

基于大数据的机电仪表自动化运行数据分析与优化

乔冠华

山西焦化集团有限公司 山西 临汾 041000

摘要: 本文聚焦大数据在机电仪表自动化运行中的应用。机电仪表运行数据具大数据特征,大数据技术为其处理提供支撑。基于此构建数据分析体系,涵盖数据采集、预处理、存储与挖掘模块。通过构建运行状态评估指标体系与故障预警模型,实现状态评估与故障预警,并可视化展示结果。同时,利用数据挖掘确定参数优化方向,构建多目标优化模型并求解,经验证与动态调整优化方案,实现机电仪表自动化运行参数的动态精准优化,提升运行性能。

关键词: 大数据; 机电仪表; 自动化运行; 数据分析; 运行优化

引言: 在工业自动化蓬勃发展的当下,机电仪表自动化运行体系面临着海量且复杂的数据挑战。这些数据来源广泛、类型多样,具备大数据典型特征,但价值密度低。大数据技术为应对此难题提供了有力支撑,从分布式存储到数据预处理,再到机器学习与数据挖掘算法,为数据利用奠定基础。基于此,构建机电仪表自动化运行数据分析体系,开展运行状态评估与故障预警,并制定运行参数优化策略,成为提升机电仪表自动化运行性能、保障工业生产稳定高效的关键所在。

1 大数据在机电仪表自动化运行中的应用基础

在机电仪表自动化运行体系里,数据呈现出爆炸式增长态势,其来源广泛且类型丰富。实时采集的生产参数数据,涵盖温度、压力、流量等关键指标,这些数据精准反映着生产过程的实时状态;仪表设备自身的运行状态数据,像电压、电流、功耗等,是判断设备健康状况的重要依据;历史故障记录数据记录了设备过往的故障类型、发生时间及处理方式等信息;环境干扰数据,如湿度、粉尘浓度等,对仪表设备的正常运行有着不可忽视的影响。这些数据具备大数据的典型特征,体量巨大、类型多样、增长速度迅猛,然而价值密度却相对较低。

大数据技术为应对这些复杂数据带来了曙光。分布式存储技术,以 Hadoop 分布式文件系统为代表,凭借其强大的扩展性和容错性,能够安全存储海量仪表数据,并实现快速访问,确保数据的完整性与可用性。数据预处理技术不可或缺,通过数据清洗可去除数据中的噪声与异常值,转换技术能将不同格式的数据统一为标准格式,集成技术则可将分散的数据整合为完整的数据集,从而提升数据质量,为后续分析奠定基础。机器学习与数据挖掘算法更是挖掘数据价值的关键利器,聚类分析可对仪表运行状态进行分类,决策树能找出影响故障的关键因素,神经网络可精准预测参数变化趋势。这些算

法从海量数据中深度挖掘出仪表运行状态与故障之间的关联规律,为机电仪表的运行监测、故障诊断以及优化调整提供坚实有力的数据支撑^[1]。

2 机电仪表自动化运行数据分析体系构建

2.1 数据采集模块

数据采集作为机电仪表自动化运行数据分析的基石,其精准性与全面性直接影响后续分析效果。(1)为实现多源数据的完整、实时采集,该模块运用分布式采集架构,将传感器、数据采集卡、工业以太网等设备进行有机整合,搭建起一张覆盖仪表运行全流程的数据采集网络。(2)对于不同类型的机电仪表,如压力仪表、流量仪表、温度仪表等,因其测量对象与工作特性不同,采集需求也存在差异。因此,需根据仪表类型定制化设置采集参数,涵盖采集频率、数据精度等关键指标。例如,对于变化频繁的压力参数,可适当提高采集频率;对于对精度要求较高的温度测量,则需设置更高的数据精度。(3)引入边缘计算技术,在数据源头对采集到的原始数据进行初步筛选与压缩处理。这不仅能有效减少不必要的数据传输,降低网络带宽压力,还能提高数据采集效率,确保后续数据分析能够基于高质量、高效率采集的数据展开。

2.2 数据预处理模块

原始采集的机电仪表数据质量参差不齐,存在诸多影响分析结果准确性的问题,像大量噪声、缺失值以及异常值等,所以数据预处理在整个数据分析体系中至关重要。(1)在异常值处理方面,运用多种异常值检测算法。基于统计的 3σ 原则,依据数据的均值和标准差来判断异常值;箱型图法则通过数据的四分位数确定异常值的范围,进而识别并剔除这些异常数据,保证数据的合理性。(2)对于缺失数据,采用不同方法处理。若缺失情况相对简单,使用线性插值、多项式插值等插值法,

