

基于大数据的输油气设备运行状态分析

李德金

国家石油天然气管网集团有限公司山东分公司

摘要: 随着输油气行业的快速发展,设备运行状态的实时监测和故障预测已成为保障系统稳定性和安全性的关键。本文基于大数据技术,研究了输油气设备运行状态的分析与优化方法。构建了一套完整的设备运行状态监测系统,包含数据采集、传输、分析与存储等模块。研究结果表明,基于大数据的分析方法不仅能够提高设备的故障预测准确性,还能显著优化设备管理效率和运行成本,为输油气行业的智能化运维提供了重要支撑。

关键词: 大数据;输油气设备;运行状态分析;故障预测;优化方案;机器学习

本文通过分析输油气设备的工作原理及常见故障,阐述基于大数据的状态分析与预测方法;设计一套设备运行状态监测系统,详细探讨数据采集、预处理及分析模型的实现;通过实际案例验证优化方案的效果,评估其在提高设备效率和安全性方面的应用价值。

1 输油气设备运行状态分析的理论基础

1.1 输油气设备工作原理

输油气设备通常由管道、泵站、调节阀、压力传感器以及流量计等组成。其主要功能是通过管道网络将油气从生产地输送到消费市场。泵站提供动力,调节阀控制流量,压力传感器和流量计实时监测设备的运行状态^[1]。油气的输送是通过压力差驱动的,在这一过程中,泵站的工作状态和管道的健康状况对输送的稳定性至关重要。设备的工作原理涉及流体力学、电气控制和自动化调节等多个领域,通过精密的仪器和技术确保输油气过程中的高效与安全。

1.2 输油气设备常见故障分析

输油气设备常见故障包括管道泄漏、泵站故障、阀门失效以及传感器故障等。管道泄漏通常由于腐蚀、机

械损伤或安装缺陷导致,可能引发环境污染或安全事故。泵站故障可能源于机械损坏、电气问题或油气流量不稳定。阀门失效可能导致流量控制不准确,影响系统运行。传感器故障则会影响数据采集的准确性,影响后续的状态分析和故障预警。针对这些常见故障的检测和预测,对于保障输油气系统的安全运行至关重要。

2 大数据分析技术与方法

2.1 大数据的基本概念与特征

大数据是指无法通过传统数据处理方法获取、管理和处理的数据集合。它通常具备5个“V”特征:体量(Volume)、速度(Velocity)、多样性(Variety)、价值(Value)和真实性(Veracity)。体量指数据量庞大,速度指数据生成和传输的速度,多样性指数据的多种格式(结构化、半结构化、非结构化数据),价值指数据中蕴含的信息价值,而真实性指数据的准确性和可靠性。大数据的这些特征使得它在设备监控和状态分析中具有强大的优势,尤其是在处理复杂、动态的设备运行数据时,大数据能够提供更为精准和及时的分析结果。

2.2 数据采集与传感器技术

数据采集技术主要依赖于传感器设备,这些传感器能够实时监测输油气设备的各项运行参数,如温度、压力、流量以及振动等。传感器的类型包括温度传感器、压力传感器、振动传感器以及流量计等,传感器的性能直接影响到数据的准确性和实时性。随着物联网技术的发展,传感器设备的智能化和网络化逐步提高,通过无线传感器网络(WSN)等技术,采集到的设备数据可以实时传输至数据处理平台,便于远程监控和实时分析^[2]。

2.3 数据预处理与清洗技术

数据预处理是对原始数据进行清洗、转换和规整化的过程,目的是消除数据中的噪声和异常值,提高数据的质量。数据清洗技术包括去重、填充缺失值、去除异常值等,确保数据的完整性和一致性。数据转换技术则包括标准化、归一化等,将数据转换为统一的格式或尺度,使其适合后续的分析。数据规整化技术用于将不同来源的数据进行合并和统一,以便进行多源数据融合分析。数据预处理的质量直接影响后续分析模型的准确性和可靠性。

3 输油气设备运行状态监测系统 系统设计

3.1 系统架构与设计原则

输油气设备运行状态监测系统的架构设计应具备灵活性、可扩展性和高可靠性。系统通常包括数据采集层、传输层、存储层、处理分析层和展示层 5 大核心模块。数据采集层负责通过传感器获取设备运行数据,如温度、压力、流量等;传输层负责将采集到的数据通过无线或有线网络传输至数据存储系统;存储层负责存储数据并进行备份;处理分析层利用大数据和机器学习算法进行数据分析与故障预测;展示层则通过图形化界面展示设备的实时状态和预警信息^[3]。设计原则应以安全性、稳定性和实时性为基础,确保系统能在复杂环境下稳定运行,同时具备良好的用户体验和可维护性。

3.2 数据采集与传输

数据采集是系统的基础环节,传感器的选择对数据采集的准确性至关重要。常见的传感器包括温度、压力、流量及振动等设备。通过智能传感器网络(WSN)将采集到的数据进行实时传输,确保每一台设备都能得到有效监控。传输层的技术可以选择基于物联网的低功耗广域网(LPWAN)技术,或通过 4G/5G 网络实现远程传输,以保证数据的实时性和传输距离。数据采集和传输的效率直接影响到后续分析的时效性,因此必须选用高精度的传感器和稳定的通信网络来确保数据的完整性和准确性。

