 DOI: https://doi.org/10.12345/cjygl.v10i2.36927

Research on Privacy Protection of E-commerce Service Platforms Driven by Digital Intelligence

Tianbao Hao[1] Zhiqin Ren[2] Li Han[2] Lisha Peng

1. College of Finance and Economics, Hebei Normal University of Science & Technology, Qinhuangdao, Hebei, 066004 China
2. Network Center, Hebei Normal University of Science & Technology, Qinhuangdao, Hebei, 066004, China

Abstract

In recent years, with the rapid development of e-commerce, privacy breaches have become increasingly frequent, posing a risk of leakage and misuse of users' personal information. As a result, the demand for privacy protection has grown, making the protection of user data privacy a crucial task for platforms. Consequently, privacy protection technologies based on e-commerce platforms have attracted widespread attention. This paper comprehensively reviews existing key privacy protection technologies, designs data processing for e-commerce platforms, and proposes a privacy-based design method, including data classification, sensitive data identification, data anonymization rules and algorithms, management of anonymization strategies, and the design of static anonymization tasks. Finally, it presents user privacy protection strategies for e-commerce platforms to safeguard user privacy and security.

Keywords

E-commerce; Privacy protection; Strategy

数智驱动电商服务平台隐私保护策略研究

郝天保[1] 任志琴[2*] 韩丽[2] 彭李沙

1. 河北科技师范学院财经学院，中国·河北 秦皇岛 066004
2. 河北科技师范学院网络技术中心，中国·河北 秦皇岛 066004

摘 要

近年来，随着电商的快速发展，隐私泄露问题频发，用户的个人信息面临泄露和滥用的风险，人们对隐私保护的需求日益增加，保护用户数据隐私成为平台的重要任务，因此基于电商平台的隐私保护技术受到了人们的广泛关注。本文全面梳理了现有隐私保护关键技术，对电商平台中的数据进行处理设计，提出了一种基于隐私保护的设计方法，包括数据分级、敏感数据的识别、数据脱敏规则算法、脱敏策略的管理和静态脱敏任务的设计。最后，给出了基于电商平台的用户隐私保护策略，保护用户的隐私安全。

关键词

电子商务；隐私保护；策略

1 引言

在数字经济蓬勃发展的当下，数智技术正以前所未有的速度渗透到各个领域，电子商务行业更是深受其影响，发生了深刻变革。然而，近年来，隐私泄露问题频发，人们对隐私保护的需求日益增加。通过大数据分析与智能推荐，电商平台能够提供个性化服务，但也使用户隐私信息暴露在风险之中。若无有效保护策略，敏感信息可能被盗用，危害用户安全。

2 国内外研究现状

随着经济全球化和信息技术的快速发展，电子商务成为新时代商业的宠儿，但是由于互联网的开放性和共享性，电子商务的信息安全问题成为制约电子商务发展的主要障碍[1]。整合来自多方的数据以实现跨机构机器学习是一项重要趋势。然而，数据共享带来的隐私风险对数据集成构成重大挑战[2]。Bing Wu 认为在大数据时代，机器学习模型的协同训练可以确保多方之间安全共享数据，同时保障用户的隐私

【基金项目】2024 年河北省高等学校科研青年基金项目（项目编号：SQ2024198）

【作者简介】郝天保（1985—），男，中国河北邯郸人，博士，从事空间数据库研究。

【通讯作者】任志琴（1984—），女，中国山西忻州人，硕士，从事空间数据库，网络安全研究。

数据 [3]。Mingjun Dai 通过加密技术保护用户的隐私，提出了一种新型的编码 FLR 框架 [4]，基于线性组合（LC）的垂直 FLR 和基于 Matdot 的垂直 FLR 两种方案，均可在边缘节点上进行并行计算和同态加密。王瀚仪等人提出一种多级 LDP 算法推荐框架。考虑服务商以及用户的需求，通过服务商和用户的多级管理实现多用户差异化隐私保护 [5]。朱骁等提出了横向联邦 PCA 差分隐私数据发布算法 [6]。一方面，保护本地数据隐私；另一方面，减少了噪声添加量，并且达到与中心化差分隐私 PCA 算法相同的噪声水平。霍炜等人着眼于各类算法与协议的不同层级安全属性，也侧重于从模块化角度剖析具体构造的内在技巧逻辑甚至缺陷。并研究了高等密码算法与协议的最新理论和技术进展、背后的发展逻辑，并深悟其中的关键技术原理 [7]。而 Qingfeng Lei 通过收集用户数据，分析用户行为序列计算用户相似度，并利用余弦相似度算法计算试题文本相似度，构建了动态用户画像 [7]，采用混合推荐机制为用户提供个性化试题推荐服务。曾德胜等人认为要制定针对性信息安全防护策略 [9]。借助数据加密技术、杀毒软件、网络监测技术以及防火墙技术来开展防护工作，切实保护好计算机网络信息的安全。

