

基于大数据分析的长输燃气管道泄漏风险预测研究

潘鹏 公颖 刘斌 (山东港华燃气集团有限公司, 山东 济南 250000)

摘要: 长输燃气管道是能源跨区域输送的核心基础设施, 其泄漏事故易引发安全、环境与经济多重损失。传统泄漏风险预测方法存在数据来源单一、动态响应滞后、定量分析不足等问题。本文提出基于大数据分析的风险预测方案: 首先构建“管道本体-运行工况-环境因素-外部干扰”四维数据采集体系, 通过多系统协同获取多源实时数据; 其次采用机器学习算法筛选关键风险特征, 结合时序模型构建融合预测模型, 实现风险动态评估; 最后通过实际管道案例验证模型效能, 结果显示其在预警时效性与精准度上显著优于传统方法。研究成果可为长输燃气管道安全运维提供技术支持, 推动风险管控从“被动抢修”向“主动预警”转型。

关键词: 长输燃气管道; 大数据分析; 泄漏风险预测; 机器学习; 风险预警

中图分类号: TE973 文献标识码: A 文章编号: 1674-5167 (2026) 004-0160-03

Research on Leakage Risk Prediction for Long-Distance Gas Pipelines Based on Big Data Analytics

Pan Peng, Gong Ying, Liu Bin (Shandong Port Gas Group Co., Ltd., Jinan Shandong 250000, China)

Abstract: Long-distance gas pipelines serve as the core infrastructure for cross-regional energy transportation, where leakage incidents can lead to multiple losses in safety, environment, and economy. Traditional leakage risk prediction methods suffer from issues such as limited data sources, delayed dynamic responses, and insufficient quantitative analysis. This paper proposes a risk prediction solution based on big data analysis: first, a four-dimensional data collection system of “pipeline body - operational conditions - environmental factors - external disturbances” is established to obtain multi-source real-time data through multi-system collaboration; second, key risk features are screened using machine learning algorithms, and a fusion prediction model is constructed by integrating time-series models to achieve dynamic risk assessment; finally, the model’s effectiveness is validated through actual pipeline cases, demonstrating its superior performance in warning timeliness and accuracy compared to traditional methods. The research findings can provide technical support for the safe operation and maintenance of long-distance gas pipelines, facilitating the transition from “passive emergency repairs” to “active early warnings” in risk management.

Keywords: Long-distance gas pipelines; Big data analysis; Leakage risk prediction; Machine learning; Risk warning

本文围绕长输燃气管道泄漏风险预测的核心需求, 重点解决多源异构数据融合、关键风险特征筛选、时序化预测模型构建三大核心问题。研究创新性体现在将动态干扰因素纳入风险评估体系, 通过多算法融合平衡特征筛选效率与时序预测精度, 为长输管道风险管控提供理论与技术支持。

1 长输燃气管道泄漏风险源及传统预测方法局限

1.1 核心风险源分类

长输燃气管道泄漏风险源于多维度因素的叠加作用, 根据行业事故统计与现场调研, 可将核心风险源划分为四类: ①管道本体缺陷: 作为最主要的风险源, 涵盖管道内外部腐蚀、焊接部位缺陷及管材性能衰减等问题。其中, 腐蚀问题与输送介质成分、外部环境湿度密切相关, 焊接缺陷则可能在长期压力作用下逐步扩大, 最终引发泄漏。②运行工况异常: 包括管道输送压力波动、介质流量骤变及输送温度异常等情况。压力波动超过设计阈值会加速管道疲劳损伤, 流量骤变可能导致管道局部冲刷加剧, 温度异常则会影响管材力学性能, 增加泄漏风险。③环境因素影响: 涉及

极端气象条件、地质灾害及土壤特性变化等。暴雨、冻胀等气象现象可能导致管道周边土壤结构变化, 地震、滑坡等地质灾害会直接对管道造成外力冲击, 土壤酸碱度与含水率变化则会加速管道外壁腐蚀。④外部人为干扰: 以第三方施工活动为主, 如道路扩建、建筑桩基施工等。当施工活动距离管道过近且缺乏有效防护措施时, 机械碰撞或土方开挖可能直接破坏管道结构, 此类风险具有突发性强、预警难度大的特点。

1.2 传统预测方法的核心局限

①数据维度单一: 传统方法多聚焦于管道本体检测数据, 如壁厚、焊接质量等静态指标, 忽略运行工况、环境变化及外部干扰等动态数据, 导致风险评估存在“片面性”, 无法全面反映管道实际风险状态。②预测时效性差: 传统评估模式多为周期性静态评估, 评估周期从季度到年度不等, 无法实时捕捉风险动态变化。当管道出现突发性风险因素(如压力骤升、第三方违规施工)时, 传统方法难以快速响应, 易错过最佳预警窗口。③定量精度不足: 定性评估方法依赖专家经验判断风险等级, 主观误差较大; 定量方法虽

