文章编号:1000-2367(2021)06-0091-08 DOI:10.16366/j.cnki.1000-2367.2021.06.013

隐私保护的网络实名制体系研究

张萍[1],蒋琳[2]

(1.广东警官学院 网络信息安全系,广州 5102301;2.哈尔滨工业大学 计算机科学与技术学院,广东 深圳 518000)

摘 要:在我国,实名制已在部分领域逐步得到了实现,但距离全面实施网络实名制体系还有大量的问题亟待研究解决,尤其是如何寻求隐私数据保护与数据价值利用的平衡.利用完全同态密码算法、ElGamal 数字签名和基于大数据的多层次异常事件检测等技术,提出了一个针对我国国情的网络实名制体系设计.此设计创新的将用户隐私数据分为高敏感度和低敏感度数据库,高敏度数据只以密文的形式存储,低敏感度数据在经过数据中心的处理后交给网络服务商,同时利用同态加密的特点实现密文层次的运算.在将用户额外操作开销最小化的同时,保障用户的个人隐私信息免遭非授权的访问,并为执法部门打击网络违法犯罪行为提供了有力的工具.

关键词:隐私保护;实名制;同态加密;电子取证;数据挖掘

中图分类号:TP302.1;TP309 **文献标志码**:A

实名制是伴随着信息化飞速发展而兴起的一种新制度,即个人在执行某些行为之前必须能够证明自己的合法有效身份.自 2017 年 6 月 1 日起实施的《网络安全法》,其中第二十四条正式以法律的形式规定了实名制的地位,针对网络虚拟空间实施网络实名登记制度、对公民的网络行为进行合理合法的监管.

网络实名制自 2007 年 7 月在韩国首次推行此制度.但韩国的网络实名制对于阻止“网络暴力”的初衷收效甚微,由于整个系统安全体制架构的不完善,韩国网络实名制引发了一系列网络安全事件,例如由于黑客的攻击致使网站的大量用户个人资料泄漏、不少网站兜售用户个人资料给第三方以进行电话或邮件推销乃至诈骗[1-2].2012 年 8 月,韩国的网络实名制被韩国宪法裁判所判决违宪而退出了韩国的历史舞台.

针对我国的网络实名制,众多学者已从经济、法制、管理、文化等方面都展开过深入的研究,虽然网络实名制在实现网络空间的长治久安方面有着独特的优势,但暂未有任何研究从技术层面探讨如何具体实现一个完善的网络实名体系.借鉴已有的实名制体系存在的问题,本文针对我国国情提出兼顾隐私保护和数据利用的安全网络实名体系设计方案.不仅关注用户的实名认证环节,同时还将如何保障用户的隐私数据与认证进行有机融合,也是首次提出认证与隐私保护环节不脱离的网络实名制设计方案,本方案还创新性引入了静态、动态数据库分离的概念和高敏感度、低敏感度数据,同时满足网络服务商对数据价值的追求和用户对隐私保障的需求.

1 网络实名制的必要性分析

网络的虚拟性和匿名性是其创建之初的一大亮点,但在缺乏信任机制的监管下,众多的安全问题也随之暴露,已经或潜在危害了每一个网络使用者,其中较为突出的问题有以下 4 点.

(1)用户信息遭受重大威胁.国内外绝大多数的网络服务提供商因各种主动或被动的原因曾经或者正在遭受到泄露用户个人隐私数据的危机.根据全球数据安全公司 Gemalto 发布调查结果显示,2018 年公开披露的数据泄露事件多达 6 500 起,涉及泄露超过 50 亿条数据记录.其中身份窃取依然是导致数据泄露的重要

收稿日期:2021-01-07;**修回日期**:2021-09-02.

基金项目:国家自然科学基金(61872109);广东省教育厅青年创新人才类项目(2018KQNCX175).

作者简介(通信作者):张萍(1986—),女,山东青岛人,广东警官学院讲师,博士,研究方向为信息安全、电子取证、网络犯罪侦查,zpecho@gdppla.edu.cn.

原因,由于用户信息的泄露已经引发了大量的安全事件,并滋生了众多赖以为生的全链条网络黑灰产业,包括网络诈骗、网络诽谤、网络赌博等等.

(2)利用网络匿名性实施网络犯罪.在互联网飞速发展、信息高度膨胀的时代,不法分子利用网络的匿名性特点,频频策划实施网络暴恐事件、网络群体性事件、网络谣言等扰乱国家安全、社会治安的网络违法犯罪事件.

(3)网络攻击更加精确且影响广泛.根据 CNNIC 发布第 43 次《中国互联网络发展状况统计报告》统计,我国网民人均周上网时间已达到 27.6 h,依据个人用户或者团体组织在网络空间的活动能够更加精准快速的对其做出画像[3-5],这使得网络攻击更加精准、更加有针对性,攻击受众面更广.例如 2017 年 5 月比特币的交易价格达到历史高点时,DDoS 攻击也同样达到了年度攻击量的最高峰值;随着物联网的普及,针对 IoT 的网络攻击量也开始攀升.

