

火电厂大数据分析平台建设与应用

刘士彩, 刘涛, 蒋俊英, 贺辽, 菅晓丽

国能山西河曲发电有限公司, 山西 忻州 036500

DOI:10.61369/ERA.2026020023

摘 要 : 火电厂作为我国能源系统的核心组成部分, 正面临能效提升、污染控制及设备可靠性优化的多重挑战。大数据分析技术的突破为火电厂运行优化提供了全新路径, 通过整合多源异构数据并构建分布式存储架构, 有效解决了传统数据孤岛问题, 为实时能耗监测与动态决策提供了技术支撑。具体而言, 基于 Spark 框架的 Mini-Batch K-means 算法与 Elbow 方法结合, 可实现锅炉燃烧效率与污染物排放的关联规律挖掘, 优化燃烧参数配置并降低煤耗; 基于时间序列数据分析的分类方法与卷积神经网络预警模型的应用, 则显著提升了汽轮机等核心设备的故障诊断能力与异常工况响应速度。这些技术手段不仅响应了国家能源战略需求, 更成为《国家重点支持的高新技术领域》及河北省火力发电节能环保工程技术研究中心的重点攻关方向。

关 键 词 : 火电厂; 大数据分析平台; 数据采集与预处理; 数据分析与挖掘

Construction and Application of Big Data Analysis Platform for Thermal Power Plants

Liu Shicai, Liu Tao, Jiang Junying, He Liao, Jian Xiaoli

CHN ENERGY Shanxi Hequ Power Generation Co., LTD. Xinzhou, Shanxi 036500

Abstract : As a cornerstone of China's energy infrastructure, thermal power plants are confronting multifaceted challenges including energy efficiency enhancement, pollution control, and equipment reliability optimization. Breakthroughs in big data analytics have pioneered innovative approaches for operational optimization. By integrating multi-source heterogeneous data and establishing distributed storage architectures, these technologies effectively resolve traditional data silo issues, providing technical support for real-time energy consumption monitoring and dynamic decision-making processes. Specifically, the integration of Spark-based Mini-Batch K-means algorithm with Elbow method enables mining of correlation patterns between boiler combustion efficiency and pollutant emissions, optimizing combustion parameter configurations while reducing coal consumption. Meanwhile, the application of time series data analysis classification methods combined with convolutional neural network (CNN) early-warning models has significantly enhanced fault diagnosis capabilities and abnormal operating condition response speed for core equipment such as steam turbines. These technologies not only respond to the national energy strategic needs, but also become the key research direction of "National Key Supported High-tech Fields" and Hebei Provincial Thermal Power Energy Conservation and Environmental Protection Engineering Technology Research Center.

Keywords : thermal power plant; big data analysis platform; data collection and preprocessing; data analysis and mining

引言

火电厂作为我国能源系统的重要组成部分, 承担着保障电力供应和推动经济发展的核心职能。随着能源转型与环保政策的深化, 传统火电企业在能效提升、污染控制及设备可靠性方面面临严峻挑战。近年来, 大数据分析技术的突破为火电厂优化提供了全新路径。通过整合多源异构数据并构建分布式存储架构, 能够有效解决传统数据孤岛问题, 为实时能耗监测与动态决策提供技术支撑。例如, 基于 Spark 框架的 Mini-Batch K-means 算法与 Elbow 方法的结合^[1], 可实现运行模式的精准划分与数据离散化处理, 从而挖掘锅炉燃烧效率与污染物排放间的关联规律。这一技术突破不仅优化了燃烧参数配置, 还显著降低了单位发电煤耗, 验证了大数据技术在火电厂经济性分析中的应用价值。

一、火电厂大数据分析平台设计

(一) 数据采集与预处理设计

火电厂大数据分析平台的数据采集与预处理设计需综合考虑数据源多样性、处理效率及模型适配性。在数据采集环节，硬件系统需建立标准化接口以实现多源数据的高效接入。通过采用模块化设计的数据接口装置，控制器可协调输入工具、存储器及建模接口，实现锅炉设计程序、汽轮机热力参数等核心数据的实时采集与格式转换。该装置通过2D建模接口与结构分析模块的联动，有效解决了异构数据在传输过程中的兼容性问题，确保锅炉钢框架分析与三维建模所需的原始数据完整性^[2]。此外，结合物联网技术构建分布式采集网络，能够整合电厂运行监测系统、环境传感器及设备状态监测装置等多源数据流，形成涵盖燃料消耗、热力循环效率、排放指标等关键参数的综合数据集。