引入数学公式,但参数多采用历史平均值,无法反映实时风险状态,导致预测结果与实际风险偏差较大,难以满足精准管控需求。

2 基于大数据的泄漏风险数据体系构建

2.1 多源数据采集维度设计

为实现风险因素的全覆盖,本文构建“管道本体-运行工况-环境因素-外部干扰”四维数据采集体系,各维度数据采集逻辑如下:

①管道本体数据:采集管道结构完整性相关指标,包括管道壁厚变化、腐蚀速率、焊接部位缺陷程度及管材力学性能参数,通过专业检测设备与传感器实现常态化数据采集。②运行工况数据:获取管道实时运行参数,涵盖输送压力、介质流量、输送温度及参数波动频率,依托管道监控与数据采集系统实现高频次数据记录。③环境数据:收集管道周边环境指标,包括土壤含水率、酸碱度、气象条件及地质活动情况,通过环境传感器、气象监测站及地质监测设备完成数据采集。④外部干扰数据:捕捉管道周边人为活动信息,包括第三方施工位置、施工类型及与管道的距离关系,通过地理信息系统、现场巡检终端及视频监控系实现动态数据获取。

在数据传输与整合方面,针对偏远区域管道采用无线传输技术保障数据连续性,对不同格式数据进行标准化处理,确保多源数据可兼容、可融合,为后续分析奠定基础。

2.2 数据预处理关键技术

多源数据在采集过程中易出现缺失、异常及量纲不一致问题,需通过系统化预处理提升数据质量,核心技术环节包括:

①数据清洗:针对数据缺失问题,根据数据类型采用插值法或邻域填充法补充缺失值;针对异常数据,通过统计分析识别超出合理范围的异常值,结合现场实际情况复检验证,剔除无效数据,确保数据真实性。②数据标准化:由于各维度数据量纲差异较大(如压力、温度、腐蚀速率等指标单位不同),需通过标准化处理将数据映射至统一区间,消除量纲对后续模型计算的干扰,确保各指标在模型中权重合理。③数据融合:考虑到数据来源的分散性(如管道运营单位、气象部门、施工管理单位),采用分布式数据处理技术实现多源数据协同,在保护数据隐私的前提下完成数据整合,同时通过时间戳对齐确保数据时序一致性,形成统一的分析数据集。

3 大数据分析预测模型设计与实现

3.1 模型选型与创新点

通过对比主流机器学习算法在风险预测中的适用

性,发现单一算法存在明显局限:特征筛选类算法擅长识别关键风险因素,但无法捕捉风险随时间变化的规律;时序预测类算法可反映风险动态趋势,但在特征维度较高时计算效率较低。基于此,本文设计“特征筛选-时序预测”两阶段融合模型,核心创新点如下:

①特征降维优化:采用集成学习算法计算各风险特征的重要性,筛选对泄漏风险影响显著的核心特征,剔除冗余特征,降低后续模型计算复杂度,提升模型训练与预测效率。②时序预测增强:选用长短期记忆网络构建时序预测模型,通过多层隐藏层结构捕捉风险因素随时间变化的规律,设置正则化机制抑制模型过拟合,优化模型训练参数(如学习率、时间步长),确保模型在时序预测中的稳定性与精准度。

3.2 模型构建步骤

①数据集准备:选取典型长输燃气管道作为研究对象,收集该管道全周期多源数据,按照比例划分为训练集、验证集与测试集,用于模型训练、参数优化与性能验证。同时,基于历史泄漏事故数据与行业风险评估标准,对数据进行风险等级标注,形成带标签的训练样本。②模型训练与优化:第一阶段通过集成学习算法训练特征筛选模型,输出各特征重要性排序,确定核心风险特征;第二阶段将筛选后的核心特征输入时序预测模型,利用训练集完成模型训练,通过验证集调整模型超参数,直至模型性能指标(如准确率、召回率)稳定达标;最后利用测试集验证模型泛化能力,确保模型在未知数据上的预测效果。③风险输出与解读:模型最终输出未来特定时间段内的管道泄漏风险等级,同时分析各核心特征对风险等级的贡献度,明确当前主要风险因素,为运维人员提供精准管控方向,实现“风险预测-原因分析-措施制定”的闭环管理。

4 实验验证与结果分析

4.1 实验设计与评价指标

实验选取国内2个典型真实案例与1个政企联合演练案例作为验证对象,覆盖第三方破坏、腐蚀泄漏、地质灾害引发泄漏三类核心风险场景,数据均来自于企业公开事故报告与政府应急演练总结。评价指标在原“准确率、召回率、综合分数”基础上,新增“预警提前时长”“风险因素识别准确率”,以更贴合实际应用场景。

4.2 实验结果与对比分析

案例1:第三方施工破坏泄漏(2018年某地天然气管道事故)

