北京金融科技产业联盟
BEIJING FINTECH INDUSTRY ALLIANCE

数字银行场景安全技术解决方案研究报告（2023 年）

北京金融科技产业联盟

2024 年 9 月

版权声明

编制委员会

编委会成员：

聂丽琴　傅宜生　祖立军

编写组成员：

张　弛　薛文哲　门小骅　陈思文　赵晓夏　方宇伦　宋鑫晶
张明虎　卢　凯　李勇攀　杜　彪　卞　凯　董　涛　夏雯君
张　游　施生燊　张　宏　勾志营　王炳辉　陈　兴　吴小平
王银燕　黄海燕　李树尉　彭俊宏　陈　波　官小波　谢世杰
龚孟旭　王李戋　董杨瑞　孙　乐　廖敏飞　吴孟晴　解　敏
李裕鹏　施妍萍　郭俊刚　廖静雅　崔正玮　严青伟　陆绍益
丁伟强　李　浩　邹长龙　战　扬　张　艺　李　东　竺铁生
袁　捷　白　慧　方绍全　曾明华　李金银　卢科兵　肖　昊
周　丹　秦旭果　焦伟哲　牟健君　薛　涛　张嘉伟　杨增宇
张宪铎　沈　超　陈　俊　杜锦文　吴　杰　吴承荣　叶家炜
张　亮　谢于明　包德伟　魏启坤　曹雅琳　岐文钰　周　楠
杨学治　冯国强

编审：

黄本涛　刘昌娟

统稿：

薛文哲

参编单位：

北京金融科技产业联盟秘书处

中国银联股份有限公司

中国工商银行股份有限公司

中国农业银行股份有限公司

中国银行股份有限公司

中国建设银行股份有限公司

中国邮政储蓄银行股份有限公司

中国民生银行股份有限公司

上海浦东发展银行股份有限公司

兴业银行股份有限公司

华夏银行股份有限公司

中国光大银行股份有限公司

渤海银行股份有限公司

广东省农村信用社联合社

复旦大学

华为技术有限公司

深圳市联软科技股份有限公司

目　录

一、研究背景

中央金融工作会议指出做好科技金融、绿色金融、普惠金融、养老金融、数字金融五篇大文章，强调优化金融服务，防范化解风险，坚定不移走中国特色金融发展之路，推动我国金融高质量发展。数字银行是基于数字技术的金融创新发展模式，通过数据和服务的共享促进跨界协作与场景互联，已成为数字金融发展的重要支撑。在拓宽金融服务渠道、丰富金融服务场景、加速数据要素流动等方面具有得天独厚的优势。既能通过更全面的“数字足迹”为科创企业、绿色企业、小微企业等降低融资门槛、提升融资效率，也能借助无处不在的“全渠道”服务能力将金融服务延伸到老年人、残障人士、农村居民等普惠群体身边，有望在数字经济时代助力金融服务更广泛、更深入地融入经济社会的方方面面。

但数字银行在广泛连接服务提供主体、场景建设主体、交易发起主体等，客观上增加了网络攻击、数据泄露风险点，扩大了风险传导范围，链条上任何一方保护存在薄弱环节都可能危及金融资金安全、信息安全。风险主要体现在以下两个方面。

一是银行侧 API 安全风险。API 是目前数字银行各方互联的主要形式，随着银行对外开放的 API 数量增多、传输的数据价值越来越高，银行 API 已成为攻击者的重点关注对象。如何准确识别 API 攻击、有效开展 API 安全防护以规避以上风险，已经成为数字银行安全合规发展亟待研究的课题。

二是应用侧数据安全风险。在数字银行业务中，银行需在用户授权下与应用方进行敏感数据交互，但通常情况下应用方并不

是持牌金融机构，不具备金融级的数据安全防护、合规处理能力，无法确保数据处理过程符合监管要求。少数情况下，部分应用方甚至将数据分享给其他第三方而造成银行数据泄露，为数字银行业务开展带来挑战、为银行声誉带来负面影响。

二、总体研究框架

面对上述问题，亟需凝聚行业合力共同构建数字银行安全“防火墙”，护航数字银行发展行稳致远。本研究课题旨在研究一系列创新技术能力，保障数字银行的 API 安全、数据安全，主要包含银行侧的 API 异常行为检测等安全技术，应用侧的数据分类分级、数据脱敏、数字水印、安全前哨等数据保护技术（如图 1 所示）。截至发稿，本研究报告涉及的 API 异常行为检测（内容异常、序列异常）、数据分类分级、数据脱敏、数字水印均已完成原型实现和验证，安全前哨完成技术方案。

配合技术研究成果，目前课题组正同步开展相关技术的标准规范研制（截止本研究报告发稿，已完成《金融 API 安全防护体系评估指南》初稿）。以标准为基础，银联既可以为相关技术提供检测认证服务，也可以作为数字银行业务的转接方，为行业各方提供可靠的安全技术能力，保证整个交易链路的安全性。

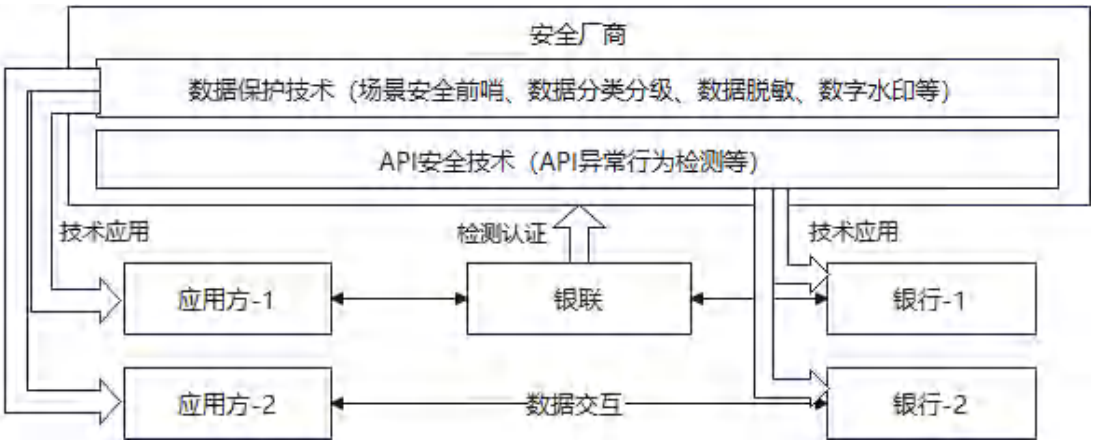


图 1 总体研究框架

三、应用程序接口异常行为检测

（一）研究背景

从全球银行业数字化转型发展来看，应用程序接口（API）是商业银行探索新金融服务模式的主要方式，第三方通过银行开放的 API 实现数据共享，基于银行的基础设施、产品和服务，在自身应用和场景中嵌入金融服务，以更好地满足客户需求。然而，银行现有的 WAF 和 API 网关等传统安全控制手段由于不了解 API 上下文并且不以正常的 API 使用为基准，通常难以对未知的、不可预测的攻击模式提供安全防护。

针对上述挑战，本章节提出基于自动编码器模型的 API 异常内容检测方法，以及基于注意力机制的双层长短期记忆（Long Short-Term Memory LSTM）网络模型对 API 的调用序列进行异常检测方法，商业银行可从原始的用户访问日志中提取用户会话内容与 API 调用序列，利用下述技术方案训练可靠的异常检测模型，检测用户会话行为是否存在异常，以提升银行开放网关的安全能力。

（二）技术实现方案

1. API 异常内容检测模型

使用深度学习模型可以更好地处理复杂的、非线性的数据模式，并能够自动地、自适应地学习数据中的特征，不需要人工预设规则或特征。因此，课题组提出了一种基于自动编码器的 API 异常内容检测方法。方法利用数据重构的思想设计模型，能自动学习日志参数的压缩表征，并通过正负样本的验证进行分界点的

动态划分，最后根据分界值进行异常日志参数的识别。此外，对于半结构化的日志，本方案采用基于Drain[1]的日志分析算法生成结构化的日志数据文件和模板文件。方案能有效解决包括但不限于以下异常事件：影子参数、请求方法异常、过度数据暴露等，从多维度保证了系统的稳定性、安全性，帮助提高性能降低风险。模型构建要点如下：

一是基于Drain的日志模板抽取：通过日志切分、分词、序列化、频繁项集挖掘算法、模板合并、日志映射，从大量的原始日志事件中识别出共同的结构，将半结构化日志文件转换为结构化的模板文件。

二是类型粒度的多模态数据表征策略：按照数据的种类对字段进行划分，将连续数值型、离散数值型、文本型的异构字段值统一转换为向量形式。

三是基于自动编码器的自动化压缩表征学习：通过编码器（Encoder）和解码器（Decoder）将样本数据先映射到低维空间再还原到高维空间，最小化重构误差使模型学习到准确的压缩表征以便更好地捕捉数据的特征和结构。