依据相邻数据点来估算缺失值；若数据关系复杂，则借助机器学习填充法，如随机森林填充，利用数据间的潜在关系填充缺失部分，确保数据的完整性。（3）对不同格式和来源的数据开展标准化转换与集成。把非结构化的故障描述文本、半结构化的 XML 格式配置文件等，统一转换为结构化数据，并规范数据格式与编码标准，形成标准化数据集，为后续深入分析提供可靠支撑^[2]。

2.3 数据存储与挖掘模块

数据存储模块构建了“分布式文件存储 + 关系型数据库 + 时序数据库”的混合存储架构，以适配不同类型数据的存储需求。分布式文件存储凭借其强大的扩展性和容错性，负责存储海量非结构化与半结构化数据，如设备日志、图像等。关系型数据库，以 MySQL 为代表，凭借其严格的数据结构和事务处理能力，用于存储仪表基础信息、故障记录等结构化数据，保证数据的完整性和一致性。时序数据库，如 InfluxDB，针对具有时间序列特征的实时监测数据，如温度、压力等随时间变化的数据，提供高效的存储和查询服务，确保数据存储的高效性与针对性。数据挖掘模块基于预处理后的标准化数据集展开深度分析。运用 K-Means 等聚类分析算法对仪表运行状态数据分类，精准识别正常运行与异常运行的状态聚类中心。借助 Apriori 等关联规则挖掘算法，挖掘故障类型与仪表运行参数、环境因素间的关联关系。同时，构建基于神经网络的预测模型，对仪表运行参数变化趋势进行预测，为故障预警提供有力的数据支持。

3 基于大数据的机电仪表运行状态评估与故障预警

3.1 运行状态评估指标体系

为实现对机电仪表运行状态的全面且精准评估，构建一套多维度、科学合理的评估指标体系至关重要。该体系包含四个核心维度，分别是数据采集精度、运行稳定性、能耗水平以及设备老化程度。在数据采集精度方面，通过测量值与真实值的偏差率、数据重复性等具体指标进行量化评估，偏差率越小、数据重复性越高，表明采集精度越高。运行稳定性则采用平均无故障运行时间、运行参数波动幅度等指标来衡量，平均无故障运行时间越长、参数波动幅度越小，说明运行越稳定。能耗水平以单位时间内的耗电量、功耗波动系数作为评估依据，耗电量低且功耗波动系数小，代表能耗水平低。设备老化程度结合仪表使用年限、关键部件磨损数据等指标综合判断。最后，运用层次分析法确定各指标权重，借助模糊综合评价法计算仪表运行状态综合得分，达成量化评估^[3]。

3.2 故障预警模型构建

依托大数据挖掘所获取的故障关联规律，精心构建机电仪表故障预警模型。该模型以经过预处理的仪表运行数据和环境数据作为输入信息，将各类故障类型，像传感器故障、电路故障、数据传输故障等设定为输出内容，选用 BP 神经网络算法搭建预警模型框架。具体实施过程中，先选取充足的历史故障数据以及与之对应的运行状态数据作为训练样本，对神经网络模型展开训练，通过不断调整优化模型的权重与阈值，使模型初步具备故障识别能力。接着，运用验证样本对训练好的模型进行全面测试，依据测试结果进一步调整模型参数，以此保障模型具备较高的预警精度以及良好的泛化能力，能在不同工况下准确预警。最后，将实时采集并完成预处理的运行数据输入到训练成熟的模型中，即可实现对潜在故障的实时预警，同时输出故障发生的概率以及可能的故障原因，为后续的故障处理提供有力依据。

3.3 预警结果可视化展示

为助力工作人员便捷、直观地掌握机电仪表运行状态与预警信息，专门构建了预警结果可视化平台。该平台运用折线图清晰呈现仪表运行参数随时间的变化趋势，柱状图直观对比不同仪表或不同时间段的运行状态评估得分，热力图则能快速展示各区域仪表的故障预警等级分布情况。在预警等级标识上，采用红、黄、绿三色进行区分。红色明确指示高风险故障即将发生或已然发生，需立即处理；黄色代表存在潜在故障风险，需密切关注；绿色表明仪表运行状态正常。此外，平台具备强大的功能，支持工作人员对故障信息进行详细查询，还能进行历史数据回溯。工作人员借此可迅速定位故障位置，精准查看故障相关的运行数据以及历史处理记录，为高效开展故障排查工作提供有力支撑。

