面向移动智能设备的多特征融合隐式鉴别机制研究

刘礼才[1,2],李锐光[3],殷丽华[1],郭云川[1],项 菲[3]
(1. 中国科学院信息工程研究所,北京 100093; 2. 北京邮电大学计算机学院,北京 100876;
3. 国家计算机网络应急技术处理协调中心,北京 100029)

摘 要: 隐式鉴别机制在解决移动智能设备的安全性与易用性冲突方面具有重要而独特的作用. 然而,已有工作通常基于单一特征或动作进行隐式鉴别,仅适合于特定动作、场景和范围. 为了解决此问题,本文利用用户使用设备时存在位置、环境、状态、生物和行为特征,提出了一种基于多特征融合的隐式鉴别方案. 该方案采集设备内置传感器、生物和行为数据,通过支持向量机方法训练和提取特征,设计多特征融合模型和构建隐式鉴别框架,计算用户身份信任水平,设计差异化安全策略并持续透明地鉴别用户身份. 实验验证了该方案的有效性,并且能够平衡安全性与易用性和资源消耗.

关键词: 隐式鉴别; 多特征融合; 移动智能设备; 支持向量机
中图分类号: TP309 **文献标识码:** A **文章编号:** 0372-2112 (2016)11-2713-07
电子学报 URL: http://www.ejournal.org.cn **DOI:** 10.3969/j.issn.0372-2112.2016.11.021

Research on Multi-feature Fusion Impact Authentication for Intelligent Mobile Device

LIU Li-cai[1,2], LI Rui-guang[3], YIN Li-hua[1], GUO Yun-chuan[1], XIANG Fei[3]
(1. *Institute of Information Engineering, Chinese Academy of Sciences, Beijing* 100093, *China*;
2. *School of Computer Science, Beijing University of Posts and Telecommunications, Beijing* 100876, *China*;
3. *CNCERT/CC, Beijing* 100029, *China*)

Abstract: Implicit authentication mechanism plays an important and unique role in addressing the collision between security and usability on intelligent mobile device. However, most of current studies are usually built on a single feature or action, while suitable for a particular action and scenes. To solve this problem, we proposed a multi-feature fusion based on implicit authentication scheme, which uses the unique feature, such as the location, environment, posture, gait, biometric and behavioral, during use device. In the scheme, the data, such as sensors, biological and behavioral data, are collected, multi-features are trained and extracted by using the support vector machine. Then, the multi-feature fusion model is designed, and the framework of implicit authentication is constructed for calculating the user confidence level. At last, personalized security policy is designed, and the scheme authenticates user continuously and transparently. Experimental results validate the effectiveness of the proposed scheme and balance the security and usability and energy consumption.

Key words: implicit authentication; multi-feature fusion; intelligent mobile device; support vector machine

1 引言

随着移动技术的快速发展和硬件性能的增强,在移动智能设备(以下称设备)上实现高敏感应用,如股票交易、金融支付、社交软件等;存储和处理高价值信息,如银行卡信息、个人隐私信息等. 因此,保证应用和数据安全,特别是防止非授权访问,尤为重要. 面向设备的用户身份鉴别方案包括 PIN 码、图形密码、数字密码、人脸、语音、指纹等,具有特点:(1)"非此即彼",即通过或拒绝用户访问请求;(2)密码方案存在弱密码、忘记和泄露密码及密码攻击等风险[1];(3)生物特征方案易受外部环境(如噪声、光线、外伤、汗水、身体状况变化等)影响而降低识别率或失效;(4)需要用户主动参与交互以完成身份鉴别. 然而,用户行为存在以下特征:

收稿日期:2015-04-09;修回日期:2015-12-21;责任编辑:马兰英
基金项目:国家高技术研究发展计划(863 计划)(No. 2013AA014002);中国科学院先导专项(No. XDA06030200)

(1)用户操作多数属短时、频繁行为,且40%操作不涉及个人隐私和敏感数据[2].(2)超过50%用户不采用安全保护措施[3].(3)56%用户在十次密码输入中存在错误[4].由这些特点和特征可知:这些方案会给用户带来不便和降低使用效率,从而降低使用意愿,即存在安全性与易用性冲突;"非此即彼"方案增加用户负担,不能满足访问不同应用和数据的多样安全需求.因此,研究更适应设备特点和用户习惯的身份鉴别方案具有重要而独特的意义.

