

电力行业数据脱敏技术研究及其在数据安全中的应用

冯在欣, 蒙泉, 黄博文

(海南电网有限责任公司信息通信分公司)

【摘要】在数字电网建设与能源互联网加速发展背景下, 电力行业数据量呈爆炸式增长, 数据泄露风险加剧, 亟需通过数据脱敏技术保障数据全生命周期安全, 文章围绕电力行业数据脱敏技术展开研究, 分析电力数据特点与安全需求, 阐述常规数据脱敏技术原理及优势, 探讨其在电力数据安全存储、传输、共享等场景的应用实践, 提出当前技术应用面临的挑战并给出应对策略, 旨在为提升电力行业数据安全防护水平提供理论与实践参考。

【关键词】电力行业; 数据脱敏; 数据安全; 隐私保护; 脱敏技术应用

中图分类号: TP309.2; TM73

DOI: 10.13655/j.cnki.ibci.2025.S2.126

Research on Data Desensitization Technology in the Power Industry and Its Application in Data Security

FENG Zai-xin, MENG Xiao, HUANG Bo-wen

(Information and Communication Branch of Hainan Power Grid Co., Ltd.)

【Abstract】Under the background of the construction of the digital power grid and the accelerated development of the energy Internet, the amount of data in the power industry has increased explosively, and the risk of data leakage has intensified. It is urgent to ensure the safety of the whole life cycle of data through data desensitization technology. This paper focuses on the data desensitization technology in the power industry, analyzes the characteristics and security needs of power data, expounds the principles and advantages of common data desensitization technology, discusses its application practice in the safe storage, transmission and sharing of power data, puts forward the challenges faced by current technology application and gives countermeasures and strategies, aiming to provide theoretical and practical reference for improving the level of data security protection in the power industry.

【Keywords】power industry; data desensitization; data security; privacy protection; desensitization technology application

1 引言

伴随着电力系统智能化和信息化程度的提高, 电力行业生产积累了大量的数据, 覆盖发电、输电、变电、配电和用电的整个业务环节。其中蕴含了用户用电信息、电网运行参数和设备状态数据, 对于电力系统的平稳运行, 优化调度和服务提升具有重要意义。数据脱敏是数据安全防护中的一项关键技术, 通过将敏感数据变形和隐匿来保证数据的可用性, 同时又保护了数据的隐私, 对于电力行业数据安全至关重要。

2 电力行业的数据特性及安全需求分析

电力行业是国家能源供应中最核心的支柱, 数据表现出明显的行业特性和复杂安全需求, 在数据类型上, 电力数据涉及生产运行数据, 用户用电数

据、设备管理数据和营销服务数据的各个领域, 生产运行数据中包含了电网实时运行参数、调度指令信息和故障监测数据, 该类数据实时性和连续性较好, 在电力系统稳定运行中发挥着至关重要的作用; 用户的电力使用数据详细记录了他们的电量消耗、缴费详情以及用电习惯等方面的信息, 这些数据涉及用户的个人隐私和商业机密; 设备管理数据中包括设备台账、检修记录、性能参数等等, 这些数据是确保电力设备正常工作及寿命管理重要基础; 与电力销售、市场交易和客户服务有关的营销服务数据, 对公司的经营策略和市场竞争能力产生了深远的影响^[1]。

3 电力行业中数据脱敏的技术原理

数据脱敏技术就是一种通过敏感数据的变形、隐藏、更换等操作来保护数据隐私与安全而又不损

害其使用价值。数据脱敏技术应用于电力行业是以各种原理为基础,目的是适应各种场景对数据安全的要求,静态数据脱敏原理,主要是对储存于数据库、文件系统和其他静态存储介质上的数据进行脱敏处理。核心方法有数据屏蔽,数据泛化,数据加密,数据替换。数据屏蔽的方法就是隐藏敏感字段的一部分,比如把用户身份证号中的中间数位替换为星号,以避免敏感信息的直接曝光;数据泛化是一种对数据进行抽象化处理的方法,目的是降低数据的准确度。

