第 8 卷◆第 7 期◆版本 1.0◆2025 年
文章类型：论文| 刊号（ISSN）：3082-8295(O) / 2630-4759(P)

金融行业客户信息隐私保护技术实践探究

夏旻
诺亚控股有限公司
DOI:10.12238/ej.v8i7.2752

[摘　要] 伴随着金融行业信息化进程的加速,客户数据隐私保护的问题越来越凸显。本研究将讨论数据加密、同态加密以及差分隐私等技术在金融行业的应用情况,通过实验分析对各种技术进行隐私保护度、计算效率、数据准确性以及系统负载等方面的评价。结果表明数据加密的隐私保护度与计算效率性能最好,适用于账户查询与在线支付的场景,同态加密为用户提供高度的隐私保护,但计算效率不高,适合数据共享与联合计算,差分隐私保持良好的计算效率又确保一定程度的隐私保护,适合大数据分析与风险评估。
[关键词] 金融行业；客户信息；隐私保护；数据加密；同态加密
中图分类号：F83　**文献标识码**：A

Practical research on customer information privacy protection technology in the financial industry

Min Xia
Noah Holdings Limited

[Abstract] With the acceleration of informatization in the financial industry, the problem of customer data privacy protection is becoming more and more prominent. In this study, we will discuss the application of data encryption, homomorphic encryption, and differential privacy technologies in the financial industry, and evaluate the privacy protection, computing efficiency, data accuracy, and system load of various technologies through experimental analysis. The results show that data encryption has the best privacy protection and computational efficiency performance, which is suitable for account query and online payment, homomorphic encryption provides users with a high degree of privacy protection, but the computational efficiency is not high, which is suitable for data sharing and joint computing, and differential privacy maintains good computational efficiency and ensures a certain degree of privacy protection, which is suitable for big data analysis and risk assessment.
[Key words] financial industry; customer information; Privacy Protection; data encryption; Homomorphic encryption

引言

在信息技术飞速发展背景下,金融行业数字化转型速度不断加快,客户信息保护工作也将遇到空前挑战。金融行业中涉及到海量敏感数据,其中包括个人身份信息、账户余额、交易历史等等,这些数据被泄露会造成严重经济损失与信任危机[1]。如何对客户隐私进行有效保护是金融机构的一项重要任务。本研究采用实证研究的方法对三种主流隐私保护技术应用于金融行业客户信息保护的效果进行评价,为金融行业选择最合适的隐私保护技术提供理论依据与实践指导。

1 理论基础

1.1隐私保护技术概述

隐私保护技术是为确保个人或者组织的数据不被非法访问、不被传递、不被处理时被泄露。伴随着信息技术的快速发展,特别是金融行业,客户数据是否安全已经逐渐成为人们关注的焦点。隐私保护技术既关系数据安全又关系法律的合规性与客户信任。常用的隐私保护技术有数据加密、同态加密和差分隐私,它们以不同方式保证敏感信息机密性与完整性,允许以隐私保护为前提,对数据作必要处理与分析[2]。

1.2数据加密技术

数据加密技术对数据进行基本而有效地保护,明文数据经过算法转换为密文来阻止未经授权的用户进入。常见的加密算法有对称加密算法(如AES)和非对称加密算法(如RSA)。数据加

密被广泛用于对客户账户信息、交易记录以及其他敏感数据进行保护。加密过程可由一个基本公式来表达：

$$C = E(K,P)$$

C为密文,E为加密算法,K为加密密钥,P为明文数据。加密后即便数据被拦截,攻击者也无法恢复原始数据,除非拥有解密密钥。

1.3同态加密技术

同态加密技术使加密后的数据不需解密就可进行操作。这使同态加密对敏感数据的处理有其独特的优势,特别是对于金融行业需分析数据同时保证数据隐私情况下的应用[3]。同态加密分为部分同态加密和全同态加密,前者支持部分运算(如加法或乘法),后者则支持任意复杂的运算。同态加密有以下几个基本计算公式：

$$E(P_1 + P_2) = E(P_1) + E(P_2)$$

$E(P)$ 表示加密后的数据, P_1和 P_2为明文数据。这意味着利用同态加密技术,金融机构可在加密数据上直接进行加法或乘法等操作,而不需解密保障客户隐私。

1.4差分隐私技术

差分隐私技术将噪声加入到数据查询当中,以保证个别数据记录的改变不影响查询结果,并保护个体隐私。差分隐私核心思想就是通过制定隐私预算(ϵ)在数据隐私保护和查询结果准确性之间进行权衡。差分隐私在大数据分析、统计数据发布等领域具有广泛的应用前景,尤其是金融行业涉及海量用户数据统计分析,差分隐私可有效地阻止敏感信息的泄漏。差分隐私数学表达式为：

$$\Pr[\mathcal{M}(D) \in S] \leq e^{\epsilon} \cdot \Pr[\mathcal{M}(D') \in S]$$

$\mathcal{M}(D)$ 表示算法在数据集D上的查询结果,S是查询结果的某个集合, ϵ 为隐私预算。公式表明差分隐私通过引入噪声,确保即使攻击者掌握部分背景知识,也无法通过查询结果推断出某个个体的敏感信息。