面向设备的用户身份鉴别方案既要保证应用和数据安全可靠,也要提高使用效率和易用性.近年来,研究者提出隐式鉴别方案[5],即从设备内置传感器的感知数据提取用户使用模式和特征,以透明的方式完成身份鉴别.Ali Fahmi 等人提出基于耳廓的隐式鉴别[6],通过接听电话时耳廓图像鉴别用户身份.该方案可提高易用性,但仅适于特定动作和应用场景.基于触摸屏手势的隐式鉴别 TIPS[7] 和 Touchalytics[8],通过操作触摸屏手势特征鉴别用户身份.然而,需要额外传感器手套采集数据,且攻击者可通过快捷键和点击绕开触摸屏手势隐式鉴别.基于传感器的隐式鉴别 SenSec[9]、SilentSense[10] 和 Gait-IA[11],通过从加速度、陀螺仪和磁场传感器获取数据并构建姿态模型或步态特征,鉴别用户身份,准确率最高时超过99%.Shi 等人提出基于用户行为的隐式鉴别[5],评估用户可信度并据此鉴别身份.Khan 等人[12]对隐式鉴别方案准确度、训练时间、检测延时和抵御攻击等进行比较和评估.这些方案可提高易用性和降低负担.但存在特征单一、仅适于特定动作和应用场景、需要额外硬件设备支持、训练时间长和消耗大量资源等问题.因此,不能广泛适用于设备身份鉴别.另外,开源可扩展隐式鉴别框架 Itus[13] 使研究者更专注于用户行为特征和分类算法研究.

针对这些问题,本文研究了面向移动智能设备的隐式鉴别问题,并提出了一种多特征融合的隐式鉴别机制.首先,该方法采集并分析用户在使用设备时的传感器、生物和行为数据,通过支持向量机方法训练、提取与用户身份相关的多种特征(包括位置、环境、姿态、步态、生物和行为特征等);然后,设计多特征融合模型并构建基于该模型的隐式鉴别框架,计算用户身份信任水平,持续透明地鉴别用户身份;针对不同安全需求的应用设置不同的访问阈值,设计差异化隐式鉴别策略;最后,实验验证了所提方法在均衡鉴别与易用性和资源消耗等方面的有效性.

2 多特征融合模型

2.1 攻击模型

假设:设备有唯一合法拥有者;在设备使用过程中,会持续产生大量使用痕迹,存在独特的、可度量的模式和特征,据此无需主动交互即可持续透明地鉴别设备使用者(也称用户)身份,防止非授权访问,抵御安全威胁.安全威胁主要有:(1)攻击者在公共场合意外或故意得到设备(如丢失、盗窃等),攻击者不了解用户任何信息,也不存在相同特征.(2)与拥有者在同一场合的攻击者在未许可下获得设备,攻击者了解用户部分信息或存在部分相同特征.(3)攻击者在用户许可情况下使用设备,但试图访问机密数据或个人隐私.由于用户许可且存在较多共同特征,可避免单一特征鉴别,所以隐式鉴别方案并不能完全抵御该类攻击.本文假设用户有义务和能力保证该情况设备安全.

2.2 多特征融合模型

在使用集成了众多传感器的设备时,会产生大量与用户身份相关的数据,可训练、提取和融合多种特征,可无需主动交互持续透明地进行用户身份鉴别.特征类型可分为:(1)位置特征,如 GPS、WiFi.(2)环境特征,如温度、相对湿度、磁场、重力传感器数据.(3)设备状态特征,如加速度、线性加速度、旋转矢量、方向、陀螺仪、光线、接近、压力传感器数据.(4)生物特征,如麦克风、指纹传感器、摄像头数据.(5)行为特征,如触摸屏、浏览器、应用程序数据.

2.2.1 时空特征

用户的日常活动是由其主要活动场所(如居住地、办公地、购物场所、学校、医院和车站等)及连接路线构成.通常,用户在特定时间处于特定地点,比如,晚上在居住地;工作时间在办公地;出行会选择相对固定路线.因此,用户日常活动存在一定的位置、轨迹和时间相结合的特征(时空特征),可用于身份鉴别.假设设备拥有者主要活动场所为 $RLoc_1,\cdots,RLoc_n$,与之相对应的正常活动时间为 $RLocTime_1,\cdots,RLocTime_n$;主要活动路线为 $RRou_1,\cdots,RRou_n$,与之相对应的正常活动时间为 $RRouTime_1,\cdots,RRouTime_n$;允许活动范围为 $RRang$.根据式(1)判断用户在时间 t 和位置 loc 时的活动是否正常.

$$IsRegular(loc,t)=\begin{cases}1, & \text{位置正常}\\ 0, & \text{位置异常}\end{cases} \tag{1}$$

通过式(2)计算当前时间 $CTime$,位置 $CLoc$ 的基于时空特征的信任水平 $LocCL$.其中,V_{Loc} 为系数,$CLocDur$ 为用户在当前位置的持续时间.用户当前与正常位置的时空距离越小,则 $LocCL$ 越大.

$$LocCL=\left(\begin{array}{l}V_{Loc}\cdot IsRegular(CLoc,CTime)/CLocDur\cdot\\ \left(1-\dfrac{Min|CLoc\cdot CTime-RLoc_i\cdot RLocTime_i|}{RRang}\right)\end{array}\right) \tag{2}$$

