软件学报 ISSN 1000-9825, CODEN RUXUEW
Journal of Software,2021,32(6):1733−1747 [doi: 10.13328/j.cnki.jos.006247]

E-mail: jos@iscas.ac.cn
http://www.jos.org.cn
Tel: +86-10-62562563

个体交互行为的平滑干预模型*

刘 霄[1], 章昭辉[1,2], 魏子明[1], 王鹏伟[1]

[1](东华大学 计算机科学与技术学院,上海 201620)

[2](上海网络信息服务工程技术研究中心,上海 201804)

通讯作者: 章昭辉, E-mail: zhzhang@dhu.edu.cn

摘 要: 基于交互行为的用户特征提取和身份认证方法是一种重要的身份识别方式,但高频用户的交互行为模式和操作习惯相对稳定,易被欺诈者模仿,使得现有模型对此类欺诈行为的误判较高.如何使得用户行为主动平滑变化且可区分,成为解决上述问题的关键.针对此问题,提出一种基于个体交互行为系统平滑干预模型:首先,根据用户历史交互行为日志从多个维度得到用户的交互行为变化趋势;然后,结合行为的稳定性和偏向性提出行为时域漂移算法(TDDA),为每个用户确定行为引导的时机;最后,基于 Petri 网提出交互行为重构系统干预模型,在系统中的非关键路径叠加行为触发因素,引导用户产生新的交互行为习惯.实验证明了提出的方法能够很好地引导用户行为平滑变化,且产生足够的区分性,使得行为伪装异常检测场景下模型的准确性显著提高.

关键词: 行为干预;交互行为;身份伪装;身份识别;Petri 网建模

中图法分类号: TP309

中文引用格式: 刘霄,章昭辉,魏子明,王鹏伟.个体交互行为的平滑干预模型.软件学报,2021,32(6):1733−1747. http://www.jos.org.cn/1000-9825/6247.htm

英文引用格式: Liu X, Zhang ZH, Wei ZM, Wang PW. Smooth intervention model of individual interaction behavior. Ruan Jian Xue Bao/Journal of Software, 2021,32(6):1733−1747 (in Chinese). http://www.jos.org.cn/1000-9825/6247.htm

Smooth Intervention Model of Individual Interaction Behavior

LIU Xiao[1], ZHANG Zhao-Hui[1,2], WEI Zi-Ming[1], WANG Peng-Wei[1]

[1](School of Computer Science and Technology, Donghua University, Shanghai 201620, China)

[2](Shanghai Engineering Research Center of Network Information Services, Shanghai 201804, China)

Abstract: User feature extraction and identity authentication methods based on interactive behavior are an important method of identity recognition. However, for high-frequency users, the interactive behavior patterns and operating habits are relatively stable, which are easily imitated by fraudsters and making the existing models have a higher misjudgment. The key point to solve the above problems is to make the user's behavior change smoothly and distinguishably. This study proposes a smooth intervention model based on an individual interactive behavior system to handle it. Firstly, according to the user history web behavior log, the change trend of user interaction behavior is obtained from multiple dimensions. Then, combined with the stability and deviation of the behavior, the time-domain drift algorithm (TDDA) is proposed to determine the behavior guidance time of each user. Finally, an intervention model for interactive behavior reconstruction systems is proposed, which superimposes behavior trigger factors on non-critical paths in the system to guide users generating new interactive behavior habits. Experiments prove that the method proposed in this study could guide the user's

* 基金项目: 上海市自然科学基金(19ZR1401900); 上海市科技创新行动计划(19511101302); 国家自然科学基金(61472004, 61602109)

Foundation item: Natural Science Foundation of Shanghai (19ZR1401900); Shanghai Science and Technology Innovation Action Plan High-Tech Field Project (19511101302); National Natural Science Foundation of China (61472004, 61602109)

本文由“形式化方法与应用”专题特约编辑田聪教授推荐.

收稿时间: 2020-08-30; 修改时间: 2020-10-26; 采用时间: 2020-12-19; jos 在线出版时间: 2021-02-07

behavior to change smoothly and produce sufficient distinction to significantly advance the model accuracy in the scenario of behavior camouflage anomaly detection.

Key words: behavioral intervention; interactive behavior; identity theft; identity recognition; Petri-net modeling

电子商务的迅速发展,使用户在享受互联网带来的高质量生活品质的同时,也在不断遭遇多种有组织的网络欺诈行为.根据 CNNIC 数据统计,在 2018 年中,30%以上的网民遭遇了个人信息泄漏,超过 25%的用户遭遇网上诈骗,23.8%的用户遭遇了病毒或木马攻击,19.2%的用户账号或密码被盗.对互联网行业的企业而言,灰色产业带来的不仅是对正常业务的干扰,更是真实的经济损失.对于很多初创互联网企业,在开拓市场之初,纷纷推出了花样繁多的优惠活动,这些优惠在吸引了众多用户改变使用习惯的同时,也成为了灰色产业从业人员的目标,大量不法分子利用[1]作弊软件与作弊硬件,通过虚假身份和模拟的虚假行为欺诈套利.同时,认知一致性理论(cognitive consistency)认为:用户的态度和行为总趋于平衡和稳定状态[2],用户的共性特征往往稳定不变且容易被欺诈者捕获,从而使得欺诈者能够绕过监控规则,获利丰厚,甚至渐渐形成了整套的灰色产业链.而这些产业因为自身的隐藏性与反侦察性,并不为社会公众所感知.如何有效且准确地验证用户身份,已经成为一个待解决的问题.

现有的大部分身份认证技术都是基于用户的账户名和密码[3].在短时间内对用户进行身份认证,之后无论用户的真实身份是什么,用户所做的一切行为都将被视为合法行为[4].然而,欺诈者通过钓鱼网站伪装成银行等金融服务提供商,骗取用户的电子邮件地址、密码等敏感信息,得以伪装身份进而实施诈骗行为.所以,拥有正确密码的用户并不意味着他是一个合法的用户.

为了弥补单一的用户名密码的身份认证模式带来的缺陷,近年来,许多学者也倾向于数据特征挖掘和行为分析方法用于身份识别领域,如对用户 Web 日志采用关联规则挖掘、隐马尔可夫过程、半马尔可夫过程、贝叶斯网络、神经网络和随机森林等方法进行行为建模和预测.尽管目前很努力地解决用户身份识别问题,但是依然面临着诸多困难.

(1) 利用机器学习相关模型,需要对训练数据进行标记.然而,由于现实中欺诈行为样本的稀疏性,会存在样本极度不均衡的情况,使得模型很难充分学习欺诈行为的特征;

(2) 由于用户的年龄、背景和爱好等相对固定,其系统交互行为模式在一定时间内相对稳定.欺诈者利用这一特性,使用盗取的部分用户行为信息模拟正常使用者的行为,绕过监控规则,从而使得身份识别模型对此类欺诈行为的误判较高;

(3) 目前,对用户交互行为和浏览行为的 Web 日志数据挖掘研究大多建立在用户对于搜索引擎、网上商城等具有大量页面浏览记录中提取特征,得到用户的浏览偏好,且研究多用于对网站结构的优化、Web 预测或推荐系统.但在功能较单一的单个系统中,在用户量较大时出现的多用户共享相似行为模式[5],使得身份识别模型的区分性降低;

(4) 如何提取用户交互行为特征构建用户行为画像和用户交互行为异常判断的标准,目前仍然面临诸多挑战.

针对以上问题,本文提出一种新的身份识别方法.与其他模型相比,本文提出的方法能够很好地缓解上述问题,并且主要有以下几点优势.

