

抗肩窥攻击的安全口令输入方法

陈舒荻, 朱友文

(南京航空航天大学, 江苏 南京 211100)

摘要: 针对手机移动端设备输入口令容易遭受肩窥攻击的问题, 提出一种抗肩窥攻击的移动端安全口令输入方法。该方法改变标准的口令输入界面, 将设备的振动信道作为辅助信道向用户传递隐蔽信息, 用户通过简单的选择输入口令。通过分析表明, 该方法能有效抵御肩窥攻击。

关键词: 肩窥攻击; 口令输入; 身份认证; 信息安全

中图分类号: TP301.6

文献标识码: A

doi: 10.11959/j.issn.2096-109x.2019010

Shoulder-surfing resistant PIN-entry method

CHEN Shudi, ZHU Youwen

Nanjing University of Aeronautics and Astronautics, Nanjing 211100, China

Abstract: Aiming at the problem that the traditional PIN entry method is vulnerable to shoulder-surfing attack, a secure PIN entry method against shoulder-surfing attack was proposed. The method changes the traditional input interface on the mobile device and uses the vibration channel of the device as the auxiliary channel to transmit the hidden information to the user. The user simply selects items to input PIN. The analysis shows that the proposed method can effectively resist shoulder-surfing attack.

Key words: shoulder-surfing attack, PIN-entry, identity authentication, information security

1 引言

基于口令的身份认证是当今使用最广泛的身份认证方式^[1]。用户在系统的登录界面上直接输入预先设置好的文本或图形口令进行认证^[1]。然

而, 输入口令的过程很容易受到肩窥攻击。肩窥攻击是指越过别人的肩膀窥探来获取信息的攻击手段^[2]。用户在公共场合输入口令时, 攻击者可以在靠近用户的地方偷看或使用设备偷拍用户输入口令的操作进行肩窥攻击。

收稿日期: 2018-12-07; 修回日期: 2018-12-28

通信作者: 陈舒荻, csdcha@foxmail.com

基金项目: 国家重点研发计划基金资助项目 (No.2017YFB0802300); 国家自然科学基金资助项目 (No.61602240)

Foundation Items: The National Key Research and Development Program (No.2017YFB0802300), The National Natural Science Foundation of China (No.61602240)

论文引用格式: 陈舒荻, 朱友文. 抗肩窥攻击的安全口令输入方法[J]. 网络与信息安全学报, 2019, 5(1): 87-93.
CHEN S D, ZHU Y W. shoulder-surfing resistant PIN-entry method[J]. Chinese Journal of Network and Information Security, 2019, 5(1): 87-93.

PIN 码即个人识别码, 由 4 或 6 个十进制数字组成, 是一种常用的口令。用户解锁手机、ATM 机取款、POS 机刷卡消费以及一些电子门禁都需要输入 PIN 码进行身份认证。由于 PIN 码长度很短以及输入方式简单, 再加上 PIN 码的输入经常在公共场合进行, 标准的 PIN 码输入方法非常容易遭受肩窥攻击, 完整的 PIN 码甚至通过一次认证会话就可能被泄露^[1]。若攻击者在获取口令后进一步盗取用户的手机或银行卡, 将会给用户造成巨大的信息和财产损失。许多用户可能在多个账户使用相同或相似的口令, 因而这一安全漏洞造成的问题可能远比想象中的严重^[2]。用户在人群拥挤或有视频监控的地方进行认证是极不安全的, 而这种情况在现实生活中又时常发生, 并且大多数人都疏于防范, 使肩窥攻击成为日常生活中最容易发生的信息安全问题。因此, 非常需要设计抗肩窥攻击的安全口令输入方法, 保护人们的身份认证安全。

文献[3-4]侧重于研究文本口令的输入, 文献[5-7]侧重于图形口令, 文献[8-10]侧重于研究 PIN 码输入。抗肩窥口令输入方案的中心思想是纳入一种间接的秘密转化方法, 将可见的口令输入过程与秘密本身分开。

一些方案仅将安全性依托于攻击者认知能力的有限性, 这类方案面对摄像攻击或交叉攻击, 安全性将大打折扣。Matsumoto 和 Imai^[11]提出了一种基于以下思想的人机认证方案: 终端和用户秘密共享一个字符串 w 和一个有序的符号集 v , 终端呈现随机字符串 $x = x_1, \dots, x_n$ 作为对用户的挑战。如果存在某个 j 使 $x_j = v_j$, 则用户将 x_j 替换为 w_j , 否则从字母表中随机选择一个符号, 由终端验证替换的正确性。与常规的输入程序相比, 记忆口令和符号集以及执行适当替换所需的脑力似乎相当大。Roth^[12]提出了被称为认知陷门游戏的口令输入方法。系统向用户显示十进制数字键盘, 键盘中一半的按键为黑色, 另一半为白色。用户单击单