1.5隐私保护技术在金融行业中的应用

金融行业中涉及到海量的顾客敏感数据,隐私保护技术显得尤为重要。数据加密在电子支付和在线银行中被广泛用作一种常规的方法。伴随着金融科技的进步,同态加密被关注于跨机构数据共享与分析,在确保隐私的前提下进行数据合作。差分隐私技术在金融数据的分析和风险评估中得到了广泛应用,尤其在大数据的背景下,它能有效地避免个人信息的外泄[4]。

2 模拟仿真实验

2.1实验设计与目标

该实验评价数据加密、同态加密以及差分隐私对金融行业客户信息保护的作用,并着重研究数据安全性和处理效率。通过对典型金融应用场景进行仿真,检验各种技术对数据加密、隐私保护度、计算效率、数据准确性和处理延迟等性能的影响。实验目的在于论证各种技术各自的优点和局限,为金融行业数据保护提供理论依据和技术支持。

2.2实验数据集与预处理

实验数据集包含10,000条客户记录,包括姓名、身份证号、账户余额、交易历史等,模拟出信用卡交易、贷款申请、账户查询的金融场景。在数据投入使用之前对其进行标准化处理以剔除缺失值及异常值,保证数据质量并降低噪声干扰,实现敏感信息的匿名化和脱敏以评价不同的隐私保护技术。

2.3实验环境与工具

实验在一台具有8核心CPU和32GB内存的计算机上进行,操作系统为Windows10,使用Python编程语言进行实验编程。为确保数据的处理效率和隐私保护技术的实施效果,实验工具包括加密库PyCryptodome、同态加密库PySEAL、差分隐私库PySyft等。为进行数据分析和实验结果的可视化,我们使用Pandas进行数据处理,Matplotlib和Seaborn用于结果展示。

2.4实验步骤与流程

2.4.1数据加载与预处理

在本实验中需加载实验数据集,并对其进行必要的数据清洗、标准化和匿名化处理,以确保数据质量和一致性。数据清洗的过程包括去除缺失值、处理异常数据和统一数据格式,以提升后续分析的准确性。标准化步骤则通过归一化或正则化方法,使数据分布更加均衡减少偏差影响。

2.4.2隐私保护技术应用

在数据预处理完成后,根据实验需求选择合适的隐私保护技术,对数据进行加密或扰动处理以提高数据安全性。在同态加密方案下,我们使用加密算法对数据进行处理,使其能在加密状态下执行计算任务避免明文暴露,确保数据在传输和存储过程中保持安全。对于差分隐私技术,则采用噪声注入的方法,通过在查询结果或数据集中添加随机噪声,使攻击者难以推测出单个数据点的信息提升隐私保护的强度。

2.4.3性能测试与数据分析

为全面评估所采用隐私保护技术的有效性和适用性,本实验对各技术方案的隐私保护度、计算效率、数据准确性以及系统负载进行深入分析。我们量化各方法在防止数据泄露方面的能力,衡量其隐私保护级别,同时评估加密处理或差分隐私方法对数据可用性的影响。

2.4.4结果展示与讨论

实验结果的可视化是分析各隐私保护技术优劣的重要手段,我们利用Matplotlib和Seaborn等数据可视化工具,对实验数据进行直观展示。通过折线图、柱状图和热力图等方式,清晰地呈现不同技术方案在隐私保护度、计算效率、数据准确性和系统负载等方面的差异。我们深入探讨隐私保护与计算效率之间的权衡关系,并结合实验结果分析不同方法的适用场景。我们记录各项关键指标,包括处理时间、计算延迟、系统资源消耗等,以支持后续的优化研究,并为未来在实际应用中选择最优隐私保护方案提供数据参考。