(4)网络犯罪打击滞后溯源困难.随着网络的高度普及,传统犯罪日渐向网络阵地进行转移,网络犯罪形势日益严峻,且呈现组织化、链条化、跨国界、低龄化等特点.因此加强网络安全管理已是当务之急,但非实名制网络只能在犯罪行为发生之后采取事后溯源、取证等工作,这大大削弱了打击网络犯罪的力度.

自网络实名制的概念被提出之后,全球尤其是我国对网络实名制就进行了多方面的探讨并从多个角度提出了一些可行性设计.已有的研究成果或仅关注有限网络区域内的实名认证,或仅关注实名认证机制,或实名认证与用户隐私保护环节进行分离.洪丹丹等[6]给出了在校园网内结合微信认证接口实现网络实名制的方案,文勇军等[7]则给出基于教育网实现实名网络认证的具体实现方案,HU 等[8]提出了一个基于 PKI 的网络实名制系统,但其易用性不高,且需要大量的后台工作,程琳等[9]则结合硬件设备 USBKey 进行个人实名认证机制,但同时也限制了网络市场经济化发展,可推性度不高.针对电信实名制,姚慧等[10]提出了基于人工智能的技术要点,但未针对用户的个人隐私进行保护设计.在保护网络用户隐私方面,XU[11]提出了基于 Shamir 秘密分享的隐私保护实名网络方案,但忽略了多个网络服务商可联合提出对数据进行解密申请的问题,李晖等[12]探讨了在移动互联网内利用虚假位置、假名证书等技术实现对用户的隐私保护方案,张梅舒等[13]则提出了多维数值型数据的隐私保护方法.

实施网络实名制是国家管理网络空间的必要手段之一,为实现可控的网络空间安全提供了强有力的工具,因此建立一个完善的实名制网络体系对实现网络空间安全有着至关重要的意义和作用.但同时实施网络实名也存在着各种困难,例如用户隐私数据的保护问题、网络服务供应商的数据盈利模式等问题有待解决.

本文提出的设计方案兼顾了公民的隐私信息和企业追求的数据价值,同时可以对网络违法犯罪行为预警,并提供有价值的合法电子证据.本设计首次将用户的隐私数据区别为高敏感度和低敏感度数据,并应用不同的处理规则,达到保障用户隐私和创造数据价值的双赢目的.同时利用用户已实名登记过的手机实现透明化使用体验,借助 ElGamal 算法的特性防止“拖库”“撞库”行为,通过与用户身份挂钩的大量网络行为数据实现分等级的异常检测预警,融合网络取证于一体实现电子证据的固定.

2 隐私保护的网络实名制体系设计

本部分将给出全面的网络实名制体系设计方案,如图 1 所示,该体系主要由 3 个功能模块构成:实名登记模块、透明化登录模块和用户行为预警模块.

本设计创新性地通过实名登记模块利用完全同态加密算法结合 ElGamal 签名算法实现用户隐私信息的分级管理,网络服务提供商无法全部获得用户的隐私信息或者只能获得经过处理的隐私信息,因此攻击者通过“拖库”或者“撞库”攻击后台数据库,也无法获得用户的隐私信息.在数据存储中心,首次采用分级别处理用户的隐私数据,高敏感度隐私数据利用完全同态加密算法实现密文层次的运算,低敏感度数据利用限制请求者获取的维数或降低数据敏感度等手段保护隐私.高敏感度隐私数据仅以密文的形式进行存储,即使数据泄露事件发生,攻击者也无法获得有用的隐私数据.透明化登录模块通过改良的 ElGamal 签名机制防止攻击者通过截获网络数据尝试“撞库”,并且将用户的额外操作降至最低,简便易用是产品商业化推广成功的必备条件.用户行为预警模块则始终工作在系统后台,通过限制数据获取的规则,既依靠网络服务商的数据处理能力进行大数据分析,对潜在的网络犯罪行为进行预警,同时避免用户数据的泄露,并且对已经发生的网

络犯罪行为能够进行迅速精准的溯源.