四是动态分界点划分：构建正负样本验证集，进行模型校验，并根据正负样本的重构误差中间值动态生成误差分界值，从而实现对不同日志定制化的阈值学习。

五是支持流水线自动运行，方案可复用性高：流程全自动化，无需配置，输入日志即可生成模型。支持对流数据的准实时判断，

[1] Drain指的是一种日志解析算法。Drain算法是一种用于日志分析的技术，它可以将大量的日志数据聚合在一起，并通过归类和过滤来发现错误和异常事件。Drain算法主要分为三个阶段：消息模板提取、消息聚类和消息推断。这个算法能够从原始日志中提取出结构化的模板，并将变化的部分用通配符代替，从而有效地识别和提取日志中的模式和异常

可以在不同行业不同业务的日志中迁移复用。

2. API 异常序列检测模型

课题组提出了一种基于自注意力机制的双层 LSTM 序列异常检测方法。通过对历史日志中 API 接口调用顺序的学习，模型可以实时检测出日志流中的异常序列调用，有效规避了开发过程中可能存在的未授权访问漏洞危害，让攻击者无法不遵循业务逻辑的调用顺序访问应用，从而降低了调用路径异常的业务逻辑风险。模型构建要点如下:

一是滑动窗口式序列生成策略: 将序列按设定的时间窗口划分形成多个子序列，依次输入包含门结构和记忆单元的网络，对序列的长短期信息进行自适应学习。

二是高维表征生成策略，长短期依赖信息动态平衡: 将序列按时间步划分，映射到高维空间，通过包含门结构和记忆单元的网络双层叠加效果，高度抽象序列中的深层次特征信息。

三是重点信息自动聚焦: 对序列的高维表示采用自注意力机制，即每个接口都需要计算与前后接口的关联度，在一定时间只关注与当前接口更相关的接口，大大增加了捕捉长距离依赖信息的能力。

四是支持手动配置异常容忍度，方案可复用性高: 通过配置预测候选集大小可以手动调节异常的容忍度，以上流程程序化，支持不同业务日志的迁移复用。

（三）测试结果

课题组根据以上技术方案开发了原型程序，并在银联沙箱环境中测试验证。首先，选取 100028 条网关应用系统上正常的用

户访问日志，经过特征提取和会话汇聚后共得到 17609 个有效的会话记录进行评估，作为正常样本。为了进一步开发和验证模型，使用 Postman[2]对 API 进行手动攻击测试，模拟包括爬虫、鉴权、SQL 注入、API Ddos 攻击等 API 攻击检测场景，将网关上记录到的日志作为异常样本。在这些样本中随机挑选 80%的正常样本数据作为训练集，各 10%的正常样本和 50%异常样本作为验证集、测试集。

根据实际场景需求，基于自动编码器的 API 异常内容检测模型对单 API 进行模板挖掘和关键参数阈值学习，从 http 方法（get、post）、主机域名（Host）、访问路径（path）、访问参数（url parameters、post body）等维度自动学习日志参数，能成功检测未知参数异常、请求方法异常、访问路径异常等内容异常，测试准确率[3]为 99.52%，召回率[4]为 99.28%，F1-score[5]为 99.40%。

基于注意力机制的双层 LSTM 网络模型基于会话的 API 序列基线，对一次完整的 API 会话行为的时间、序列、交互信息进行异常检测，测试准确率为 91.045%，召回率为 88.4%，F1-score 为 89.706%。

测试结果显示，课题组研制出的 API 异常行为检测算法，实现了在 API 内容数据与调用序列层面识别异常风险，解决传统方式中规则库需要不断更新维护的问题，降低了误报漏报率，能够保障数字银行开放 API 的安全。

[2] Postman 是一个功能强大的工具，它不仅可以用于 API 的开发和测试，还可以用于 API 的安全性测试和验证。

[3] 准确率表示的是所有预测异常中真实异常的百分比，用于衡量模型的查准能力；

[4] 召回率测量的是所有异常中真实异常的百分比，用于衡量模型的查全能力；

[5] F1-score 为查准率与召回率两者的谐波平均值，用于衡量模型的查全查准能力。

四、场景安全前哨

（一）研究背景

在数字银行业务中，银行将数据通过开放 API 的方式传输给应用方。根据《商业银行应用程序接口安全管理规范》应用安全责任章节要求“应用方不应将通过商业银行应用程序接口获得的金融服务能力与数据以任何方式转移、共享或分包给其他第三方”。为保证数字银行业务合规，商业银行在与应用方开展业务时，应确保应用方有意识和能力履行数字银行数据防泄漏及其他数据安全和隐私保护的要求。

目前各商业银行大多通过协议约定的方式，要求应用方执行数据安全的相关要求，但是银行对应用方的实际执行情况缺少有效审查。少部分银行会安排分行对应用方进行定期巡检，但是一方面，分行本身就是业务拓展方，又要做安全审计，无法平衡各方利益；另一方面，定期巡检也不是常态化的监控手段，且对巡检人员的专业性要求极高，分行员工往往难以执行到位。

为了有效解决商业银行在开展数字银行业务中面临的数据安全困境，课题组提出数字银行场景安全前哨的概念，希望通过技术手段配合管理制度，降低银行在开展数字银行业务的数据安全合规风险。数字银行安全前哨是部署在应用方的数据安全系统。为保障数据在应用方的安全性与合规性，安全前哨监控并记录应用方对银行数据的访问、存储、使用、传输的全过程。借助应用方在风险第一触点的优势，安全前哨可以建立起数字银行数据安全的关键防线。

安全前哨兼具了传统企业 DLP 解决方案的基础功能，在以下

方面更加精准地适配了数字银行的合规要求:

一是安全前哨是银行对应用方数据合规的监控设备，它的安全控制策略需要经过银行审核和检查，应用方的风险事件自动向银行报告;

二是安全前哨基于金融数据分类分级技术底座，针对不同数据类别、敏感等级配置精细化的管控策略;

三是安全前哨实现对数字银行共享数据在接收、存储、处理多环节、多系统交互过程的数据防泄漏，可封禁或监控的泄露点范围更加全面。

商业银行可向相关厂商采购符合技术标准要求的安全前哨解决方案产品，并部署于与其开展业务合作的应用方。对于一些自身有较强数据安全技术能力的应用方，商业银行也可要求其依据技术标准升级自身安全策略，以满足数字银行业务的数据安全要求。

（二）技术解决方案

1.整体技术框架

应用方对数字银行数据的处理架构可以抽象为数据接收、数据存储、数据处理三个环节，各处理环节间明文传输数据，安全前哨整体技术框架见图 2 所示:

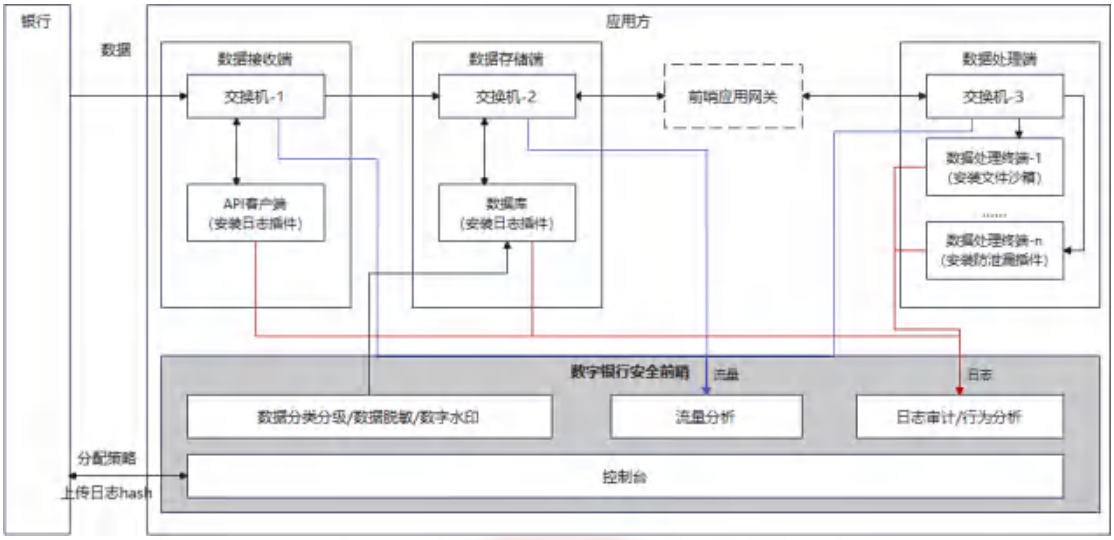


图 2 场景安全前哨技术方案

数据接收端：通常是应用方的 API 客户端，可以从银行开放 API 接收数据，并发送到数据存储端保存，或发送到数据处理终端做数据分析。数据接收端需要关闭数据接收端主机除网络之外的其他物理外发通道，如 USB、串口等，且接收数据和发送数据需要记录操作日志并上传日志审计模块。

数据存储端：通常是应用方的数据库系统。数据存储端需要记录具体的数据操作日志（如：XX 时间 XX 账号对 XX 数据进行了增/删/改/查的操作）并上传日志审计模块。根据数据敏感级别和业务诉求，在一些业务中应用方可能不存储数据，此时数据接收端直接将数据发送至数据处理端做业务处理（如向用户展示）。

