

Design of big data-driven oil and gas pipeline operation status monitoring and early warning system

Jiawei Yan Bize Yang

State Petroleum and Gas Pipeline Network Group Co., Ltd. Northeast Branch, Shenyang, Liaoning, 110167, China

Abstract

This study focuses on the big data-driven monitoring and early warning system of oil and gas pipeline operation status, and deeply explores the application basis of big data technology in oil and gas pipeline safety management. We systematically elaborated the overall architecture, key technologies and innovative strategies of the system, aiming to build an accurate and intelligent monitoring and early warning system to achieve timely identification and effective response of oil and gas pipeline operation risks, so as to ensure the safe and stable operation of the pipeline. The research results provide a new technical path and solid theoretical support for the safety management of oil and gas pipelines, and strongly promote the development process of the oil and gas pipeline industry to the direction of intelligence, which has important practical significance and theoretical value.

Keywords

big data; oil and gas pipeline; monitoring and early warning; system design

大数据驱动油气管道运行状态监测预警系统设计

严佳伟 杨璧泽

国家石油天然气管网集团有限公司东北分公司, 中国·辽宁 沈阳 110167

摘要

本研究专注于大数据驱动的油气管道运行状态监测预警系统, 深入探索大数据技术在油气管道安全管理中的应用基础。我们系统阐述了该系统的整体架构、关键技术以及创新策略, 旨在构建一个精准、智能的监测预警体系, 以实现油气管道运行风险的及时识别与有效应对, 从而确保管道的安全稳定运行。本研究成果为油气管道的安全管理提供了全新的技术路径与坚实的理论支撑, 有力推动了油气管道行业向智能化方向发展的进程, 具有重要的实践意义与理论价值。

关键词

大数据; 油气管道; 监测预警; 系统设计

1 引言

油气资源作为现代工业的核心支柱, 油气管道则是其安全、高效输送的关键设施, 其安全稳定运行对能源供应和工业发展至关重要。随着油气管道规模不断扩张, 运行环境日益复杂, 传统监测手段在数据处理能力、实时性和准确性等方面的局限性逐渐凸显, 已难以满足当前严格的安全管理需求。大数据技术的兴起, 凭借其强大的数据处理和分析能力, 为油气管道监测预警领域带来了新的发展机遇。开展大数据驱动的油气管道运行状态监测预警系统设计研究, 对于提升管道安全保障水平、推动油气行业智能化转型升级具有重要的现实意义和学术价值。

2 大数据技术在油气管道监测预警中的应用基础

2.1 数据采集

数据采集是大数据驱动监测预警系统的基石。在油气管道监测预警系统中, 需部署多种类型传感器, 对管道及其周边环境进行全方位实时监测, 以获取海量数据。温度传感器用于监测管道壁和介质的温度变化, 压力传感器负责测量管道内介质压力值, 位移传感器则用于跟踪管道的位移情况^[1]。在管道的连接处、阀门附近以及穿越复杂地质区域的管段等关键节点布置传感器, 能够及时捕捉潜在异常。此外, 环境参数如雨量、气温等对管道运行安全也有重要影响。雨量传感器记录降水情况, 大量降雨可能引发地质灾害威胁管道安全。气温传感器监测环境温度, 极端气温变化会影响管道材料性能。这些运行参数和环境参数共同构成了反映油气管道运行状态的数据集合, 为后续数据处理和分析提供了丰富素材。

【作者简介】严佳伟(1990-), 男, 中国宁夏人, 本科, 工程师, 从事油气储运研究。

2.2 数据处理与分析

采集到的原始数据通常包含噪声和异常值，若直接用于分析，会严重影响结果的准确性。因此，必须对数据进行预处理，主要包括数据清洗和归一化操作。数据清洗通过设定合理的阈值和规则，剔除明显异常的数据。归一化处理则将不同量纲的数据统一映射到特定区间，增强数据的可比性。完成预处理后，借助数据挖掘和机器学习技术对数据进行深度分析。数据挖掘算法能够发现数据中隐藏的模式和规律，例如找出管道压力、流量、温度等参数之间的内在联系。机器学习算法如决策树、神经网络等，则可对管道运行数据进行学习和分析，识别异常模式，进而评估管道存在的泄漏、爆炸等风险。

