

基于大数据分析的电厂设备故障诊断与维护决策

崔文博

宁夏东部热电股份有限公司 宁夏 银川 750000

摘要：通过深入探讨基于大数据分析的电厂设备故障诊断与维护决策方法，概述了大数据分析基础及电厂设备数据来源，并详细分析了数据的采集、存储、清洗与预处理流程。然后介绍了故障诊断方法，提出设备健康状态评估、维护模式选择、维护计划制定及决策优化等策略。本文旨在利用大数据技术提高电厂设备故障诊断的准确性和维护决策的科学性，确保电厂的安全稳定运行，为电力行业的高效发展提供保障。

关键词：大数据分析；电厂设备；故障诊断；维护决策

1 大数据分析基础与电厂设备数据概述

1.1 大数据分析基础

大数据以其Volume（规模庞大）、Velocity（高速流转）、Variety（类型多样）及Value（潜在价值）四大特性而著称。其处理流程全面覆盖了从数据采集、高效存储、精细清洗、深度分析至直观可视化的各个环节。数据采集阶段侧重于从众多不同来源和格式的数据渠道中汇聚信息；存储方面，则依赖于分布式文件系统（例如HDFS）、列式数据库（如HBase）以及时序数据库（诸如InfluxDB）等技术，为庞大的数据集提供高效的存储解决方案。数据清洗步骤致力于消除噪声、填补缺失数据并纠正异常值，以提升数据质量。进入分析阶段，运用数据挖掘算法（如关联规则挖掘、聚类分析）、机器学习模型（例如支持向量机、随机森林）以及深度学习技术（如卷积神经网络、循环神经网络）深入发掘数据的潜在价值^[1]。最终，这些分析结果以可视化的形式直观展现，为决策制定提供强有力的辅助。大数据分析技术凭借对海量数据的深度挖掘与高效处理能力，能够揭示数据内在的模式与关联性，为众多行业的战略决策提供坚实的依据。

1.2 电厂设备数据来源

电厂设备运行过程中产生的数据来源广泛且复杂，分散控制系统（DCS）和数据采集与监视控制系统（SCADA）实时记录设备运行参数，如温度、压力、流量、转速等，为设备运行状态监测提供基础数据。各类传感器，如振动传感器、温度传感器、压力传感器等，直接采集设备关键部位的物理量数据，可用于分析设备的机械性能和工作状态。设备台账数据包含设备型号、规格、安装时间、使用年限等信息，是了解设备基本情况和历史运行状态的重要依据^[2]。电厂的生产管理系统、能源管理系统等也会产生与设备运行相关的数据，这些

数据共同构成了电厂设备大数据的来源。

2 电厂设备大数据来源与处理

2.1 数据来源分析

电厂设备数据的多源性决定了其复杂性和多样性，从设备类型来看，发电机组、锅炉、汽轮机等核心设备是数据的主要产生源。以发电机组为例，其运行过程中产生的电参数（电压、电流、功率）、机械参数（振动、位移）和热工参数（温度、压力）等数据，反映了设备的运行状态和性能。锅炉设备的数据则主要包括炉膛温度、蒸汽压力、烟气成分等，这些数据对于保障锅炉的安全稳定运行至关重要。汽轮机的转速、轴向位移、润滑油温度等数据是监测其运行状态的关键指标。除核心设备外，辅助设备如风机、水泵、阀门等也会产生大量数据，这些数据同样对电厂的整体运行具有重要影响。从数据产生环节来看，设备的设计、制造、安装、运行、维护等全生命周期过程都会产生数据，这些数据相互关联，共同构成了电厂设备大数据的完整体系。

2.2 数据采集与存储

电厂数据采集结合实时与周期性方式，利用高速数据采集系统捕捉关键设备参数变化，如发电机转速、锅炉蒸汽压力等，而周期性采集适用于缓慢变化参数，如设备温度，以降低成本。工业以太网、无线网络等技术实现数据快速稳定传输。存储方面，结构化数据采用关系型数据库存储，便于查询管理；半结构化和非结构化数据则利用分布式文件系统或NoSQL数据库存储，满足大数据和高并发需求。时序数据库针对时间序列数据提供高效存储查询，同时建立数据备份和容灾机制，确保数据安全和可靠性^[3]。

