challenges [17], [19].

Let us very briefly present the *Undercover* scheme, created by Sasamoto et al. [19], that uses graphical passwords. As with almost any other schemes, they made use of (visible) graphical challenges, but made further use of separate (invisible) tactile challenges, delivered through a specially designed haptic device. The user placed one hand on a trackball to sense its direction of spin or vibration. This tactile challenge was mentally combined with the graphical challenge to create a temporary identifier for the user's secret image, which was submitted to the system with the other hand. Our TictocPIN scheme, which uses short vibration signals, relies on the innovative idea set forth by Undercover in that we also exploit the security enhancements made possible by a hidden haptic channel. Needless to say, TictocPIN, which provides sufficient security against even the camera-based recording attackers, is clearly also a security-strengthened extension to the creation of [18].

Our work may also be seen as extending the work of Kwon et al. [11], which dealt with the human shoulder surfer attacking a IOC BW system. By exercising selective cognitive attention, a trained attacker was able to conduct a perceptual grouping of colored patterns to single out a PIN digit. Furthermore, through parallel motor operations, each of his finding could be written down without impeding his ability to identify the next PIN digit, thus completely breaking the IOC BW method. As noted previously, some of the analysis results given in our work provides insights as to why such an attack was possible. Furthermore, our work extends the work of Kwon et al. by providing security shortcoming of all versions of the BW method, through both theoretical and experimental means.

Finally, we remark that the lessons learned from the timing attacks of [17] has allowed us to specify for the display of the input color pad to be delayed and randomized with TictocPIN.

VI. Conclusion

The BW method, proposed by Roth. et al. [18], was an impressive pioneering work which created a simple indirect PIN entry method that leveraged the intuitiveness of interacting with colors over multiple rounds. In this work, we first analyzed the BW method both experimentally and theoretically, and uncovered multiple inconspicuous problems, such as round redundancy, unbalanced key presses, frequent system errors, and recording non-resilience. We then strengthened the BW method into the TictocPIN scheme, so that the previous problems are resolved and the basic ideas of the BW method survive in a more viable PIN entry method. The TictocPIN method requires a smaller number of rounds than the original BW method and utilizes vibrotactile signals to inform the user as to which of multiple displayed challenges are to be considered as valid. We analyzed the security of TictocPIN and further conducted a user study involving 24 volunteers in a 2×2 within-subject design to evaluate its usability.

We have shown that TictocPIN remains secure against camera-based recording attacks *for any number of sessions*, as long as the adversary is unable to access the haptic channel, and TictocPIN holds two measures that can prevent information leakage of the haptic channel through the audio channel. Our experiences with rudimentary implementations of the two methods for obfuscating the 30 msec vibration sounds have given us confidence that these defenses are sufficiently effective against even the strongest adversaries with directional microphones. However, a more systematic experimental study of the audio channel obfuscation, under various realistic and extreme physical conditions, is outside the scope of this paper and is left to a future work.

Note that our threat model, as with the original BW scheme paper and many related works, rule out the possibility of malware. We believe this assumption to be quite reasonable, since a malware installed on a smartphone is capable of doing much more than just obtaining access to the haptic channel. However, given the rapid growth of mobile malware, it would be necessary to consider the threat of malware in future studies. In fact, the gyroscopes commonly found in smartphones are sufficiently sensitive to vibration to enable even rudimentary reconstruction of speech [15] and smartphone apps do not require any permission from current Android or iOS to access the gyroscope data. It would be interesting to see if the audio signals over headphones can be used as a possible alternative to vibrotactile channels in such a stronger threat model, but we leave this as a subject of future study.

One limitation of TictocPIN is with its PIN entry time. Although our user study survey indicated that TictocPIN is usable for seldom-used security-sensitive applications, such as on-line banking in public places, the requirement of approximately 15 seconds per session is still uncomfortable. We hope for the analysis arguments, findings, and the approach of this work to lead to sufficiently secure and more usable future PIN entry methods.

Appendix
Proofs of Lemmas

This section provides all the proofs of lemmas that were omitted in Section III. We ask the reader to recall the concepts 5+5 split, 5-digit set, partition, cell, and i-cell that were introduced at the beginning part of Section III before reading these proofs.