独的黑白按键指示 PIN 码的颜色。系统通过用户 4 轮黑白颜色的选择确定一位数字, 因此 4 位 PIN 码用户需要 16 轮选择完成。该方案不能抵抗摄像攻击。De-Luca^[8]提出的 ColorPIN 不能抵抗交叉攻击。Lee^[10]提出的 LIN₄ 同样无法抵抗交叉攻击。

有一些方案使用辅助信道, 文献[12]的 Vibrapass 用于输入文本口令, 系统随机产生振动序列, 用户每次输入字符前可能会收到一个信号, 当收到振动信号时输入错误字符, 未收到时输入正确字符。由于每次认证都会输入口令所有的字符, Vibrapass 不能抵抗交叉攻击。文献[13]的 PhoneLock 用于输入 PIN 码, 系统键盘不显示数字, 通过声音传递按键数字, 这个方案需要耳机才能保证信道安全, 给未携带耳机的用户带来不便。文献[14]的 TimeLock 使用不规律的振动来计数, 以按压按键开始振动计数、松开停止计数的方式输入口令, 为了抵抗时序攻击, 方案将口令可输入的数字范围缩小为 1~5, 这使口令空间大幅缩小, 降低了安全性。

本文提出了一种抗肩窥攻击口令输入方法, 在设计认知陷门游戏的基础上利用设备自身的振动功能作为向用户传递信息的辅助信道, 达到了有效抵抗肩窥攻击的目的。

2 抗肩窥的安全口令输入方法

抗肩窥口令输入方法不能让用户直接输入口令, 而是输入其他与口令有关的信息, 通过这些信息能唯一确定用户的口令。本文沿用文献[2]中认知陷门游戏让用户二选一的方式。文献[2]中用户进行一轮 4 次选择能确定一位数字, 这样 4 位 PIN 码需要 4 轮共 16 次选择, 本文考虑若能一轮确定 2 位数字, 最少需要 7 次选择 ($\lceil \lg 100 \rceil = 7$), 这样 4 位 PIN 码所需选择次数缩减为 2 轮共 14 次选择。由此本文提出了 FR (five rows) 方法和 FRV (five rows vibrate) 方法, FR 方法实现了 14 次选择

输入 4 位 PIN 码, FRV 方法在 FR 方法的基础上使用振动信道传递信息, 具有更高的安全性。

2.1 FR 方法

本文提出了 FR 方法, 通过用户的 6 次回答从 50 个候选值中筛选出用户想要输入的候选值。用户进行回答前需要观察屏幕上出现的候选值, 若正确的候选值在其中, 就选择 yes; 若不在, 就选择 no。本文令 A 为 0~49 的候选值, 首先将 A 平分成两个 5×5 的候选集, 5×5 的候选集中十位数为 0、1、2、3、4 的数字各 5 个。系统随机显示其中一个候选集。用户回答后得到其中正确的一个, 系统由此生成新的 5×5 的候选集, 继续让用户选择, 最后系统确定出候选值。

FR 算法流程如下, 其中集合 A 为最初要筛选的集合, γ_ϕ 表示从集合 A 中选出 5×5 个候选值元素, γ_L 表示从集合中选出元素与集合 L 组成 5×5 个候选值元素。 g 表示将集合平分成两个集合, 令十位数相同的数字个数为 q , 集合 L 随机分配 $\left\lfloor \frac{q}{2} \right\rfloor$ 或 $\left\lceil \frac{q}{2} \right\rceil$ 个对应数字, 相应的集合 R 包含另外的 $\left\lfloor \frac{q}{2} \right\rfloor$ 或 $\left\lceil \frac{q}{2} \right\rceil$ 个数字。