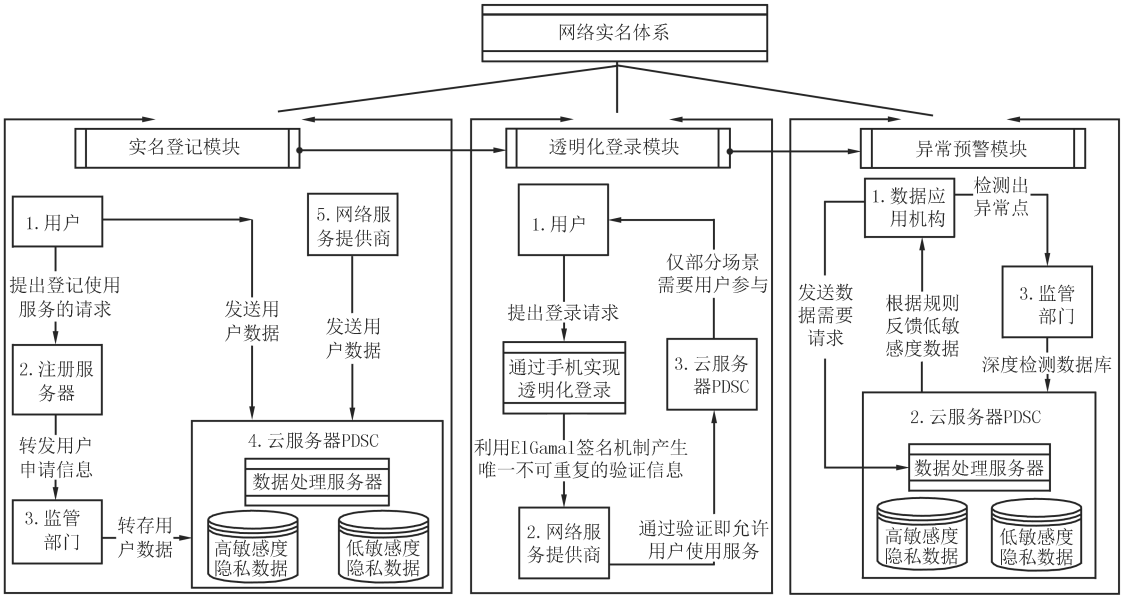

图1　网络实名体系架构

Fig.1　Framework of Internet real-name system

2.1　用户实名登记模块

实名制体系中普通用户最关心的是个人隐私数据的安全性，因各种主动或被动手段导致个人或集团用户隐私数据泄露的案例层出不穷.因此，此模块主要研究如何高效、安全地保障用户的隐私数据不被其他用户或组织盗用、滥用.

依据对实名制体系的安全性需求，本模块总共有以下多个参与方涉及用户隐私信息的监管或使用.

· 用户 User(U)：普通网络服务使用者；

· 服务提供商 Service Provider(SP)或其他用户：希望验证用户个人信息的参与者；

· 独立第三方 Independent Authority(IA)：起到独立监管的作用，存有用户的部分关键个人信息，可类比于现实中的公安机关；

· 注册中心 Registration Center(RC)：用户需要首先在注册中心进行注册，设置为云服务器；

· 个人信息存储中心 Private Data Storage Center(PDSC)：用于存储用户的个人隐私信息 Private Information(PI)，设置为云存储服务器，本方案将用户的个人数据分为两类分别存储于不同的数据库中，一类是用户的关键性隐私信息 Critical Private Information(CPI)，另一类是用户的一般性隐私信息 General Private Information(GPI).其中 CPI 数据库中的数据是可以唯一性指向该用户的高敏感度隐私数据，例如用户的身份证号码、手机号码、详细住址、社交平台账号、邮箱账号、生物特征值等，此类数据必须以密文形式进行存储.而 GPI 又细分为静态用户数据库(Static User Information Database，SUID)和动态用户数据库(Dynamic User Information Database，DUID)，其中 SUID 用于保存用户较少进行更新的个人数据，例如性别、年龄、工作类型等，DUID 则存储用户的动态网络数据，例如用户的上网轨迹、用户的位置轨迹.GPI 内的数据单独一项都不唯一性的指向某一个用户，因此一般以明文的形式进行存储，但如果针对某一用户的此类数据全部提取后分析，可以对此用户进行人物画像，从而带来隐私泄露的风险.因此 GPI 数据属于低敏感度，可在降低敏感度之后提供给网络服务提供商，或者交由 DPC 处理；

· 数据处理中心 Data Processing Center(DPC)：用户隐私数据在此进行计算并将处理结果返回给请求者.

其中用户的个人隐私信息 PI＝{CPI，GPI}，其中 CPI＝$\{c_1,\cdots,c_n\}$，GPI＝$\{g_1,\cdots,g_n\}$，c_i 是用户的某一项高敏感度隐私数据，g_i 是用户某一项低敏感度隐私数据.本模块的处理流程如图 2 所示.

①U 首次使用网络服务提供商 SP_n 的服务时，通过 RC 进行服务账户 $Account_{SP_n+U_{id}}$ 的申请，将 U 的个人身份信息数据(属于关键性隐私数据)经过 U 的公钥加密处理后，再通过 IA 的公钥加密 $E_{PK_{IA}}(U_{id}\ ||$

$\{E_{PK_U}(PI)\}$,$Sign(U_{id})=E_{PK_{IA}}(U_{id}\ ||\ E_{PK_U}(c_1)\ ||\ \cdots\ ||\ E_{PK_U}(c_n)$,$Sign(U_{id}))$ 发送给 RC,其中 U_{id} 代表 U 的身份证号码,Sign(·)代表 ElGamal 签名算法,$\{E_{PK_U}(PI)\}$代表对 U 的 CPI 中的元素进行加密的集合;