A. Proof of Lemma 1

Let us consider an adversary that has completely missed either the 1-st or the 2-nd round and calculate the probability for her to be still successful in uniquely identifying the key digit, using only the information gathered from the remaining rounds.

After the 2-nd round, the system knows that the key digit belongs to either a certain 2-cell or a certain 3-cell, but the adversary has narrowed down the key only to a set of 5 digits.

Let us first consider the system's 2-cell case. Since the 3-rd round 5+5 split will separate the key and non-key digits belonging to the 2-cell, the 3-rd round user input will allow the adversary to eliminate the single non-key belonging to the 2-cell. The adversary still needs to remove the 3 non-keys from his remaining pool of 4 possible keys.

In the 3-rd round, the system holds 8 eliminated keys and will color 4 of these in the color opposite to the key. In the 4-th round, the system has already identified the key and will color 5 of the 9 eliminated keys differently from the key. The probability for these two colorings and the corresponding user inputs to disclose the remaining 3 non-keys held by the adversary is $\frac{\binom{3}{3}\binom{5}{1}}{\binom{8}{4}}\frac{\binom{9}{5}}{\binom{9}{5}}+\frac{\binom{3}{2}\binom{5}{2}}{\binom{8}{4}}\frac{\binom{8}{4}}{\binom{9}{5}}+\frac{\binom{3}{1}\binom{5}{3}}{\binom{8}{4}}\frac{\binom{7}{3}}{\binom{9}{5}}+\frac{\binom{3}{0}\binom{5}{4}}{\binom{8}{4}}\frac{\binom{6}{2}}{\binom{9}{5}}=\frac{257}{588}$.

To treat the system's 3-cell case, we must further break it down into two sub-cases. At the 3-rd round, the system will color either 1 or 2 digits among his pool of 3 digits containing the key in the anti-key color.

When all 2 non-keys from the 3-cell are colored differently from the key, through arguments similar to those given before, we can state that the observer can identify the key with probability $\frac{\binom{2}{2}\binom{5}{1}}{\binom{7}{3}}\frac{\binom{9}{5}}{\binom{9}{5}}+\frac{\binom{2}{1}\binom{5}{2}}{\binom{7}{3}}\frac{\binom{8}{4}}{\binom{9}{5}}+\frac{\binom{2}{0}\binom{5}{3}}{\binom{7}{3}}\frac{\binom{7}{3}}{\binom{9}{5}}=\frac{34}{63}$.

The remaining case is when just 1 of the 2 non-keys belonging to the system's 3-cell is colored differently from the key digit in the 3-rd round. In this case, we can argue that the probability for the key to be completely disclosed to the adversary is $\frac{\binom{2}{2}\binom{5}{2}}{\binom{7}{4}}\frac{\binom{8}{4}}{\binom{8}{4}}+\frac{\binom{2}{1}\binom{5}{3}}{\binom{7}{4}}\frac{\binom{7}{3}}{\binom{8}{4}}+\frac{\binom{2}{0}\binom{5}{4}}{\binom{7}{4}}\frac{\binom{6}{2}}{\binom{8}{4}}=\frac{59}{98}$.

Finally, since the probabilities for the above three separate cases to occur are $\frac{2}{5}$, $\frac{3}{5}\frac{1}{3}$, and $\frac{3}{5}\frac{2}{3}$, respectively, we can claim that $\frac{2}{5}\frac{257}{588}+\frac{3}{5}\frac{1}{3}\frac{34}{63}+\frac{3}{5}\frac{2}{3}\frac{59}{98}=\frac{2309}{4410}$ is the probability for the observer to be able to uniquely identify the key without knowledge of the 1-st round.