数据处理端：通常是应用方的业务系统或个人电脑，能够访问和操作（增删改查）数据库中的数字银行数据。数据处理终端需要针对处理的数据敏感等级进行分级管理，需要关闭主机上除网络之外的其他物理外发通道，如 USB、串口等，需要对数据处理过程记录操作日志（如：XX 时间对 XX 类型数据做了 XX 处理

并发送给 XX 对象，数据量 XXM）并上传日志审计模块。终端分级管理要求如表 1 所示：

表 1 终端分级管理及安全措施

终端安全等级	可处理的数据等级[6]	安全措施
一级	1 级	**确保终端满足最基本的安全保护要求。** 安全基线：补丁更新检测、防病毒安装与更新检测、木马查杀软件安装与更新检测、弱口令、共享目录、Guest 账号。
二级	1-3 级	**监控所有外发数据行为。** 安全基线：补丁更新检测、防病毒安装与更新检测、木马查杀软件安装与更新检测、弱口令、共享目录、Guest 账号。 文件操作监控：支持对文件的读取、写入、复制、剪切、删除、另存为、新建、重命名的动作进行记录； 即时通讯外发文件监控：支持 QQ、企业 QQ、微信、企业微信、钉钉，飞书、skype 等外发文件的行为记录。 Email 发送文件监控：支持记录 email 外发文件行为（文件路径、发件人、收件人、抄送人）。 阻断外部连接：禁止光驱、优盘、蓝牙、红外、FTP 等网络共享、网盘上传等。 泄露溯源：截屏水印、打印水印、屏幕暗水印。 剪切板监控：复制粘贴模式下，可记录源文件的设备类型、文件类型、路径、文件名、文件大小和目的文件的设备类型、路径。
三级	1-5 级	**使用文件沙箱隔离数字银行敏感数据。** 具备文件沙箱功能：文件沙箱具备落地文件加密、内外网络隔离、剪切板隔离、进程保护、屏幕水印等安全保护等能力。数字银行数据文件下载到终端后即被加密，只可在沙箱里被读取、编辑或者其他处理。沙箱保护的数据外发需要经过严格审批。

场景安全前哨由如下功能模块组成：

数据分类分级：数据分类分级管理是数据安全的基础，此功能模块将数字银行数据依据《金融数据安全 数据安全分级指南》（JR/T 0197-2022）进行标准化的分类分级，帮助应用方实现数据的精细化管控，如更准确的安全级别防护。

[6] 根据行标《金融数据安全 数据安全分级指南》

数据脱敏：安全前哨可以根据配置的安全管理策略，对数据进行智能脱敏，帮助应用方在满足业务诉求的前提下进行适度脱敏的要求。

数字水印：安全前哨支持为银行数据添加数字水印，可以在发生数据泄露后进行溯源追责。

流量分析：由于安全前哨需要支持多系统数据防泄漏的能力，因此确定网络关系并能清晰洞察系统运行过程中的网络变化至关重要。流量分析支持检测网络通道、对端系统是否符合业务预先设定的配置。

日志审计：日志审计通过全面收集数据接收、存储、处理各系统的操作日志（包括运行、告警、操作、消息、状态等）并进行存储、审计、分析，能记录并显示数字银行数据的每一步操作和流动状况，并发现潜在安全事件与安全风险。同时，日志审计模块会将日志的哈希值实时上传至银行并保存最近 180 天的哈希值，以确保应用方日志的完整性和真实性。

行为检测：基于用户长期行为进行画像，通过偏离画像的异常行为发现数据处理人对数据的异常访问。

文件沙箱/防泄露插件：二级终端需要安装防泄漏插件，插件监控所有外数据行为。三级数据处理终端需要安装文件沙箱，下载数据强制保存在沙箱。沙箱与本地环境隔离，加密存储，杜绝拆卸硬盘、PE 启动系统恶意拿走数据。

应用网关：安全前哨通过应用网关实现数据库的访问控制。应用网关会校验终端用户身份和环境，符合要求的终端才能通过应用网关代理访问数据库，达到“读取数据先装插件，不装插件

数据库隐身”的效果。

控制台：由银行指导应用方配置安全前哨各功能模块的执行策略。

2.应用方数据处理流程

1）数据接收端：

一是流量分析模块配置数据接收端合法连接（如数字银行、应用方数据库等业务系统）的 IP 段和 port[7]信息作为白名单。

二是 API 客户端插件持续监控主机上物理通道使用情况，并上传日志信息给日志审计模块。日志审计模块判断是否有被禁物理通道向外部传输数据，如有，则记录告警日志。

三是流量分析模块通过交换机镜像流量采集报文信息，如时间、报文数、字节数、目的 IP、Port 等，并对报文负荷做敏感数据识别，获取数据类型、数据安全级别等信息上报给日志审计模块。

四是日志审计模块对报文信息和安全信息进行检测，判断报文目的 IP 和 port 是否在白名单范围内，如否，则记录告警日志。

五是日志审计模块分析接收流量的业务系统是否接收了超过本系统安全等级的数据，如是，则记录告警信息。

2）数据存储端：

一是日志审计模块收集各终端插件上报的终端 IP 及授权安全等级。

二是流量分析模块通过交换机镜像流量采集报文信息，如时间、报文数、字节数、目的 IP、Port 等，并对报文负荷做敏感

[7] port 指的是计算机网络通信中的端口号

数据识别，获取数据类型、数据安全级别等信息上报给日志审计模块。

三是日志审计模块对报文信息和安全信息进行检测，判断报文目的 IP 和 port 是否在白名单范围内，如否，则记录告警日志。

四是日志审计模块分析接收流量的业务系统是否接收了超过本系统安全等级的数据，如是，则记录告警信息。

3）数据处理端:

一是数据处理终端通过应用网关接入数据库获取数字银行数据，数据出库时根据数据处理终端的用户身份添加对应水印。对于三级管控终端，获得的数据需要保存在文件沙箱。

二是数据处理终端插件持续监控主机上物理通道使用情况，如发现被禁物理通道向外部传输数据，则作为告警日志上报给日志审计模块。

三是数据处理终端插件持续监控终端对数字银行数据的操作，如邮件外发、即时通讯、文件改名、进程访问及进程使用的 IP/Port 等，并上报终端行为日志给日志审计模块。

四是流量分析模块通过交换机镜像流量采集报文信息，如时间、报文数、字节数、目的 IP、Port 等。

五是对非加密流量，流量分析模块对报文负荷做敏感数据识别，获取数据类型、数据安全级别等信息上报给日志审计模块;对于加密流量，根据源 port 确定是哪个进程，把本次进程外发信息上报给日志审计模块。

六是日志审计模块分析数字银行密级数据报文是否发送给

了合法接收人（外网，无权限内网终端），如否，则记录告警日志。

七是日志审计模块分析目的终端接收了超过本终端安全等级允许接收的数据，如是，则记录告警日志。

八是日志审计模块长期收集数据操作行为进行数据泄露行为建模。如终端行为日志偏离模型基线，则记录告警日志。

3.自身安全措施

为了保证安全前哨的正常运行，安全前哨需要连接数字银行服务方的 IT 系统定时通报运行情况。连接需要通过 https 双向 SSL 证书认证，上报数据执行两个原则：

一是监督安全前哨在应用方的正常运行情况并上报一些严重的异常事件。

二是上报异常事件不应泄露应用方敏感信息，比如只上报发生时间、事件类型，不上报具体数据、人员、设备等信息，包括但不限如下内容：

——定时心跳，每隔 1 小时上报一次。

——日志 hash 数据。

——安全前哨和银行的网络连接被中断。

——安全前哨的运行被中断，比如断电等。

——终端插件的运行被中断，比如被卸载、进程被杀死等。

——严重数据泄露行为事件。

——违禁端口被打开，如 API 客户端所在服务器或者终端上的 USB、串口被打开。

五、智能化数据分类分级算法

（一）研究背景

数据分类分级管理是数据安全的基础。只有在准确识别数据类别及其安全级别的前提下，才能进一步明确数据保护对象，有的放矢地处理数据，避免数据保护资源的浪费。

现有数据分类分级方法主要存在如下三方面局限：一是采用传统人工定级的方法，通常参与人员多，耗时周期长；二是基于预置规则的自动化分类分级方法，受专家知识和历史数据的限制，分类规则较少，覆盖度和识别准确率不高；三是现有数据分类分级方法的拓展性不高，处理多源数据的能力较差。例如，不同机构的数据字典往往各不相同，造成同一数据在不同机构和不同业务系统中命名不同的普遍现象。由于各机构数据字典一般不对外公开，预置一个全面整合所有机构的数据规则是不现实的。如何构建一个相对完整、合规的分类分级规则库，并且能够在处理多源数据中自动拓展规则库，从而持续提升数据分类分级的性能和准确率，是现有数据分类分级方法迫切需要解决的技术难题。

课题组提出了相较于主流水平更加完善的敏感数据分级分类规则库，其优势是联合行业各方共享共建，规则库更全、准确率更高。数字银行的参与方如商业银行、应用方，可在数据传输、存储、处理等全生命周期实施统一的数据分类分级管理，更好地落实数据安全监管要求。