数据预处理阶段主要包含清洗、降噪与特征标准化三类核心流程。针对汽轮机点云数据中存在的噪声干扰，采用基于双边滤波的多级预处理方法可有效提升模型精度。具体流程包括：首先通过空间分割算法对原始点云数据进行区域划分，继而利用双边滤波器在保留几何特征边界的前提下消除随机噪声，最终通过边缘校正技术确保曲面重构的拓扑连贯性。实验表明，该方法将汽轮机三维建模的均方误差控制在10以下，显著优于传统k-means聚类与混合滤波算法。对于锅炉设计中的结构分析数据，需通过异常值检测与插值算法修复缺失值，确保钢框架应力分布、热膨胀系数等参数的连续性。此外，为满足卷积神经网络等机器学习模型的输入要求，需对温度场、压力分布等时序数据进行标准化处理，通过Z-score变换消除量纲差异，并采用滑动窗口技术构建时序特征矩阵，为后续分析奠定数据基础。

(二) 数据分析与挖掘设计

火电厂大数据分析平台的构建以数据分析与挖掘为核心，其设计需结合算法模型的科学性和实现过程的可操作性。在算法选择上，基于燃料特性与燃烧效率的关联分析，采用统计建模和机器学习方法对关键参数进行深度挖掘。例如，针对燃煤电厂飞灰中未燃烧碳(UBC)的优化问题，研究通过分析燃料热值、挥发分、水分及灰分等参数与UBC含量的关联性，发现燃料热值在5000-6000 Kcal/kg区间时，UBC含量稳定在5%-10%^[3]，表明燃料特性对燃烧效率具有显著影响。此类分析为燃烧过程的动态优化提供了量化依据，同时也验证了数据驱动模型在复杂工业场景中的适用性。

在模型设计方面，网络物理系统(CPS)的集成被认为是提升模型可靠性的关键路径。通过将物理模型与大数据分析相结合，能够有效规避单一数据驱动建模的局限性。例如，物理模型可描述锅炉燃料流量、汽轮机调速阀开度等变量间的机理关系，而大数据分析则补充了实际运行数据的统计特征。这种融合模式在火电厂协调控制系统的非线性分析中得到应用，研究基于机组负荷、汽轮机压力及汽包水位的多变量微分方程模型，结合高斯统计高阶谱分析不同负荷工况下的非线性程度，结果表明低负荷工况下的非线性效应更为显著，为控制策略的参数优化提供了理论支撑。

二、数据收集与分析方法

(一) 数据收集途径

火电厂大数据分析平台的数据收集体系构建需兼顾多源异构数据的高效整合与质量保障，其核心在于建立覆盖生产全流程的多维度数据采集网络。数据来源涵盖发电机组核心设备、辅助系统及外部环境三大类：锅炉、汽轮机、发电机等主设备的运行参数通过传感器和智能仪表实时监测；脱硫、除尘、输煤等辅助系统提供设备状态与能耗数据；同时集成气象监测、燃料供应、电网调度等外部数据源，形成完整的生产决策支持基础^[4]。在数据采集方式上，采用分层异构集成策略，SCADA系统通过OPC协议实现毫秒级实时数据采集，DCS系统以固定周期(通常为5-15秒)导出结构化历史数据，而智能仪表则通过Modbus/TCP等工业协议进行周期性轮询。针对非结构化数据，部署图像识别系统对设备巡检视频和热成像图片进行结构化解析，同步构建基于MQTT协议的物联网数据接入通道，实现边缘计算网关与云端平台的可靠传输。数据采集频率遵循梯度化设计原则：关键运行参数(如主蒸汽温度、转子振动值)采用秒级连续采集，常规监测指标(如辅机功率、润滑油压力)设置为分钟级采样，管理类数据(燃料消耗量、发电负荷计划)则按小时或日频次收集。

(二) 数据分析方法

火电厂大数据分析平台的数据分析方法需结合多源异构数据特征与业务需求，构建系统化的技术路径。在数据预处理阶段，需通过清洗、标准化与特征工程实现数据质量的提升。首先，采用异常值检测技术(如 3σ 准则、孤立森林算法)识别并修复传感器数据中的噪声与缺失值，同时利用时间对齐与插值算法处理不同采样频率的时序数据。其次，针对热工参数、设备状态、环境监测等多维度数据，建立统一的元数据管理框架，通过数据标准化与归一化消除量纲差异^[5]。特征工程部分需结合领域知识提取关键特征，例如对锅炉燃烧效率进行傅里叶变换提取频域特征，或对机组振动信号进行小波包分解提取时频特征。

在分析方法选择上，基于问题导向构建多元模型体系。对于设备状态监测与预测性维护，采用监督学习中的随机森林与长短期记忆网络(LSTM)构建故障预测模型，通过历史振动、温度、压力数据预测轴承磨损与阀门卡涩风险。在能耗优化领域，应用灰色关联分析与偏最小二乘回归(PLSR)挖掘煤质、负荷率与热效率间的非线性关系，结合遗传算法进行多目标优化。排放控制方面，建立基于支持向量机(SVM)的多污染物排放预测模型，通过特征重要性分析识别NO_x、SO₂排放的关键影响因素^[6]。此外，针对机组运行优化问题，采用蒙特卡洛模拟进行不确定性分析，结合响应面法构建多变量优化模型，实现燃料配比与运行参数的动态调整。