B. Proof of Lemma 2

After the 2-nd round, both the system and the observer are left with either a 2-cell or a 3-cell that is known to contain the correct key digit. In the 2-cell case, the 3-rd round identifies the key digit to the system, and 5 of the 9 eliminated digits are colored in the anti-key color. This will remove all of adversary's ambiguity with probability $\binom{8}{4}/\binom{9}{5}=\frac{5}{9}$. In the 3-cell case, the 3-rd round may either leave the system with a uniquely identified key or a 2-cell. In the former case, there is probability $\binom{7}{3}/\binom{9}{5}=\frac{5}{18}$ for the observer's ambiguity to be removed, and the said probability is $\binom{7}{3}/\binom{8}{4}=\frac{1}{2}$, in the latter case. In summary, the observer can unique identify the key digit with probability $\frac{2}{5}\frac{5}{9}+\frac{3}{5}\frac{1}{3}\frac{5}{18}+\frac{3}{5}\frac{2}{3}\frac{1}{2}=\frac{43}{90}$, even without observing the 3-rd round.

C. Proof of Lemma 3

After the 2-nd round the system will see the key digit as belonging to a 2-cell with probability $\frac{2}{5}$ and to a 3-cell with probability $\frac{3}{5}$. If the key belongs to a 2-cell, it is uniquely identified by the system from the 3-rd round user input. On the other hand, if the key belongs to a 3-cell, it will be uniquely identified at the 3-rd round with probability $\frac{1}{3}$ only. Hence, one can claim $\frac{2}{5}+\frac{3}{5}\frac{1}{3}=\frac{3}{5}$ to be the probability for a key digit to be uniquely identified through the first 3 rounds of inputs.

D. Proof of Lemma 5

The random 5+5 split of the 1-st round implies that B is pressed with probability $\frac{1}{2}$. The 5 remaining possible keys are divided into groups of 3 and 2 in the 2-nd round, with the larger group colored B, so that B is pressed with probability $\frac{3}{5}$.

If the 2-nd round was W, there are only 2 possible keys remaining, so that B and W are pressed with equal likelihood in the 3-rd round. However, this must be followed by B on the 4-th round, since the key has been identified uniquely.

If the 2-nd round was B, there are three possible keys remaining. These are divided into 2 Bs and 1 W, so that B is pressed with probability $\frac{2}{3}$ and W with $\frac{1}{3}$. If B is pressed, the 4-th round can be either B or W with equal probability, and if the W is pressed, the 4-th round must be B.

Gathering the discussed information, the probability for B presses averaged over the four rounds can be calculated as $\frac{1}{4}\left\{\frac{1}{2}+\frac{3}{5}+\frac{2}{5}\left(\frac{1}{2}+\frac{1}{1}\right)+\frac{3}{5}\left(\frac{2}{3}+\frac{2}{3}\frac{1}{2}+\frac{1}{3}\frac{1}{1}\right)\right\}=\frac{5}{8}$. Furthermore, the probabilities for B presses for each of the 1-st through 4-th rounds can be stated as $\frac{1}{2}$, $\frac{3}{5}$, $\frac{2}{5}\frac{1}{2}+\frac{3}{5}\frac{2}{3}=\frac{3}{5}$, and $\frac{2}{5}\frac{1}{1}+\frac{3}{5}\left(\frac{2}{3}\frac{1}{2}+\frac{1}{3}\frac{1}{1}\right)=\frac{4}{5}$.

E. Proof of Lemma 6

After the 2-nd round, the system will have narrowed down key digit k to either a 2-cell or a 3-cell, with respective probabilities $\frac{2}{5}$ and $\frac{3}{5}$.

Let us first treat the 3-cell case, naming the two non-key digit in the 3-cell as x and y. The system fails to identify k if and only if at least one of x or y is given the same B/W colors as k in both the 3-rd and 4-th round 5+5 splits. To construct the 5+5 split for the 3-rd round, the system will divide the 5-digit set containing k, x, and y into two parts of sizes 2 and 3. Of the $\binom{5}{2} = 10$ such division, one will contain k, x, and y, all in the same part, three will contain k and x in the same part without y, and another three will contain k and y in the same part without x.