（二）技术实现方案

算法主要包括三大模块：分类分级规则库的构建模块、数据识别模块以及自动增广模块。

1. 数据分类分级规则库构建模块

为保证规则库在金融行业的合规性和普适性，依托人民银行制定的行业标准《金融数据安全 数据安全分级指南》（JR/T 0197-2022）构建规则库。该标准根据金融业机构数据安全性遭受破坏后的影响对象和所造成的影响程度，将数据安全级别划分为 5 级。重要数据以及一旦安全性遭受破坏将直接影响到国家安全、社会秩序、公众利益与金融市场稳定的金融数据，其安全等级应不低于 5 级。

该标准中的附录 B 给出了“金融业机构典型数据的定级规则参考表”，包括 4 个一级子类、13 个二级子类、71 个三级子类和 279 个四级子类。我们利用文本挖掘技术从四级子类的内容描述中提取规则库的规则名称。例如，四级子类“个人基本概况信息”的安全级别为 3，从其内容描述“指个人基本情况数据，如个人姓名、性别、国籍...”中提取出姓名、性别、国籍三个规则，三个规则对应的分类设置为个人基本概况信息、安全级别设置为 3 级。将文本挖掘的所有规则组合在一起就构成了数据分类分级规则库，结构如表 2 所示。

表 2 数据分级分类规则库示意

规则编号	规则名称	数据分类	安全等级	特征项	敏感词	关键词

除了“规则名称”“数据分类”“安全等级”三个参数外，为提高数据识别的效率，规则库还设计了三个参数：（1）“特征项”即规则的特征，比如规则“入学日期”“毕业日期”“证件生效日期”拥有共同的特征项“日期”；（2）“关键词”用于匹配结构化数据中的字段名，命名相对规范；（3）“敏感词”用于定位非结构化数据中的敏感信息。

对于数据形式具有明显规律的特征项，进一步构建了特征项库，结构如表 3 所示。其中，“正则项”是用于匹配该特征项的正则表达式；“字典”是用于匹配该特征项的词典的名称，如机构名称词典、银行名称词典等；“校验和”是一个布尔值，用于说明该特征项是否有公开的校验函数，例如身份证的校验。

表 3 特征项库示意

特征编号	特征项名称	正则表达式	特征词典	校验和

2. 敏感数据识别模块

敏感数据识别模块在规则库、特征项库的基础上，借助命名实体识别（Named Entity Recognition, NER）模型识别敏感数据，最终输出敏感词、敏感类别和敏感等级。金融敏感数据分类分级的流程如图 3 所示。

图 3 分类分级规则库处理流程

当待识别的数据字段与规则库的中的关键词不直接匹配时，

参考文献 Shen et al. (2019)[8]和何文竹（2009）[9]通过构建分类器计算待识别数据字段隶属不同规则的概率（公式 1）和离散信息量（公式 2），综合判断其对应的规则类别。

$$P(t_{ij} \in S_i) = \phi(q(S_i \cup t_{ij}) - q(S_i)) \quad \text{（公式 1）}$$

$$H(X) = -\sum_{x \in A} P(x) \log P(x) \quad \text{（公式 2）}$$

公式中：关键词集为S_i，候选词为t_{ij}，$q(\cdot)$是集合的量化函数，$\phi(\cdot)$是 sigmoid 函数。$0 \leq P(x) \leq 1$，$\sum_{x \in A} P(x) = 1$，$P(x)$是每个离散信息发生的概率

算法中的 NER 模型包括三步：特征表示、特征编码和标签解码。其中，特征表示采用的是 word embedding（Word2Vec），特征编码采用的是双向长短期记忆网络（Bi-LSTM），标签解码是采用的是条件随机场（CRF）。

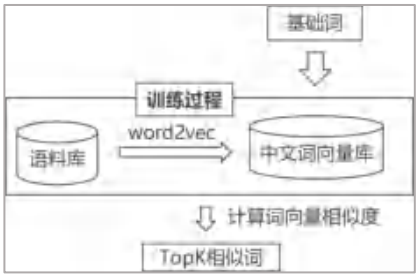
3. 利用 NLP 技术的增广模块

为解决同一数据命名不同（比如薪水可以被命名为薪资、薪酬、工资等多种方式）、存在缩略词以及特征词典覆盖面不全等造成的规则匹配困难的问题，提出了基于同义词库、基于上下文语义和基于模式的规则库自动增广技术（如图 4 所示）。综合三种方法来实现规则库在面对多源数据时规则的自动增广，极大节省人工拓展规则的成本。

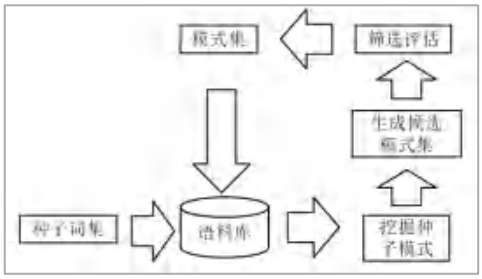
[8] Shen, J., Lyu, R., Ren, X., Vanni, M., Sadler, B., & Han, J. (2019, July). Mining entity synonyms with efficient neural set generation. In Proceedings of the AAAI Conference on Artificial Intelligence (Vol. 33, No. 01, pp. 249-256).

[9] 何文竹. 敏感数据的智能识别算法及自适应保护模型研究. 2020. (硕士论文, 贵州大学).

(a) 基于同义词库



(b) 基于上下文语义



(c) 基于模式

图 4 规则库自动增广方法

算法设计中用到的同义词库包括：公开的金融数据集，如监管机构提供的统计数据中的字段名、银行公开的接口信息，哈工大词林扩展版和 HowNet[10]。同义词增广简单便捷，但是忽略了词汇所在语料的上下文信息，基于语义相似度的词向量法可以弥补这一不足。词向量是将文本特征映射为数值向量的函数映射关系，从而将词汇之间的相似性通过词向量间的距离体现。算法设计中主要用到的是 word2vec[11]实现对规则的增广。另外，对于专业词汇，算法设计了基于模式的增广方式，即构建种子词集挖掘本名和别名在百科语料库中出现的模式，生成模式集，再通过模式集去语料库中挖掘新的本名别名词对，如此反复迭代达到增广的目的。

（三）测试结果

课题组根据以上技术方案开发了原型程序，并联合工行（涉及技术部、软件开发中心两个部门）、农行、建行、中行、邮储、民生、广东农信进行了测试验证，测试数据类别广泛覆盖银行存

[10] HowNet 是指一个大规模的中英文双语语义知识库。

[11] word2vec 是一种用于生成词向量的技术。它通过在大规模文本数据上训练神经网络模型，将单词转换为高维空间中的向量表示，这些向量能够捕捉单词之间的语义和语法关系。

款、贷款、外汇的业务数据、经营数据、监管数据以及用户敏感信息等多种数据类别，测试情况具体如表 4 所示：

表 4 部分银行测试结果

银行	测试场景	数据量	测试结果
中国工商银行（一轮）	个人征信平台：客户数据、贷款等业务数据、经营管理数据、监管数据	总字段 192	准确率>91.67% 召回率>92.31%
中国工商银行（二轮）	银企数据、缴费数据、银证数据、客户数据，票据数据、国库数据、ALTAS 支付数据	总字段 894	准确率>90.9% 召回率>95.5%
中国银行	外汇数据、存款数据、卡数据、进口数据	总字段 248	准确率>88.9% 召回率=100%
邮储银行	用户订单交易数据、合作方数据等	总字段 108	准确率>95% 召回率>95%
广东农信	客户数据	总字段 389	准确率=100% 召回率>98.73%
中国建设银行	客户认证数据	总字段 98	准确率=100% 召回率=100%
中国民生银行	贷记来账表、客户金融资产信息表，投保人信息表	银行测试，未反馈具体测试数据信息	准确率>90% 召回率>90%
中国农业银行	贷款申请人信息，缴费记录、交易记录	银行测试，未反馈具体测试数据信息	准确率>90% 召回率>90%

测试结果显示，课题组研制的敏感数据识别算法（目前支持 1000 个以上的字段类别）能广泛识别银行各种数据类型，且测试准确率（除中行外均在 90%以上）、回调率（均在 90%以上）基本能达到 90%以上的理想水平，属于行业领先水平。

从性能上来看，识别具有 150 个数据字段、20000 条记录的数据集，耗时约为 500 秒，性能高低主要是受数据集中数据字段规模大小的影响，如图 5 所示。

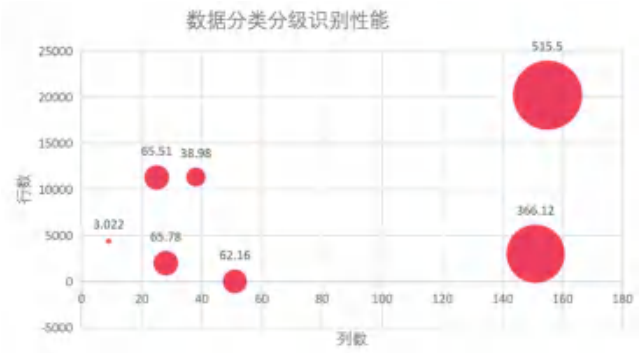


图 5 数据分类分级识别性能

六、数据脱敏效果综合评估体系

（一）研究背景

数字银行的业务和产品中产生了大量包含个人敏感信息的数据，根据《GB/T 35273-2020 信息安全技术 个人信息安全规范》《JR/T 0197-2020 金融数据安全 数据安全分级指南》等标准对数据安全的要求，在对数据进行使用时需要对个人信息进行去标识化处理，通过脱敏等算法或规则进行数据的变形，保护隐私数据的安全。