(1) 从用户的角度出发,持续关注用户的 Web 使用日志,构建用户的交互行为画像,能够很好地规避样本不均衡这一问题;

(2) 考虑了用户之间的差异,提出了行为漂移引导模型,根据每个用户的历史交互行为记录,综合考虑交互行为的稳定性和偏向性,为每个用户确定各自的行为干预时机,有效避免了复杂性不高且功能单一的系统中出现多用户共享相似的行为模式问题;

(3) 提出了系统行为集合的定义,并划分为系统关键行为集合和非关键行为集合,采用系统内部触发的方法,通过在不破坏系统运行逻辑的前提下叠加新的非关键行为流程的方式,非强制性约束用户行为流程,使得用户行为能够顺应引导机制,从而逐步培养用户产生新的交互行为习惯,与原始交互行为保

持一定的行为差异,以对抗已存在的身份伪装和欺诈行为.

本文的研究是基于一种常见的归纳偏差,即行为稳定性偏差.这种偏差表现在 3 个方面:首先,短时间内用户的交互行为不会发生明显的变化;其次,当一个常用账号被行为伪装或身份冒用后,欺诈者的系统交互行为与合法账号持有者保持较高的相似性;最后,由于欺诈行为的稀疏性,持续的行为引导机制将改变正常用户的交互行为习惯,而身份伪装的欺诈者仍保持原始场景的交互行为.

综上所述,本文的主要贡献如下:一是提出了一种刻画用户交互行为方法,从系统登录时间、登录时间间隔等多个维度考虑用户的系统交互行为特征;二是提出了行为漂移引导模型,为每个用户确定各自的行为干预时机和手段,使得前后行为保持一定的差异性和平滑性;三是设计实验验证了该模型在行为伪装异常检测中的可行性和准确性.本文第 1 节详细介绍相关工作.第 2 节讨论本文提出的模型方法.第 3 节介绍数据来源和本文的实验结果以及对比实验.第 4 节总结本文研究成果以及对未来的展望.

1 相关工作

近年来,基于用户交互行为进行异常检测也逐渐得到了重视,现有研究主要分为用户对计算机外设的操作行为建模和对系统交互浏览日志进行行为建模两种研究方向.在计算机外设操作行为分析方面,Roth 等人[6]提出一种通过连续打字对用户进行身份认证的方法.Ma 等人[7]提出一种基于动态软键盘上的鼠标行为认证方法.在用户与系统交互产生的浏览日志建模等方面,Liu 等人[8]针对开放网络环境下的身份认证和行为认证分离问题,使用系统浏览时间和导航路径来确定用户行为的正常性和行为访问控制方法.Zhao 等人[9]使用行为马尔可夫模型描述用户的逻辑行为,提出了基于序列和首选项的身份验证方法.Zheng 等人[10]定义了基于信息熵的多样性系数和 BP(LGBP)逻辑图,用以表征用户交易行为的多样性;同时提出一种新的[11]基于行为证书(BC)的信用卡欺诈检测系统(FDS),利用用户的行为证书对用户交易进行识别.Chen 等人[12]设计了一种综合多种因素的移动终端 APP 浏览行为认证系统架构,该架构适合于日常生活中使用移动终端 APP 的用户.Zhong 等人[13]提出一种基于 Web 浏览行为认证的方法.Zhang 等人[14]提出一种基于关联规则挖掘的 Web 浏览行为取证方法.Liu 等人[15,16]提出了安全互模拟的概念,首次采用形式化的方法,利用二值化理论区分具有相同功能但不同安全性的两个系统.Zhang 等人[17]通过 DBSCAN 聚类算法,通过迁移当前交易组的行为和交易状态的方法扩充低频用户的数据记录,使低频用户的行为刻画更加准确.

对用户交互行为进行合理的引导和干预,促使用户改变其使用习惯也开展了一定的研究.Toledo 等人[18]基于 fogg 行为模型,结合模糊规则和模糊推理建立了一套说服式系统建模框架,用于分析用户触发学习活动的动机和能力水平.Hamper[19]和 Nishiyama[20]等人针对用户的 TTM 和 FBM 特征设计了基于劝导模型的应用程序的方法及理论结构,并结合激励机制完成了促进用户进行体育健康锻炼应用程序理论建模;Zhang 等人[21]尝试通过视觉和听觉刺激来诱发行为变化,基于行为更改支持系统(BCSS)理论,通过分析当前的环境状态,自适应变更系统中通知图标的样式.尽管对用户交互行为的引导和干预在交互设计的某些领域有了部分应用与输出,但尚未形成相对完整的理论体系和实践方法,且大多研究停留在分析和理论完善的阶段,缺少对行为变化前后差异性和平滑性的定性定量分析和效果验证.

2 模型方法

本节将会介绍一种个体交互行为干预模型,该模型主要包括两个部分:第 1 部分是用户交互行为画像的生成,第 2 部分是行为漂移引导模型的建立.行为漂移引导模型分为行为时域漂移算法和交互行为重构模型.用户交互行为画像的生成会根据每一个用户历史正常系统交互行为,从多维度为用户生成交互行为画像;行为漂移引导模型首先基于用户历史交互行为画像确定行为干预的时间,其次根据得到的干预时间,通过系统内部触发的方式,通过叠加新的行为路径改变系统行为流程来达到非强制性约束用户行为的目的.

2.1 用户交互行为画像的生成

本文假设用户的个人访问记录可以被服务器记录和检索,并可以进行连续监控生成用户的历史 Web 交互

行为日志.本节将介绍用户交互行为基准的生成过程,根据用户历史 Web 交互日志数据生成用户交互行为画像.

定义 1(交互行为记录). 用户在系统中的一条交互行为记录 i 包含 m 个属性,记作:

$$i=\{a_1,a_2,a_3,\ldots,a_m|a_1\in A_1,a_2\in A_2,a_3\in A_3,\ldots,a_m\in A_m\}.$$

本文中,交互记录属性包括用户标号、会话编号、登录时间、页面编号、进入页面时间、页面持续时间,即交互行为 $i=\{a_{id},a_{ses_{no}},a_{time},a_{page\_no},a_{start},a_{duration}\}$.即:给定一个用户标号 u,则该用户的交互行为日志是其截至当前日期的历史交互记录集合,记为 $R_u=\{i_1^u,i_2^u,i_3^u,...,i_n^u\}$,其中,$n$ 是该用户的交互记录条数,即 $n=|R_u|$.用户的交互行为记录日志中的正常行为即 $T_u=\{t\in R_u|label=\text{true}\}$,其中,$n_{ut}=|T_u|$.对于用户的正常交互行为,我们需要进一步分析处理得到用户的交互行为画像.将用户的交互行为画像的各个属性定义如下:

$$\begin{cases}A_1^u=\{a\in A_1\mid\exists r\in r_u:a\in r\}\\A_2^u=\{a\in A_2\mid\exists r\in r_u:a\in r\}\\\cdots\\A_m^u=\{a\in A_m\mid\exists r\in r_u:a\in r\}\end{cases},$$

其中, $A_1^u\subseteq A_1,A_2^u\subseteq A_2,...,A_m^u\subseteq A_m$.不失一般性,我们定义 $A_1^u=\{a_1^i,a_1^i,...,a_1^i\}$.