算法 1 FR

输入 用户的选择 $choices$

输出 筛选结果 L

1) yes=0

2) no=1

3) $L \leftarrow \gamma_\phi(A)$

4) $R \leftarrow A - L$

5) $P = \phi$

6) $S = \phi$

7) For $i=1, \dots, 6$

8) $P \leftarrow \gamma_L(S)$

9) b 随机为 0 或 1

10) If $b=0$

11) 系统显示 $Q=L \cup P$

12) Else

13) 系统显示 $Q=S \cup R - L \cup P$

14) End if

15) 用户输入当前选择 $choice \in \{yes, no\}$

16) If $choice \oplus b = yes$

17) $S \leftarrow S \cup R$

18) $(L, R) \leftarrow g(L)$

19) Else

20) $S \leftarrow S \cup L$

21) $(L, R) \leftarrow g(R)$

22) End if

23) End for

24) Return L

一轮要输入两位口令, 则有 100 个候选值 (0~99)。将 0~99 分为两个集合 A_1 、 A_2 (0~49 和 50~99) 分别执行 FR 算法。系统显示如图 1 所示, 屏幕上显示 5×5 的方格, 每个方格第一排显示 A_1 对应的候选集, 第二排显示 A_2 对应的候选集, 根据用户的回答同时对 A_1 、 A_2 进行筛选。用户回答的前 6 个问题为所选数字是否在屏幕出现, 如图 1 中用户要输入 25, 用户观察十位数为 2 的那一行的 5 个数字是否包含 25 后回答 yes。当用户选择 6 次后, 得到两个候选值 L_1 、 L_2 ($0 \leq L_1 \leq 49, 50 \leq L_2 \leq 100$)。第 7 个问题为所选数字是否小于 50, 如图 2 所示, 确定唯一的候选值。图 2 中的 DEL 按键用于删除输入的两位口令, 进度条提示用户问题回答进度, 进度条下方的后退按键可以让用户重新回答之前的问题。

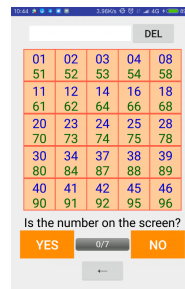


图 1 用户前 6 次回答根据屏幕上对应行的数字做出选择

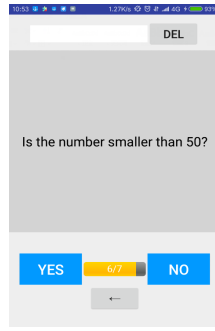


图 2 用户第 7 次作答界面

2.2 FRV 方法

由于 FR 方法中系统提供的挑战只通过视觉信道展示给用户, 攻击者也可以只通过视觉获取的方式得到用户与系统完整的交互过程。攻击者记录下每次屏幕显示的候选值, 结合用户每次的选择进行排除, 能和系统一样得到用户要输入的口令, 因此 FR 方法不能抵抗攻击强度较强的摄像攻击。摄像攻击能获得视觉信道的所有信息, 若系统将所有的交互信息都通过没有遮挡的屏幕显示, 任何口令输入方法都无法抵抗摄像攻击, 因此需要构建辅助信道向用户传递一些隐蔽信息来改变用户的输入。攻击者得不到隐蔽信息就不能得到用户输入与屏幕展示信息的关联, 也就无法用观察到的信息转换成用户的口令。

本文设计了 FRV 方法, 使用设备的振动信道向用户传递信息, 振动信道拥有其固有的安全性, 需要用户用触觉感知。系统每次显示候选值时会随机决定是否振动, 若设备振动, 则用户此次选择相反的选项 (如正确的候选值出现在屏幕, 而用户选择未出现)。方法也可以使用其他方式传递隐蔽信息, 如使用手掌遮挡屏幕, 屏幕才会显示是否输入相反回答的信息等。FRV 算法流程如下。其中, vib 表示系统是否振动。