4 机电仪表自动化运行参数优化策略

4.1 基于数据挖掘的参数优化方向确定

在机电仪表自动化运行过程中，借助数据挖掘技术对海量运行数据进行深度剖析，是确定参数优化方向的关键途径。详细分析不同运行参数，像采样频率、校准周期、阈值设定等，与仪表运行效率、数据采集精度以及能耗水平之间的内在关联关系。（1）以采样频率为例，通过数据挖掘可发现，采样频率过高虽能提升数据采集精度，但会大幅增加能耗；采样频率过低则会导致数据采集精度下降，影响仪表对运行状态的准确反映。据此，可找到兼顾精度与能耗的最优采样频率区间。对于校准周期，分析其与数据偏差率的关联规律，若校准周期过长，数据偏差率会显著上升，影响仪表测量准确性，基于此确定合理的校准周期调整方案。（2）结合仪

表运行状态评估结果与故障预警信息,对于频繁出现异常的仪表,深入分析其关键运行参数,找出参数设置不合理之处,并重点进行优化,从而有效解决因参数设置不当引发的运行问题,提升机电仪表自动化运行的整体性能与稳定性。

4.2 多目标优化模型构建与求解

为全面提升机电仪表自动化运行性能,以提升运行可靠性、降低能耗、提高数据采集精度作为核心优化目标,构建多目标优化模型。(1)将仪表运行中的关键参数,如采样频率、校准周期、阈值参数等明确为决策变量,这些参数的调整会直接影响仪表的各项性能指标。同时,把各优化目标的量化指标,如数据偏差率反映数据采集精度、单位时间能耗体现能耗水平、平均无故障运行时间衡量运行可靠性,作为目标函数。(2)充分考虑仪表设备的物理约束,例如参数的取值范围需在设备允许的区间内,运行环境限制如温度、湿度等对参数的影响,以及性能约束,如设备最大负载能力等,以此建立完整的多目标优化数学模型。(3)采用非支配排序遗传算法(NSGA-III)对构建的优化模型进行求解。该算法通过初始化种群开启优化过程,接着进行快速非支配排序、拥挤度计算,随后依据计算结果进行选择、交叉、变异等操作,经过逐步迭代,最终得到 Pareto 最优解集。从该解集中,紧密结合实际生产需求与仪表运行特点,筛选出综合性能最优的参数组合,作为最终的优化方案。

4.3 优化方案的验证与动态调整

为切实保障优化方案能切实发挥作用、达成预期效果,需选取具有代表性的典型机电仪表开展现场验证试验。(1)把优化后的运行参数精准应用于实际运行的仪表,在仪表运行过程中持续、稳定地采集各类关键数据,涵盖数据采集精度、能耗水平、故障发生率等重要指标。将优化前后的这些指标进行细致对比分析,若优化效果未达预期目标,需深入结合验证过程中新生成的

数据,对优化模型的参数以及约束条件进行重新审视与调整。随后再次运用合适的算法求解,从而得到全新的优化方案,并持续重复此过程,直至达到理想的优化成效^[4]。(2)鉴于机电仪表运行环境复杂多变、生产需求也处于动态调整之中,必须建立一套科学合理的优化方案动态调整机制。通过实时、精准地监测仪表运行状态数据以及优化效果指标,一旦察觉参数优化效果出现下滑趋势,或者运行环境发生显著改变,立即自动触发参数重新优化流程,基于最新采集到的运行数据对优化模型进行更新完善,输出契合新工况的优化参数,实现机电仪表运行参数的动态、精准优化调整。

结束语

综上所述,大数据为机电仪表自动化运行带来了深刻变革。通过构建涵盖数据采集、预处理、存储与挖掘的完整数据分析体系,实现了对机电仪表运行状态的精准评估与故障预警,可视化平台更助力工作人员高效掌握信息。同时,基于数据挖掘确定参数优化方向,构建多目标优化模型并求解,经现场验证与动态调整,确保优化方案切实有效。未来,随着大数据技术的持续发展,机电仪表自动化运行将更加智能、高效、稳定,为工业生产的高质量发展提供坚实有力的支撑,推动行业迈向新的发展阶段。

参考文献

- [1]张中超,江志敏,程梦成.基于大数据的电力设备运行数据分析研究[J].中国设备工程,2024,(02):8-10.
- [2]潘飞,潘树昌,周识远.大数据背景下的电力信息化智能管控平台设计[J].华东科技,2023,(11):101-103.
- [3]刘晔,蒋道环,伍江瑶.基于大数据分析的电力信息化智能管控平台设计[J].自动化与仪器仪表,2022,(09):173-176+185.
- [4]夏盛海,金宇,杨攀,等.基于大数据分析技术的电网智能调控分析方法[J].电力大数据,2022,25(03):51-58.