定义 2(系统登录时间属性). 用户 u 的系统登录时间属性定义为 n 个时间段内用户登录概率的 n 元组,记为 $LTA^u=(time_1,time_2,\ldots,time_n)$.在用户 u 的正常交互行为日志 R_u 中,取出 A_{time}^u 作为该用户的登录系统的时间集合,且 $n_{time}^u=|A_{time}^u|$.为了区分不同用户的登录时间偏好和习惯,我们将登录时间划分为 12 个区间,计算 A_{time}^u 中的每个元素 a_i^{time} 对应的时间区间,并为各个元素打上标签,得到以下子集:

$$lta_1=\{a_i^{time}\in A_{time}^u\mid 0\leqslant\log intime<2\},$$
$$lta_2=\{a_i^{time}\in A_{time}^u\mid 2\leqslant\log intime<4\},$$
$$lta_3=\{a_i^{time}\in A_{time}^u\mid 4\leqslant\log intime<6\},$$
$$...,$$
$$lta_{12}=\{a_i^{time}\in A_{time}^u\mid 22\leqslant\log intime<24\}.$$

以此求出 $time_1=\frac{|lta_1|}{n_{time}^u},time_2=\frac{|lta_2|}{n_{time}^u},...,time_{12}=\frac{|lta_{12}|}{n_{time}^u}$,从而得到该用户的登录时间属性:

$$LTA^u=(time_1,time_2,time_3,\ldots,time_{12}).$$

定义 3(工作时间登录属性). 用户 u 的工作时间登录属性定义为 $WTA^u=(isworktime,noworktime)$,表示该用户的登录时间是否发生在工作日的工作时间的概率,其中,工作日不包含双休日和法定节假日.根据集合 A_{time}^u 中的每一个元素,判断其是否属于工作时间.依据判断结果为每个元素打上 T 和 F 的标签,代表工作时间登录和非工作时间登录,从而得到两个子集如下:

$$wta_1=\{a\in A_{time}^u\mid label=T\},$$
$$wta_2=\{a\in A_{time}^u\mid label=F\}.$$

因此可以得出 $isworktime=\frac{|wta_1|}{n_{time}^u},noworktime=\frac{|wta_2|}{n_{time}^u}$,从而得到该用户的工作时间登录属性:

$$WTA^u=(isworktime,noworktime).$$

定义 4(登录间隔属性). 用户 u 的登录间隔属性定义为 $LIA^u=(period_1,period_2,\ldots,period_n)$,表示该用户的登录时间间隔发生在各区间的概率,反映用户登录系统的交互行为习惯.根据集合 A_{time}^u 中的每一个元素,依次计算其登录时间间隔,得到集合 A_{period}^u,其中, $n_{period}^u=|A_{period}^u|$.对于集合 A_{period}^u 中各元素的计算公式如下:

$$a_i^{period}=a_i^{time}-a_{i-1}^{time}.$$

求出集合 A_{period}^u 的第一四分位数 Q_1、第二四分位数 Q_2、第三四分位数 Q_3 和上限 $Q_{\max}$、下限 $Q_{\min}$,将集合分为 5 个子集,即:

$$lia_1=\{a^u_{period}\in A^u_{period}\,|\,Q_{\min}\leqslant a^u_{period}<Q_1\},$$
$$lia_2=\{a^u_{period}\in A^u_{period}\,|\,Q_1\leqslant a^u_{period}<Q_2\},$$
$$lia_3=\{a^u_{period}\in A^u_{period}\,|\,Q_2\leqslant a^u_{period}<Q_3\},$$
$$lia_4=\{a^u_{period}\in A^u_{period}\,|\,Q_3\leqslant a^u_{period}<Q_{\max}\},$$
$$lia_5=\{a^u_{period}\in A^u_{period}\,|\,a^u_{period}\leqslant Q_{\min},a^u_{period}\geqslant Q_{\max}\}.$$

因此可以得出 $period_1=\frac{|lia_1|}{n^u_{period}},period_2=\frac{|lia_2|}{n^u_{period}},...,period_5=\frac{|lia_5|}{n^u_{period}}$.即,该用户登录间隔属性:

$$LIA^u=(period_1,period_2,period_3,period_4,period_5).$$

定义 5(关键页面停留时间属性). 用户 u 的关键页面停留时间属性定义为 $KSA^u=(distance_1,distance_2,\ldots,distance_n)$,表示用户在行为引导触发因素对应的载体页面的停留时间长短的概率.在用户 u 的正常交互行为日志 R_u 中,取出该用户历史关键页面 $a_{page\_no}=key$ 的停留时间集合 $A^u_{distance}$,其中,$n^u_{distance}=|A^u_{distance}|$.为了区分用户在关键页面交互时间的偏好和习惯,根据集合 $A^u_{distance}$ 中的所有元素,求出集合第一四分位数 Q_1、第二四分位数 Q_2、第三四分位数 Q_3 和上限 $Q_{\max}$、下限 $Q_{\min}$,将集合分为 5 个子集,即:

$$ksa_1=\{a^u_{distance}\in A^u_{distance}\,|\,Q_{\min}\leqslant a^u_{distance}<Q_1\},$$
$$ksa_2=\{a^u_{distance}\in A^u_{distance}\,|\,Q_1\leqslant a^u_{distance}<Q_2\},$$
$$ksa_3=\{a^u_{distance}\in A^u_{distance}\,|\,Q_2\leqslant a^u_{distance}<Q_3\},$$
$$ksa_4=\{a^u_{distance}\in A^u_{distance}\,|\,Q_3\leqslant a^u_{distance}<Q_{\max}\},$$
$$ksa_5=\{a^u_{distance}\in A^u_{distance}\,|\,a^u_{distance}\leqslant Q_{\min},a^u_{distance}\geqslant Q_{\max}\}.$$

因此可以得出 $distance_1=\frac{|ksa_1|}{n^u_{distance}},distance_2=\frac{|ksa_2|}{n^u_{distance}},...,distance_5=\frac{|ksa_5|}{n^u_{distance}}$.即,该用户登录间隔属性为

$$KSA^u=(distance_1,distance_2,distance_3,distance_4,distance_5).$$

定义 6(交互行为特征). 令 $IBC^u=(LTA^u,WTA^u,LIA^u,KSA^u)$为用户 u 的交互行为特征,根据登录时间、交互行为是否发生在工作时间、登录时间的时间间隔、系统关键页面停留时间等属性,将用户的交互行为定义为一个 24 维的特征向量 IBC^u 来描述用户的交互行为,其中,

(1) $LTA^u=(time_1,time_2,\ldots,time_{12})$表示登录时间属性,其中,$time_k(k=1\sim12)$分别表示用户在各时间段发生登录行为的概率;

(2) $WTA^u=(isworktime,noworktime)$表示工作时间登录属性,其中,$isworktime$ 和 $noworktime$ 分别表示用户在工作日和非工作日产生交互行为的概率;

(3) $LIA^u=(period_1,period_2,\ldots,period_n)$表示登录间隔属性,其中,$period_k(k=1\sim5)$表示用户与上一次登录行为的间隔时间在 lia_k 内的概率;

(4) $KSA^u=(distance_1,distance_2,distance_3,distance_4,distance_5)$表示页面停留时间属性,其中,$distance_k(k=1\sim5)$表示用户在系统关键页面的停留时间在 ksa_k 内的概率.

2.2 行为漂移引导模型的建立

一般来说,每个人都有相对不变的行为习惯,这由人的性格、年龄、职业等决定的,如内向的人在暴露信息时比较谨慎,老年人操作比较慢,计算机专业的人操作大多比较快,等等.这些共性的特征也往往容易被欺诈者捕获.同时,每个人的行为习惯事实上也是可以相对改变的,但这种改变大多来自于外界施加的条件.在多数交易系统中,由于系统行为只注重业务逻辑和业务功能,交易数据所蕴含的用户行为是通过交易系统行为实现的.在系统不对用户主动干预下,数据所蕴含的行为是一般用户自身形成的行为.而一个用户施加于系统的行为通常具有一定的不变性.交易频率越高的用户,其行为也就比较稳定.因此,当用户数据被盗取并被欺诈者模拟行为后,那么原有的合法用户行为模型将无法区分欺诈行为.另外,如果用户交易频率过低,相当于无用户行为,这种情况使得欺诈者更易模拟用户行为.因此,为了使得用户行为能主动对抗欺诈行为并保持良好的交互体验,需

要对原有的用户行为进行平滑漂移.

本节提出交互行为集中性系数 *CS* 和交互行为偏向性系数 *CP* 对用户行为进行量化,并基于 *CS* 和 *CP* 提出行为时域漂移算法 TDDA(time-domain drift algorithm)确定触发因素的出现时机,并实现了 TDDA 作用下的交互行为重构访问控制模型.

2.2.1 行为时域漂移算法

为了使得漂移后的引导时域与用户原始登录时域存在差异性,且充分考虑变化的平滑性,我们对用户的历史交互行为记录进行分析,采用分位数分析方法和四分位距(IQR)来描述不同用户登录时间序列的离散程度.分位数是将总体的全部数据按从大到小顺序排列后,处于各等分位置的变量值.求出用户登录时间记录的第一四分位数 Q_1、第二四分位数 Q_2 即中位数、第三四分位数 Q_3,并求出上限 $Q_{\max}$、下限 $Q_{\min}$.