2.3 数据清洗与预处理流程

电厂原始数据集内含的噪声、缺失值及异常值等瑕疵，会对数据分析的精确性与可信度构成威胁，故数据

清洗与预处理环节显得至关重要。噪声数据,源于传感器误差或信号干扰等因素导致的误读,需借助滤波技术(诸如均值滤波、中值滤波)实施平滑处理,以剔除不必要的干扰。面对缺失值问题,需依据数据特征及其散布态势,采取恰当的填补策略。若缺失数据数量有限,可利用均值、中位数或众数进行估算填充;反之,若缺失规模较大,则可考虑采用机器学习算法(如决策树模型、随机森林算法)进行预测填补。至于异常值,即那些显著偏离正常取值范围的数据点,则需借助统计手段(例如 3σ 原则)或机器学习算法(如孤立森林模型)加以识别,并视具体情况采取修正、删除或作为特例深入分析的处理方式。数据标准化与归一化构成了数据预处理的关键步骤。标准化旨在将数据转换为均值为零、标准差为一的标准正态分布形式,以此消除变量间的量纲差异,进而提升算法收敛速度与精度。而归一化则是将数据映射至 $[0,1]$ 或 $[-1,1]$ 区间,确保数据尺度一致,便于后续的比较与分析^[4]。经此数据清洗与预处理流程,数据质量得以显著提升,为后续故障诊断及维护决策奠定了坚实的数据基石。

3 基于大数据分析的电厂设备故障诊断方法

3.1 数据预处理

针对设备运行参数,实施特征工程是关键一环,它涵盖特征构造与特征选择两大方面。特征构造基于原始数据的物理意义及相互关联,创新构建新的特征变量,例如设备的负荷率与效率等,旨在全方位体现设备运行状态。而特征选择,则是从海量特征中精挑细选,识别出对故障诊断最具贡献度的特征,旨在降低数据维度,简化计算流程,进而优化模型效能。此过程借助相关性分析、主成分分析(PCA)等手段,精准剔除低相关性或非故障相关特征,确保关键特征得以保留。同时,对数据进行离散化处理,将连续的数值型数据转化为离散的类别数据,以更好地适配分类算法,如将设备温度细分为低温、正常及高温区间,为后续故障分类奠定坚实基础^[5]。

3.2 特征提取与精选机制

信号处理技术在设备振动、电流信号特征提取中占据主导地位。快速傅里叶变换(FFT)作为时频转换工具,能够将时间域信号转化为频率域信号,深入分析信号的频率构成,精准捕捉故障特征频率。小波变换则在时频双域内展现多尺度分析能力,尤其擅长处理非平稳信号,有效提炼故障信号的时频特性。进入机器学习范畴,递归特征消除(RFE)策略通过逐步剔除次要特征,精心挑选出最优特征子集。而以随机森林为代表的基于

树的特征选择方法,依据特征在决策树构建进程中的权重,筛选出对分类结果具有显著影响的特征子集。这一系列特征提取与选择手段,能够自海量数据中提炼出最能映射设备故障状态的关键特征,为故障诊断模型的搭建奠定坚实基础。

3.3 故障诊断模型构建策略

机器学习与深度学习技术在故障诊断模型构建中大放异彩。支持向量机(SVM)作为机器学习领域的分类算法翘楚,通过构建最优分类超平面,对不同故障类型的数据样本实施精准分类。面对电厂设备故障诊断中样本稀缺的挑战,SVM凭借出色的泛化能力脱颖而出。随机森林(RF)算法则通过集成多个决策树,展现出处理高维数据及非线性问题的强大实力,同时对噪声与异常值展现出优异的鲁棒性。在模型训练阶段,随机森林巧妙运用Bagging抽样与随机特征选择,显著提升了模型的稳健性与精确度。深度学习领域,卷积神经网络(CNN)以其独特的卷积层、池化层及全连接层架构,自动挖掘图像或信号数据的深层特征,尤其适用于基于设备振动图像或声音特征的故障诊断。在模型构建实践中,需根据设备故障数据的特性及诊断需求,审慎选择算法并进行细致的参数调整,以确保模型诊断精度的最大化^[6]。

3.4 模型评估与优化流程

进入模型评估与优化环节,采用一系列关键指标对故障诊断模型实施全面而深入的考量,这些指标囊括准确率、召回率及F1值等。其中,准确率直观展现了模型的整体预测精度,召回率则精准量化了模型在识别故障样本方面的效能。而F1值,作为准确率与召回率的和谐统一,提供一个更为综合的性能评价视角,有助于全方位洞悉模型的实际表现。为了确保评估结果的公正性与精确性,运用交叉验证策略,将数据集合理划分为训练集与测试集,历经多次模型训练与测试循环后,取其性能指标的平均值作为最终评估基准^[7]。通过这一持续不断的评估与优化进程,致力于提升故障诊断模型的精确度与泛化能力,力求使其在电厂设备故障诊断实践中展现出更大的价值,为电厂的安全稳定运营提供坚实的技术支撑与保障。