式(2)中位置替换为路线可得式(3),计算当前时

间为 $CTime$,路线为 $CRou$ 的基于时空特征的信任水平 $RouCL$. 其中,V_{Rou}为系数,$CRouDur$ 为持续时间.

$$RouCL=\left(\begin{array}{l}V_{Rou}\cdot IsRegular(CRou,CTime)/CRouDur\cdot\\ \left(1-\dfrac{Min|CLoc\cdot CTime-RRou_i\cdot RRouTime_i|}{RRang}\right)\end{array}\right) \tag{3}$$

2.2.2　环境特征

与位置关联,设备所处环境也存在一定规律,如在特定位置,其温度、湿度、磁场强度和重力在特定范围. 假设设备在位置 $RLoc_i$ 的正常温度范围为 $TPLow_i \sim TPHigh_i$,则当前温度为 TP 与正常温度范围的差异如式(4).

$$TPDiff=\left(\frac{|TP-TPLow_i|^2+|TPHigh_i-TP|^2}{(TPHigh_i-TPLow_i)^2}\right) \tag{4}$$

类似地,当前湿度 HM 与正常湿度范围为 $HMLow_i \sim HMHigh_i$ 的差异如式(5).

$$HMDiff=\left(\frac{|HM-HMLow_i|^2+|HMHigh_i-HM|^2}{(HMHigh_i-HMLow_i)^2}\right) \tag{5}$$

类似地,当前磁场强度为 MFI 与正常磁场强度范围为 $MFILow_i \sim MFIHigh_i$ 的差异如式(6).

$$MFIDiff=\left(\frac{|MFI\text{-}MFILow_i|^2+|MFIHigh_i\text{-}MFI|^2}{(MFIHigh_i-MFILow_i)^2}\right) \tag{6}$$

类似地,当前重力为 GR 与正常重力范围为 $GRLow_i \sim GRHigh_i$ 的差异如式(7).

$$GRDiff=\left(\frac{|GR\text{-}GRLow_i|^2+|GRHigh_i\text{-}GR|^2}{(GRHigh_i-GRLow_i)^2}\right) \tag{7}$$

由此,可以计算由温度、湿度、磁场和重力等环境作为影响因素的信任水平 $EnvCL$,如式(8). 其中,V_{TP},V_{HM},V_{MFI},V_{GR}为系数.

$$EnvCL=\left(\begin{array}{l}V_{TP}\cdot TPDiff+V_{HM}\cdot HMDiff\\ +V_{MFI}\cdot MFIDiff+V_{GR}GRDiff\end{array}\right) \tag{8}$$

2.2.3　状态特征

通过内置传感器可判断设备状态特征(如设备本身所处状态及用户步态),进而鉴别用户身份. 将设备状态定义为:(1)静止状态:传感器数据显示设备静止,光线和压力不明显改变,屏幕锁定并且无其他物体接近. 当设备处于静止状态时,其安全性不变;(2)在手中使用:传感器数据显示设备有规律运动,处于活跃状态且用户在使用设备;(3)在包中:传感器数据显示设备有规律运动,屏幕锁定. 假设设备在手中未使用、口袋中和包中均为在包中状态;(4)其他状态:设备处于无规律运动状态.

当设备随用户运动,蕴含着用户活动类型(如步行、跑步和乘车等)和步态特征(如加速度、减速度和频率等). 通过采集传感器数据,基于 Frank 等人[14]的方法可训练和提取活动类型和步态特征 GM. 假设拥有者 $GM_A=(GM_{A1},GM_{A2},\cdots,GM_{An})$,而用户 $GM_T=(GM_{T1},GM_{T2},\cdots,GM_{Tn})$,则两者间距离如式(9).

$$GDist(T,A)=\sum_{i=1}^{n}|GM_{Ti}-GM_{Ai}| \tag{9}$$

基于设备状态和用户步态,可计算由该因素决定的信任水平 $StaCL$,如式(10). 其中,V_{Hand},V_{Poc},V_{Un}为系数,$StaDur$ 为设备在该状态的持续时间.

$$StaCL=\begin{cases}1, & \text{设备静止}\\ V_{Hand}\times GDist(T,A)\times StaDur, & \text{在手中}\\ V_{Poc}\times GDist(T,A)\times StaDur, & \text{在包中}\\ V_{Un}\times StaDur, & \text{其他状态}\end{cases} \tag{10}$$

2.2.4　生物特征

随着硬件技术发展,设备能识别的生物特征越来越多,如声音、人脸和指纹等. 通过采集麦克风数据,基于 Lu 等人提出 SpeakerSense 方法[15]利用声音鉴别身份. 假设拥有者声音特征集 $VM_A=(VM_{A1},VM_{A2},\cdots,VM_{An})$,而用户声音特征集 $VM_T=(VM_{T1},VM_{T2},\cdots,VM_{Tn})$,则两者间的距离如式(11).