3.3 实时监控与数据存储

实时监控是输油气设备监测系统的核心功能之一,通过设定关键参数的阈值,当设备参数超出设定范围时,系统应及时发出预警。实时监控不仅要求系统具备快速响应能力,还需要处理大规模、多维度的数据。数据存储则需要处理和保存大量的历史数据,通常采用分布式存储系统,以提高数据的存储能力和访问效率。通过云计算平台,系统

能够对存储的数据进行灵活地检索和分析,同时还需要采用备份机制确保数据不丢失。实时监控还需支持多终端展示,方便操作人员在不同位置对设备状态进行监控。

4 输油气设备运行状态数据分析与结果

4.1 数据采集与样本分析

数据采集是输油气设备运行状态监测中的第一步,确保数据的准确性和代表性是分析的基础。设备通过传感器实时监测多个指标,如压力、温度、流量及振动等,采集的数据应覆盖设备的正常运行状态和可能的异常情况。数据采集不仅要考虑采样频率,还需确保数据的质量。通过数据清洗和预处理,剔除噪声和异常值后,可以得到更为精确的样本数据,这些数据将作为后续分析的基础^[4]。在样本分析中,采用统计分析方法对数据的分布、趋势和相关性进行研究,从而为后续模型构建和状态评估提供可靠的输入。

4.2 设备运行状态分类与识别

在输油气设备运行状态监测中,分类模型的选择对故障检测和异常识别具有重要影响。本研究选取了多种常

见且实用的机器学习模型,包括决策树(Decision Tree)、支持向量机(SVM)、随机森林(Random Forest)、神经网络(NN)以及长短期记忆网络(LSTM),并分别对它们在准确率、召回率、精确率及 F1 分数等指标上的表现进行了评估。如图 1 所示,LSTM 由于能够捕捉时序数据的长期依赖关系,其在准确率和 F1 分数等关键指标上表现出明显优势,可达到约 93% 的准确率和 90.5% 的 F1 分数;而神经网络(NN)在同样具备较高预测能力的同时,也保持了较好的泛化能力。随机森林在抗过拟合和整体稳定性方面表现良好,准确率和 F1 分数均接近 88% 以上。支持向量机(SVM)和决策树的性能略低一些,但它们训练和推理速度较快,易于实现,对于某些对实时性要求较高或数据规模相对较小的场景,仍有其应用价值。LSTM 在本次研究的分类任务中性能最佳,但针对不同应用需求,仍需综合考虑模型的复杂度、训练成本以及可解释性等因素来选择合适的分类算法^[5]。

4.3 故障预测与预警模型

在输油气设备的故障预测与预警中,常见方法包括线性回归、随机森林回归、支持向量回归以及基于深度学习

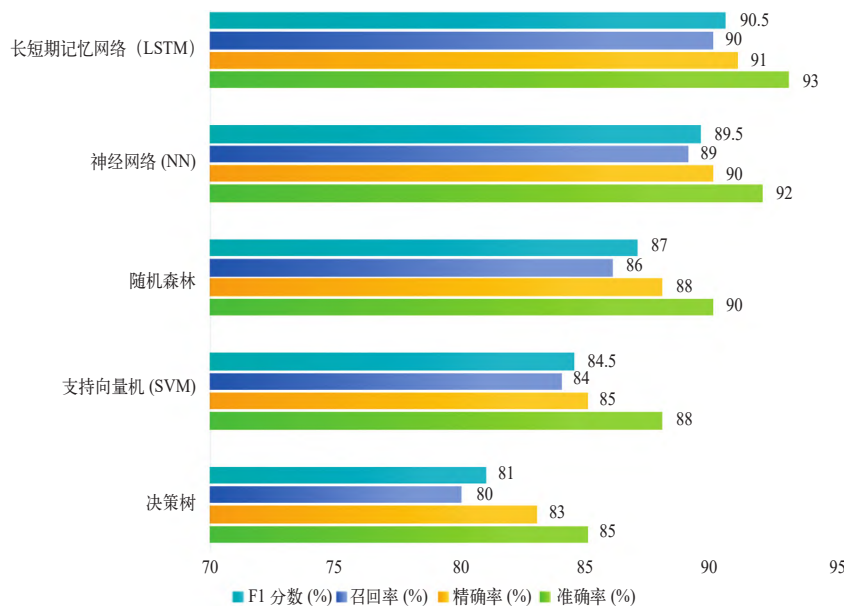


图 1 设备运行状态分类模型性能比较

的模型。为了准确评估模型性能，本文选取了平均绝对误差（MAE）、均方根误差（RMSE）和决定系数（R²）作为主要评价指标，并分别对各模型在不同数据集上的预测精度进行了对比分析。