There is probability $\frac{1}{10}$ for k, x, and y to be colored the same in the 3-rd round, in which case, the 4-th round will fail to separate k from at least one of x or y with probability $\frac{1}{10} + \frac{3}{10} + \frac{3}{10} = \frac{7}{10}$. On the other hand, if k is separated from one of x and y, but not from the other, in the 3-rd round, then the 4-th round will again fail to separate k from the partnered x or y with probability $\frac{2}{5}$, by our previous argument. The 3-cell case fails to separate k from the other two digits with probability $\frac{1}{10}\frac{7}{10} + \frac{3+3}{10}\frac{2}{5} = \frac{31}{100}$.

The case of k belonging to a 2-cell, after the 2-nd round, can be treated as above, except that this case is slightly easier than the 3-cell case. Let us only state that, in the 2-cell case, one may fail to separate k from the single other digit in the 2-cell with probability $\frac{4}{25}$.

In all, there is probability $\frac{2}{5}\frac{4}{25} + \frac{3}{5}\frac{31}{100} = \frac{1}{4}$ for the 4-rounds to fail in separating k from the other 9 digits through at least one opposite coloring.

References

[1] H. J. Asghar, S. Li, J. Pieprzyk, and H. Wang, "Cryptoanalysis of the convex hull click human identification protocol," in *Proc. the 13th International Conference on Information Security*, 2010, pp. 24–30.

[2] H. J. Asghar, S. Li, R. Steinfeld, and J. Pieprzyk, "Does counting still count? revisiting the security of counting based user authentication protocols against statistical attacks," in *the 20th Internet Society NDSS (Network and Distributed System Security) Symposium*, 2013.

[3] X. Bai, W. Gu, S. Chellappan, X. Wang, D. Xuan, and B. Ma, "PAS: predicate-based authentication services against powerful passive adversaries," in *Proc. IEEE Annual Computer Security Applications Conference.*, 2008, pp. 433–442.

[4] A. Bianchi, I. Oakley, and D. S. Kwon, "Counting clicks and beeps: Exploring numerosity based haptic and audio pin entry," *Interacting with Computers*, vol. 24, pp. 409–422, 2012.

[5] A. De Luca, K. Hertzschuch, and H. Hussmann, "ColorPIN - Securing PIN entry through indirect input," in *Proc. ACM CHI Conference on Human Factors in Computing Systems*, 2010, pp. 1103–1106.

[6] A. De Luca, E. von Zezschwitz, and H. Hussmann, "Vibrapass - secure authentication based on shared lies," in *Proc. ACM CHI Conference on Human Factors in Computing Systems*, 2009, pp. 913–916.

[7] P. Dunphy, A. P. Heiner, and N. Asokan, "A closer look at recognition-based graphical passwords on mobile devices," in *Proc. the 6th Symposium on Usable Privacy and Security*, 2010, pp. 1–12.

[8] P. Golle and D. Wagner, "Cryptanalysis of a cognitive authentication scheme," in *Proc. IEEE Symposium on Security and Privacy.*, 2007, pp. 66–70.

[9] N. Hopper and M. Blum, "Secure human identification protocols," *Advances in Cryptology-ASIACRYPT 2001*, pp. 52–66, 2001.

[10] D. Kim, P. Dunphy, P. Briggs, J. Hook, J. W. Nicholson, J. Nicholson, and P. Olivier, "Multi-touch authentication on tabletops," in *Proc. ACM SIGCHI Conference on Human Factors in Computing Systems (CHI '10)*, 1093-1102, Ed., 2010.

[11] T. Kwon, S. Shin, and S. Na, "Covert attentional shoulder surfing: Human adversaries are more powerful than expected," *IEEE Trans. on Systems, Man, and Cybernetics: Systems*, vol. 44, no. 6, pp. 716–727, 2014 (Online publication on August 2013).

[12] S. Li, H. J. Asghar, J. Pieprzyk, A.-R. Sadeghi, R. Schmitz, and H. Wang, "On the security of PAS (predicate-based authentication service)," in *Proc. the 2009 Annual Computer Security Applications Conference*, 2009, pp. 209–218.

[13] J. Long and J. Wiles, *No Tech Hacking: A Guide to Social Engineering, Dumpster Diving, and Shoulder Surfing*. Syngress, 2008.