四分位数间距即第三四分位数 Q_3 与第一四分位数 Q_1 之差,即 $IQR=Q_3-Q_1$,其数值越大,反映变异度越大;反之,变异度越小.为了减少极端值对用户行为稳定性衡量的影响,用户行为的上限、下限计算公式如下:

$$Q_{\max}=Q_3+\alpha IQR,$$

$$Q_{\min}=Q_1-\alpha IQR,$$

其中,α为异常程度的权重.若α越大,则更多偏离点会被接受;反之,α越小,则更多的偏离点会被排除.

定义 7(交互行为集中性系数). 用户 u 的交互行为集中性系数定义为 CS^u,表示用户行为的稳定和聚集程度,为样本中部的观察值 Q_1 与 Q_3 的极差与样本上下限极差的比值.计算公式如下:

$$CS^u=\frac{IQR}{Q_{\max}-Q_{\min}}=\frac{R\left[\frac{3(n+1)}{4}\right]-R\left[\frac{n+1}{4}\right]}{Q_3+\alpha IQR-(Q_1-\alpha IQR)},$$

其中,R 是对于总体排序后的集合,n 是集合中元素的个数.因此,依据 *CS* 值的大小,我们可以衡量每个用户交互行为事件的离散程度.以用户的登录时间属性为例,其登录时间属性下对应的 *CS* 值越大,反映用户的历史登录时间较为集中,行为离散程度较小;反之,对应的 *CS* 值越小,用户的登录时间较为分散,行为离散程度较高.在此基础上,不同时间区间内用户的 *CS* 值的变化情况也反映了用户交互行为的集中性趋势.在衡量用户稳定和聚集程度的基础上,还需要考虑用户交互行为的偏向性.

定义 8(交互行为偏向性系数). 用户 u 的交互行为偏向性系数定义为 CP^u,表示用户行为偏好和偏向程度,为样本均值与第二四分位数的差值.计算公式如下:

$$CP=\bar{R}-Q_2=\frac{1}{n}\sum_{i=1}^{n}r_i-Q_2,$$

其中,$\bar{R}$ 为总体集合的均值,r 为集合 R 中的全部元素.以用户登录时间为例,不同用户登录系统的时间与自身习惯和工作性质等有关:若用户登录系统的时间更偏向于均值左侧,即 $CP<0$;相反,用户登录系统的时间更偏向于均值右侧,即 $CP>0$.不同时间区间内 CP^u 值的变化情况,也体现了用户交互行为的偏向性趋势.

用户行为的引导需要充分结合用户自身的行为能力和习惯,不给用户体验带来较大负担,用户才会有更高的接受度.本文结合用户的历史交互行为习惯,在交互行为集中性系数 CS^u 和交互行为的偏向性系数 CP^u 的基础上,充分考虑用户历史行为,提出交互行为时域漂移算法(TDDA).算法 1 给出了用户登录行为时域漂移算法的伪代码,如下所示:

算法 1. Time domain drift algorithm on login behavior.

Input: *R*—The list of user login time log; *n*—The size of user login time log;

Output: *drift_start*—The guide mechanism start time; *drift_stop*—The guide mechanism end time.

1: *threshold*:=0.4;

2: *sort*(*R*);

3: calculate Q_1, Q_2, Q_3, $\bar{R}$;

4: $IQR=Q_3-Q_1$

5: $CS=IQR/(Q_{\max}-Q_{\min})$

6: $CP=\bar{R}-Q_2$

7: **if** *CS*>*threshold* **then**

8: **if** *CP*>0 **then**

9: *drift_start*= $\bar{R}$; *drift_stop*=Q_3;

10: **else**

11: *drift_start*=Q_1; *drift_stop*= $\bar{R}$;

12: **else**

13: *drift_start*=Q_1; *drift_stop*=Q_3

14: **return** *drift_start*, *drift_stop*;

2.2.2 交互行为重构模型

本节使用 Petri 网作为系统建模工具,首先将系统行为集合分为系统关键行为集合和非关键行为集合,定义了交互行为 Petri 网并给出了交互行为 Petri 网行为轮廓的定义;在此基础上提出交互行为重构的系统 Petri 网模型,并在在线信贷交易系统实例中验证模型的有效性.

Petri 网作为并发、分布式系统的建模和分析工具,对系统的性质和行为分析具有强大的理论基础支持[22],在业务流程建模分析和优化方面也有广泛的应用.在计算机软件系统领域,可以利用 Petri 网进行[23,24]完整的建模与相关结构及性质分析.同时,由于软件系统必然存在着与用户的系统交互,而用户与系统平台的交互往往反映出用户对平台提供服务的兴趣、关注程度以及交互行为习惯,因此,用户访问行为分析无疑对评价和优化平台的业务流程、服务设置等有直接且重要的价值.

定义 9(系统行为集合). 令 $S_A=\{s_1,s_2,\ldots,s_n\}$为系统正常运行期间能够触发的行为事件的全部集合.进一步将系统行为集合分为系统关键行为集合 S_A^* 和非关键行为集合 S_A' .

定义 10(系统关键行为集合). 令 $S_A^*=\{s_1,s_2,...,s_p\}$ 为系统正常运行期间能够触发的关键行为事件的全部集合,其中,$p<n$.关键行为集合对应着系统的核心功能页面,承担系统关键功能的运行.关键行为流程 cp^*即 S_A^* 集合中元素的特定排列,由于关键行为流程反映着系统核心功能的运行逻辑,因此具备一定的业务逻辑顺序,即:

$$|cp^*|<p!$$

定义 11(系统非关键行为集合). 令 $S_A'=\{s_1,s_2,...,s_q\}$ 为系统正常运行期间能够触发的非关键行为事件的全部集合,其中,$q<n$.非关键行为集合对应着系统的次要功能页面,承担对于系统关键功能的补充作用.非关键行为流程 cp'即 S_A' 集合中元素的特定排列,即$|cp'|<q!$.

关键行为集合对应着系统的核心功能页面,如在线信贷业务系统中的提交借款申请、信贷信息核验、借款协议签署等关键功能页面;非关键行为集合往往包含敏感度较低的其他系统功能业务页面,如银行卡信息页、信贷业务浏览、帮助中心、个人中心等.如图 1 所示,用户行为集合 $U_A=\{u_1,u_2,\ldots,u_m\}$,即用户能发生的全部行为事件的集合,且 $U_A\subseteq S_A$,用户的操作流程即 U_A 集合中所有元素的全排列,用户行为序列有 $m!$种.图中 *A* 表示系统关键行为集合 S_A^*,*B* 表示系统非关键行为集合 S_A',*C* 表示用户行为集合 *Ru*,*A*+*B* 表示系统行为集合 S_A.

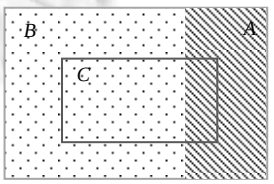


Fig.1 System behavior set

图 1 系统行为集合划分

触发因素是促使用户做出某行为的诱导因素,可分为外部触发和内部触发:外部触发往往由用户所处外部环境所决定;而内部触发则嵌入于产品和系统,是引起行为变化的关键.内部触发以友好的交互方式将下一步行

动清楚地传达给用户,经由用户频繁地将触发因素与系统使用行为相联系,使其发展出新的交互行为习惯.因此,我们通过内部触发的方式,通过叠加新的非关键行为流程改变系统行为流程来非强制性约束用户行为流程.