$$VDist(T,A)=\sum_{i=1}^{n}|VM_{Ti}-VM_{Ai}| \tag{11}$$

类似地,假设拥有者人脸特征集 $FM_A=(FM_{A1},FM_{A2},\cdots,FM_{An})$,而用户人脸特征集 $FM_T=(FM_{T1},FM_{T2},\cdots,FM_{Tn})$,则两者间距离如式(12).

$$FDist(T,A)=\sum_{i=1}^{n}|FM_{Ti}-FM_{Ai}| \tag{12}$$

由此,可计算由声音和人脸等生物特征作为影响因素的信任水平 $BioCL$,如式(13). 其中,V_{VDist},V_{FDist}为系数.

$$BioCL=V_{VDist}\cdot VDist(T,A)+V_{FDist}\cdot FDist(T,A) \tag{13}$$

2.2.5　行为特征

用户使用设备时存在特定特征和兴趣,如触摸屏幕位置和速度、感兴趣应用和内容等,可鉴别用户身份. 基于 Frank 等人[8]提出的方法可训练和提取用户触摸屏幕行为特征 TM. 假设拥有者 $TM_A=(TM_{A1},TM_{A2},\cdots,TM_{An})$,而用户 $TM_T=(TM_{T1},TM_{T2},\cdots,TM_{Tn})$,则定义两者间距离如式(14).

$$TDist(T,A)=1-\frac{\sum_{j=1}^{n}\sum_{i=1}^{n}p(TM_{Ti}TM_{Aj})\log p(TM_{Ti}/TM_{Aj})}{\sum_{k=1}^{n}p(TM_{Tk})\log p(TM_{Tk})} \tag{14}$$

类似地,通过采集浏览应用和内容数据,利用兴趣挖掘算法挖掘用户兴趣 IM. 假设拥有者 $IM_A=(IM_{A1},$

$IM_{A2},\cdots,IM_{An}$)；而用户 $TM_T=(TM_{T1},TM_{T2},\cdots,TM_{Tn})$. 其中 $IM_i=(Int_i,Deg_i)$，Int_i 为兴趣，Deg_i 为兴趣度，则定义两者间距离如式(15).

$$IDist(T,A)=1-\frac{\sum_{j=1}^{n}\sum_{i=1}^{n}p(IM_{Ti}IM_{Aj})\log p(IM_{Ti}/IM_{Aj})}{\sum_{k=1}^{n}p(IM_{Tk})\log p(IM_{Tk})} \tag{15}$$

由此，可计算由用户行为特征作为影响因素的信任水平 $BehCL$，如式(16). 其中，V_{TDist}，V_{IDist} 为系数.

$$BehCL=V_{TDist}\cdot TDist(T,A)+V_{IDist}\cdot IDist(T,A) \tag{16}$$

最终，将时空、环境、状态、生物和行为等多种特征融合，得到可用于持续透明地鉴别用户身份的信任水平 $UserCL$ 如式(17). 其中，W_{Loc}，W_{Rou}，W_{Env}，W_{Sta}，W_{Bio} 和 W_{Beh} 为各类型特征权重系数，可由用户选择的特征类型和可信程度确定.

$$UserCL=\begin{pmatrix}W_{Loc}\times LocCL+W_{Rou}\times RouCL+\\W_{Env}\times EnvCL+W_{Sta}\times StaCL+\\W_{Bio}\times BioCL+W_{Beh}\times BehCL\end{pmatrix} \tag{17}$$

3 基于多特征融合模型的隐式鉴别框架

3.1 隐式鉴别框架

为了无需主动交互地、持续透明地鉴别用户身份，构建基于多特征融合模型的隐式鉴别框架，如图 1 所示. 隐式鉴别过程包括：

(1)数据采集：为了获得用户身份特征，采集使用设备时的传感器、生物和行为信息. 对数据预处理和敏感性清洗，满足特征提取和隐私保护要求.

(2)特征提取与模型训练：由于在解决小样本的机器学习问题和非线性可分问题的良好性能，选择支持向量机(SVM)方法，从采集数据中提取用户身份特征并进行模型训练. SVM 方法是 20 世纪 90 年代初 Vapnik 等人根据统计学习理论提出的一种机器学习方法[16]. 其基本思想是通过非线性映射将输入向量映射到高维特征空间，构造能将两类样本正确分开且使分类间隔最大的最优分类超平面. 最优分类函数如式(18)，其中，$K(x_i,x_j)$ 为非线性映射的核函数(内积函数).

$$\begin{aligned}f(x)&=\mathrm{sgn}((w^*)^T\varphi(x)+b^*)\\&=\mathrm{sgn}\left(\sum_{i=1}^{n}a_i^*y_iK(x_i,x)+b^*\right)\end{aligned} \tag{18}$$

在形式上，SVM 类似于神经网络，输出是若干中间层节点的线性组合，每个中间层节点对应一个支持向量，如图 2 所示.