然而上述规范仅要求在特定场景下做数据脱敏，但没有对脱敏的程度进行明确要求。从行业调研来看，银行内对数据脱敏算法已有了较充分的工作投入与技术储备，但还缺少多维度的评价方法来对脱敏的效果进行评估。不同使用场景和不同用户对数据使用的业务需求与安全需求也不尽相同。使用数据的业务方期望在满足合规要求下只经过低程度的脱敏处理，不希望复杂的脱敏导致数据不可用；管理数据的技术方期望确保数据安全，最好能彻底脱敏，完全不能识别出个人。这是数据可用性和数据安全性之间的权衡，也是业务方与技术方之间的权衡。由于目前业界主要是对匿名化数据的重识别风险进行研究，缺少能够对去标识化

信息进行差异化评价的、多维度的综合评价脱敏效果的量化的方法体系。

针对上述挑战，本课题提出一种数据脱敏效果评估的方法，从有效性、安全性、可用性与合规性四个维度定量评估，结合了量化距离计算、重识别风险模型与数据集信息论模型等，帮助数字银行业务的参与方如商业银行及应用方实现对数据脱敏效果进行定量评估，在平衡业务需求和安全需求下选择合适的脱敏算法对数据脱敏。

（二）技术实现方案

本方案综合考虑对数据脱敏的目的与可能的数据挖掘使用目的，设计了包含有效性、安全性、可用性与合规性四个维度定量评估体系。从有效性维度评估脱敏效果可以衡量脱敏算法对数据集的处理程度，从安全性维度评估脱敏效果可以衡量数据集的个人身份信息泄露风险，从可用性维度评估脱敏效果可以衡量经过脱敏后的数据集可使用价值，从安全性维度评估脱敏效果保障满足相关法律法规要求。以上四个维度的评估指标能更好地帮助企业在不同的业务场景下与合规要求下选择适合的脱敏算法，对数据进行监控与管理。

1. 数据脱敏有效性评估指标

数据脱敏是对敏感数据进行变形处理，其目的是保护隐私数据等信息的安全，因此定义数据脱敏的有效性为数据敏感性的去除程度。通过量化评估字段数据项差异进行评估，计算得到的距离越大，代表脱敏程度越大。例如同样对手机号使用掩码算法进行脱敏，掩盖后四位和后八位所得到的数据涵盖的敏感性并不相

同。为了量化脱敏前后数据字段间的距离，针对不同脱敏算法及字段类型本方法提出了基于莱文斯坦距离的脱敏程度量化、基于汉明距离的脱敏程度量化、基于泛化树的脱敏程度量化、基于差值的脱敏程度量化与基于公共子序列的脱敏程度量化，有效性指标的计算流程如图 6 所示。

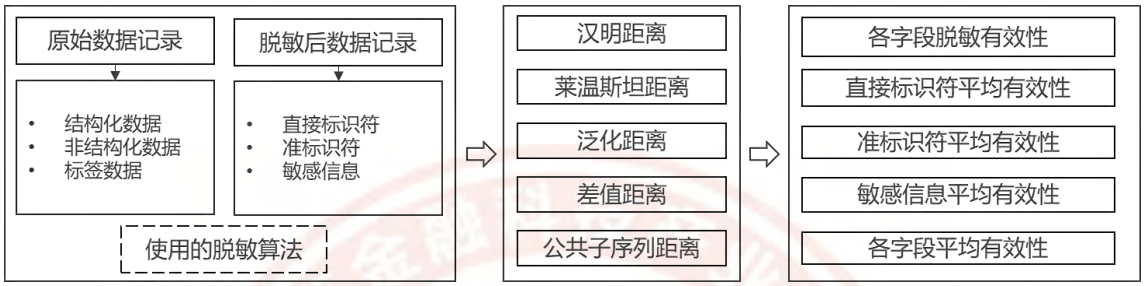


图 6 有效性指标计算流程

本方法中针对字段的差异性评估方法如表 5 所示，对于每个字段的计算结果均进行了归一化以提高可比性，计算结果越接近 1，代表字段的脱敏程度越大，反之越小。

表 5 脱敏效果有效性评估方法

脱敏算法	字段类型	Y 有效性评估方法
掩码	结构化字符串/非结构化字符串：姓名，号码，邮箱，地址……	Dis=汉明距离(x,y)/len(x) [取值范围 0~1]
局部混淆/随机替换	结构化字符串/非结构化字符串：姓名，号码，邮箱，地址……	Dis=汉明距离(x,y)/len(x) [取值范围 0~1]
替换/重写/加密	标签类数据：证件类型……	Dis=莱文斯坦距离(x,y)/len(x) [取值范围 0~1]
泛化	数值类型：重量，数量，金额……	Dis=泛化后（right-left）/原数据集（max-min）
	非结构化字符串：职位，街道……	脱敏时定义好泛化规则和距离计算规则（利用泛化树）
偏移	数值类型：重量，数量，金额……	Dis=|num1-num2|/num1 [取值范围 0~1]
	日期类型：日期，期限……	Dis=1 - 最长公共子序列/len(x) [取值范围 0~1]
取整	数值类型：重量，数量，金额……	Dis=|num1-num2|/num1

		[取值范围 0~1]
	日期类型：日期，期限……	Dis=1 - 最长公共子序列/len(x) [取值范围 0~1]
规整	数值类型：重量，数量，金额……	Dis=泛化后（right-left）/原数据集（max-min）

一是基于汉明距离的脱敏程度量化

针对脱敏前后数据长度未改变的结构化或者非结构化数据类型，通过遮盖或者改变替换原信息的脱敏情况设计了基于汉明距离的脱敏有效性评估。汉明距离（Hamming distance）为两个等长字符串之间的汉明距离是两个字符串对应位置的不同字符的个数。可以评估字符串类型数据的脱敏程度，计算公式如（公式 3）所示:

$$d(x,y)=\frac{\sum x_i \oplus y_i}{length(x)} \quad \text{（公式 3）}$$

其中 x 为 i 位脱敏前的字段，y 为 i 位脱敏后的字段，length 为计算字符的长度，⊕代表异或计算。

二是基于莱温斯坦距离的脱敏程度量化

针对脱敏前后数据长度改变的结构化或者非结构化数据类型，通过替换等方法改变原信息的脱敏情况设计了基于莱文斯坦距离的脱敏有效性评估。莱文斯坦距离（Levenshtein Distance）为两个字符串之间莱文斯坦距离指的是将一个字符串变为另一个字符串需要进行编辑操作最少的次数。其中，允许的编辑操作有替换、插入和删除。可以评估字符串类型数据的脱敏程度，计算公式如（公式 4）和（公式 5）所示:

$$d(x,y)=\frac{lev_{x,y}(i,j)}{length(x)} \quad \text{（公式 4）}$$

$$lev_{x,y}(i,j)=\begin{cases}\max(i,j) & if\ \min(i,j)=0\\ min\begin{cases} lev_{x,y}(i-1,j)+1 \\ lev_{x,y}(i,j-1)+1 \\ lev_{x,y}(i-1,j-1)+1_{(x_i\neq y_j)}\end{cases} & otherwise\end{cases}$$ （公式 5）

其中 x 为 i 位脱敏前的字段，y 为 j 位脱敏后的字段，length 为计算字符的长度。

三是基于泛化树的脱敏程度量化

对于使用泛化进行脱敏的数值类型字段设计了计算其泛化后区间占比的方法进行有效性评估。对于使用泛化进行脱敏的非结构化类型字符字段设计了基于泛化规则构建泛化树的有效性评估方法，计算公式如（公式 6）：

$$d(x,y)=\frac{h(i,j)}{H}$$ （公式 6）

其中 x 为泛化前字段，y 为泛化后字段，i 和 j 对应 x 和 y 在泛化树中的节点位置，h（i，j）为以 i 和 j 结点为最小公共祖先为根的子树高度，H 为泛化树高度，以地区泛化为例的泛化树结构如图 7 所示。