根据用户交互行为日志的数据特点及 Petri 网的定义,定义系统行为 Petri 网和用户交互行为 Petri 网如下:

定义 12(用户交互行为 Petri 网). 设某平台在线系统的系统行为 Petri 网 $PN=(S;T;F)$,则用户在该系统的交互行为 Petri 网 $IPN=(IS;IT;IF)$,其中,

(1) $IS\subseteq S$,为用户执行相关的输入输出对应的库所元素;

(2) $IT\subseteq T$,为用户执行交互行为对应的变迁元素;

(3) $S\cap T=\varnothing$;

(4) $IF\subseteq F$,为变迁元素之间在系统 Petri 网 PN 中的流关系,即 $IF\subseteq(IS\times IT)\cup(IT\times IS)$.

定义 13(系统行为轮廓). 令系统行为 Petri 网 $PN=(S;T;F)$,集合 $MB_l=\{\Rightarrow,\ominus,\circledcirc\}$是 Petri 网 $PN=(S;T;F)$的行为轮廓.对任给的变迁对$(t_1,t_2)\in(T\times T)$满足如下关系之一:

(1) 顺序关系$\Rightarrow$:若 $\tau(t_1,t_2)=\{\Rightarrow\}$,则 $t_1\succ t_2$ 且 $t_2\not\succ t_1$;

(2) 平行关系$\ominus$:若 $\tau(t_1,t_2)=\{\ominus\}$,则 $t_1\not\succ t_2$ 且 $t_2\not\succ t_1$;

(3) 循环关系$\circledcirc$:若 $\tau(t_1,t_2)=\{\circledcirc\}$,则 $t_1\succ t_2$ 且 $t_2\succ t_1$.

如图 2 所示,系统行为 Petri 网中,T_2 与 T_4 为平行关系记作 $T_2\ominus T_4$.T_6,T_7,T_8 是依次发生的,具有严格的顺序关系,记作 $T_6\Rightarrow T_7\Rightarrow T_8$.$T_9$ 与 T_{10} 为循环关系,记为 $T_9\circledcirc T_{10}$.基于系统的 3 种行为轮廓,我们给出了不同行为轮廓下的交互行为重构方法,如图 3~图 5 所示.

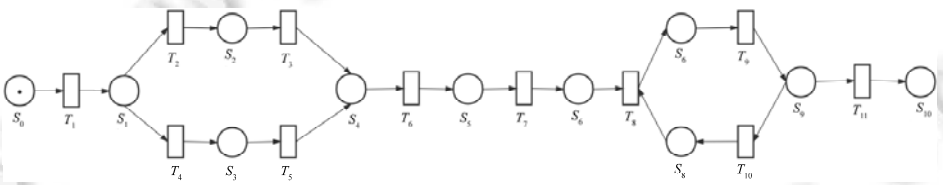


Fig.2 Example of system behavior Petri net

图 2 系统行为 Petri 网示例

Fig.3 Reconstruction of sequence relationship

图 3 顺序关系重构

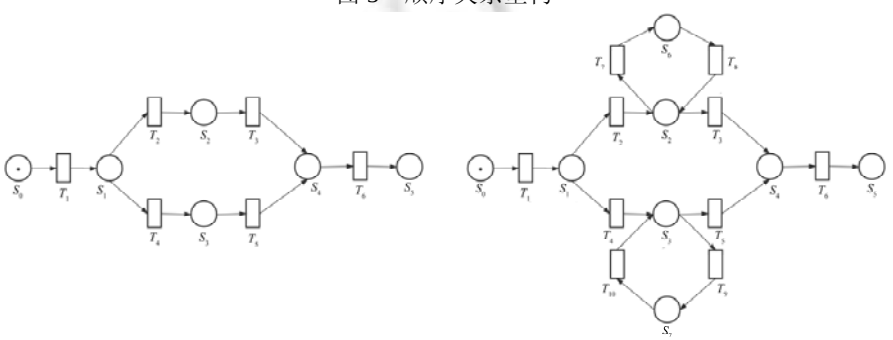


Fig.4 Reconstruction of parallel relationship

图 4 平行关系重构

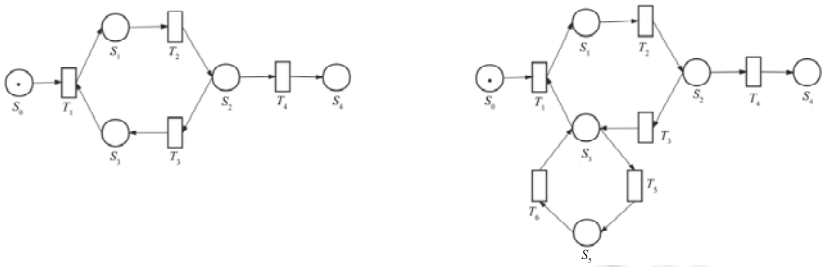


Fig.5 Reconstruction of circular relationship

图 5 循环关系重构

建立交互行为重构模型,是为了使得合法用户行为能够平滑变化主动对抗欺诈行为,并且不破坏系统原有的交互逻辑和使用体验.因此,我们在基本交互行为的轮廓中叠加新的非关键行为循环结构,这样既保持系统的基本业务逻辑和流程不受影响,又形成了新的系统行为集合.由于欺诈者在操作系统时的目的十分明确,为了尽快取得非法收益,往往在系统中的关键行为集合中的页面保持较高的优先级,而对于非关键行为集合保持较低的兴趣;而正常用户的操作行为则在两个集合间保持均衡的优先级.所以采取上述的交互行为重构方法可以使得正常用户的交互行为习惯产生变化,在持续的行为重构方法引导下产生新的交互行为习惯,通过自身行为的变化逐渐与欺诈者的欺诈行为产生对抗.

本文使用的系统为实验室搭建的在线信贷交易系统,其系统业务流程 Petri 网模型如图 6 所示.

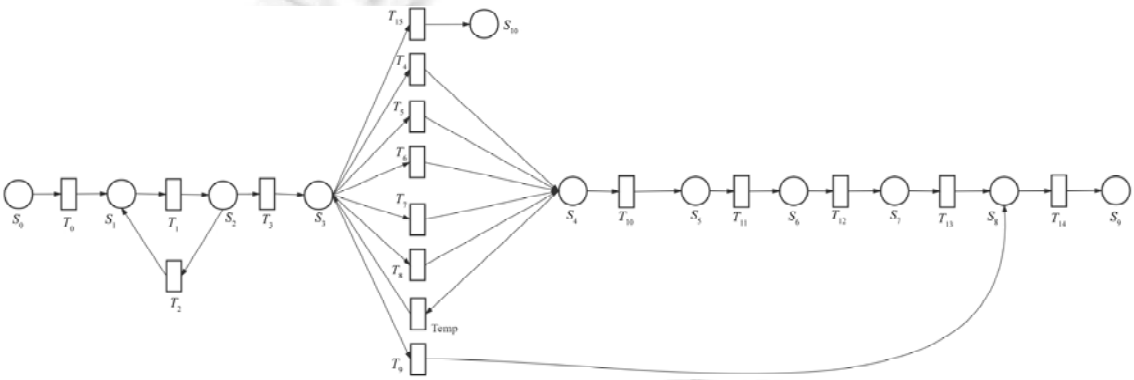


Fig.6 Petri net model of credit trading system

图 6 信贷交易系统 Petri 网模型

该平台分为客户端和服务端,其主要业务操作见表 1 中变迁标识说明所示,主要包括注册、登录、浏览个人中心、修改银行卡信息、浏览信贷项目、评估授信额度、还款、提交借款申请等功能.

本文在系统中增加了交互行为重构的流程和相应的控制模块,重构的变迁即系统中的控制模块触发条件,如图 7 所示.在用户成功登录系统首页触发 T_3 后,将会同时激活交互行为重构的控制模块 T_{16},交互行为重构的激活条件依赖于上一节中提到的 TDDA 算法的输出结果,即判断 T_{17} 能否执行:若满足激活条件,则 T_{18} 被激活;若用户在操作系统时激活 T_4,则使得交互行为重构叠加的行为流程 T_{21},T_{22} 被激活,从而产生新的行为流程.