[14] T. Matsumoto and H. Imai, "Human identification through insecure channel," in *Proc. Advances in Cryptology-EUROCRYPT91*. Springer, 1991, pp. 409–421.

[15] Y. Michalevsky, D. Boneh, and G. Nakibly, "Gyrophone: Recognizing speech from gyroscope signals," in *Proc. USENIX Security Symposium*, 2014, pp. 1053–1067.

[16] T. Perkovic, M. Cagalj, and N. Rakic, "SSSL: Shoulder surfing safe login," in *International Conference on Software, Telecommunication and Computer Networks - (SoftCOM'09)*, 2009.

[17] T. Perkovic, A. Mumtaz, Y. Javed, S. Li, S. A. Khayam, and M. Cagalj, "Breaking undercover: Exploiting design flaws and nonuniform human behavior," in *Proc. the 7th Symposium on Usable Privacy and Security*, 2011.

[18] V. Roth, K. Richter, and R. Freidinger, "A PIN-entry method resilient against shoulder surfing," in *Proc. ACM CCS (Computer and Communications Security)*, 2004, pp. 236–245.

[19] H. Sasamoto, N. Christin, and E. Hayashi, "Undercover: authentication usable in front of prying eyes," in *Proc. ACM SIGCHI Conference on Human Factors in Computing Systems (CHI '08)*, 2008, pp. 183–192.

[20] C. Wang, T. Hwang, and J. Tsai, "On the Matsumoto and Imai's human identification scheme," in *EUROCRYPT'95*. Springer-Verlag, 1995, pp. 382–392.

[21] D. Weinshall, "Cognitive authentication schemes safe against spyware," in *Proc. IEEE Symposium on Security and Privacy*, 2006.

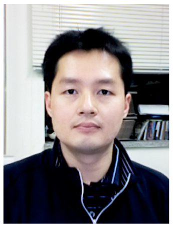
[22] S. Wiedenbeck, J. Waters, L. Sobrado, and J. C. Birget, "Design and evaluation of a shoulder-surfing resistant graphical password scheme," in *Proc. ACM International Working Conference on Advanced Visual Interfaces*, 2006, pp. 177–184.

[23] G. T. Wilfong, "Method and apparatus for secure PIN entry," *In Lucent Technologies, Inc., Murray Hill, NJ, U.S. Patent, Ed. United States*, 1999.

[24] Q. Yan, J. Han, Y. Li, and R. H. Deng, "On limitations of designing leakage-resilient password systems: Attacks, principles and usability," in *the 19th Internet Society NDSS (Network and Distributed System Security) Symposium*, 2012.

[25] Q. Yan, J. Han, Y. Li, J. Zhou, and R. H. Deng, "Designing leakage-resilient password entry on touchscreen mobile devices," in *ASIA CCS'13*, 2013.

[26] H. Zhao and X. Li, "S3PAS: A scalable shoulder-surfing resistant textual-graphical password authentication scheme," in *Proc. IEEE International Conference on Advanced Information Networking and Applications Workshops.*, vol. 2, 2007, pp. 467–472.



Taekyoung Kwon received the BS, MS, and PhD degrees in computer science from Yonsei University, Seoul, Korea, in 1992, 1995, and 1999, respectively. He is currently an associate professor of information at Yonsei University, Seoul, Korea. From 1999 to 2000, he was a postdoctoral research fellow at the University of California, Berkeley. From 2001 to 2013, he was a professor of computer engineering at Sejong University, Seoul, Korea. From 2007 to 2008, he visited the University of Maryland, College Park, for sabbatical. His research interests include information security and privacy, applied cryptography, cryptographic protocol, Internet of Things, usable security, and human-computer interaction.

Jin Hong received the BS, MS, and PhD degrees in mathematics from Seoul National University, Korea, in 1994, 1996, and 2000, respectively. After working as a postdoctoral fellow at Korea Institute for Advanced Study, studying quantum groups and crystal bases, he moved to the National Security Research Institute (Korea) as a senior researcher, where he started working on cryptography. He is currently an associate professor at Seoul National University. His research interest lies mainly with symmetric key cryptography.