算法 2 FRV

输入 用户的选择 $choices$

输出 筛选结果 L

1) $yes=0$

2) $no=1$

3) $L \leftarrow \gamma_\phi(A)$

4) $R \leftarrow A - L$

5) $P = \phi$

6) $S = \phi$

7) **For** $i = 1, \dots, 6$

8) $P \leftarrow \gamma_L(S)$

9) b 随机为 0 或 1

10) vib 随机为 0 或 1

11) **If** $b=0$ and $vib=0$

12) 系统显示 $Q=L \cup P$ 并发出仿振动噪声 50 ms

13) **Else if** $b=1$ and $vib=0$

14) 系统显示 $Q=S \cup R - L \cup P$ 并发出仿振动噪声 50 ms

15) **Else if** $b=0$ and $vib=1$

16) 系统显示 $Q=L \cup P$ 同时振动 30 ms 并发出仿振动噪声 50 ms

17) **Else**

18) 系统显示 $Q=S \cup R - L \cup P$ 同时振动 30 ms 并发出仿振动噪声 50 ms

19) **End if**

20) 输入用户当前选择 $choice \in \{yes, no\}$

21) **If** $choice \oplus b \oplus vib = yes$

22) $S \leftarrow S \cup R$

23) $(L, R) \leftarrow g(L)$

24) **Else**

25) $S \leftarrow S \cup L$

26) $(L, R) \leftarrow g(R)$

27) **End if**

28) **End for**

29) **Return** L

为了防止攻击者获取到设备振动时从声音信道泄露的信息, 无论展示的挑战是否包含振动, 系统都会发出仿振动噪声以隐藏振动的声音或扰乱攻击者对声音的判断。通过这样的调整,

攻击者无法分辨系统是否振动,从而无法得知用户选择的答案是否正确。针对用户忘记或不确定当前系统是否振动,界面设置了重放键,如图 3 所示。

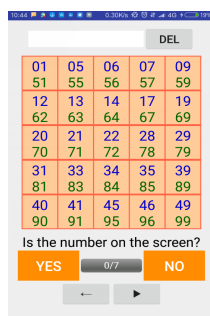


图 3 FRV 方法口令输入界面

3 安全性分析

第 2 节已提及 FR 方法不能抵抗摄像攻击,因此本节只对 FRV 方法的安全性进行分析,考虑攻击者使用普通的猜测攻击和肩窥攻击。

3.1 猜测攻击

4 位 PIN 码的口令空间大小为 10 000, FRV 方法并未缩小口令空间,因此猜测攻击理论上攻击成功率为 $\frac{1}{10\,000}$ 。

3.2 肩窥攻击

由于 FRV 方法使用振动信道传递信息,攻击者不能完整地获知系统发起的挑战,不能确定用户的应答是否正确,所以用户的应答对攻击者而言没有价值。对于肉眼肩窥攻击,由于人类短时记忆量很有限,普通人短时只能记忆 5~9 个符号^[15],一般的攻击者连一轮的候选集都不能完全记下,更无法从多轮的候选集中挖掘有用信息,攻击者不能实施肉眼肩窥攻击,只能采用摄像攻击才能获得会话信息。

攻击者通过摄像攻击拍摄会话后,由于不能判断系统的振动情况,无法从用户的应答获取信息,只能考虑从屏幕获取信息。攻击者有两种方

式可以从屏幕中显示的候选值获取有用信息。

第一种方式是攻击者通过观察每次显示的候选集,采用取交集的方式进一步获得筛选的候选值。因为 FRV 方法每次显示的 0~49 的候选集 Q_1 和 50~99 的候选集 Q_2 都是由未筛选的候选值(简称为筛选值)和被排除的候选值(简称为排除值)组成。以 Q_1 为例,因为每次都要平分上一次得到的筛选值并随机显示其中一个筛选值集合,所以筛选值出现的比率为 $\frac{1}{2}$ 。与筛选值组成 5×5 候

集的排除值每行显示的个数为 5 减去该行筛选值的个数。最初十位数相同的筛选值 10 个,第一次筛选后,十位数相同的筛选值为 5 个,对应的排除值为 5 个。第二次显示的候选集中,上一次的筛选值出现 2 或 3 个,排除值出现比率为 $\frac{3}{5}$ 或 $\frac{2}{5}$,约为 $\frac{1}{2}$ 。第三次显示的候选集中,若上一次的筛选值为 2 个,则排除值为 8 个,排除值出现比率为 $\frac{5-1}{8} = \frac{1}{2}$ 。若上一次的筛选值为 3 个,排除值出现比率为 $\frac{5-1}{7} = \frac{4}{7}$ 或 $\frac{5-2}{7} = \frac{3}{7}$,约为 $\frac{1}{2}$ 。之后,该行的筛选值可能为 1 或 0 个,排除值出现比率为 $\frac{4}{9}$ 或 $\frac{5}{10}$ 。可以得到每次显示的候选集中排除值出现比率约为 $\frac{1}{2}$,即候选集中候选值为排除值或筛选值的概率都为 $\frac{1}{2}$ 。因此攻击者每次观察到的

候选集相当于没有规律的数字,不包含有用的信息。

第二种方式是攻击者假定系统的振动情况,模拟用户真实应答从候选集中得到正确口令。因为候选集中的数值等概率出现,攻击者不能通过候选集缩小候选值的范围,所以无论每次的候选集数值是哪些,100 个候选值都能通过相应的应答被输入,且不同候选值被输入的概率相同,因而攻击者模拟所有的应答情况可以得到 100 个候选值,模拟两轮输入能得到 10 000 个 4 位候