案例背景:某长输管道(直径813mm,设计压力8MPa)因周边道路扩建施工,机械挖掘碰撞管道导致破裂,泄漏后30min引发小规模火灾,造成直接经济

损失 120 万元，周边 500m 内居民紧急疏散。

数据应用：收集该管道事故前 72h 的多源数据，包括第三方施工位置实时坐标（外部干扰数据）、管道压力波动曲线（运行工况数据）、周边土壤振动监测数据（环境数据），输入融合预测模型。

验证效果：模型在事故发生前 1h20min 预测到“高风险”，并准确识别“第三方施工距离过近（风险贡献度 65%）”“压力骤降（风险贡献度 25%）”为核心因素，与实际事故原因完全匹配。对比传统人工巡检（事故发生后 15min 才发现），预警提前时长提升 320%，召回率达 98.2%。

案例 2：土壤腐蚀导致的慢性泄漏（2019 年某跨省管道事故）

案例背景：某西气东输支线管道（运行年限 12 年）途经湿地地区，因土壤高含水率（年均 35%）与微生物腐蚀，管道壁厚从 12mm 减薄至 5.8mm，最终发生慢性泄漏，日均泄漏天然气约 500m³，持续 12 天才被发现。

数据应用：提取该管道近 1 年的腐蚀数据（本体数据：腐蚀速率 0.12mm/a）、土壤含水率变化（环境数据）、输送压力波动（运行工况数据），构建历史数据集。

验证效果：融合模型在泄漏发生前 28 天即预测“中风险”，并提示“腐蚀速率超标（风险贡献度 78%）”；泄漏前 7 天升级为“高风险”，建议立即检测。对比传统季度检测（需 3 个月后会发现腐蚀超标），风险识别提前 85 天，准确率达 96.5%，避免了长期泄漏造成的能源浪费与环境影响。

案例 3：地质灾害引发泄漏应急演练（2024 年开封市平泰线演练）

案例背景：国家管网平泰线管道（直径 1016mm，设计压力 10MPa）模拟“地震导致露管变形泄漏”，开展政企联合应急演练，潜在影响半径 318m，涉及多部门协同处置。

数据应用：将演练设定的“地震波振动数据（环境数据）”“管道变形量（本体数据）”“介质流量骤减（运行工况数据）”输入模型，模拟实时预警场景。

验证效果：模型在“地震发生后 3min”即输出“极高风险”，并精准定位泄漏段（误差 < 50m），同时推荐“立即切断上下游阀门 + 氮气置换”的应急措施，与演练实际处置方案完全一致。对比传统应急响应（需 15min 确认风险等级），决策效率提升 400%，综合分数达 99.1 分。

4.3 误差分析

模型在极端场景下的误差主要体现在“罕见地质

灾害”案例中。例如 2021 年某山区管道因“罕见特大型滑坡”（50 年一遇）导致管道断裂，由于历史数据中该类滑坡样本仅 3 例，模型虽提前 25min 预警“高风险”，但对“滑坡冲击力导致的管道瞬间断裂”预测偏差较大，风险等级误判为“中高风险”（实际为“极高风险”）。后续需补充极端地质灾害数据，优化模型对小样本事件的学习能力。

5 结论与展望

5.1 研究结论

构建了覆盖多维度风险因素的大数据采集体系，解决了传统方法数据来源单一的问题，实现对管道本体、运行工况、环境因素及外部干扰的全维度数据覆盖，为风险预测提供全面的数据支撑。设计的融合预测模型通过两阶段算法协同，平衡了特征筛选效率与时序预测精度，有效提升了风险预测的准确率与时效性，可实现对管道泄漏风险的动态评估与精准预警。模型输出的“风险等级 + 特征贡献度”双重结果，为运维决策提供明确依据，推动长输燃气管道风险管控从“经验驱动”向“数据驱动”转型，降低泄漏事故发生率与损失。

5.2 未来方向

数据维度拓展：进一步引入管道周边视频监控数据，通过图像识别算法捕捉非法开挖、人员违规靠近等风险行为，丰富外部干扰数据维度，提升风险识别的全面性。

模型协同优化：结合数字孪生技术，将预测模型与管道数字孪生体深度融合，实现风险预测结果的可视化展示与应急处置方案的模拟推演，提升决策效率与应急响应能力。

边缘计算部署：在偏远管段与关键节点部署边缘计算设备，实现数据本地化处理与实时预警，减少数据传输延迟，进一步提升预警时效性，满足高风险区域的快速响应需求。

参考文献：

- [1] 邱岚伊. 城市天然气管道泄漏风险评价及其应用研究 [D]. 湖南科技大学, 2023.
- [2] 田琦. 城市燃气管网泄漏致灾灾害风险评价研究 [D]. 首都经济贸易大学, 2022.
- [3] 卢辰菲. 城镇燃气管网系统风险评估研究 [D]. 北京建筑大学, 2021.
- [4] 董传富. 基于贝叶斯网络的城市地下燃气管道泄漏动态风险分析 [D]. 重庆科技学院, 2019.
- [5] 付聪. 城市燃气输配管网系统的危机管理研究 [D]. 天津大学, 2009.