2.3 预警与决策支持

基于数据分析结果，监测预警系统能够实时发出预警信号。当系统判定管道存在安全隐患时，会通过声光报警、短信通知、系统弹窗等多种方式及时告知管理人员。同时，系统具备决策支持功能，它会依据管道的实时运行状态、风险类型以及历史数据经验，运用智能算法为管理人员提供应对策略。例如，在管道发生轻微泄漏时，系统会综合考虑泄漏位置、周边环境、管道输送介质特性等因素，推荐合适的抢修方案，助力管理人员快速做出科学决策，有效降低事故发生概率及其造成的影响。

3 大数据驱动油气管道监测预警系统设计研究

3.1 系统架构设计

数据采集层作为系统的“感知触角”，在管道关键节

点部署高密度、多类型的传感器网络。这些关键节点一般选择在管道的连接处、阀门附近、穿越复杂地质区域的管段等易出现故障的部位。传感器网络以高频率实时采集数据，确保能够及时捕捉到管道运行状态的细微变化。为保障数据采集的准确性和可靠性，采用冗余设计，对重要参数设置多个传感器进行监测，当某个传感器出现故障时，其他传感器仍能正常工作，维持数据的连续性^[2]。

数据传输层是连接数据采集层与数据中心的“高速通道”。利用4G、5G等高速无线网络以及光纤通信技术，搭建稳定可靠的数据传输链路。针对不同的数据传输需求，制定差异化的传输策略。对于实时性要求极高的关键数据，如管道压力、流量的突变数据，优先采用低延迟的5G网络进行传输。对于批量的历史数据或非紧急数据，则通过光纤网络进行稳定传输，确保数据在传输过程中不丢失、不损坏，实现数据的实时、准确汇聚至数据中心。

数据处理层面对海量的采集数据，采用分布式计算框架Hadoop和Spark。Hadoop的分布式文件系统(HDFS)能够将大规模数据分散存储在多个节点上，实现数据的可靠存储和高效管理。MapReduce编程模型将数据处理任务分解为多个子任务，并行运行在不同节点上，大幅提高处理效率。Spark基于内存计算，相较于Hadoop的磁盘I/O操作，能够更快地处理数据，尤其适用于迭代式的机器学习算法(见表1)。通过这两种框架的协同工作，可对海量数据进行高效清洗、转换和分析，为后续的数据分析提供高质量的数据支持。

表1 Hadoop与Spark框架功能对比

框架名称	分布式文件系统	数据处理模型	处理速度特点	适用场景
Hadoop	HDFS	MapReduce	磁盘I/O多，速度相对慢	大规模数据存储与离线批处理
Spark	可集成HDFS	内存计算相关模型	内存计算，速度快	迭代计算、交互式与实时流处理

数据分析层运用先进的机器学习算法。随机森林算法通过构建多个决策树，对数据进行分类和预测，利用多个决策树的投票机制提高模型的准确性和稳定性，能够有效识别管道运行中的各类异常模式。深度学习算法，如卷积神经网络(CNN)和循环神经网络(RNN)，在处理复杂数据和时序数据方面具有独特优势。CNN能够自动提取图像化数据(如管道内部的超声检测图像)的特征，用于管道缺陷的识别。RNN则擅长处理具有时序特征的管道运行数据，预测管道未来的运行趋势，提前发现潜在风险。

3.2 关键技术研究

数据预处理是保障数据质量的核心环节。在数据清洗方面，除了运用简单的阈值判断，还引入基于统计分析的方法，如3 σ 准则，更精准地识别和去除数据中的异常值。针对数据缺失问题，依据数据的特点和分布情况，选择合适的填充方法，如均值填充、中位数填充或基于机器学习模型的预测填充。在归一化处理中，针对不同类型的数据，分别采

用线性归一化、Z-score标准化等方法，充分发挥数据在后续分析中的价值。

特征提取技术旨在从原始数据中挖掘出最能反映管道状态的关键特征。基于信号处理的方法，对管道振动信号进行傅里叶变换、小波变换等，将时域信号转换到频域，提取出反映管道运行状态的频率特征。从数据的统计特征入手，计算数据的均值、方差、偏度、峰度等统计量，这些统计量能够从不同角度刻画数据的分布特征，为管道状态评估提供丰富信息。利用机器学习中的特征选择算法，如递归特征消除(RFE)和随机森林的特征重要性评估，筛选出对管道状态分类或预测贡献最大的特征，提高模型的训练效率和准确性。

模型优化技术是提升系统性能的核心手段。网格搜索通过在指定的参数空间内遍历所有可能的参数组合，找到使模型性能最优的参数设置，这种方法虽然计算量大，但能够保证找到全局最优解。贝叶斯优化则基于贝叶斯理论，通过