②RC 在 IA 的协助下,例如通过公安人口数据库对该用户 U 的个人信息数据再次利用 U 的公钥进行加密,进行对比后将 U 身份符实与否的反馈发回给 RC;

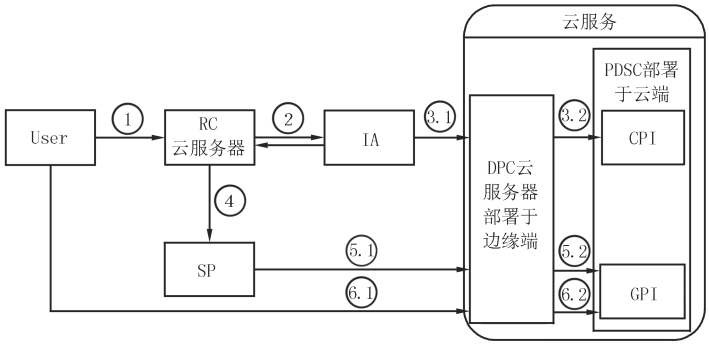

图2 用户注册模块流程

Fig.2 Registration flow in real-name system

③如 U 的身份符实 IA 将会将加密后的 U 的个人身份信息数据集 $\{E_{PK_U}(PI)\}=\{E_{PK_U}(c_1),\cdots,E_{PK_U}(c_n)\}$ 转存至云端数据库的 CPI 中,包括 $E_{PK_U}(U_{id})$,CPI 数据库已有的数据无须重复发送以节省通信成本并减少密文出现的次数,以降低暴力攻击或重放攻击的概率;

④如 U 的身份符实则 RC 将通知 SP_n 为 U 开设服务账户 $Account_{SP_n+U_{id}}$;

⑤SP 能够获取到 U 的隐私数据仅限于一般性隐私数据,此类数据将通过 DPC 被存储于云端数据库的 GPI;

⑥用户 U 产生的低敏感度隐私数据也将通过 DPC 自动上传于云端数据库的 GPI 中.

本环节中所有需要涉及用户的高敏感度隐私信息 CPI 的过程都将重复图 2 中的①~③步骤,且整个流程中用户的 CPI 数据从未以明文的形式出现,即使在 IA 处进行 CPI 的核对也无需对$\{E_{PK_U}(PI)\}$进行解密,仅需核对密文.此外,由于 Elgamal 签名算法[15]的特性每次选取一个随机数 k,因此每次生成的签名数值都不一样,流程①~③的通信传输中出现的密文每次也就都不一样,可以抵抗暴力攻击.

需要说明的一点是本文中 U 使用的加密算法选取的是同态公开加密算法[16],既同时满足以下几个条件:

$$\begin{cases}E(x)+E(y)=E(x+y), & (1)\\ E(x)\cdot E(y)=E(x\cdot y). & (2)\end{cases}$$

2.2 用户透明化登录模块

易用性和人性化设计是一个产品得以成功的关键因素之一,除保护用户的隐私信息之外,还应简化用户注册和登陆的手续,提高本系统的易用度,使得用户透明使用网络实名登记体系,最大限度降低用户使用网络实名登记体系的额外工作.

本模块的设计方案利用用户已经实名登记过的手机 SIM 卡和使用的手机来省略用户再次登记的手续,当用户使用手机进行所有网络行为时,用户身份自动通过识别手机的 IMEI 号、SIM 卡的 IMSI 号和 ICCID 号确认此手机号码实名登记的公民身份,同时可启用双重因子身份认证,利用面部识别或指纹识别增加安全性.本环节的具体流程如图 3 所示.

①用户 U 已经通过 3.1 模块在 SP 开通了账户后,每次使用 SP 的服务时需验证 U 的身份,首先选取一个 CPI 的子集 $CP=\{c_i\mid c_i\in CPI\}$,U 将选取的 c_i 的参数索引 i 和数字签名 $Sign_U(E_{PK_U}(\sum CP))$发送给 SP;

②SP 将接收到的信息转发给数据处理中心 DPC;

③DPC 从 CPI 数据库中根据参数索引 i 抽取 $E_{PK_U}(c_i)$,且根据完全同态加密运算规则验证 $\sum E_{PK_U}(c_i)$与 $E_{PK_U}(\sum CP)$是否相等;

④DPC 将验证结果反馈给 SP;

此模块的步骤将根据用户登录的环境选取不同的

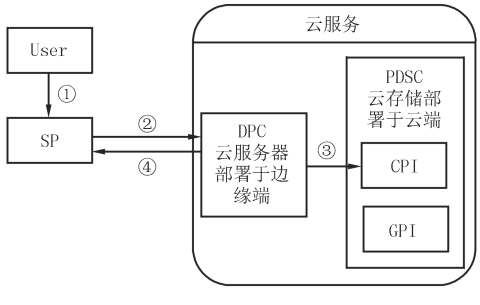

图3 透明化登录模块流程

Fig.3 Login flow in real-name system

CPI 子集:

(1)在 U 首次登录、长时间未使用 SP 的服务或涉及支付手续等场景时,将从 CPI 中选取除手机特征信息之外 c_i 的进行验证;

(2)除前述场景之外本环节将依靠 Apps 自动获取 U 使用手机的 IMEI,IMSI,ICCID 信息作为选取的 c_i,这样 U 可以无须输入任何信息实现透明化登录;

(3)同时,本系统会随机启动用户生物特征验证,或者当用户的行为偏离正常阈值也会启动生物特征验证,即 c_i 选取为 U 的生物特征信息,利用双重因子认证(2FA)加强系统安全性.