选口令。用户可能的应答数有 2^7 种, 输入 4 位 PIN 码应答数为 $2^7 \times 2^7 = 16\,384$, 4 位 PIN 码口令空间大小为 10 000, 攻击者尝试所有的应答序列也只能排除部分应答序列, 而没有缩小可能的口令空间, 因此模拟用户应答攻击者也不能获得口令。

通过上述分析, 攻击者通过摄像攻击不能获得有用信息来缩小口令可能的范围, 因此对于基于拍摄多次会话的交叉攻击同样不能缩小口令范围。

FRV 方法的安全性基于振动信号的保密性。攻击者若不能知晓振动信号就无法从其余的交互过程中获得任何有效信息, 一旦攻击者获知振动信号, 便确定了用户的应答, 也就获得唯一的正确口令。振动信息除通过触觉获得以外, 还可能通过听觉获得, 因此 FRV 方法使用仿振动噪声掩盖振动信息的传递。30 ms 的振动产生的声音十分微小, 使用平时消息或来电提示音量大小的仿振动噪声就能很好地掩盖, 即使攻击者在非常安静的环境下使用麦克风录制音频也难以分辨, 在公共场合下更是难以做到。

4 可用性分析

本节通过实验比较了 FR 方法、FRV 方法和标准的口令输入方法 (以下简称 REG, regular) 的可用性。本文在安卓手机红米 Note3 (5.5 寸, 1920×1080 分辨率, 403 每寸像素) 上实现了 FR 方法、FRV 方法以及 REG 方法。

4.1 实验设置

本文进行了 3 个组内实验比较 FR 方法、FRV 方法以及 REG 方法的可用性。实验条件分别为使用 FR、FRV 和 REG 登录。每次登录用户都有 3 次机会, 若 3 次机会都失败则判定为此次登录失败。实验招募了 10 名在校生进行实验, 所有的参与者都有多年使用移动手机的经验, 同时对标准的口令输入机制 REG 非常熟悉。

4.2 实验步骤

笔者先对参与者进行指导, 向参与者解释 FR 方法和 FRV 方法的使用, 并向他们演示使用 FR 方法和 FRV 方法登录的过程。之后笔者向参与者说明实验流程并要求他们提交 3 个便于记忆的 PIN 码, 分别用于 FR、FRV 和 REG 的输入。首先参与者进行口令输入的练习阶段。练习阶段要求参与者分别使用 FR、FRV、REG 成功登录 3 次, 在此之前参与者将不断地尝试登录, 之后在研究人员的观察下成功登录一次。练习阶段后, 进行测试阶段。测试阶段要求参与者分别使用 FR、FRV 以及 REG 登录成功一次, 最多 3 次机会, 登录成功的输入用时和使用机会数将被记录用于之后的分析。

4.3 输入用时

输入用时是从界面显示出候选集开始计时, 到用户输入完毕停止计时。实验只记录了测试阶段中登录成功的会话所需输入时间。图 4 展示了实验中 3 种方法的输入用时。单因素重复测量方差分析结果表明不同输入方法间存在显著性差异 ($F=125.073, p<0.001$)。根据这一结果, 可以得出 FR 方法和 FRV 方法的输入时间明显比 REG 方法长的结论。这是由于 FR 方法和 FRV 方法中系统展示多轮的界面需要占去部分时间, 而 FRV 方法中使用振动信道确保安全性的同时, 一定程度上增加了用户的判断时间。

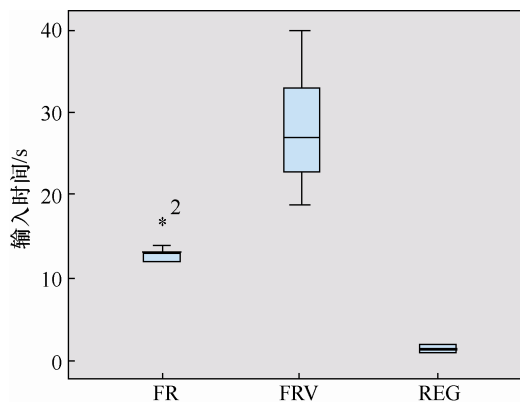


图 4 一次成功登录所需的输入时间

4.4 错误率

测试阶段中实验记录了用户是否能在3次机会内登录成功和登录成功使用的机会数。对于REG方法,所有的参与者都能一次登录成功;对于FR方法,所有的参与者也都一次登录成功;对于FRV方法,两名参与者进行第2次尝试时登录成功,其余参与者一次登录成功。对于3种方法,没有哪个参与者登录失败。不同输入方法并无显著性差异($F=2.250$, $p=0.168$),FR方法和FRV方法相比REG方法错误率不会升高。