4 基于大数据分析的电厂设备维护决策策略

4.1 设备健康状态评估

建立基于大数据的设备健康状态评估指标体系,综合考虑设备运行参数、故障历史、维护记录等多源数据。从设备的性能指标(如效率、能耗)、可靠性指标(如故障率、平均无故障时间)、安全性指标(如振动

幅值、温度超限次数)等方面选取评估指标。采用层次分析法(AHP)确定各指标的权重,通过构建判断矩阵,计算指标的相对重要性。模糊综合评价法用于处理评估过程中的不确定性和模糊性,将定量数据和定性分析相结合,对设备健康状态进行量化评估。根据评估结果,将设备健康状态划分为良好、正常、注意、故障四个等级,为维护决策提供直观的依据^[8]。

4.2 维护模式选择

根据设备健康状态评估结果,科学选择维护模式。对于健康状态良好且运行稳定的设备,采用预防性维护模式,按照预定的维护计划进行定期检查、保养和更换易损件,以降低设备故障发生的概率。对于存在潜在故障隐患或健康状态下降的设备,采用预测性维护模式,利用大数据分析和故障诊断模型,预测设备故障的发生时间和类型,提前制定维护方案,避免非计划停机。对于故障已经发生的设备,采用事后维修模式,尽快修复设备,恢复生产^[9]。在实际应用中,综合考虑设备的重要性、维护成本、生产计划等因素,灵活选择维护模式,实现维护资源的合理配置。

4.3 维护计划制定

制定基于大数据分析的维护计划,需要考虑设备健康状态、维护成本、生产计划等多方面因素。利用优化算法(如遗传算法、粒子群优化算法)对维护时间、维护方式和维护资源进行优化配置。在确定维护时间时,综合考虑设备故障预测结果和生产计划,选择对生产影响最小的时间段进行维护。对于维护方式,根据设备故障类型和严重程度,选择合适的维修方法,如局部维修、整体更换等。在维护资源配置方面,合理安排维修人员、维修设备和零部件,确保维护工作的顺利进行。

4.4 维护决策优化

在维护决策实施过程中,不断收集和分析维护数据,对维护决策进行优化。通过对比维护前后设备运行指标(如故障率、可用率)、成本指标(如维护成本、停机损失)等,评估维护决策的效果。利用大数据分析技术,挖掘维护数据中的潜在信息,总结维护经验和教训,为后续维护决策提供参考。例如,分析不同维护方

式对设备性能恢复的影响,找出最有效的维护方法;研究维护周期与设备故障率之间的关系,优化维护周^[10]。同时,引入人工智能技术,建立维护决策智能优化模型,根据设备运行状态和历史维护数据,自动调整维护决策策略,实现维护决策的智能化和自适应化。

结束语

综上所述,基于大数据分析的电厂设备故障诊断与维护决策方法具有显著的优势和应用前景。通过深入挖掘设备运行数据中的有价值信息,可以实现对设备健康状态的准确评估,为科学选择维护模式和制定维护计划提供有力依据。未来,随着大数据技术的不断发展和完善,相信电厂设备故障诊断与维护决策将更加智能化、自动化,为电力行业的高质量发展注入新的动力。

参考文献

- [1]生俊阳.基于智能技术的发电设备故障诊断与维修策略分析[J].集成电路应用,2024,41(03):396-397.
- [2]付鹏飞,张国攀.基于大数据平台的煤矿机电设备数据集成管理研究[J].科学技术创新,2024,(13):69-72.
- [3]刘光,李俊鑫.基于大数据技术的煤矿机电设备故障诊断与维修策略研究[J].模具制造,2024,24(06):259-261.
- [4]王涛.大数据技术在煤矿设备监测中的应用研究[J].矿业工程,2021,50(2):25-30.
- [5]李明.基于大数据分析的煤矿设备故障诊断与维修优化[J].煤炭技术,2020,39(4):10-15.
- [6]何洪,阮航.基于数据分析的设备故障预警与诊断技术在燃气电厂的应用[J].石油科技论坛,2023,42(1):61-66.
- [7]吴晓翠,游超.数字化技术在水电厂设备状态监测与故障诊断中的应用[J].数字技术与应用,2025,43(2):223-225.
- [8]艾思维,赵俊杰,吴秀峰,等.基于大数据分析的智慧火电厂电动机状态智能监测与诊断应用[J].今日自动化,2021(8):11-12.
- [9]水琳,张国茹,赵海鹏.风力发电设备润滑需求和标准发展现状与展望[J].石油商技,2023,41(04):4-10.
- [10]李泽英,路发志.风力发电设备的运行管理措施分析[J].电子技术,2023,52(06):210-211.