本模块通过验证用户的高敏感度隐私信息,在特殊情况下需要用户参与登录过程,其他多数场景将由系统自动完成无须用户的参与,对用户来说是完全透明的,因此极大程度上提高系统的实用性.同时,本系统设计中用户的隐私信息都保存在可信任第三方,而非每个网络服务商处,由于 SP 不保存用户的个人隐私信息,因此用户的个人信息不存在被非法利用的危险.

此外,本模块的签名算法 Sign(·)与 2.1 节的算法保持一致都为 ElGamal 算法,从而保证即使对同一个信息进行签名得到的数字签名也是不同的,防止不法分子利用“撞库”获得用户的信息.

2.3　用户行为预警模块

此模块针对的是云存储器内的用户数据处理方案,在当下数据资本的时代,数据不仅仅已成为不容忽视的企业资本,更在潜移默化地改变着整个社会的经济格局.如一味地注重用户隐私而限制对数据的利用是违背“数据资本”时代的发展步伐的,因此本模块的设计原则是寻求隐私保护与数据价值利用的平衡.图 4 是本模块的流程图,具体操作的流程步骤分为:

①、② U 和 SP 产生的一般性隐私数据都通过 DPC 预处理后直接存储于 GPI 数据库中;

③ 对于非政府或非公益机构想获取 GPI 中的用户一般性隐私数据进行商业/非商业化研究,首先必须同意执行步骤②,即机构需要将自己获取的用户数据分享给云服务器,才允许其将数据的利用请求 Request 发送给 DPC;

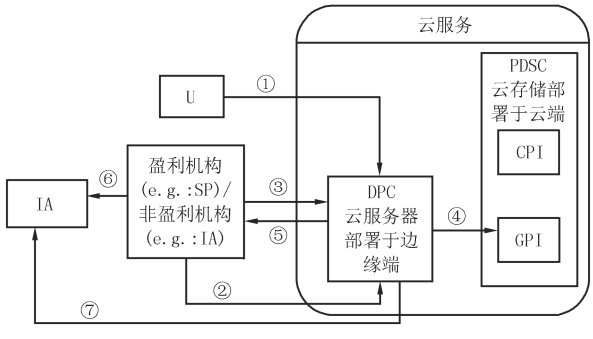

图4　数据处理与预警模块流程

Fig.4　Pre-alarm flow in real-name system

④ DPC 根据 Request 从 GPI 中调取数据;

⑤ 将脱敏后的数据或数据处理结果 Reply 反馈给请求者;

⑥ 针对 Reply 进行用户行为分析,若出现异常点则将结果反馈给 IA,由 IA 再次进一步检测 GPI 进行预判;

⑦ DPC 根据 IA 设定的规则对单独用户的 GPI 数据进行预警处理,若出现异常点则反馈给 IA.

此流程中对于 GPI 的数据处理请求 Request 有以下几个限制条件,其中 GPI 数据库中的每一条记录标记为 $\mathrm{GPI}_{\mathrm{U}_{id}}(g_1,\cdots,g_i,\cdots,g_n)$,任一属性标记为 G_i:

规则 1　当且仅当 Request=$\mathrm{GPI}_{\mathrm{U}_{id}}$,DPC 拒绝处理,即不得获取某一个用户的全部 GPI 数据;

规则 2　当且仅当 Request=G_i,DPC 将 G_i 返回给请求者;

规则 3　Request=$\{G_i\}$,当且仅当 Card(Request)<ln(Card(GPI)/2),DPC 将$\{G_i\}$返回给请求者;

规则 4　Request=$\{G_i\}$且 Card(Request)≥ln(Card(GPI)/2),DPC 首先对 G_i 进行脱敏处理,对于数值型属性值转化为范围值,对于描述性数值则转化为关键字,再将脱敏处理后的结果返回给请求者.

以上处理限制的原则是避免请求者通过获取的 Reply 对用户进行用户网络画像,规则 3 和规则 4 依据生日悖论限制请求者获取的隐私数据数量,达到保护用户的隐私目的.