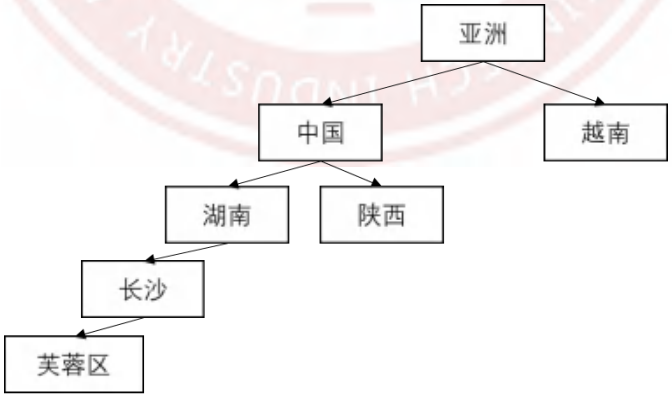


图 7 地区泛化树实例

四是基于公共子序列的脱敏程度量化

针对经过偏移或者取整的日期类型字段，设计了基于最长公共子序列的有效性评估方法，在结合语意信息的基础上进行脱敏

程度的计算。最长公共子序列（Longest-Common-Subsequence）是两个字符串中共同的最长子序列长度。

五是基于差值的脱敏程度量化

对于经过泛化、偏移或取整等脱敏算法的数值类型字段，设计了基于计算改变的差值比例评估脱敏程度。

在对各字段计算脱敏程度有效性后，企业可根据数据集实际使用场景或者数据集特征选择输出特定字段组合的平均有效性。

2. 数据脱敏安全性评估指标

安全性指标计算首先确定准标识符的等价类集合，结合使用场景计算 K 匿名重识别风险。风险值越小代表安全性越高，反之越低。

使用 K 匿名隐私保护模型，具有相同准标识符的记录构成一个等价类，确定等价类集合 J，以及每个子集的大小 f_j，则每个子集的重识别风险为:

$$\theta_j = \frac{1}{f_j} \quad \text{（公式 7）}$$

其中每个子集内记录的准标识符相同。计算数据集的重识别风险，可用总体风险、最大值或者平均值来代表。

$$R_a = \frac{1}{|J|}\sum I(\theta_j > \gamma) \quad \text{（公式 8）}$$

$$R_b = \max(\theta_j) \quad \text{（公式 9）}$$

$$R_c = \frac{1}{|J|}\sum \theta_j \quad \text{（公式 10）}$$

针对数据集可能面临的不同风险攻击情形，可以选择使用不同的重识别风险指标计算方式。

3. 数据脱敏可用性评估指标

金融行业内数据集通常用于数据挖掘、建模分析或特征提取等用途，经过数据脱敏后的数据集相比原数据集会损失一定的信息量和可用性。当数据集的脱敏程度很大改变了原数据集的分布特征时数据集的可用价值就变小了，本方法基于优化信息熵理论来评估数据集脱敏后的可用价值，信息熵可以识别单字段包含的信息量，但脱敏的数据集各字段间通常还包含关联关系，因此设计了基于联合条件熵的方法对数据集的可用性进行评估。数据脱敏过程中信息的变化如图 8 所示。

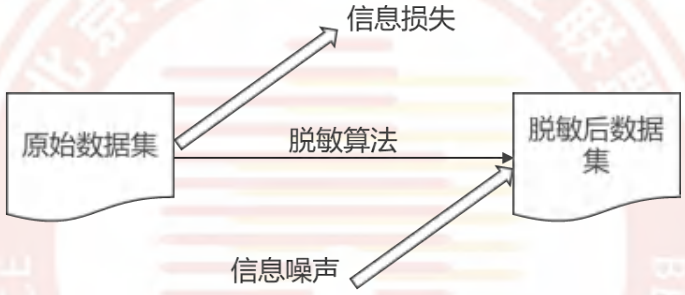


图 8 数据脱敏信息变化

信息量度量数据集携带的信息，熵是可能携带的信息量的期望，通过计算信息熵可用评估数据集的信息使用价值，事件的信息熵计算公式如（公式 11），p_i为事件 i 的概率。

$$H = -\sum_{i=0}^{n} p_i log p_i \qquad （公式 11）$$

对于评估包含多个字段（例如：准标识符）的数据集信息熵首先计算其联合熵，数据集的联合熵计算公式如（公式 12，其中 A、B 为数据集中的字段，$p(a_i, b_j)$为 A 与 B 分别取值为a_i, b_j的概率，联合熵能表示出不同字段数据项间的可用关联信息。

$$H(A,B) = -\sum_{i=0}^{m}\sum_{j=0}^{n} p(a_i, b_j) log p(a_i, b_j) \qquad （公式 12）$$

评估数据脱敏效果的可用性需要排除数据集本身的影响，企业可根据实际使用场景或数据集特征选择条件熵，熵变比来评估脱敏效果，令 X 代表脱敏前数据集，Y 代表脱敏后数据集，条件熵为 H(X|Y)，熵变比为 H(Y)/H(X)。其中条件熵是在已知信息或数据集分布的基础上获取另外一个信息或数据集时所获得的信息量，条件熵的计算公式如（公式 13。

$$H(X|Y) = -\sum_{y \in Y} p(y)H(X|Y=y) \quad \text{（公式 13）}$$

本方案针对脱敏数据集包含多个字段，字段间可能存在关联关系的特点设计了联合条件熵来对脱敏后数据集的可用性进行评估。

4. 数据脱敏合规性评估指标

对于金融行业去标识化场景，相关法规要求在不借助外部信息的情况下不能识别出特定的自然人，因此需要对脱敏后的数据集进行二次评估，判断直接标识符和准标识符是否全部完成脱敏。利用金融行业敏感数据识别工具和标识符识别工具判断脱敏后的数据集是否合法合规。

（三）测试结果

测试方法为课题组提供测试程序，渤海银行、建设银行自主进行测试。

渤海银行使用加密算法脱敏后的数据进行测试，测试结果见表 6，结果显示在加密算法进行脱敏时，脱敏评估体系中的有效性指标能够较好体现出脱敏前后数据的差异性改变，可用性指标由于加密算法从而计算指标均为 1，符合评估算法的设计逻辑。

表 6 渤海银行测试结果

有效性（直接标识符）	安全性（重标识风险）	可用性（准标识符）
1	0.01	1

建设银行使用测试环境的订单数据进行测试，测试结果见表7，分别在轻度脱敏、中度脱敏和重度脱敏场景下进行结果评估，总体评估结果符合预设评估算法的逻辑。

表 7 建设银行测试结果

脱敏程度	有效性（直接标识符）	安全性（重标识风险）	可用性（准标识符）
轻度	0.1	1	1
中度	0.52	1	1
重度	0.95	1	1

评估指标中的有效性能较好的反映出脱敏算法对数据的改变程度，改变程度越大则计算出的有效性指标越大。

可用性指标与原始数据集数据分布程度关联较大，可以在原始数据集维度上评估不同脱敏算法或者参数下对数据集的熵分布的改变，但无法建立通用的普适性的基线。

安全性与数据集的关联程度较大，脱敏后只要准标识符存在一个唯一值，则风险为 1。

七、基于语义分析的开放文档格式隐式水印算法

（一）研究背景

与国外办公软件的发展次序类似，我国在 20 世纪 90 年代之前主要是流式软件。2010 年前后，国家软件版权保护政策连续落地，开始自主版式文档标准的编制工作。2016 年开放文档格式（Open Fixed Document Format OFD）自主版式文件的国家标准正式发布《电子文件存储与交换格式－版式文档》（GB/T 33190-2016），OFD 在的应用由此开始加速发展（见表）。2021 年，财政部会同国电联办起草《电子凭证-银行回单标准》，明确电子回单版式文件应使用 OFD 存储；2022 年，财政部联合多部委开

展了电子凭证会计数据标准试点工作，其中《银行电子凭证技术规范》（征求意见稿）中鼓励将 OFD 作为银行电子凭证文件的格式，并提出了相关技术安全要求防止数据篡改和泄露，我国 OFD 标准格式推进过程如表 8 所示。

表 8 我国 OFD 标准格式推进过程

时间	内容
2011.08	形成《电子文件存储与交换格式 文书类版式文档》（OFD 标准）草案并在全国试点试用
2016.10	OFD 作为国家标准正式发布
2016.12	《党政机关电子公文格式规范》中明确要求“电子公文的承载格式为 OFD”
2018.11	《电子证照系列国家标准》中明确规定“电子证照使用 OFD 格式”
2020.01	《关于增值税发票综合服务平台等事项的公告》中，明确指出“增值税电子普通发票版式文件格式为 OFD 格式”
2022.02	《关于加快推进电子证照扩大应用领域和全国互通互认的意见》
2022.09	《银行电子凭证技术规范》金融行业标准（征求意见稿）中鼓励将 OFD 作为银行电子凭证文件的格式。

从行业调研来看，目前 OFD 主要通过电子签名技术防止数据被篡改，在防止数据泄露方面则依赖信息环境的安全。数字水印技术研究能有效提升 OFD 版权保护与数据泄漏后溯源追责能力，具有较高的创新性和前瞻性。

数字水印是指将特定的信息嵌入文本、音频、图片或是视频等载体中，当拷贝分发带有数字水印的数据时，嵌入的水印信息也会被拷贝，以起到版权保护、秘密通信、数据文件的真伪鉴别和产品标志等作用。数字水印可分为显式和隐式两种。显式水印可以起到直观告知数据使用者数据所有权、使用范围等信息，但也容易被识别和去除。隐式水印利用特定算法将水印信息通过不可见的方式隐藏于数字载体中，不容易被人察觉，也不会破坏原

数据使用价值与视觉效果。本课题研究的是隐式水印。

在 OFD 中添加数字水印信息包括数据发送者和接收者信息、分发目的、数据用途、版权归属等信息。与电子签名不同的是，电子签名关注发送者的身份认证，用于保证信息传输的真实性、完整性，防止伪造、抵赖、冒充、篡改；数字水印更关注接收者身份及数据用途，解决数据泄露后溯源而非数据完整性问题，水印还可以在接收方转发数据的过程持续叠加新的水印信息，记录传输链中所有数据接收方。