用户行为最终是要通过系统行为来表现,因此,要改变用户行为就必须改变系统行为集合来劝导性约束用户行为集合.在通过时域漂移算法确定了用户登录行为触发因素出现时机的基础上,需要通过重构系统行为将触发因素的实现路径映射到系统行为序列中.交互行为重构是在保持交易系统的基本业务逻辑(即关键系统行为路径)不变的前提下叠加新的触发因素实现路径,从而形成新的系统行为集合.

当用户在周期性重复的 TDDA 算法和交互行为重构模型的引导下交互操作,则会逐渐形成新的用户行为.且由于欺诈行为的稀疏性和明确的目的性,持续的行为引导机制将无法改变欺诈者的交互行为习惯,使得身份

伪装的欺诈者仍保持原始场景的交互行为,从而与引导后的合法用户行为产生差异.

Table 1 Corresponding names of transition indicator

表 1 Petri 网变迁对应名称

变迁	变迁含义
T_0	注册
T_1	登录
T_2	登录失败
T_3	进入首页
T_4	浏览个人中心
T_5	查看及修改银行卡信息
T_6	浏览信贷项目
T_7	评估授信额度
T_8	查看借款详情
T_9	还款
T_{10}	提交借款申请
T_{11}	信贷条款签署
T_{12}	个人信息核验
T_{13}	平台放款
T_{14}	更新与恢复额度
T_{15}	注销

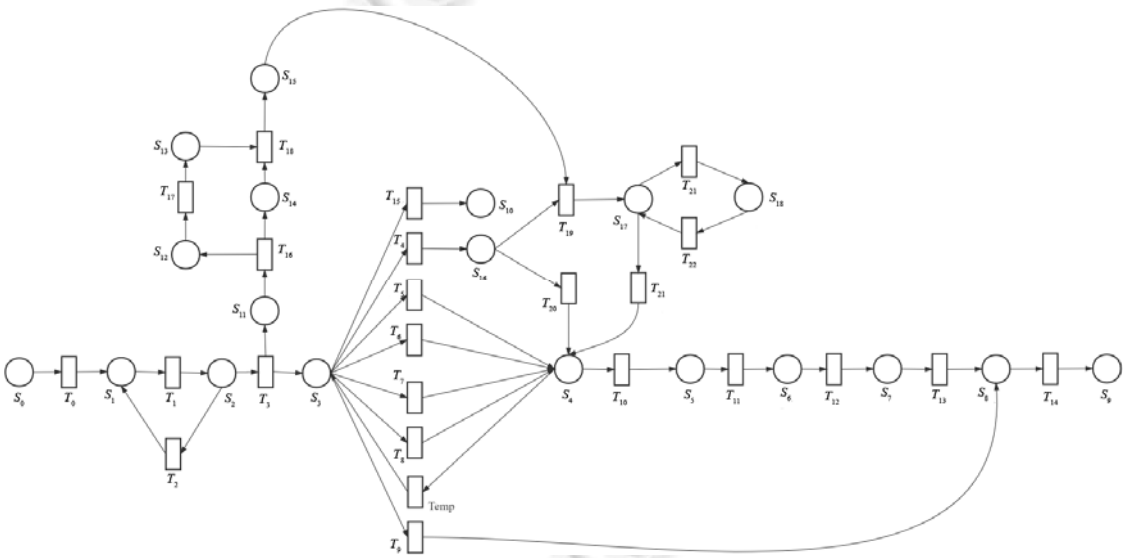


Fig.7 Interactive behavior restructuring Petri net model of credit trading system

图 7 交互行为重构信贷交易系统 Petri 网模型

3 实验结果

本节中,我们将通过实验验证本文提出的劝导式交互行为引导的身份识别方法的效果,首先介绍本次实验所采用数据集,然后说明行为引导实验的稳定性和区分性效果以及在身份伪装识别行为异常检测中的效果.

3.1 实验数据集

因为目前关于行为引导和行为改变的研究较少,且大多研究停留在分析和理论完善的阶段,论文的效果多以分析性结论,缺少定性定量指标衡量其变化差异,在调研中尚未找到连续时间周期下行为变化前后的公开数据集,所以本文的研究数据来源于实验室搭建的在线信贷交易系统.该系统根据信贷产品设计的任务流,在系统各页面通过 SDK 收集用户在页面中的操作数据,记录用户操作页面序列数据,具体包括用户名、用户手机号、

sessionID、操作的页面编号、页面名称、进入页面时间(时间戳)、离开页面时间(时间戳).本数据集持续收集了 5 位用户在 2018 年 9 月~2019 年 4 月间的 1 017 次系统登录行为和 16 741 条页面操作数据.其中:2018 年 9 月~2018 年 12 月期间的记录为系统原始场景下的用户交互行为数据;2019 年 1 月~4 月,根据本文提出的 TDDA 算法对各用户在系统的非关键行为页面“个人中心页”叠加行为引导机制,在此期间所收集到的数据作为行为引导后的用户交互行为记录.

由于本文研究的是行为伪装场景下的身份识别,身份伪装欺诈行为与正常交互行为具有高度相似性,且由于真实交易环境中黑样本的稀疏性,我们在原始数据集中对每个用户的行为记录随机标记了其总量 20%的交互行为数据作为该用户身份伪装欺诈的黑样本,其余交互记录作为该用户的白样本.并在行为引导后实施伪装欺诈,将这些身份伪装欺诈的黑样本按照时间线补充至引导后对应的用户数据集中.

3.2 对比实验

3.2.1 行为引导漂移模型

图 8~图 12 分别反映了在系统中施加行为漂移引导模型前后用户登录系统的时间频率变化.

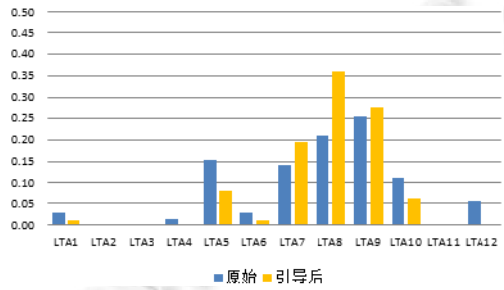


Fig.8 U1 Frequency change

图 8 用户 1 登录时间频率变化

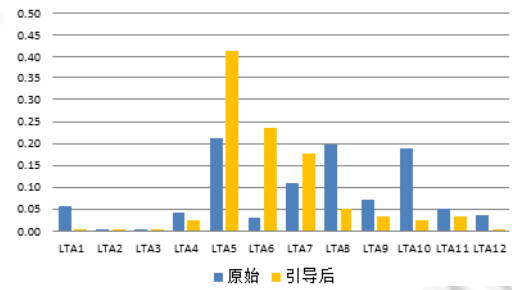


Fig.9 U2 Frequency change

图 9 用户 2 登录时间频率变化

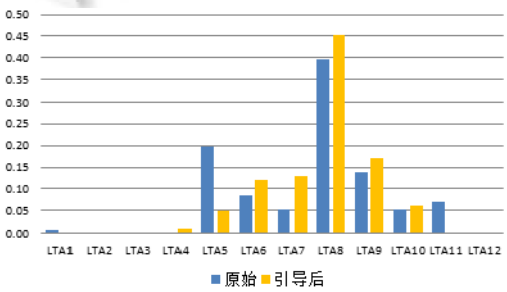


Fig.10 U3 Frequency change

图 10 用户 3 登录时间频率变化

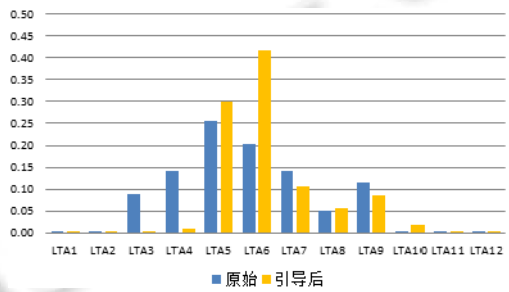


Fig.11 U4 Frequency change

图 11 用户 4 登录时间频率变化

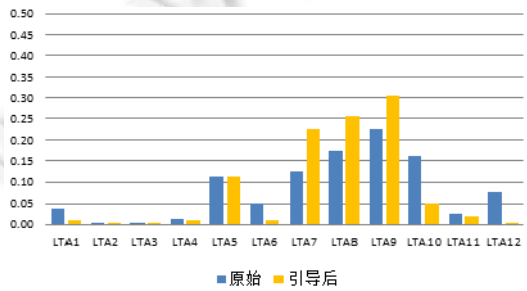


Fig.12 U5 Frequency change

图 12 用户 5 登录时间频率变化

为了验证引导前后用户交互行为具备一定的变化差异,我们引入评价指标:群体稳定性指标 PSI(population stability index),用来衡量行为引导前后用户登录概率分布之间的变化差异.在风控建模领域,PSI 常用来验证样本在各分数段的分布与建模样本分布的稳定性,判断建模时期样本与当前样本的偏差,用以衡量用户群体稳定性和适应性.计算公式如下:

$$psi=\sum_{i=1}^{n}(A_i-E_i)\times\ln\left(\frac{A_i}{E_i}\right),$$

其中,A_i 为实际占比,E_i 为预期占比.PSI 数值越小,两个分布之间的差异就越小,代表越稳定.一般来说,PSI 所反映的差异性见表 2.

Table 2 Meaning of PSI value

表 2 PSI 值的含义

PSI 值的范围	稳定性	建议事项
0~0.1	较稳定	没有变化或很少变化
0.1~0.25	略不稳定	有一定变化,注意后续变化
>0.25	不稳定	发生明显变化,进行模型分析

我们分别计算了各个用户在引导前后登录时间概率所对应的 PSI 值,见表 3.

Table 3 PSI value of each user

表 3 各用户的 PSI 值

用户	Before guidance	Before and after guidance	After guidance
1	0.105 303	0.466 226	0.146 624
2	0.169 659	1.556 132	0.166 411
3	0.124 616	0.628 511	0.167 358
4	0.109 631	0.973 336	0.100 913
5	0.142 602	0.658 965	0.165 745

可以看出:

- 在系统中施加行为漂移引导模型前,各用户的登录时间概率分布对应的 PSI 值变动均低于 0.17,其中大部分用户低于 0.15,说明用户的登录行为在不受系统行为额外干预的状态下保持了相对的稳定性,因此,行为伪装欺诈成为可能;
- 在系统施加了漂移引导模型的激励机制后,用户在引导前后登录行为概率分布对应的 PSI 均发生了不同程度的变化,各用户的 PSI 均大于 0.45,即:在激励机制下,使得用户的登录时间概率分布较原始场景发生了较为明显的变化;
- 在持续施加行为引导激励机制后,各用户的 PSI 趋于稳定保持在 0.16 左右,与原始行为的稳定性相比略有降低.

实验证明:在系统中施加行为漂移引导模型的激励机制后,用户的登录时间分布发生了明显的变化;且在本文提出的激励机制引导下,用户行为变化后产生的新登录行为模式相对稳定.与原始登录时间分布相比,既保持了一定的区分性也保证了行为变化的平滑性.

3.2.2 伪装行为异常身份识别

本文采用的用户身份识别方法为基于系统交互行为的 LTA^u,WTA^u,LIA^u,KSA^u 合并的 24 维向量进行用户交互行为建模,采用前期研究[25]中提出的超球体检测模型(UR)分别对各个用户在引导前后两组数据集上进行身份伪装异常行为检测.

由于本文是对异常行为的识别,所以对混淆矩阵稍作修改,重点放在异常交互行为的判别,即修改后的混淆矩阵:TP(true positive)是模型将真实异常行为判断为异常行为的数量,FP(false positive)是真实正常行为被模型判断为异常行为的数量,TN(true negative)是真实正常行为被模型判断为正常行为的数量,FN(false negative)是真实异常行为被模型判断为正常行为的数量.

为了能让比较结果更加有说服力,本实验使用欺诈识别领域文献中常见的几个指标作为本论文实验的评

估指标.一共 4 项评估指标,分别为准确率、召回率、精确率和 *F*1 值,计算方式见表 4:准确率表示该模型判定结果中正确判断的数量占总体行为数量的百分比;精确率是模型判断的真实异常行为与判断为异常行为的比率;召回率是模型判断的真实欺诈交易数量占所有异常行为数量的百分比;*F*1 值为衡量模型性能的综合指标,是准确率和召回率的调和平均值.

Table 4 Indicator calculation method

表 4 PSI 值的含义

Indicator name	Calculation method
Accuracy	$(TP+TN)/(TP+TN+FP+FN)$
Precision	$TP/(TP+FP)$
Recall	$TP/(TP+FN)$
F-measure	$2\times Precision\times Recall/(Precision+Recall)$

从实验结果图 13~图 16 可以看出:采用超球体模型(UR)进行的用户身份伪装行为异常检测实验中,采用 TDDA 行为引导后的各项指标均不同程度高于未采用 TDDA 引导机制.其中,

- 准确率结果如图 13 所示,指标平均提升 15.12%,说明在伪装行为识别中采用 TDDA 引导机制能够更准确地判断出用户的正常行为和伪装异常行为;
- 精确率如图 14 所示,虽然在存在波动,但指标平均仍高出 14.10%,说明绝大部分用户中,采用 TDDA 引导机制在判断异常行为方面更为出色;
- 召回率如图 15 所示,平均提升 10.87%,说明采用 TDDA 行为引导机制后,在精准判断异常行为的同时,对于正常交易的误判情况也较少;
- *F*1 值如图 16 所示,*F*1 值反映模型的整体性能,可以看出:采用 TDDA 行为引导机制后,模型的 *F*1 值平均提升 28.76%.即:采用 TDDA 行为引导机制后,模型的整体性能均优于未使用行为引导的原始场景.

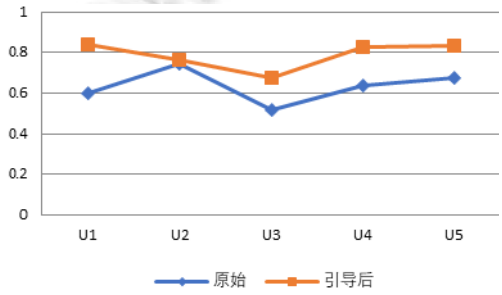


Fig.13 OM model *accuracy*

图 13 OM 模型检测准确率

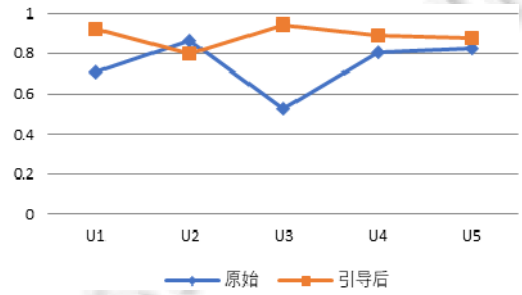


Fig.14 OM model *precision*

图 14 OM 模型检测精确率

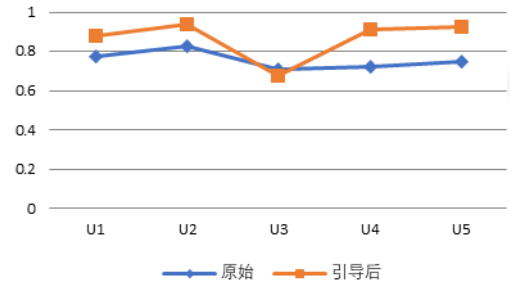


Fig.15 OM model *recall*

图 15 OM 模型检测召回率

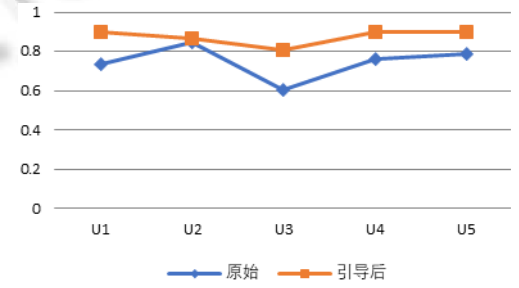


Fig.16 OM model *F*1-value

图 16 OM 模型检测 *F*1 值

通过对实验结果的分析可以看出:采用 TDDA 行为引导机制后,对于行为伪装异常检测的结果要比原始场景有更好的整体性能.主要原因有以下几点:一是从用户的历史交互行为出发,基于用户行为的稳定性和偏向

性,使得本文提出的交互行为重构系统干预模型能够平滑的引导用户行为发生变化并使得引导前后具备一定的行为区分性,为身份伪装欺诈检测提供了新的解决思路;二是对于行为稳定性和偏向性系数的刻画时采用1.5IQR的异常值分析方法,能够更好避免异常值对于引导模型平滑性的干扰.因此,采用TDDA行为引导机制能够在不改变原有欺诈检测方法的情况下使得正常行为和伪装的欺诈行为产生区分,使得模型的准确率、精确率等指标相对于原始场景取得显著提升,对于模型在伪装行为的判断方面具有更好的整体性能.