如图 2 所示，对比线性回归与随机森林回归可以发现，随机森林在预测精度上表现更优，MAE 与 RMSE 较低且 R² 值更高，说明其具备更强的拟合与泛化能力。而线性回归虽然易于解释且计算开销较低，但在设备故障这一类可能存在非线性关联的场景中表现相对不足。在对输油气设备进行故障预测时，若数据规模和多样性较大且对预测精度要求高，可以优先考虑随机森林或其他非线性模型。需要结合设备运行实际情况以及数据采集质量，综合权衡模型的可解释性、计算成本与预测准确度，选择更契合工程需求的预测模型。

4.4 运行效率与安全性分析

设备的运行效率和安全性分析是对设备状态监测和管理效果的重要评价指标。运行效率分析通常包括对设备的实际运行时间、停机时间、能源消耗等进行量化分析，评估设备的工作效率。通过计算单位时间内的流量和能耗，分析设备的工作效率是否符合预期。安全性分析则侧重于设备的故障率、事故发生频率、应急响应时间等安全相关指标的评估。大数据分析技术能够帮助识别设备的运行瓶颈、隐患区域和高风险环节，为设备的维护决策提供数据支持，确保设备运行过程中的安全性和稳定性。

5 基于大数据的输油气设备运行状态优化方案

5.1 运行状态优化的目标与意义

运行状态优化旨在通过大数据技术提升输油气设备的整体性能和可靠性，降低运营成本，确保能源运输的安全与高效。具体目标包括实时监测设备运行状态，准确预测潜在故障，优化维护计划，以及提高设备的运行效率和使用寿命。通过大数据分析，可以深入理解设备的运行模式和故障机制，从而实现精细化管理。优化运行状态不仅能够减少意外停机和维修费用，还能提升能源输送的连续性和稳定性，保障能源供应链的畅通。优化方案有助于环境保护，减少因设备故障导致的泄漏和污染事故，提升企业的社会责任形象。

5.2 优化方案实施的效果评估

优化方案实施后的效果评估是确保方案有效性和持续改进的重要环节。设定明确的评估指标，如设备运行效率、故障率、维护成本、能源消耗及安全事故频率等，通过定量和定性分析，全面衡量优化方案的实施效果。采用前后对比的方法，比较优化前后的关键指标变化，评估优化措施的实际效果。利用数据可视化工具，展示优化成果，便于管理层和有关人员理解和决策。进行案例分析，选取典型设备或项目，深入分析优化方案在实际应用中的表现和收益。还需收集用户反馈和运维人员的意见，了解优化方案在实际操作中的可行性和改进空间。基于评估结果，进行方案的

调整和优化，确保持续提升设备运行状态。通过系统的效果评估，不仅能够验证优化方案的有效性，还能为后续的优化工作提供数据支持和经验借鉴，推动输油气设备管理的持续进步。

6 结语

本文从输油气设备运行状态的监测与分析入手，结合大数据技术，提出了一套完整的设备运行状态监测与优化方案。研究表明，通过传感器数据的实时采集与分析，可有效识别设备的异常状态，并基于机器学习与深度学习模型实现精准的故障预测与预警。通过优化方案的实施，设备的运行效率和安全性得到了显著提升，突发故障率和维护成本显著降低。通过实际案例的验证，进一步证明了基于大数据技术的监测系统在实际应用中的可行性和有效性。然而，本研究仍存在一些局限性，如部分模型在处理数据噪声时的稳定性，以及对复杂环境下多源数据融合和优化需求。在未来的研究中，可进一步引入更加复杂的深度学习模型和增强型数据融合技术，以提高设备状态分析的精准度和鲁棒性。

参考文献：

[1] 于洪乾. 浅析石油化工电气安全合规性管理[J]. 流程工业, 2024(09): 22-26.
[2] 刘文华, 张晓松, 杜华东, 等. 基于 GIS 的三维可视化输油气站场设备设施管理系统设计与实现[J]. 信息系统工程, 2021(09): 50-55.
[3] 胡庆有. 长输油气管道维抢修设备集装化建设实践应用[J]. 石化技术, 2020, 27 (11): 76-79.
[4] 陈凯. 长输油气管道工艺设备的自动控制技术研究[J]. 工程建设与设计, 2020(02): 277-280.
[5] 陈镜宇. 输油气管道关键设备储备定额研究[J]. 中国物流与采购, 2017(03): 76-77.

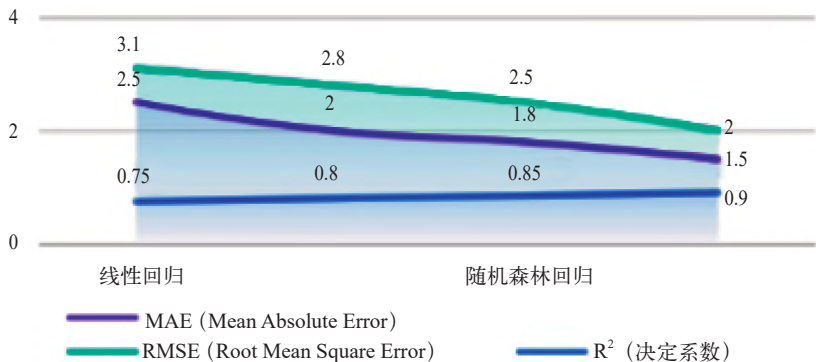


图 2 故障预测模型性能评估