在数字银行中，银行数据在多个第三方合作机构之间共享使用，一旦发生数据泄露，银行可以通过解析水印及时确定泄露途径、泄密机构，有效解决了目前只能依靠合同协议约束应用方保护数据，出现安全事件难以取证追责的困难。

（二）技术实现方案

OFD 标准基于可扩展标记语言（Extensible Markup Language，XML）对版式进行描述。OFD 采用“容器+文档”的方式描述和存储数据，文档的内容由 zip 包内的多个文件共同决定。一个 OFD 文件的内部基础构成如图 9 所示。

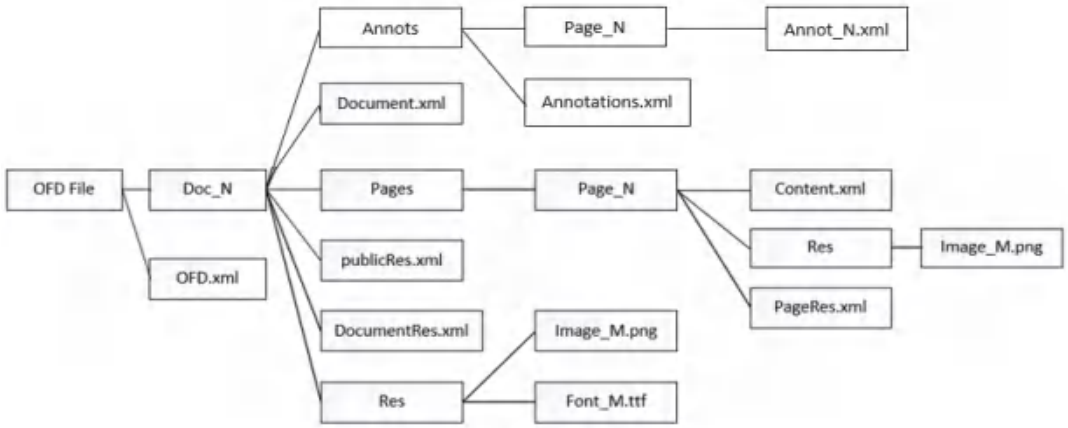


图 9 OFD 文件结构

本算法首先需要解析待添加水印的 OFD 文件，提取所有文件

夹名、文件名；解析 OFD 结构中 xml 文件字段，提取结构体及属性名。

然后，利用词向量模型中的连续词袋模型训练上述文件夹名、文件名、结构体及属性名，得到语料高位空间分布，根据距离输出与输入特征相近的语料词，作为仿真语料。

接着，根据 OFD 结构文件中原本的语句特征，生成伪文件夹、伪文件（包含仿真语句的结构体，如图 10 所示）。水印信息经过转换后将写入仿真语句中。

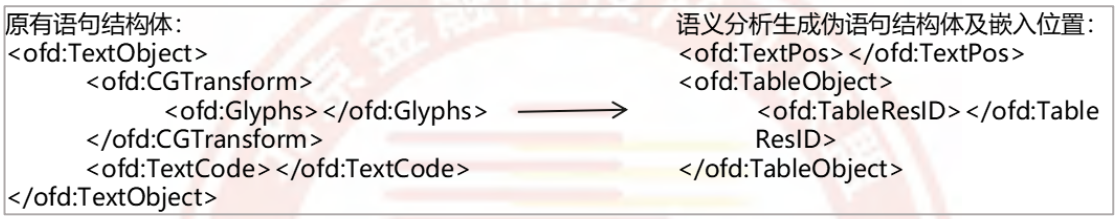


图 10 生成 OFD 的伪造结构体

图展示了基于伪结构体的 OFD 隐式水印算法嵌入与提取流程，如图 11 所示。

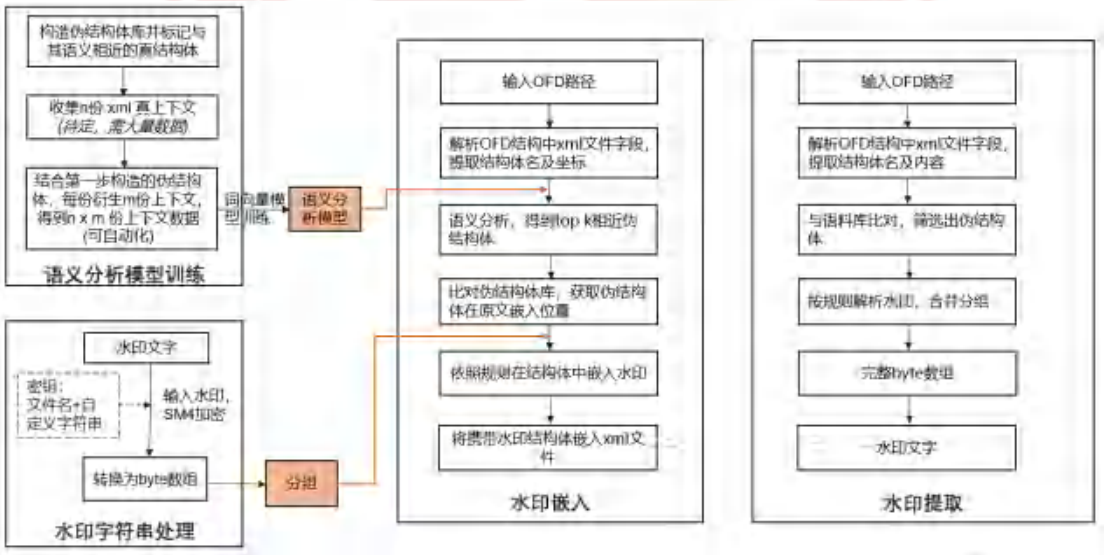


图 11 基于伪结构体的 OFD 隐式水印算法流程

此外，还可以将水印信息依照零宽字符[12]规则表，转换为零宽

[12] 零宽度字符是一种字节宽度为 0 的不可打印的 Unicode 字符，在浏览器和一般的文本编辑器中是不可见。

字符水印编码，嵌入水印位。图展示了基于零宽字符的 OFD 隐式水印算法嵌入与提取流程，如图 12 所示。

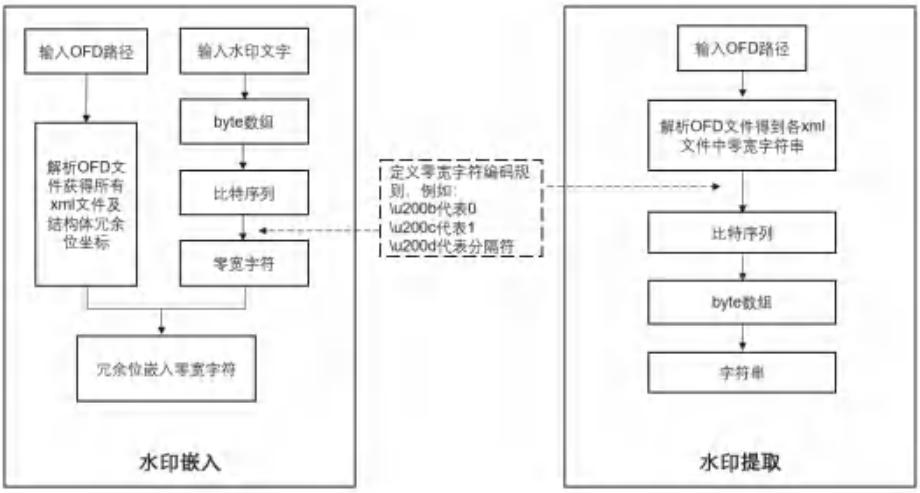


图 12 基于零宽字符的 OFD 隐式水印算法流程

（三）测试结果

课题组在银联数字银行网络平台上线了“OFD 文件添加水印”和“OFD 文件验证水印”接口，并使用数字银行业务中传输的 OFD 格式电子凭证文件进行了测试。测试过程中，通过调用接口在文件中添加了隐式水印“中国银联授权-测试”，添加水印之后的文件能正常打开、验签。调用水印验证接口能正常读取“中国银联授权-测试”的水印信息。测试结果如图 13 所示：