对用户不法行为的及时预警是一个安全的网络实名体系需要提供的功能之一,在互联网飞速发展、信息高度膨胀的时代,网络已成为恐怖组织和恐怖分子组织、策划、实施恐怖犯罪活动的重要工具.近年发生的暴力恐怖活动中,绝大部分犯罪分子是收到网络上传播暴恐思想的文字、音视频资料后实施犯罪的,一些重大

案件甚至在境外通过互联网策划、远程指挥实施的.因此加强网络安全管理已是当务之急,网络实名制的部署可以为网络反恐的实施提供有力有效的平台.上述流程的步骤⑥、⑦的设计就是满足系统的预警需求,主要利用聚类和频繁机挖掘算法对用户的异常行为进行预警.其中⑥是针对全局用户,⑦则是针对单一用户,数据异常点的处理如图 5 所示,具体处理步骤如下.

(1)通过 DPC 获取用户 GPI 中的网络行为类型、网络行为来源、持续时间、访问深度、访问频率、关联跳转等,随后利用预剪枝技术将每个用户的无意义或分析意义较低的数据进行排除,以降低数据分析的数据量,提高效率.

(2)针对全局用户的网络行为,由于数据维数高且数据量庞大,因此首先结合 PCA 和 TSNE 技术对用户行为进行降维处理,即先进行 PCA 降维再利用 TSNE 降维[17],可大幅度提高降维效率同时提升降维质量.再利用聚类分析技术对用户的网络行为属性进行聚类,最后利用改进的 iForest 算法对每一个聚类进行异常点挖掘[18-19],实现对用户的异常行为进行高准确率的预警.

(3)针对每个用户的自身网络行为,由于样本量单一且较为稳定,因此选用 One Class SVM 算法进行异常点检测[20],实时检测单用户的异常行为并启动预警.

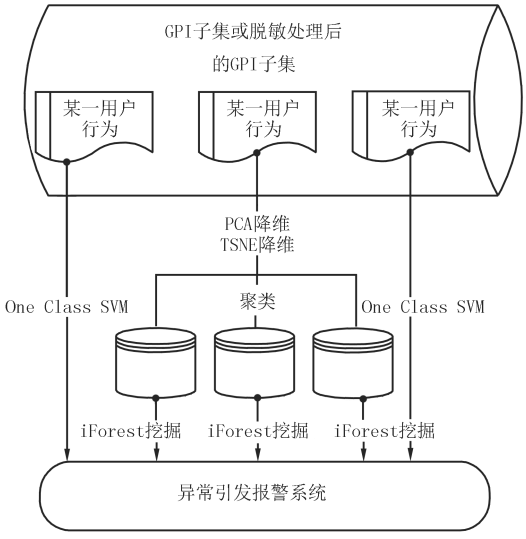

图5 基于GPI数据异常点检测

Fig.5 Anomaly detection with GPI data

与此同时,在定位异常行为的同时自动将涉及异常点的 CPI 和 GPI 数据进行统一,这为动态网络取证的证据固定提供了有利条件.在危害网络行为被确认后,执法机关在符合调取用户个人资料的授权下,相关执法人员可以通过后台固定的电子证据查找到用户的相关信息和证据.

3 安全性能分析

本设计重点针对非实名制网络出现的用户数据无法得到保障,网络违法、违规实践溯源困难等特点,在设计方案中给出对应的对策,并且在实用性上不额外增加用户的操作,达到透明化登陆的目的.对于本设计方案的安全性分析主要有以下几点:

3.1 存储状态的数据机密性分析

在本设计中,用户的隐私数据根据敏感的等级分为 CPI 和 GPI 全部保存在云端数据库中.其中 CPI 中存储的用户高敏感度隐私信息,且以密文的形式 $E_{PK_U}(\cdot)$ 的形式存在,因此针对存储状态的数据的攻击所需的时间复杂度等同于破解完全同态公开密码算法 $E(\cdot)$ 的时间复杂度.

而 GPI 是用户一般性隐私数据,此类数据与行为人的对应关系是多对多的映射 $n:m$,若对此类数据进行加密处理将为用户行为预警模块带来很高的计算复杂度,因此本设计中并无针对此类数据机密性保护.

分等级数据库保障用户隐私数据的设计保障在发生数据泄露事件之时,攻击者无法获得有用的隐私数据,可以有效防止攻击者获得用户的隐私数据进行违法犯罪活动.

3.2 传输状态的数据机密性分析

传输状态下的数据容易受到攻击者截获,常见的攻击手段有暴力攻击(Brute-Force Attack)、中间人攻击(The-Man-in-the-Middle Attack)和重放攻击(Replay Attack).以下从传输加密数据的两个模块分别分析传输状态的数据机密性.

(1)用户实名登记模块中,步骤①~③传输的均为单重加密或双重加密状态的 $E_{PK_U}(c_i)$ 或 $E_{PK_{IA}}(E_{PK_U}(c_i))$,透明化登录模块中,步骤①~③传输的均为加密单重或签名状态的 $E_{PK_U}(c_i)$,因此以上阶段 Brute-Force 攻击的时间复杂度破解完全同态公开密码算法 $E(\cdot)$ 的时间复杂度.

