

市政地下管网运维大数据预处理技术的研究与应用

李 琪

(保定市市政工程项目服务中心, 河北 保定 071000)

摘 要 市政地下管网是维系城市正常运转的“生命线”, 承担着供水、排水、燃气等核心保障职能。这类管网设施长期隐蔽铺设于地下, 随着使用年限增长, 老化漏损、运维管理滞后等问题愈发突出。伴随智慧化改造工作的持续推进, 物联网终端设备的应用范围逐步扩大, 管网运维环节产生的数据量也随之大幅攀升。但当前各类运维数据存在质量参差不齐、格式标准不统一的问题, 极大地制约了数据价值的落地发挥。在此背景下, 高效可行的大数据预处理技术成为破解管网运维难点、推动管理模式由被动抢修向主动预判转型的核心抓手。

关键词 市政地下管网; 运维大数据; 预处理技术

中图分类号: TU990.3; TP274.2

文献标志码: A

DOI: 10.3969/j.issn.2097-3365.2026.16.011

0 引言

城市化进程持续加快, 城市地下管网的铺设规模也在同步扩张。传统人工巡检模式工作效率低, 应急响应速度迟缓, 早已无法适配精细化运维的现实需求。大数据技术的融入, 为管网运维管理开辟了全新路径; 而预处理作为数据挖掘的前置环节, 其成效直接决定后续分析结果的精准性与可靠性。现阶段常用的预处理技术存在针对性不足、操作流程不规范等短板, 难以适配管网运维数据的复杂特性。因此, 构建科学完备的预处理技术体系并推动实际应用, 对于提升管网运维效能、筑牢城市运行安全防线, 具有重要的现实价值。

1 市政地下管网运维大数据核心特征

1.1 多源异构性突出

市政地下管网运维大数据的多源异构特征极为显著, 数据覆盖范围广、呈现形式繁杂。这类数据的采集渠道多元分散, 既包含管网沿线布设传感器传回的实时监测数据, 也涵盖运维人员现场巡检登记的文字台账, 还有管网设备全生命周期留存的各类档案资料。从数据形态划分, 既有温度、压力、流量这类可量化的数值型数据, 也有巡检工况描述、故障现场纪实这类非结构化的文本数据; 不同渠道采集的数据格式互不兼容, 缺乏统一的规范标准, 这无疑加大了数据整合与预处理的难度, 也是该类数据最直观的核心属性。

1.2 海量性与高增长性显著

海量存储与高速增长是管网运维大数据不容忽视的重要特征。随着市政管网建设体系日趋完善, 各类智能监测传感器的布设密度持续提升, 当前运维大数据的总体量已突破数十亿条。加之管网覆盖范围持续拓展、监测精度不断升级, 数据总量仍在逐年大幅攀升, 扩张态势十分迅猛^[1]。这一特征对数据存储设备的承载容量、数据处理系统的运行效率都提出了严格的标准。只有搭建高性能的存储与处理架构, 才能保障后续运维数据的高效调用, 为管网常态化运维工作筑牢数据支撑。

1.3 质量不均且兼具时空关联性与实时性

管网运维大数据具备两大鲜明特点: 一是数据质量参差不齐, 二是时空关联紧密且实时性要求极高。传感器设备故障、数据传输链路受外界干扰等问题, 极易引发数据缺失、记录偏差、重复冗余等情况, 直接拉低了数据的实际应用价值。与此同时, 所有运维数据都与管网具体点位、对应时间节点深度绑定, 尤其在管网故障应急处置场景下, 必须在短时间内完成实时数据的解析处理, 才能为故障排查、抢修决策提供及时支撑。

2 市政地下管网运维大数据预处理技术体系构建

2.1 数据采集与接入预处理技术

数据采集与接入预处理是整套技术体系的根基, 核心目标是实现多源数据的高效汇聚, 再完成基础规

作者简介: 李琪 (1997-), 男, 本科, 助理工程师, 研究方向: 市政道路。

整处理。管网运维领域的数据本身具有分散性、异构性特点,来源杂乱且格式不统一。针对这一难点,依托多协议适配技术,全面兼容NB-IoT、5G等主流通信协议,顺利实现传感器实时监测数据、管网历史档案数据、人工巡检台账数据的统一接入^[2]。与此同时,构建完善的接入校验机制,针对数据采集时间、格式规范、取值区间开展前置核查,自动过滤剔除显性异常数据。此外,引入边缘计算架构,在数据采集终端侧完成初步筛选与压缩处理,削减无效数据传输量,既缓解了云端算力负荷,又保障了接入数据的时效性与完整性,为后续深度预处理筑牢根基。