3 实验结果与分析

3.1计算开销与系统负载分析

表1 计算开销对比

方案	计算时间(s)
伪匿名化	0.02
差分隐私(ε=1.0)	0.08
差分隐私(ε=0.5)	0.15

从实验结果可看出,伪匿名化的计算开销最低(0.07),计算时间和CPU占用率也最低,适用于高并发、低计算资源需求的业务场景。差分隐私的计算开销随ϵ下降而上升,表明更强的隐私保护需更高的计算资源。同态加密的计算开销最高(13.14),计算时间长达5.32s,CPU负载高达78.1%,这意味着在实时交易等高计算效率需求的场景下其应用受限。

3.2数据准确性分析

表2 数据准确率对比

方案	数据准确率(%)	方案
伪匿名化	98.5%	伪匿名化
差分隐私(ε=1.0)	92.3%	差分隐私(ε=1.0)
差分隐私(ε=0.5)	85.6%	差分隐私(ε=0.5)

同态加密的准确率最高(100%),因数据在加密状态下进行计算不会影响数据的完整性,但其计算成本较高。伪匿名化方法的准确率也较高(98.5%),主要是其仅替换部分数据字段,对数据分析的影响较小。差分隐私方法引入的噪声导致数据准确性下降,ϵ=1.0时准确率为92.3%,但ϵ=0.5进一步降低至85.6%,这表明更强的隐私保护会影响数据的分析精度。

3.3适用场景分析

表3 适用场景对比

方案	适用场景	主要优势	主要劣势
伪匿名化	低敏感度数据处理(如交易流水分析)	计算成本低、处理速度快	隐私保护能力较弱
差分隐私(ε=1.0)	统计分析、风控模型	提供隐私保护,影响数据可用性较小	计算成本适中,数据准确率受影响
差分隐私(ε=0.5)	高安全性统计分析	更强隐私保护	数据准确率下降,计算成本较高
同态加密	高敏感度数据计算(如多方安全计算)	最高隐私保护	计算成本极高,不适合实时计算

伪匿名化适用于日常金融数据处理,如交易流水、客服数据分析等,其计算成本低,隐私保护能力有限。差分隐私(ϵ=1.0)适用于风控建模、用户行为分析等场景,提供较好的隐私保护,同时保持较高的数据可用性。当ϵ=0.5时,适用于高安全性要求的统计分析,但数据准确率下降。同态加密适用于需最高隐私保护的应用,如金融机构间联合建模,但计算成本极高不适用于实时业务。

4 讨论

本研究对比三种主要隐私保护技术在金融行业中的应用,包括同态加密、差分隐私和伪匿名化,并评估它们在隐私保护、安全性、计算成本、数据准确性和适用场景方面的差异。实验结果表明不同技术在隐私保护能力与计算效率之间存在明显权衡。同态加密方案提供最强的隐私保护能力,其隐私泄露概率最低(0.05),但计算开销高,CPU负载达78.1%,不适合实时交易场景。差分隐私通过调整噪声参数ϵ影响隐私保护度和数据准确性,当ϵ=0.5时,隐私保护度达72%,但数据准确率下降至85.6%,适用于风控建模和用户行为分析等场景。伪匿名化方法计算成本最低(0.07),准确率接近98.5%,但隐私保护能力较低,仅适用于低敏感度数据处理。

5 结论

该研究对金融行业中数据加密、同态加密、差分隐私等技术进行试验分析。结果表明数据加密的隐私保护、数据准确性以及计算效率等方面的性能最好,适合账户查询以及在线支付等功能,但系统负载较大,将制约大范围数据处理。同态加密具有很高的隐私保护度,但计算复杂且处理延迟较大,适用于跨机构的数据共享与联合计算[5]。差分隐私虽然保护度不高,但其计算效率高、系统负载小,适合大数据分析及金融风险评估等领域,兼顾隐私和效率。数据加密和同态加密都适合在高隐私要求的场景下使用,差分隐私则更加适合在信用评分这样的大规模数据中使用。

[参考文献]

[1]赵腊梅.区块链技术在金融领域的隐私保护与安全性研究[J].电脑知识与技术,2024,20(33):91-93.

[2]吴迪,何林华.数字时代商业银行个人金融信息利用问题的法律研究[J].辽宁经济,2023(1):62-68.

[3]张铭杰.隐私计算技术赋能金融数据安全共享[J].中国农村金融,2023(4):97-98.

[4]刘雨佳.大数据背景下金融隐私权保护的困境与对策[J].西部学刊,2023(6):116-119.

[5]Díaz B S,Coussement K,Caigny D A. From collaborative filtering to deep learning: Advancing recommender systems with longitudinal data in the financial services industry [J]. European Journal of Operational Research,2025,323(2):609-625.

作者简介：

夏旻(1983--),男,汉族,上海人,本科,研究方向：信息安全管理创新型实践。