在隐式鉴别框架中，采集并标记多个用户 $Y=(y_1,y_2,\cdots,y_s)$ 使用设备时的数据样本 $X=(x_1,x_2,\cdots,x_n)$，

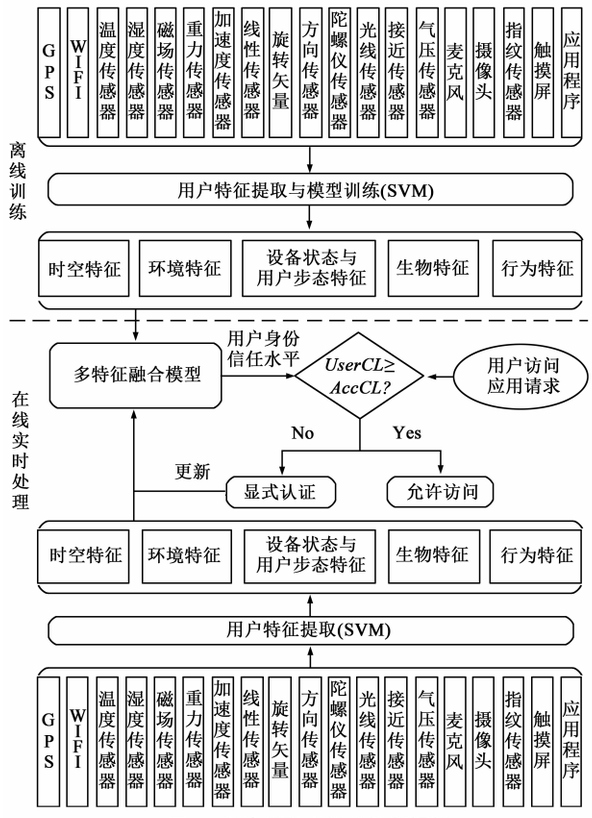
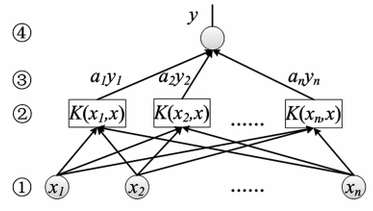


图1 多特征融合隐式鉴别框架



①. 输入向量 $\boldsymbol{x}=(x_1,x_2,...,x_n)$

②. $K(x_i,x)$ 为基于 s 个支持向量向量 $x_1,\ ...,x_s$ 的非线性变换

③. 权值 a_iy_i

④. 输出(决策规则) $y=\mathrm{sgn}(\sum a_iy_iK(x_i,x)+b)$

图2 支持向量机示意图

采用核函数为径向基函数(RBF)的 SVM 进行训练，如式(19)，可以得到相应分类器.

$$K(x,x_i)=\exp\left\{-\frac{|x-x_i|^2}{\sigma^2}\right\} \tag{19}$$

(3)用户身份鉴别：由于不同应用有不同安全需求，且使用频率不同，如天气、播放器等一般应用安全需求较低，但使用频率很高；信息、相册和社交软件等敏感应用安全需求较高，使用频率较高；而网银安全需求极高，但使用频率很低. 为保证应用安全和降低用户鉴别负担，设计差异化安全策略，针对不同类别应用设置不

同的用户身份信任水平阈值 $AccCL$,如式(20).

$$AccCL=\begin{cases}V_C\times ConfF\times CL_T, & \text{机密应用}\\ V_S\times\dfrac{ConfF}{UsabF}\times CL_T, & \text{敏感应用}\\ V_N\times\dfrac{ConfF}{UsabF}\times EnerF\times CL_T, & \text{一般应用}\end{cases}\tag{20}$$

其中,CL_T 为设备的用户身份信任水平阈值,$ConfF$,$UsabF$ 和 $EnerF$ 分别为机密性因子、可用性因子和能量因子. 可调节 CL_T 以平衡安全性、可用性和能量消耗. 当 $UserCL\geqslant AccCL$ 时,隐式鉴别用户身份并允许其访问应用;否则,显式鉴别用户身份,如输入密码.

3.2　用户隐私保护

在隐式鉴别中,隐私安全是用户关注的重要问题. 由于设备有唯一合法拥有者,只鉴别用户身份特征,不识别真实身份,不存在用户身份信息与真实身份的一一对应关系. 另外,采取有效措施保障隐私安全:(1)不将通话、信息、联系人等作为特征进行分析和融合. (2)对采集数据进行清洗和脱敏. (3)用特征值代替包含用户隐私的信息,如用时空特征值代替用户位置和轨迹. (4)采用加密措施保证隐私安全.