3 隐私保护设计方法

3.1 敏感数据分级

根据数据的隐私程度对数据进行分级划分，将数据分为公开级数据、内部级数据、敏感级数据、高级敏感级数据。

公开级数据包括对外公开且不涉及用户隐私或业务机密的信息，例如平台公开的商品评价、用户昵称（非实名）、已脱敏的统计数据等。此类数据安全级别最低，但仍需防范篡改或滥用风险。

内部级数据仅限企业内部使用或特定部门共享的数据，如员工操作日志、内部业务流程记录、未公开的运营分析报告等。通过访问权限控制（如 RBAC 角色权限模型）防止非授权访问。

敏感级数据是指直接关联用户隐私或企业核心利益的数据，例如个人姓名、联系方式、订单地址、浏览行为记录等。采取加密存储、匿名化处理、最小化收集原则，并符合《个人信息保护法》等法规要求。

高级敏感级数据具有极高风险的数据，一旦泄露可能导致严重法律或财务后果，例如身份证号、银行卡信息、生物识别特征（指纹、面部识别数据）、企业核心商业机密（如供应链策略、未公开的财务数据）等。数据需最高级别保护，包括端到端加密、多因素认证、严格审计日志记录及物理隔离存储。

3.2 敏感数据识别

防止敏感信息泄漏威胁的首要步骤是定义敏感信息，通过建立敏感信息样本库，定义数智电商的敏感信息的具体特征。敏感信息库内置各类敏感信息的识别规则，包括但不限于：身份证号码、手机号码、生日、信用卡号码…

数智电商平台的敏感数据识别通过多源数据库适配联邦查询技术，打破异构数据源（如 MySQL、Oracle、Hadoop 等）间的壁垒，实现跨库、跨系统的统一查询与分析。基于动态化的识别规则组配置（如正则表达式、关键词匹配、机器学习模型），系统自动扫描结构化与非结构化数据，精准识别身份证号、银行卡信息、生物特征等高敏感字段，并结合数据分级标准（公开级、内部级、敏感级等），实时更新数据表或字段的敏感级别标签，确保分级与风险动态匹配。例如，当检测到某数据表新增用户实名信息时，系统自动将其标记为“高级敏感级”，并触发加密存储、访问权限收紧等策略。通过智能调度策略（如定时任务、事件驱动、优先级队列），灵活配置全量扫描、增量监测或实时触发机制，在低峰期执行批量敏感数据发现任务，或针对高频交易场景实时监控数据流动，兼顾效率与资源消耗。不仅提升敏感数据识别的覆盖率和时效性，还通过自动化流程减少人工干预，助力企业快速响应《个人信息保护法》等合规要求，在保障隐私安全的同时，支撑精准营销、风控建模等业务场景的数据高效利用。

4 隐私保护数据脱敏设计

在数智电商平台的隐私保护体系中，基于数据分级分类结果（如敏感级、高级敏感级数据），需设计动态化、差异化的脱敏技术方案，并结合大模型能力实现智能化升级。根据不同的隐私信息采用具体规则算法。

4.1 内部脱敏规则

表 1 内部脱敏规则

脱敏规则	描述
手机号	中间四位用“*”号代替，例如：138***1234
银行卡号	中间四位用“*”号代替，例如：6228********1234
身份证号	前四位和后四位用“*”号代替，例如：5101**********4321
姓名	只显示姓氏或者替换为“*”号代替，例如：张或者 * 三
地址	只显示省市或者区县，例如：四川省成都市或者北京市海淀区
车牌号	只显示前两位和后两位，例如：川 A1234
IP 地址	只显示前三位和后一位，例如：192.168.1.*
…	…

敏感数据的识别对于防止敏感信息泄漏威胁的重要步骤，定义敏感信息是敏感数据识别的第一步，我们是通过建立敏感信息样本库，定义数智电商的敏感信息的具体特征。内部脱敏规则内置各类敏感信息的识别规则，包括但不限于：身份证号码、手机号码、生日、信用卡号码等信息。同时敏感信息规则支持用户自定义各类敏感信息规则在不同应用场景中允许用户进行规则扩展。

4.2 自动识别算法

表 2 自动识别算法

识别算法	描述
替换算法	将敏感信息替换为一些无意义的符号或者其他信息
加密算法	使用加密算法对敏感信息进行加密，只有具备相应密钥的人才能解密
哈希算法	将敏感信息通过哈希算法转化为固定长度的字符串，从而保护敏感信息的隐私
K- 匿名算法 [10]	通过对数据进行分组，保证每个组内的数据无法区分个体，从而保护隐私
L- 多样性算法 [11]	在 K- 匿名的基础上，保证每个组内有至少 L 个不同的敏感信息值，从而提高隐私保护水平
T- 接近匿名算法 [12]	在 K- 匿名的基础上，保证每个组内的敏感信息值与其他组内的值相差不超过 T，从而提高隐私保护水平
D- 差分隐私算法 [13]	通过在数据中添加噪音，保证个体数据不被泄露，从而保护隐私
…	…