(2)用户实名登记模块中的步骤①～②和透明化登录模块中的步骤①～②可能受到 Replay 攻击,依靠 ElGamal 的特性,每次计算 Sign(U_{id})和 Sign($E_{PK_U}(c_i)$)选取的随机数 k 不同,因此生成的签名数值也不同,换言之,本设计利用随机数抗击重放攻击.

3.3　应用状态的数据机密性分析

本设计中的用户行为预警模块涉及应用状态的数据,在图 4 所示的步骤③～⑤中,通过设置处理规则,禁止针对某单一用户提取 GPI 信息,以防止攻击者利用 $GPI_{U_{id}}(g_i,\cdots,g_i,\cdots,g_n)$实施单一用户网络画像,杜绝间接泄露用户隐私数据的隐患.

GPI 中的某单一属性数据无法与用户的身份进行一一对应,因此 3.3 节中的规则 2 统一为请求者提供单一属性的 G_i 数值,同时也减轻了 DPC 的处理负担.

虽然 GPI 中的数据是用户一般性隐私数据,但如果获取了足够多的 GPI 同样存在将网络行为与行为人进行关联的风险,因此 3.3 节中的规则 3 和 4 就限制了请求者对多维 GPI 数据的获取权限.若同时请求的 GPI 属性数量超过了阈值 j,DPC 需对 GPI 数值进行脱敏处理后,仅反馈 g_i 所在的范围,以保障用户的隐私.

3.4　其他安全特性分析

本设计利用 RC,IA,PDSC 将用户的隐私数据使用权和管理权进行分离,同时网络服务提供商只能得到部分一般性隐私数据,从而杜绝因不同网络服务供应商的各种主观或客观原因造成的用户数据泄漏危机.

PDSC 中云存储服务器的数据始终以密文形式保存、传输和处理,就有效地避免了攻击者通过“拖库”或者“撞库”攻击后台数据库,也无法获得用户的隐私信息.此外本设计虽然将用户的额外操作降为最低,但随机启动的双重用户生物特征识别可提高系统的安全性,提高用户便捷使用实名制网络的同时开启完善的保障.

4　结　语

一个完善、安全的网络实名体系具有重大的实用价值,但网络本身的虚拟性和易失性增加了网络实名登记制度的工作难度,同时用户对隐私信息的保护和企业对数据价值的最大化利用更是一个看似无法调和的矛盾.本设计利用完全同态公开密码算法和分级存储隐私信息的思路保障对用户的隐私数据进行分级保护,以此达到用户隐私保障和数据价值利用双赢的平衡.此外网络实名制体系无法大规模实现的另一个难点就是实用性,本方案利用用户已实名登记过的手机实现透明使用系统,具有强可移植性的设计方便了网络实名体系推广至全国范围应用.此外借助 ElGamal 算法防止“拖库”“撞库”行为,通过与用户身份挂钩的网络行为进行数据分析,借助多层次异常检测实现网络行为异常预警,融合网络取证于一体实现电子证据的固定.完善的实名制体系可以帮助公安执法人员便利地维护网络治安环境、及时预警网络违法犯罪行为、迅速精确定位跟踪网络犯罪嫌疑人,对于打击和减少网络空间的违法犯罪行为提供了科学的平台和有力的工具,对于深化国家安全工作信息化的进程具有重大的推动作用.

参　考　文　献

[1]　杨婷.韩国网络实名制发展脉络研究及反思[J].西南政法大学学报,2018,20(2):102-111.
YANG T.Internet Real-Name System in S.Korea and Its Implications for China[J].Journal of Southwest University of Political Science and Law,2018,20(2):102-111.

[2]　董俊祺.网络治理的多方博弈:韩国网络实名制的历史回顾[J].南京理工大学学报:社会科学版,2016,29(3):87-92.
DONG J Q.A Multi-player Game in Internet Governance:An Analysis to South Korea's Internet Real-Name System[J].Journal of Nanjing University of Science and Technology(Social Sciences Edition),2016,29(3):87-92.

[3]　ADNAN M,LIMA A,ROSSI L et al.The Uncertainty of Identity Toolset:Analysing Digital Traces for User Profiling[C]//Proceedings of the 7th International Conference on Security of Information and Networks.New York:[s.n.],2014.DOI:10.1145/2659651.2659741.

[4]　JIE W,HAI-YAN L,BIAO C,et al.Application of educational data mining on analysis of students'online learning behavior[C]//2017 2nd International Conference on Image,Vision and Computing(ICIVC).Chengdu:[s.n.],2017.

[5]　SALDIVAR A A F,YUN L,CHEN W N,et al.Industry 4.0 with cyber-physical integration: A design and manufacture perspective[C]//International Conference on Automation & Computing.[S.l.:s.n.],2015.DOI:10.1109/IConAC.2015.7313954.

[6]　洪丹丹,罗军峰,冯兴利,等.基于 RSA 与 MD5 签名的实名制微门户设计[J].微电子学与计算机,2016(9):36-41.