4 电力行业数据脱敏技术研究及其在数据安全中的应用方法

4.1 对用户信息管理进行脱敏

在电力行业的用户信息管理情景下,覆盖了居民及企业等多类用户名称、身份证号、联系方式、用电地址以及缴费记录等敏感信息。根据数据显示,某省级电网公司每天新增的用户信息记录超过10万条,其中涉及身份证号、银行卡号等高度敏感的字段的数据比例高达35%。为了保护用户隐私安全采取了静态脱敏和动态脱敏相结合的技术方案,历史存量数据在静态脱敏环节进行不可逆脱敏,在处理身份证号时,采纳“前面的六个保留加上中间的八个哈希混淆再加上后面的四个”的方法,经过哈希处理后,数据的熵值增加到了256位,这大大降低了身份信息被破解的风险;在银行卡号的设置上,采纳了“前6位保留+中间10位替换为固定字符+后4位保留”的策略,从而使得敏感信息的泄露风险减少了超过90%^[2]。在动态脱敏中,布放脱敏中间件并根据访问者的权限在用户信息调用后进行差异化脱敏。一般客服人员在询问的时候,电话号码只显示前面三位和后面四位,而中间四位则用星号代替;高级管理人员授权时可以看到完整的编号,并对整个运行过程进行日志记录,做到脱敏策略和权限管控准确匹配。实际测试结果显示,该方案成功地将数据泄露检测的误报率从15%减少到了2%,从而显著提高了用户信息安全的防护水平。

4.2 操作数据传输脱敏

电力系统的运行数据传输特点是实时性强、连续性高,覆盖了电网的电压、电流、功率等核心参数,某特高压输电线路的每秒传输监测数据量高达

5MB,数据传输中断的容忍时间必须控制在50ms以内,为了保证数据在传输时的机密性和完整性,采用端到端加密和动态脱敏相结合的技术方案,在数据加密方面,选择了国密SM4对称加密技术进行处理,其密钥的长度为128位,并且加密的处理速度可以达到1.2Gbps,这完全可以满足电网高频数据的传输要求^[3]。结合SM2的非对称加密技术来传递对称的密钥,这使得密钥的协商成功率高达99.99%,从而确保了密钥的安全分发。动态脱敏环节用于将实时数据在传输过程中脱敏,例如,在向外部系统传递电压数据的过程中,通过在 $\pm 5\%$ 的误差范围内对实际电压值进行随机扰动,成功地将数据精度从0.1kV降低到了0.5kV,不仅保证了外部系统能够进行趋势分析而且避免了核心参数的泄漏。引入区块链技术建立数据传输存证系统对每一条数据产生唯一哈希值进行上链存证处理,将数据篡改检测耗时降低到秒级,提高了数据传输安全性与可追溯性。

4.3 营销数据共享脱敏

电力营销的数据共享涉及多个参与方,这包括政府机构、合作伙伴和第三方组织等,其数据内容覆盖了用户的用电习惯、市场的交易行为以及电力负荷的预测等方面,在某一特定区域的电力市场,每年都会向外界共享高达2TB的数据量,其中,敏感数据所占的比例大约是40%。为了达到数据“可供看不见”共享的目的,使用了联邦学习和同态加密相结合的脱敏技术方案,该方法以联邦学习框架为基础实现了多方数据协同建模,且原始数据不会外泄。以电力用户的用电行为为研究对象,电力公司、电信服务提供商和互联网公司都在其本地保存了初始数据,并通过加密参数的交换来进行模型的培训,这种方法使得模型的准确性超过了传统集中训练方法的95%。同态加密方法主要应用于处理共享的数据,它采纳了CKKS同态加密技术,并支持浮点数的计算,与传统方法相比,其加密效果提高了30%。当数据共享完成后,用户用电负荷数据同态加密完成后,第三方机构可以通过密文状态求和和平均值的计算,并将计算结果返回给电力企业解密验证以保证数据始终处于共享状态。通过实际应用的检验,这一方案成功地将数据共享的效率提高了40%,同时也有力地维护了用户的个人隐私和企业的商业机密。