2.2 数据清洗预处理技术

数据清洗是提升数据质量的核心环节,主要攻克数据缺失、误差、冗余三大难题,为管网运行数据分析、模型构建及运维决策提供可靠支撑。面对数据缺失问题,需结合管网实际运维场景,灵活选用均值填充、线性插值等方法,补齐传感器故障引发的缺失数据;针对关键节点的缺失项,结合时空关联特征开展精准推算,避免单一填充方式导致的数据失真。针对数据误差问题,搭建专项异常检测模型,对照管网常规运行阈值,筛查超标数据与传输畸变数据,结合管网历史运行台账完成修正校准,确保数据贴合实际工况。针对数据冗余问题,借助专用去重算法清除重复采集、重复录入的数据,同步压缩冗余字段、优化存储结构,提升数据处理效率。

2.3 数据标准化预处理技术

数据标准化的核心目标是打通各系统间的信息壁垒,让多源异构、格式各异的数据实现统一规整。针对市政地下管网运维大数据,首先要构建专属标准体系,明确数据编码规则、字段定义规范、格式约束要求等核心内容,将管网基础台账、实时监测数值、日常运维记录等全品类数据全部纳入管控范畴。针对巡检文本、故障上报报告这类非结构化数据,借助自然语言处理技术,将零散的文字信息转化为标准结构化数据^[3];面对格式不一的异构数据,通过格式适配、单位统一等操作,实现全量数据格式的一致性。除此之外,还要建立动态数据更新机制,确保标准规范调整后数据能够同步迭代,让标准化数据在跨系统、跨部门场景下实现高效流转共享,为后续的数据融合分析筑牢根基。

3 市政地下管网运维大数据预处理技术应用实践

3.1 供水管网运维中的应用实践

在供水管网日常运维工作中,预处理技术精准破解了漏损检测效率偏低、数据价值难以发挥的难点,

是实现管网精细化运维的关键支撑。依托数据采集与接入预处理环节,实时对接管网压力、流量等各类传感器传回的数据,同步开展基础校验工作,及时筛除无效异常数据,从源头规避因数据失真引发的漏损误判问题。借助数据清洗技术,修正传感器自身误差导致的压力数据异常情况,补齐巡检台账中缺失的关键信息,让管网运行的数据流更连贯完整^[4]。通过统一管网节点编码规则与监测数据格式,实现供水管网全维度数据的集中管控,打通各环节数据壁垒、破除数据孤岛。经过层层预处理的优质数据,搭配漏损识别模型协同运作,可精准定位管网微小漏损点位,相较传统运维模式,既能大幅压降管网漏损率,又能压缩应急处置响应时长,减少水资源流失与经济损耗,切实提升供水管网运维的整体效率与运行可靠性,筑牢城市供水安全防线。

3.2 排水管网运维中的应用实践

在排水管网运维场景中,预处理技术为城市防涝减灾、管网常态化疏通养护提供了硬核保障,有效解决了极端天气下数据处理滞后、风险预警不及时难题。暴雨天气下,管网监测端会产生海量数据,依托边缘计算开展前端初步预处理,快速甄别液位异常数据,保障数据处理的时效性,为防汛防涝决策部署抢占先机。运用数据清洗技术,剔除雨水冲刷冲击引发的传感器误报信息,补充各区域管网历史积水相关资料,让数据更贴合实际、精准度更高。通过多维度数据融合手段,整合气象预报、管网液位、地形地貌等多元数据,搭建专属排水防涝分析模型。依托预处理后的高质量数据支撑,可提前预判积水风险点位,联动排水泵站开展智能调度^[5];面对台风、强降雨等极端天气,能最大限度降低城市内涝造成的损失,推动排水管网运维向精细化升级,守护城市排水安全,保障居民日常出行平稳顺畅。

3.3 燃气管网运维中的应用实践

在燃气管网运维工作中,预处理技术有效提高了泄漏预警的准确性和及时性,为城市燃气安全筑牢防线。通过数据接入预处理,24小时实时采集、校验燃气浓度、压力等传感器数据,搭建稳定的数据传输通道,确保数据不丢失、不延迟。采用数据清洗技术,修正传感器漂移带来的错误数据,剔除环境干扰产生的多余数据,避免因数据偏差出现误预警、漏预警的情况。通过标准化和融合处理,统一管网权属信息、监测数据格式,结合地理空间数据,搭建燃气泄漏预警模型,实现泄漏风险的精准识别。此外,经过预处理的数据,