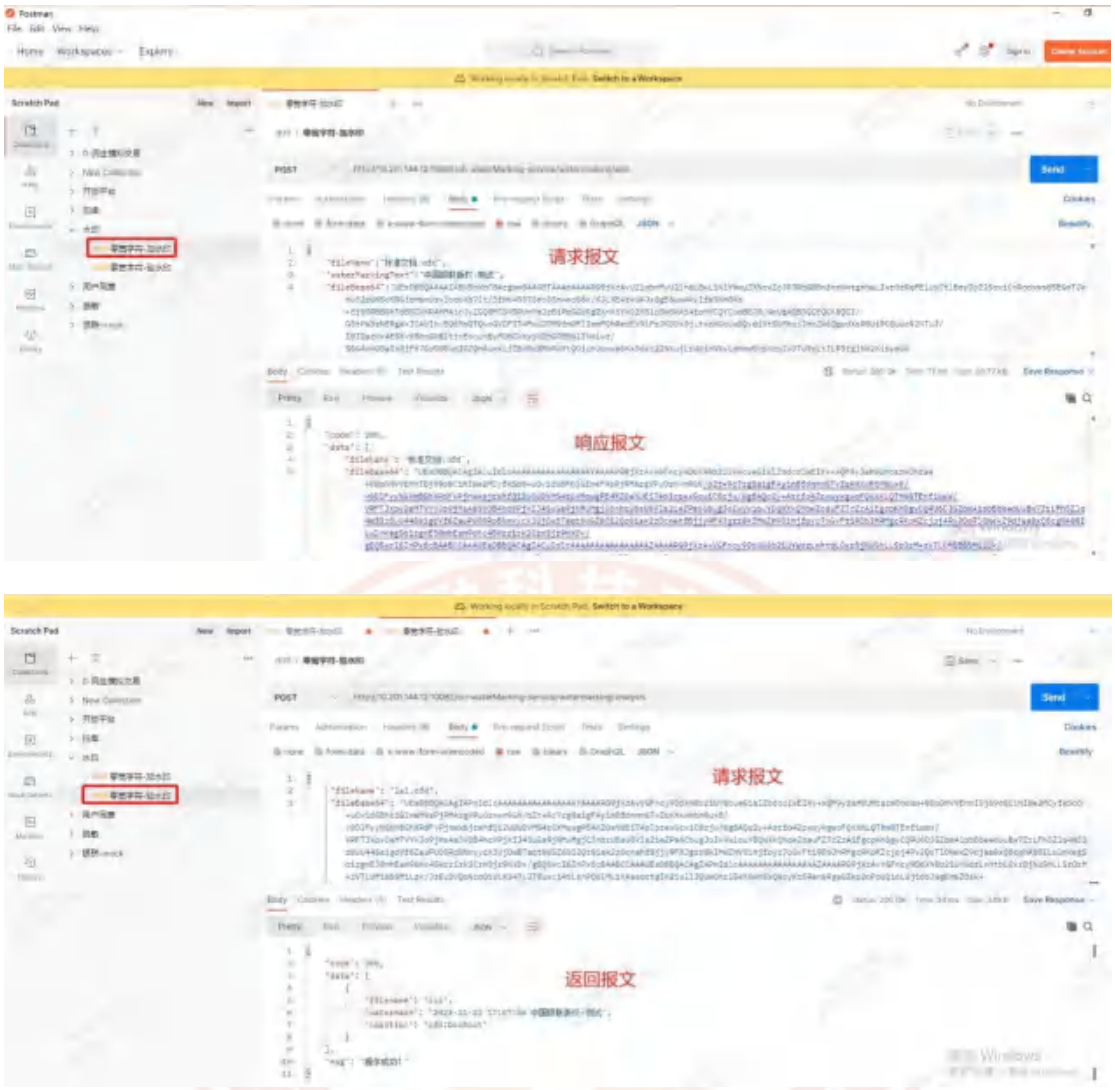


图 13 OFD 水印测试结果

八、总结和建议

（一）继续深入数据安全相关技术及标准研究

本研究报告涉及的技术研究课题均是数据安全基础性的技术课题，能充分反应行业的迫切诉求。目前，虽然几项技术均已有相关成果，部分成果还在银行的联合测试中取得了非常优秀的测试结果，但是本课题组的研究工作并未就此结束，仅目前的几个研究方向就仍有不少需要深耕的技术难题：

1. 半结构化、非结构化数据识别算法研究

为了补齐数据分类分级的能力，后续研究工作将继续深入半

结构化数据、非结构化数据的识别技术研究。

半结构化数据指单一数据字段的内容是包含了各种数据类型的文本，如交易查询流程需要提交的表单有“需求内容”字段，其内容本身是一段需求内容的文本信息，但内容会包含“姓名”“卡号”“联系方式”等敏感信息。

非结构化数据指一些敏感数据在企业内不是以数据表的形式存在，比如已经签署的合同，其上可能会包含账户信息。

2. 智能脱敏技术研究

目前行业有“根据业务使用诉求对数据进行适当脱敏”要求，但是针对此要求行业没有形成统一认识，因此脱敏在行业中是一项对工作经验要求颇高的工作。本研究报告涉及的脱敏效果评估体系研究，是为了解决脱敏结果的度量问题，为了在自动化、智能化、标准化方面继续优化脱敏工作现状，需要进一步探索论证基于业务特征智能推荐脱敏算法及配置的可行性。

3. 结构化数据水印研究

数据水印是数据发生泄露后，追溯责任方、亡羊补牢的最终手段，目前无论是针对显式水印或是隐式水印，行业中针对图片、PDF 等文件类型均有较多研究，但是数字银行存在很多在报文中直接传输、数据表中存储的结构化数据，这类结构化数据如何添加水印用于溯源，需要进一步探讨。

4. 完善数据安全规范体系

目前数字银行产业的相关方特别是中小银行和应用方尚未完善地应用相关数据安全技术能力，下一步需要通过完善数据安全技术规范体系加速相关技术的应用推广：

一是完善技术规范，行业需要通过优秀的技术成果编制示范性规范，以提升行业的整体研究能力。同时，诸如智能脱敏等技术只有通过了行业公认的标准认证才具备推广应用的基础。

二是完善技术评价规范，诸如 API 异常行为检测等技术，由于使用了基于人工智能的分析技术，人工智能的描述性、稳定性问题导致不同厂商的技术能力难以被横向比较，因此有必要编制诸如《金融 API 安全防护体系评估指南》的评价规范帮助产业相关方遴选符合要求的技术能力。

（二）数字银行场景安全需要加强管理

虽然，国家已经颁布了《中华人民共和国数据安全法》《商业银行应用程序接口安全管理规范》《金融数据安全 数据生命周期安全规范》《金融数据安全 数据安全分级指南》等制度规范，但是行业对于相关制度规范的执行情况仍处于摸索、难以完全执行到位的状态，需要金融管理部门加强政策引导:

一是建议出台与数字银行直接相关的或是针对现行相关要求更细致的指导文件，进一步明确各方的权责义务，指导行业相关机构对相关要求形成统一认识。

二是数据安全相关制度规范需要更强的执行力度来支持相关技术解决方案的落地应用，如安全前哨等应用于数字银行应用方的安全技术。

三是在数字银行业务中，中小银行作为数据安全能力的弱势群体，难以依靠自身能力构建完善的数据安全能力体系，建议在制度规范制定中应更加关注中小银行的数据安全诉求，如支持有资质背书的转接清算机构在为中小银行提供接口转接服务的同

时配套提供数据安全能力。

（三）加强自律管理完善标准体系

本研究课题旨在从技术角度提出解决方案，以解决数字银行的数据安全问题。技术是行业生态开展安全管理的基础，然而，金融行业自律体系有待健全，可能导致其他行业利用行业风险敞口不断向金融行业渗透。因此，需要完善自律体系加强自律管理：

一是通过行业自律型机构定期开展检测认证和安全审计，以排除技术监控手段被旁路的风险。

二是单一行业相关机构的技术能力和风险信源有限，通过行业自律性管理机构实现技术能力互通和风险信源共享，将有助于行业相关机构以更小的成本投入获得更有效的安全技术能力成果。

三是数字银行的应用方通常是业务导向的，有必要通过有效的管理体系来规避安全方面存在的“劣币驱逐良币”的现象。

附录：数据安全法律规范

本文参考的数据安全法律规范清单

法律/规范名称	颁发机构/状态	生效时间	对数字银行实践的影响
《中华人民共和国数据安全法》	国家/已生效	2021-06-10	该法律是我国第一部有关数据安全的专门法律。是数据领域的基础法律，不仅有助于维护我国的数据安全，更为促进数字经济的健康发展，提供了核心法制依据。
《金融数据安全数据安全分级指南》	中国人民银行、国内众多银行/已生效	2020-09-23	《指南》给出了金融数据安全分级的目标、原则和范围，以及数据安全定级的要素、规则和定级过程。因此可以更清晰地指导金融业机构开展电子数据安全分级工作，也可以指导第三方评估机构等单位开展数据安全检查与评估工作。
《金融数据安全数据生命周期安全规范》	中国人民银行/已生效	2021-04-08	《规范》首次明确定义了数据安全的原则，包括合法正当原则、目的明确原则、选择同意原则、最小够用原则、全程可控原则、动态控制原则、权责一致原则，为金融机构安全建设提供参考。
《商业银行应用程序接口安全管理规范》	中国人民银行/已生效	2020-02-13	《规范》是监管部门发布的首份数字银行监管政策和行业标准，能够为金融行业在数字化经济转型中提供更多参照与指导。为从事或参与商业银行应用程序接口服务的银行业金融机构、集成接口服务的应用方开展相关工作以及第三方安全评估机构等单位开展安全检查与评估工作提供了重要的参考。
《信息安全技术—个人信息安全规范》	全国信息安全标准化技术委员会/已生效	2020-10-01	《规范》对个人信息控制者在收集、存储、使用、共享、转让、公开披露等信息处理环节中的相关行为做出了规范，对遏制个人信息非法收集、滥用、泄漏等乱象提供了依据，最大程度地保障了个人的合法权益和社会公共利益。