5 结束语

针对生活中极易发生的肩窥攻击,本文提出了一种抗肩窥攻击的移动端安全口令输入方法FRV,改变了以往的口令输入界面,同时利用设备的振动功能传递信息,用户以回答的问题的方式间接输入口令。本文对FRV方法理论上进行安全性分析,结果表明新方法能有效抵抗攻击强度较高的摄像攻击以及交叉攻击。进一步招募参与者进行组内实验评估了方法的可用性,结果表明相对标准的口令输入方法,新方法输入时间较长,错误率并无差异,在保证安全性的基础上具有可用性。

参考文献:

- [1] KWON T, HONG J. Analysis and improvement of a PIN-entry method resilient to shoulder-surfing and recording attacks[J]. IEEE Transactions on Information Forensics and Security, 2015, 10(2): 278-292.
- [2] ROTH V, RICHTER K, FREIDINGER R. A PIN-entry method resilient against shoulder surfing[C]//ACM Conf Comput Commun Secur (CCS). 2004: 236-245.
- [3] BAI X, GU W, CHELLAPPAN S, et al. Pas: predicate based authentication services against powerful passive adversaries[C]//Computer Security Applications Conference. 2008: 433-442.
- [4] ZHAO H, LI X. S3PAS: a scalable shoulder-surfing resistant textual-graphical password authentication scheme[C]//International Conference on Advanced Information Networking & Applications Workshops. 2007: 467-472.
- [5] WEINSHALL D. Cognitive authentication schemes for unassisted humans, safe against spyware[C]//IEEE Symposium on Security &

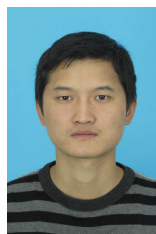
Privacy. 2006.

- [6] WIEDENBECK S, WATERS J, SOBRADO L, et al. Design and evaluation of a shoulder-surfing resistant graphical password scheme[C]//The Working Conference on Advanced Visual Interfaces. 2006: 177-184.
- [7] SUN H, CHEN S, YEH J, et al. A shoulder surfing resistant graphical authentication system[J]. IEEE Transactions on Dependable and Secure Computing, 2018, 15(2): 180-193.
- [8] DE LUCA A, HERTZSCHUCH K, HUSMANN H. ColorPIN—securing PIN entry through indirect input[C]//ACM CHI Conf Human Factors Comput. Syst. 2010: 1103-1106.
- [9] PERKOVIC T, CAGALJ M, RAKIC N. SSSL: Shoulder surfing safe login[C]//Int Conf Softw Telecommun Comput Netw (SoftCOM). 2009: 270-275.
- [10] LEE M K. Security notions and advanced method for human shoulder-surfing resistant PIN-entry[J]. IEEE Transactions on Information Forensics & Security, 2017, 9(4): 695-708.
- [11] MATSUMOTO T, IMAI H. Human identification through insecure channel[C]//International Conference on Theory & Application of Cryptographic Techniques. 1991.
- [12] DE LUCA A, VON ZEISCHWITZ E, HUSMANN H. Vibrapass: secure authentication based on shared lies[C]//International Conference on Human Factors Incomputing Systems. 2009: 913-916.
- [13] BIANCHI A, OAKLEY I, KOSTAKOS V, et al. The phone lock: audio and haptic shoulder-surfing resistant pin entry methods for mobile devices[C]//International Conference on Tangible, Embedded, and Embodied interaction. 2011: 197-200.
- [14] BIANCHI A, OAKLEY I, KWON D S. Counting clicks and beeps: exploring numerosity based haptic and audio PIN entry[J]. Interact Comput, 2012, 24(5): 409-422.
- [15] MILLER G. The magical number seven, plus or minus two: some limits on our capacity for processing information[J]. Psychological Review, 1956, 63(2): 81-97.

[作者简介]



陈舒荻(1995—),女,福建福州人,南京航空航天大学硕士生,主要研究方向为信息安全。



朱友文(1986—),男,安徽阜南人,博士,南京航空航天大学副教授,主要研究方向为信息安全。