能够精准发现微量泄漏,提前发出预警并及时调度巡检人员处理,有效防范燃气爆炸等安全事故,保障城市公共安全和居民生命财产安全,推动燃气管网运维向智能化、安全化发展。

4 安全防护措施

4.1 全流程数据加密防护

针对管网运维大数据的传输、存储、使用这三个关键环节,构建全域覆盖、层级分明的多层次加密防护体系,实现数据全生命周期加密管控,筑牢数据安全底层防线。

数据传输环节,运用 SSL/TLS 协议做端到端加密,防止数据在传输过程中被人窃取或者篡改;在数据存储环节,采用 AES-256 加密算法,对结构化和非结构化数据都进行加密处理。其中,像管网核心位置、设施参数这类核心敏感数据,运用同态加密技术,能在数据不解密的情况下直接进行运算分析,从根本上避免了解密时可能出现的安全风险。与此同时,还会定期更新加密密钥,建立完善的密钥管理体系,明确密钥的生成、分发、销毁等流程,确保加密机制一直有效,从源头守住数据安全,满足管网运维数据在不同场景下的传输和存储需求。

4.2 精细化访问控制与操作溯源

以标准 RBAC 权限模型为核心底座,结合市政管网运维岗位分工、职责边界与业务场景,构建精细化、颗粒化的权限管控体系,严格恪守最小权限原则,实现权限分配精准化、权限管控刚性化。根据运维人员岗位职责、工作范畴进行权限细分,限定不同岗位人员仅可访问履职必需的数据资源与系统功能,杜绝越权访问、越级操作、非授权数据调取等违规行为。为强化身份准入安全,引入多因素认证机制,融合静态密码、动态验证码、设备唯一绑定、生物特征核验等多重验证方式,构建多层次身份校验屏障,有效防范密码泄露、暴力破解、仿冒登录等非法入侵行为,提升系统登录与访问的安全性^[6]。同步构建全维度操作日志审计体系,对数据访问、编辑、修改、删除、导出等全类型操作进行全程留痕,精准记录操作人身份、操作时间、操作终端、操作内容、操作结果等关键溯源信息,实现每一项数据操作均可追溯、可审计、可核查。

4.3 敏感数据脱敏与安全合规保障

对于管网运维数据中的敏感信息,运用数据脱敏技术处理。例如:管网位置、设施核心参数,运维人

员个人隐私等敏感信息分类实施差异化脱敏处理:管网核心点位采用坐标模糊化、范围化处理,设施核心参数进行脱敏加密与字段遮蔽,人员隐私信息隐藏身份证号、联系方式等关键字段,确保脱敏后数据仅用于管网运维分析、决策支撑等合规场景,杜绝敏感信息外泄。构建“本地实时备份+异地容灾备份”双重数据备份机制,明确备份频率、存储介质、校验规则、恢复流程等标准规范,定期开展数据备份与完整性校验。针对硬件故障、自然灾害、网络攻击、系统崩溃等突发应急场景,可快速启动数据恢复预案,最大限度降低数据丢失风险,保障管网运维业务连续性。在合规层面,严格按照《数据安全法》《个人信息保护法》等相关法律,贴合市政管网的行业标准和地方要求,制定完善的数据安全管理制度。

5 结束语

大数据预处理技术是破解市政地下管网运维数据难题、释放数据核心价值的关键抓手。这套技术体系的构建与落地应用,精准弥补了传统运维模式的短板,推动管网管理逐步迈向精细化、智能化方向。依托多环节预处理技术的协同配合,管网运维数据的整体质量与利用效率得到显著改善,也为供排水、燃气等各类管网的日常运维工作提供了扎实可靠的科学支撑。未来,要结合技术发展,优化预处理算法,提高应对复杂场景的能力,不断推动这项技术落地见效,为城市地下管网的安全稳定运行保驾护航,助力智慧城市建设。

参考文献:

- [1] 姜卓玮.地下管网管理系统建设实践与数字化转型探索[J].中国市政工程,2025(06):68-71,139.
- [2] 李琴文,郑璐.BIM技术在市政管网三维参数化设计中的应用[J].工程机械与维修,2025(06):110-112.
- [3] 杨可亮,刘家珣.基于AR+BIM的地下管网运维应用研究[J].工程质量,2025,43(04):57-61.
- [4] 赵起祥.工程测量技术在市政地下管网建设中的应用研究[J].科技与创新,2025(03):227-229.
- [5] 吴坚.浅析GPS-RTK在市政地下排水管网测量中的应用[J].城市建设理论研究(电子版),2024(33):170-172.
- [6] 郭兴根,刘帅,赵威.市政地下管网运维大数据预处理技术的研究与应用[J].城市建设理论研究(电子版),2023(28):211-213.