敏感数据识别可以通过多源数据库适配联邦查询技术，打破异构数据源间的壁垒，实现跨库、跨系统的统一查询与分析。基于动态化的识别规则组配置，系统自动扫描结构化与非结构化数据，精准识别身份证号、银行卡信息、生物特征等高敏感字段，并结合数据分级标准，实时更新数据表或字段的敏感级别标签，确保分类与风险动态算法匹配。

5 隐私保护策略

第一，隐私保护制度。基于“合规性 - 可操作性 - 可持续性”原则，结合国内外隐私保护法律法规与电商平台业务特性，设计分层化、模块化的隐私保护管理制度。数据隐私保护应采取积极防御、综合防范的方针，坚持保障数据隐私保护与促进信息化发展相协调、管理与技术统筹兼顾的原则，实行统一协调、分级管理、分工负责。数据隐私保护和信息化工作应当同步规划、同步建设、同步实施、同步发展。

第二，用户隐私保护画像重构。用户画像隐私模型构建基于“分层加密 - 联邦学习 - 大模型增强”三位一体的用户画像隐私模型，通过技术创新实现隐私保护与数据价值的平衡。模型覆盖数据采集、处理、应用全流程，结合大模型技术提升隐私保护能力与业务效果。

第三，隐私保护平台。建立基于用户隐私保护的电商平台是利用数据智能技术（Data Intelligence）来优化电商运营的平台。整合大数据分析、人工智能、机器学习等技术，在通过智能化手段提升用户体验、优化供应链管理、并增强市场竞争力。平台是一个简单的交易平台，提供个性化服务、精准营销和高效管理的综合解决方案。实现脱敏管理，根据数据敏感级别与业务场景，制定分层的脱敏策略。高级敏感数据：强制全字段脱敏，仅限授权角色访问原始数据；敏感数据：部分脱敏，支持特定业务场景的模糊查询；业务数据：保留格式但替换真实内容，确保数据分析与测试环境安全。

6 结语

本研究深入剖析了数智驱动下电商服务平台隐私保护这一关键议题，介绍了现阶段电子商务中隐私保护方面在国内外的发展现状，并提出了隐私保护的设计方法，通过将数据分为四级进行敏感数据处理，对敏感数据进行了脱敏算法采用内部脱敏处理和自动识别算法处理，最终结合当前的实际情况给出制度、用户画像、隐私保护平台 3 个方面的策略分析，以此来保障电子商务平台的个人隐私的安全性。

参考文献

[1] 孙艳. 电子商务信息安全问题研究[J].探索与观察, 2022,37-39.

[2] JIANPING CAI, XIMENG LIU, ZHIYONG YU. Efficient Vertical Federated Learning Method for Ridge Regression of Large-Scale Samples [J]. IEEE Transactions on EMC, 2022, 7, 511-526.

[3] Bing Wu, Xiaolei Dong, Jiachen Shen, Zhenfu Cao. Enhancing Model Performance via Vertical Federated Learning for Non-Overlapping Data Utilization [C]. ISPDS, 2023.

[4] Mingjun Dai, Ziying Zheng, Zhaoyan Hong, Shengli Zhang. Edge Computing-Aided Coded Vertical Federated Linear Regression [J]. IEEE TRANSACTIONS ON COGNITIVE COMMUNICATIONS AND NETWORKING, 2022, 8, (3), 1543:1551.

[5] 王瀚仪，李效光，毕文卿，陈亚虹，李凤华，牛犇. 多级本地化差分隐私算法推荐框架[J].通讯学报，2022, 43(8), 52-63.

[6] 朱骁，杨庚. 横向联邦学习中 PCA 差分隐私数据发布算法[J].计算机应用研究, 2022, 39(1),236-240.

[7] 霍炜，郁昱，杨糠，郑中翔. 隐私保护计算密码技术研究进展与应用[J]. 中国科学:信息科学, 2023, 58(9),1688-1733.

[8] Qingfeng Lei, Hua Yang, Taiyong Deng, Jiangrong Liu, Falin Jiang. Research on online test question recommendation technology based on multi-dimensional user dynamic port [C]. ICSCAS, 2022.

[9] 曾德胜，何健，宁建飞，欧国成. 大数据时代计算机网络信息安全防护策略分析[J]. 软件, 2022, 43(9), 64-66.

[10] 吴响，臧昊，俞啸. 基于抽样路径的K-匿名隐私保护算法[J].电子技术应用, 2016,42(12):115-118.

[11] 韩雪. 一种基于聚类的语义l-多样性隐私保护算法[M].哈尔滨工程大学. 2014:25-28.

[12] 杨静，张冰，张健沛，谢静. 基于敏感等级划分的(1, t)-相近性匿名算法[J], 华中科技大学学报, 2014, 42(8), 12-17.

[13] 郭雅馨，范启天，严利民. 一种基于差分隐私的联邦学习隐私保护算法[J].复旦学报, 2025,64(04),385-394.