HONG D D,LUO J F,FENG X L,et al.Design of a Real-name Wechat Portal Based RSA and MD5 Signature[J].Microelectronics & Computer,2016(9):36-41.

[7] 文勇军,李程,王键,等.基于教育电子身份号的关联认证[J].计算机科学与应用,2016,6(5):265-270.
WEN Y J,LI C,WANG J,et al.Education Electronic Identity Based on the Related Certification[J].Computer Science and Application,2016,6(5):265-270.

[8] HU W X,CHANG J J.Design and implementation of the Internet real-name authentication system based on public key infrastructure [C]//2010 International Conference on Management and Service Science.Wuhan:[s.n.],2010.

[9] 程琳,李明桂.网络实名制的隐私保护研究[J].信息安全与通信保密,2013,11(11):84-88.
CHENG L,LI M G.Privacy Protection in Internet Real-name System[J].Information Security and Communications Privacy,2013,11(11):84-88.

[10] 姚慧,马思研.人工智能在电信实名认证中的关键技术及应用[J].电信科学,2019,35(5):51-58.
YAO H,MA S Y.Key technologies and application of artificial intelligence in telecom real-name system[J].Telecommunications Science,2019,35(5):51-58.

[11] XU F,YAU K,ZHANG P,et al.A Privacy-Preserving Encryption Scheme for an Internet Realname Registration System[C]//Advances in Digital Forensics XI.[S.l.:s.n.],2015.DOI:10.1007/978-3-319-24123-4_7.

[12] 李晖,牛犇,李维皓.移动互联网服务的隐私保护机制[J].中兴通讯技术,2015(3):16-22.
LI H,NIU B,LI W H.Privacy Mechanism in Mobile Internet[J].ZTE Technology Journal,2015(3):16-22.

[13] 张梅舒,徐雅斌.多维数值型敏感属性数据的个性化隐私保护方法[J].计算机应用,2020,40(2):491-496.
ZHANG M S,XU Y B.Personalized Privacy Protection Method for Data with Multiple Numerical Sensitive Attributes[J].Journal of Computer Applications,2020,40(2):491-496.

[14] ELGAMAL T.A Public Key Cryptosystem and a Signature Scheme Based on Discrete Logarithms[J].1985,31(4):469-472.

[15] GENTRY C B.Fully Homomorphic Encryption[EB/OL].[2020-12-23].https://xueshu.baidu.com/usercenter/paper/show? paperid=efe7f4a1def57e42a8fba41ac66907a7.

[16] 魏世超,李歆,张宜弛,等.基于 E-t-SNE 的混合属性数据降维可视化方法[J].计算机工程与应用,2020,56(6):66-72.
WEI S C,LI X,ZHANG Y C,et al.Dimension reduction and visualization of mixed-type data based on E-t-SNE[J].Computer Engineering and Applications,2020,56(6):66-72.

[17] 杨先圣,姜磊,彭雄,等.基于大数据的异常检测方法研究[J].计算机工程与科学,2018,40(7):1180-1186.
YANG X S,Jiang L,Peng X,et al.A New Outlier Detection Method Based on Large Data[J].Computer Engineering&Science,2018,40(7):1180-1186.

[18] 吴昊.云性能监控中的智能采样方法与异常事件检测技术[D].杭州:浙江大学,2017.
WU H.Intelligent Sampling Method and Abnormal Event Detection Technology in Cloud Performance Monitoring[D].Hangzhou:Zhejiang University,2017.

[19] 刘垚磊,杨瑞,杨艺.基于 iForest 的虚拟机异常检测机制[C]//第 33 次全国计算机安全学术交流会论文集.广州:[出版者不详],2018.

[20] 黄谦,王震,韦韬,等.基于 One-class SVM 的实时入侵检测系统[J].计算机工程,2006,32(16):127-129.
HUANG Q,WANG Z,WEI T,et al.A Real-time Intrusion Detection System Based on one-class SVM[J].Computer Engineering,2006,32(16):127-129.

Research on privacy-preserving internet real-name system

Zhang Ping[1], Jiang Lin[2]

(1. Department of Network and Information Security, Guangdong Police College, Guangzhou 510230, China;
2. School of Computer Science and Technology, Harbin Institute of Technology, Shenzhen 518000, China)

Abstract: In China, the real-name system has been gradually realized in some fields, but a lots of problems are still to be studied and solved before the full implementation of the internet real-name system, especially the way to find the balance between privacy data protection and data value utilization. In this paper, using secret sharing algorithm, ElGamal digital signature and multi-level anomaly detection based on big data technology, we put forward a design of internet real-name system according to the condition of China. This design guarantees that the user's personal privacy information is protected from unauthorized access while minimizing additional operational overhead of the user, and provides a convenient tool for law enforcement authorities to combat Internet crime.

Keywords: privacy preserving; real-name system; Homomorphic eneryption; digital forensics; data mining

[责任编校 陈留院 赵晓华]