4　实验结果与分析

4.1　实验设置

为评估本文隐式鉴别方案性能,进行实验和分析,主要验证目标:(1)安全性,即可接受的隐式鉴别准确性 IAR,定义如式(21),其中,TA 为真接受,即设备拥有者通过隐式鉴别、允许访问应用;FR、TR 和 FA 分别为假拒绝、真拒绝和假接受. 保证安全性的同时减少显式鉴别次数,降低用户负担. (2)适用性,即适用于更多场景和动作的隐式鉴别. (3)平衡安全性、易用性和能量消耗,即用户根据安全需求、易用性和能量调节访问应用的用户身份信任水平阈值;(4)可接受的鉴别延时和能量消耗.

$$IAR=\frac{TA+TR}{TA+FR+TR+FA}\tag{21}$$

在实验中,采集了 11 个用户(8 个男士,3 个女士)使用设备 Samsung Galaxy S4 GT-I9502 一周内的传感器数据、生物特征和行为数据. 实验设备系统为 Android OS 4.2.2,CPU 为 Exynos 5410 1.6GHz,RAM 为 2GB.

4.2　实验结果

(1)方案适用性

在已有隐式鉴别方案中,大多基于单一特征鉴别身份,仅适于特定动作和场景. 通过对比可知,多特征融合隐式鉴别方案具有良好的适用性,如表 1 所示.

表 1　各类隐式鉴别方案的适用性比较

方案 \ 场景	地理位置	环境	打电话	触摸屏幕	运动	声音	人脸	应用兴趣
多特征融合	√	√	√	√	√	√	√	√
PA				√	√	√	√	
SilentSense				√	√			
SenSec					√			
Gait-IA					√			
Shi-IA	√		√					√
TIPS				√				
Touchalytics				√				
EarShape			√					

(2)准确率分析

由于各类隐式鉴别方法基于的特征和适应场景不一样,其准确率没有可比性. 比如,Gait-IA 通过步态特征鉴别用户身份,在运动状态下可接近 100% 准确率,但不适应其他场景. 因此,仅对本方案准确率进行分析. 由式(20)可知,通过设置 CL_T 来调节用户访问应用时需要满足的 $AccCL$. 图 3 是 IAR、假接受率(FAR)、假拒绝率(FRR)和机密应用假接受率(FAR)随 CL_T 的变化情况. FAR 随着 CL_T 增加而降低,即通过升高设备的用户身份信任水平阈值来防止非设备拥有者访问应用,增加安全性. 但会增加 FRR 而使得设备拥有者需要更多的显式鉴别,降低 IAR,增加负担. 因此,合适的 CL_T(图中点 A)可保证可接受的 FAR 和 FRR,保证设备和应用安全同时保证 IAR. 需要指出的是,针对机密应用,可设置较高访问阈值,使其假接受率维持低水平,防止机密应用被访问.

(3)平衡安全性、易用性和能量消耗

由于增加安全性会降低 IAR 和增加 FRR,即降低易用性和增加鉴别负担和能量消耗. 因此,平衡安全性、易

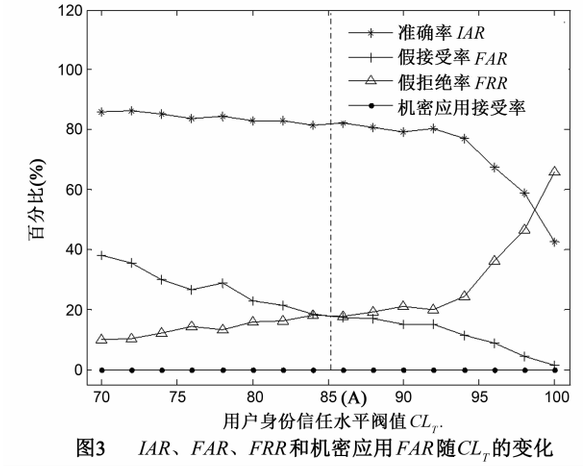


图3　IAR、FAR、FRR和机密应用FAR随CL_T的变化

用性和能量消耗对于隐式鉴别方案至关重要. 由式(20)和图4可知,通过调节 *ConfF* 可影响 *AccCL*,机密性越高,*AccCL* 越高,从而增加机密应用安全性. 但是,*FRR* 会随之而增加,从而降低 *IAR* 和增加显式鉴别率(*EAR*). 因此,设置合适的 *ConfF*(图中点 *B*)可保证可接受的 *IAR* 和 *FRR*,保证隐式鉴别准确率的同时保证安全性.