4 结 论

本文提出了一种个体交互行为的平滑干预模型,该模型从用户的历史交互行为出发,考虑了用户之间的差异,分析用户交互行为的稳定性和偏向性系数,并提出了交互行为时域漂移算法,为每个用户确定各自的行为干预时机.同时提出了交互行为重构的系统实现方法,采用Petri网对业务系统进行建模,给出了不同系统行为轮廓下行为流程的重构方式,在确保不破坏系统基本业务逻辑的前提下,使得合法用户的行为能够平滑变化,且与原始行为特征具备一定的区分性.实验证明:在行为伪装欺诈检测场景中,使欺诈者模拟的用户行为失效,使得检测模型的准确率、精确率等指标相对于原始场景提升 10%以上.证明了本文提出的行为干预方法的有效性,为身份伪装场景的欺诈检测提供了一个全新的解决思路和视角.在下一步的工作中,将持续关注如何衡量行为漂移程度与良好交互体验之间的平衡,以及如何从形式化的角度论证该策略的有效性.

References:

[1] China Information and Communication Research Institute. Mobile Digital Finance and Electronic Commerce Anti-fraud White Paper (in Chinese).

[2] Aiken LR. Attitude and Behavior. Beijing: China Light Industry Press, 2008 (in Chinese).

[3] Nenadic A, Zhang N, Barton S. A security protocol for certified e-goods delivery. In: Proc. of the Int'l Conf. on Information Technology: Coding and Computing (ITCC 2004), 2004. 22–28.

[4] Zhong J, Yan C, Yu W. *et al*. A kind of identity authentication method based on browsing behaviors. In: Proc. of the 2014 7th Int'l Symp. on Computational Intelligence and Design. 2014. 279–284.

[5] Zhao P, Yan C, Jiang C. Authenticating Web user's identity through browsing sequences modeling. In: Proc. of the 2016 IEEE 16th Int'l Conf. on Data Mining Workshops (ICDMW). 2016. 335–342.

[6] Roth J, Liu X, Metaxas D. On continuous user authentication via typing behavior. IEEE Trans. on Image Processing, 2014,23(10): 4611–4624.

[7] Ma L, Yan C, Zhao P, *et al*. A kind of mouse behavior authentication method on dynamic soft keyboard. In: Proc. of the 2016 IEEE Int'l Conf. on Systems, Man, and Cybernetics (SMC). IEEE, 2016.

[8] Liu C, He J. Access control to Web pages based on user browsing behavior. In: Proc. of the 2017 IEEE 9th Int'l Conf. on Communication Software and Networks (ICCSN). 2017. 1016–1020.

[9] Zhao P, Yan C, Jiang C. Authenticating Web user's identity through browsing sequences modeling. In: Proc. of the 2016 IEEE 16th Int'l Conf. on Data Mining Workshops (ICDMW). 2016. 335–342.

[10] Zheng L, Liu G, Yan C, Jiang C. Transaction fraud detection based on total order relation and behavior diversity. IEEE Trans. on Computational Social Systems, 2018:796–806.

[11] Zheng L, *et al*. A new credit card fraud detecting method based on behavior certificate. In: Proc. of the 2018 IEEE 15th Int'l Conf. on Networking, Sensing and Control (ICNSC). 2018. 1–6.

[12] Chen D, Ding Z, Yan C, *et al*. A behavioral authentication method for mobile based on browsing behaviors. In: Proc. of the Institute of Electrical and Electronics Engineers Inc. 2019.

[13] Zhong J, Yan C, Yu W, Zhao P, Wang M. A kind of identity authentication method based on browsing behaviors. In: Proc. of the 2014 7th Int'l Symp. on Computational Intelligence and Design. 2014. 279–284.

[14] Zhang Y, Chen G. A forensics method of Web browsing behavior based on association rule mining. In: Proc. of the 2014 2nd Int'l Conf. on Systems and Informatics (ICSAI 2014). 2014. 927–932.

[15] Liu GJ, Jiang CJ. Behavioral equivalence of security-oriented interactive systems. IEICE Trans. on Information and Systems, 2016,E99-D: 2061–2068.

[16] Liu GJ, Jiang CJ. Secure bisimulation for interactive systems. In: Proc. of the 15th ICA3PP. LNCS 9530. 2015. 625–639.

[17] Zhang Z, Chen L, Liu Q, *et al*. A fraud detection method for low-frequency Trans. IEEE Access, 2020:25210−25220.

[18] Toledo FPD, Devincenzi S, Kwecko V, Mota FP, Botelho SSDC. A framework for modeling persuasive technologies based on the fogg behavior model. In: Proc. of the 2018 IEEE Frontiers Education Conf. (FIE). 2018. 1−5.

[19] Hamper A, Wendt J, Zagel C, *et al*, Behavior change support for physical activity promotion: A theoretical view on mobile health and fitness applications. In: Proc. of the 2016 49th Hawaii Int'l Conf. on System Sciences (HICSS). 2016. 3349−3358.

[20] Nishiyama Y, Okoshi T, Yonezawa T, *et al*. Toward health exercise behavior change for teams using lifelog sharing models. IEEE Journal of Biomedical and Health Informatics, 2016,775−786.

[21] Zhang Z, Arakawa Y, Oinas-kukkonen H. Design of behavior change environment with interactive signage having active talk function. In: Proc. of the 2019 IEEE Int'l Conf. on Pervasive Computing and Communications Workshops (PerCom Workshops). Kyoto, 2019. 796−801.

[22] Zhang ZH, Cui J. An agile perception method for behavior abnormality large-scale network service systems. Chinese Journal of Computers, 2017,40(2):505−519 (in Chinese with English abstract).

[23] Pan L, Ma B, Wang Y. The similarity calculation of E-commerce user behaviors with Petri net. 2017.

[24] Weidlich M. Behavioural profiles: A relational approach to behaviour consistency. Journal of Biological Chemistry, 2011,269(36): 22847−22852.

[25] Chen L, Zhang ZH, Liu QW, *et al*. A method for online transaction fraud detection based on individual behavior. In: Proc. of the ACM Turing Celebration Conf.-China. ACM, 2019. 119.

附中文参考文献:

[1] 移动数字金融与电子商务及欺诈白皮书,2019..

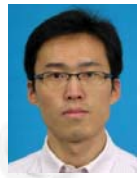
[2] 态度与行为.北京:中国轻工业出版社,2008.

[22] 章昭辉,崔君.大规模网络服务系统行为异常的敏捷感知方法.计算机学报,2017,40(2):505−519.

刘霄(1995－),男,硕士,主要研究领域为交互行为异常检测.

魏子明(1995－),男,硕士,主要研究领域为交互行为异常检测.

章昭辉(1971－),男,博士,教授,博士生导师,CCF 专业会员,主要研究领域为大数据,行为分析.



王鹏伟(1984－),男,博士,副教授,CCF 专业会员,主要研究领域为云计算与边缘计算,服务计算,数据挖掘与分析.