4.4 设备运维存储脱敏

电力设备运维数据主要包括设备台账、检修记录和故障数据,具有巨大的存储规模,某大型电力公司的设备运维数据库容量超过 500TB,其中涉及设备的核心参数、设计图纸等敏感数据的占比高达 30%。为了确保存储数据的安全性,本文采用了分级分类脱敏和访问控制的技术方案,根据数据的敏感程度,设备数据分为公开、内部、敏感和机密 4 个层次。对公开级数据制作格式标准化;内部级数据利用数据泛化技术对设备型号等关键参数进行模糊化;对于敏感级别的数据,使用了可逆的加密技术,例如 AES-256 加密,并定期更新密钥,这样加密后的数据恢复的成功率高达 99.9%;在机密级的数据处理中,使用了多重加密技术和碎片化的数据存储方式,将数据划分为多个部分并存放在不同的物理位置,当需要访问时,需要同时获取这些部分并进行解密处理。

4.5 数据分析隐私脱敏

涉及用户用电行为模式、负荷预测和电网故障诊断等电力大数据分析情景下的核心操作,其数据处理规模往往达到 PB 级别,由于在分析过程中需要调用海量原始数据,很容易造成隐私泄露的危险。为了在数据价值挖掘和隐私保护之间进行权衡,以差分隐私技术作为核心并结合数据匿名化提出了解决方案,差分隐私是在个体数据中加入满足拉普拉斯或高斯分布噪声来达到隐藏个体数据目的,以用户用电量统计分析为例,在聚合查询结果中添加噪声 ϵ (隐私预算),其噪声添加量 b 遵循拉普拉斯机制公式:

$$b = \epsilon / \Delta f \quad (1)$$

式(1)中: Δf 为查询函数的敏感度。假定某个地区的用户使用电量的统计查询函数的敏感度为 $\Delta f=100$,在设置 $\epsilon=0.1$ 的条件下,加入噪声 $b=1000$,保证了即使对特定个体数据进行删除或者修改,不同的查询结果仍然处于噪声干扰下,攻击者不能推断出个人用电信息。

5 电力行业数据脱敏技术研究及其在数据安全中的应用趋势

数据脱敏技术通过对敏感信息进行变形、屏蔽、替换等处理,在保障数据可用性的同时,有效防止敏

感信息泄露,成为电力数据安全防护体系的重要一环。当前,电力行业数据涵盖用户用电信息、电网运行参数、设备运维记录等多源敏感内容,传统规则化脱敏已难以满足复杂场景需求。新兴的人工智能驱动脱敏技术,如基于深度学习的语义理解脱敏、差分隐私算法应用,正成为研究热点。未来,随着电力物联网、智能电网的发展,数据脱敏技术将向智能化、自动化方向演进,实现动态实时脱敏;同时,与隐私计算、区块链等技术融合,构建更完善的数据安全生态,确保电力数据在采集、存储、共享等全生命周期的安全可控,为电力行业数字化发展筑牢安全屏障。

6 结语

在电力行业中,数据脱敏技术对于确保数据安全和保护用户隐私具有十分重要的意义。从电力行业的数据特性,安全需求以及脱敏技术等方面进行研究和应用分析表明,虽然目前数据脱敏技术已经在电力行业中取得了一定的成果,但是仍然面临着很多挑战。今后需要在技术创新,管理优化和标准完善方面多维度努力,不断提高数据脱敏技术适用性和有效性,实现电力行业数字化、智能化的发展,构筑数据安全防线,促进电力行业安全、稳定和高效发展。

参考文献

- [1] 邢汇笛, 龚钢军, 翟明岳, 刘雪松, 王浩淼, 杨爽. 电力数据共享安全防护与隐私保护综述[J]. 综合智慧能源, 2024, 46 (5): 30-40.
- [2] 李学龄, 崔焱, 李文俊. 电力数据资产脱敏汇集管控方法[J]. 微型电脑应用, 2024, 40 (6): 160-163.
- [3] 乔勇, 邱辛泰, 程熹. 电网企业敏感数据保护研究[J]. 河南科技, 2024, 51 (14): 19-26.