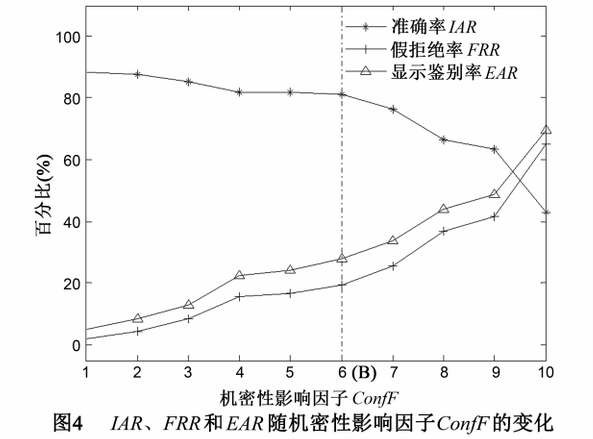


图4 *IAR*、*FRR* 和 *EAR* 随机密性影响因子 *ConfF* 的变化

由式(20)和图5可知,通过调节 *UsabF* 可影响 *AccCL*,易用性越高,*AccCL* 越低. 但 *FAR* 会随之而增加,降低设备和应用的安全性. 因此,设置合适的 *UsabF*(图中点 *C*)可保证可接受的 *IAR* 和 *FAR*,即保证易用性的同时保证安全性.

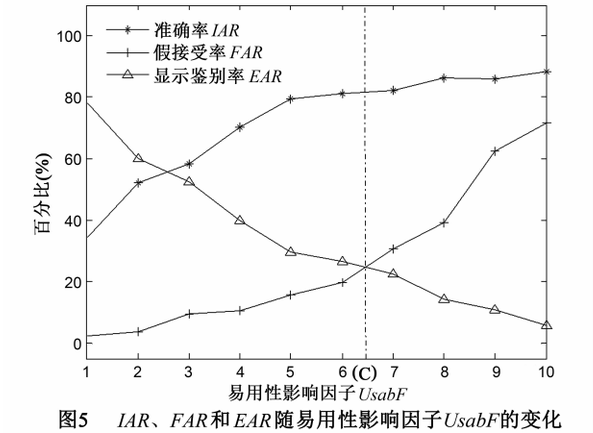


图5 *IAR*、*FAR* 和 *EAR* 随易用性影响因子 *UsabF* 的变化

由式(20)和图6图可知,通过调节 *EnerF* 可以影响 *AccCL*,能量消耗越高,*AccCL* 越高. 但 *FRR* 会随之而增加,从而降低 *IAR* 和增加 *EAR*. 因此,设置合适的 *EnerF*(图中点 *D*)可保证可接受的 *IAR* 和 *FRR*,保证能量消耗合理同时保证隐式鉴别准确率.

因此,通过设置合适的 *ConfF*、*EnerF* 和 *UsabF* 可以平衡安全性、易用性和能量消耗.

最后,由于多特征融合模型的训练和特征提取是在离线完成;然后在设备实时计算 *UserCL*,身份鉴别过程可在秒级时间内完成. 通过运行、统计和分析多特征融合隐式鉴别系统的使用情况,可知其平均能量消耗约为设备电量的10%,处于用户可接受的范围内.

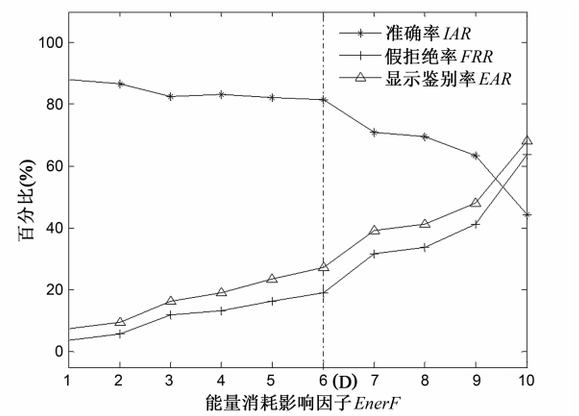


图6 *IAR*、*FRR* 和 *EAR* 随能量消耗影响因子 *EnerF* 的变化

5 结束语

针对鉴别与易用性的均衡问题,本文研究了面向移动智能设备的隐式鉴别问题,并提出一种基于多特征融合的隐式鉴别机制. 该方法在采集和分析用户使用设备时的传感器、生物和行为数据基础上,通过支持向量机方法训练、提取与用户身份相关的多种特征(包括位置、环境、姿态、步态、生物和行为特征等). 其次,设计多特征融合模型并建立基于该模型的隐式鉴别框架,计算用户身份信任水平,实现持续透明地鉴别用户身份. 然后,针对用户安全需求的多样性,设计差异化的隐式鉴别策略. 最后,实验验证了所提方法能够提供可接受的安全保护和提高易用性,能够平衡安全性、易用性和能量消耗.