三、平台应用与效果评估

(一) 效果评估指标与方法

火电厂大数据分析平台的建设成效评估需结合技术性能与业

务价值的多维视角,通过科学的指标体系与评估方法实现综合量化分析。在指标设计方面,需覆盖数据质量、平台功能、业务影响及经济价值四个核心维度。数据质量评估聚焦于数据采集的完整性、实时性、一致性及准确性,通过建立数据缺失率、异常值占比、时间延迟阈值等指标,结合数据校验算法验证数据清洗效果。平台功能评估则需从系统稳定性、响应效率及功能完备性入手,采用系统可用性(如年均无故障运行时间)、查询响应时间、任务吞吐量等硬性指标,同时通过用户操作日志分析模块使用频率与功能覆盖率,衡量平台功能的实用性和易用性^[7]。

(二) 效果评估结果与分析

本平台通过多维度数据采集与智能分析模型的应用,在运行效率提升、成本控制优化及设备可靠性增强等方面取得了显著成效。在运行效率方面,平台基于实时运行参数与历史数据的深度学习,实现了机组运行参数的动态优化。数据显示,实施平台后机组平均热效率提升1.2%~1.8%,其中通过燃烧调整优化使单位发电煤耗降低4.3g/kWh,综合厂用电率下降0.5个百分点。负荷响应速度提升效果尤为突出,平台预测负荷偏差率由传统方法的8.7%降至3.2%^[8],显著提高了机组对电网调度的适应能力。特别在锅炉燃烧优化模块中,通过氧量、煤粉浓度等200余项参数的协同优化,实现nox排放与煤耗的帕累托最优,为机组在环保限值内寻找经济运行区间提供了量化依据。

四、结论

本研究针对火电厂大数据分析平台的建设与应用展开了系统性研究,通过理论分析、技术开发与实践验证,形成了具有创新性和实用价值的成果体系。首先,在平台架构设计层面,构建了覆盖数据采集、存储、处理与分析的全生命周期技术框架。该框架基于Hadoop和Spark等分布式计算技术,实现了多源异构数据的高效整合与实时处理,有效解决了传统火电厂数据孤岛与存储瓶颈问题。数据采集与存储模块通过边缘计算与云边协同机制,确保了传感器、SCADA系统、DCS系统等设备数据的高可靠性传输,数据清洗算法将异常值识别准确率提升至98%以上,为后续分析奠定了可靠基础。

在核心分析技术方面,研究团队开发了基于机器学习与深度学习的智能分析模型,包括锅炉燃烧优化模型、汽轮机振动预测模型和机组故障诊断系统。其中,结合LSTM网络的燃烧效率预测模型将预测误差控制在2%以内,较传统方法提升了35%;基于改进型随机森林算法的设备故障预警系统,成功识别了9类典型故障模式,预警响应时间缩短至10秒级,显著增强了机组运行安全性。同时,通过构建多维度可视化分析界面,实现了机组性能指标、能耗数据及环保参数的动态监测与趋势分析,为生产调度提供了直观决策支持。

参考文献

- [1] 吴青云,高景辉,李昭,谭祥帅,郭云飞,姚智,赵威,赵如宇,蔺奕存,刘世雄,王涛,王林.基于Adam优化卷积神经网络的火电厂大数据平台预警模型测试与应用[J].科学与技术,2023,23(35):15075-15083.
- [2] 曾庆华,冉鹏,董坤,刘旭.基于大数据平台的SO₂排放GWO-N-BEATS预测算法[J].热能动力工程,2024,39(03):125-131.
- [3] 杨剑,刘振祥.火电厂辅机智能监测平台研究与应用[J].工业控制计算机,2021,34(01):4-6.
- [4] 刘永利,蒋春雷,刘志超.火力发电厂数字化发展现状以及向智能化电厂转型分析[J].轻松学电脑,2021,000(003):P.1-1.
- [5] 庄茗贺.基于物联网、大数据、人工智能技术的智能电厂平台建设应用[J].节能技术,2025,43(03):256-264.
- [6] 陈文.基于大数据分析的火力发电厂锅炉性能优化研究[J].新疆钢铁,2025,(02):228-230.
- [7] 张书舟.基于数字化技术的火电厂电气监控系统方案实际应用研究[J].软件:教育现代化(电子版),2013(5):2.
- [8] 李雪波.基于大数据的火电厂汽轮机故障诊断与预测维护[J].电力设备管理,2024,(24):96-98.