构建概率模型来预测参数的最优值,相比网格搜索,它能够更智能地选择参数进行试验,减少搜索次数,提高优化效率。还可以采用集成学习的方法,将多个不同的模型进行融合,如将多个神经网络模型进行堆叠,综合它们的预测结果,提升模型的泛化能力和稳定性。

系统集成技术关乎整个系统的稳定性和可扩展性。在硬件集成方面,充分考虑不同设备的兼容性和性能匹配,确保传感器、服务器、通信设备等硬件组件能够协同工作。在软件集成上,采用面向服务的架构(SOA),将系统划分为多个独立的服务模块,每个模块通过标准的接口进行通信和交互,便于系统的维护和扩展。引入容器化技术,如Docker,将每个服务模块封装成独立的容器,实现环境隔离和快速部署,提高系统的部署效率和可维护性。

3.3 创新策略

3.3.1 多源数据融合

传统的油气管道监测往往仅依赖单一类型的数据,这种方式具有很大的局限性,难以全面掌握管道的实际运行状况。本系统的多源数据融合策略有效打破了这一局限。它整合了来自压力、温度、位移等不同传感器的数据,从多个维度反映管道的运行状态。压力数据能体现管道内部介质的受力情况,温度数据可辅助判断介质特性及管道是否存在异常摩擦生热,位移数据则能反映管道结构的稳定性,这些数据相互补充,提供了更全面的信息。同时,结合气象数据、地质数据等外部环境数据,进一步丰富了数据的维度^[1]。例如,暴雨天气可能引发地质灾害威胁管道安全,通过融合气象数据中的雨量信息和地质数据中的地形地貌、土壤特性等信息,可以提前预判风险。此外,利用卡尔曼滤波、D-S证据理论等数据融合算法对多源数据进行综合处理,能够有效消除数据中的噪声和矛盾信息,提高监测的全面性和准确性,从而更精准地识别管道运行中的潜在风险。

3.3.2 实时动态更新

油气管道的运行状态处于不断变化之中,为了使监测预警系统能够及时适应这些变化,实时动态更新策略至关重要。系统采用在线学习算法,如随机梯度下降(SGD)。在新数据到来时,模型能够实时更新参数,不断学习新的模式和规律。这种机制让模型始终保持对管道运行状态的敏锐

感知,不会因为运行情况的变化而滞后。当管道运行工况发生变化,比如输送介质调整、运行压力改变,或者出现新的异常情况时,动态数据更新机制会及时将新数据纳入模型训练。通过这种方式,模型能够快速适应新的运行状况,始终保持良好的性能,大大提高了预警的及时性和准确性,避免因模型陈旧而导致的预警失误。

3.3.3 用户交互设计

以用户需求为导向的用户交互设计,旨在提升用户体验,提高管理效率。系统采用可视化技术,将复杂的管道运行数据、预警信息、决策建议等以图表、地图等直观形式呈现给用户。通过图表,用户可以清晰地看到管道各参数的变化趋势。借助地图,能够直观了解管道的地理位置分布以及异常发生的具体位置,方便快速获取关键信息。同时,系统设置了丰富的交互功能。用户可以通过界面实时查询管道历史数据,为分析管道运行规律提供依据。能够根据实际情况调整预警阈值,使预警更加符合实际需求。还可以反馈决策执行效果,帮助系统不断优化决策建议,增强了用户与系统的互动性,有效提高了决策效率和管理水平。

4 结论

本研究围绕大数据驱动的油气管道监测预警系统展开了全面深入的设计探讨。构建的系统有望实现对油气管道运行状态的精准监测和预警,为油气管道安全管理提供有力支持,推动油气行业向数字化、智能化转型。然而,在实际应用过程中,仍面临数据安全与隐私保护、系统性能优化等诸多挑战。未来需持续关注相关技术的发展动态,不断完善系统设计,以保障油气管道安全可持续运行,促进油气行业的稳健发展。

参考文献

- [1] 艾力群,吴冈刚,李承相,杨陈国,金山.基于大数据的油气管道雷电监测预警系统的设计与应用[J].设备管理与维修,2025(1):150-155.
- [2] 何栋楠,陈胜利.油气田储罐腐蚀在线监测与预警系统设计[J].化工设计通讯,2024,50(12):29-3135.
- [3] 卢艳.油气管道应力监测系统研究与现场应用[J].石油天然气学报,2023,45(2):183-194.