参考文献

[1] Threatpost: Samsung android lockscreen bypass[EB/OL]. http://threatpost. com/lock-screen-bypass-flaw-found-samsung-androids-030413/77580.

[2] Hayashi E, Riva O, Strauss K, Bernheim Brush A J, Schechter S. Goldilocks and the two mobile devices: Going beyond all-or-nothing access to a device′s applications [A]. Proceedings of the Symposium on Usable Privacy and Security[C]. New York: ACM, 2012. 1 – 11.

[3] Riva O, Qin C, Strauss K, Lymberopoulos D. Progressive authentication: deciding when to authenticate on mobile phones[A]. Proceedings of the 21 st USENIX Security Symposium[C]. Berkeley: USENIX Association, 2012. 301 – 316.

[4] Sasse M A, Steves M, Krol K, Chisnell D. The great au-

thentication fatigue-and how to overcome it[A]. Proceedings of 6th International Conference on Cross-Cultural Design, Held as Part of HCI International 2014[C]. Berlin: Springer, 2014. 228 – 239.

[5] Shi E, Niu Y, Jakobsson M, Chow R. Implicit authentication through learning user behavior[A]. Proceedings of the 13th international conference on Information security Information Security[C]. Berlin: Springer, 2011. 99 – 113.

[6] Negara A F P, Yeom J, Choi D. A study on multibiometrics derived from calling activity context using smartphone for implicit user authentication system[J]. International Journal of contents, 2013, 9(2): 14 – 21.

[7] Feng T, Yang J, Yan Z, Tapia E M, Shi W. TIPS: Context-aware implicit user identification using touch screen in uncontrolled environments[A]. Proceedings of the 15th Workshop on Mobile Computing Systems and Applications[C]. New York: ACM, 2014. 91 – 96.

[8] Frank M, Biedert R, Ma E, Martinovic I, Song D. Touchalytics: On the applicability of touchscreen input as a behavioral biometric for continuous authentication[J]. IEEE Transactions on Information Forensics and Security, 2013, 8(1): 136 – 148.

[9] Zhu J, Wu P, Wang X, Zhang J. SenSec: Mobile security through passive sensing[A]. Proceedings of 2013 International Conference on Computing, Networking and Communications[C]. Washington, DC: IEEE Computer Society, 2013. 1128 – 1133.

[10] Bo C, Zhang L, Jung T, Han J, Li X Y, Wang Y. Continuous user identification via touch and movement behavioral biometrics[A]. Proceedings of the IEEE International Performance Computing and Communications Conference[C]. Washington, DC: IEEE Computer Society, 2014. 1 – 8.

[11] Youn I H, Choi S, May R Le, Bertelsen D, Youn J H. New gait metrics for biometric authentication using a 3-axis acceleration[A]. Proceedings of the IEEE 11th Consumer Communications and Networking Conference[C]. Los Alamitos: IEEE, 2014. 596 – 601.

[12] Khan H, Atwater A, Hengartner U. A comparative evaluation of implicit authentication schemes[A]. Proceedings of the 17th International Symposium on Research in Attacks, Intrusions and Defenses[C]. Berlin: Springer, 2014. 255 – 275.

[13] Khan H, Atwater A, Hengartner U. Itus: An implicit authentication framework for android[A]. Proceedings of the 20th ACM Annual International Conference on Mobile Computing and Networking[C]. New York: ACM, 2014. 507 – 518.

[14] Frank J, Mannor S, Precup D. Activity and gait recognition with time-delay embeddings[A]. Proceedings of the National Conference on Artificial Intelligence[C]. Palo Alto, California: American Association for Artificial Intelligence, 2010. 1581 – 1586.

[15] Lu H, Bernheim Brush A J, Priyantha B, Karlson A K, Liu J. SpeakerSense: Energy efficient unobtrusive speaker identification on mobile phones[A]. Proceedings of the 9th International Conference on Pervasive Computing[C]. Berlin: Springer, 2011. 188 – 205.

[16] Cortes C, Vapnik V. Support-vector networks[J]. Mach. Learn, 1995, 20(3): 273 – 297.

作者简介

刘礼才　男,1986 年出生,江西赣州人,博士,现就职于中国网络空间研究院,主要研究方向为网络安全、隐私保护、物联网安全.
E-mail: liulc_r@163.com

李锐光　男,1979 年出生,山西阳泉人,硕士,国家计算机网络应急技术处理协调中心高级工程师,主要研究方向为网络与信息安全.
E-mail: lrg@cert.org.cn

殷丽华(通讯作者)　女,1973 年出生,黑龙江哈尔滨人,博士后,硕士生导师,中国科学院信息工程研究所副研究员,主要研究方向为信息安全、物联网安全.
E-mail: yinlh_iie@163.com