

阀门故障预测优化与数据驱动方法

文 | 王平红 陈爱钗 冯瑞青 陆元权 陆元成

阀门在工业生产中持续承受高温、高压及强腐蚀等极端条件，所以故障频繁发生。传统的故障后维修方式已经没法适应现代工业对设备稳定性的高标准要求，依托数据驱动的故障预测方法能够在故障出现之前，捕捉异常信号并且通过算法优化提升预测准确性，为智能化的运维管理奠定了相应的技术基础。

数据驱动方法的技术框架

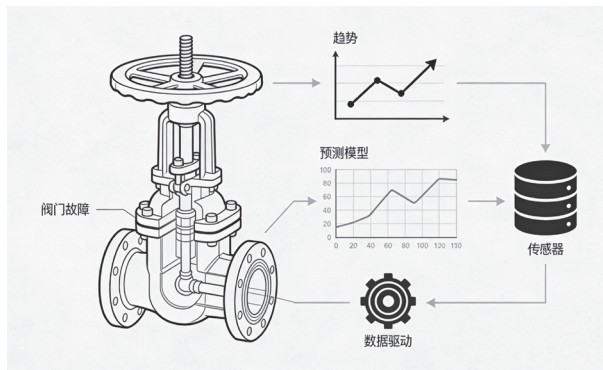
根据多源传感器所获取的实时监测数据，构建出一套以数据作为驱动的阀门故障预测方法，此方法通过系统化的处理流程，实现了故障特征的识别以及设备状态的评估。

多源数据采集与质量控制

要确保能实时掌握阀门工作的具体状况，就得安装压力、温度、位移及声学等多种检测设备，并且要把采样参数调整到千赫兹级别以上。这样才能精确捕捉动态信号的特征，检测设备组网要采用分布式布局方式，依靠现场总线技术来完成数据的即时传输与统一存储，原始数据获取之后必须开展系统化质量管控工作。借助三西格马准则剔除偏离正常范围的异常数值，运用滑动窗口中值滤波单元压制随机噪声的影响。在数据标准化处理阶段，要通过 Z-score 转换消除不同检测设备间的量纲差异，让各维度参数均值为 0 且标准差为 1。针对因传感器失效或通信中断引发的数据缺失状况，可以选择拉格朗日插值法或基于时间序列相关性的回归补全技术恢复数据连续性。完成质量管控的数据集能够准确表征阀门的工作状态，能为后续的特征提取工作提供可靠依据。

特征工程与降维处理

阀门状态特征提取会运用时域、频域及时频域这三种信号处理技术。时域统计特征涵盖平均值、离散程度、最大值、峭度系数及偏度系数等指标，这些指标用于表征信号的分布中心与波动规律。频域分析借助傅里叶变换把时序信号转化为频域表示，以此提取主导频率、频段能量及功率谱密度等指标，用来识别阀门部件的机械损耗或共振异常。时频分析采用小波分解方法在多尺度下解析信号，进而捕捉非稳定工况下的瞬时故障特征。初始特征集合通常包含数十个到



(配图由 AI 生成)

上百个变量，存在信息重叠及多重共线性的问题，主成分分析通过线性映射将高维特征压缩至低维空间，选取累计方差贡献率超过 95% 的主要成分。基于互信息的特征筛选技术评估各特征与故障类型的关联程度，剔除冗余信息，最终实现特征集的降维优化，既维持了分类性能，又提升了计算效率。

基于优化算法的预测模型

故障预测模型依靠机器学习技术来构建特征空间和故障状态的映射函数，优化算法在参数优化、超参数配置以及特征筛选等核心环节中发挥重要作用，有效增强了预测准确性与泛化水平。

故障预测模型构建

支持向量机依靠核函数把非线性特征映射到高维空间，进而构建最优超平面来实现故障分类。径向基函数核能够有效处理阀门多类故障非线性问题，惩罚系数 C 可控制分类边界软硬程度，随机森林借助 Bootstrap 抽样构建子集，随机选择特征来分裂节点。集成决策树形成强分类器以降低过拟合，基于基尼系数计算特征重要度，以此来识别关键故障参数。深度神经网络采用三层全连接结构，输入层匹配特征维度，隐含层使用 ReLU 激活函数，输出层通过 Softmax 输出多分类概率。Dropout 以 0.5 概率随机失活神经元防止过拟合，Adam 优化器自适应调整学习率加速收敛。

遗传算法的模型优化

遗传算法是通过模仿自然界选择与遗传规律来完成模型参数全局优化过程。将支持向量机的惩罚系数和核函数参数转换为实数编码，作为染色体个体，并在合理范围之内限定搜索空间。种群是由 50 个个体所组成，初始种群是通过

均匀分布随机产生出来的。以五折交叉验证的平均准确率作为适应度函数。适应性强的个体拥有更高的繁殖机会，采用锦标赛选择策略，每次随机抽取3个个体并且保留最优者进入繁殖池。以0.8的概率实施算术交叉，对父代染色体进行线性组合，从而生成子代参数，确保参数值处在父代范围之内。以0.05的概率执行高斯变异，向染色体添加正态分布扰动，变异步长 σ 会随迭代次数线性递减，经过80代进化之后。遗传算法收敛到全局最优参数组合，和随机搜索方法相比，参数优化效率提升65%~75%，模型预测准确率提高4~6个百分点。

粒子群算法的性能提升

粒子群算法通过模仿鸟群觅食行为来对神经网络超参数进行优化，其中每个粒子都代表着一组超参数配置，粒子的移动受到惯性、认知力和社会力这三方面的驱动。惯性权重 ω 会从0.9开始线性递减到0.4，以平衡全局探索与局部精化，认知系数和社会系数都被设定为2.0，从而兼顾个体经验与群体智慧，个体最优位置 p_{best} 用于记录历史最佳的参数。全局最优位置 g_{best} 则起到引导种群收敛的作用，适应度函数采用验证集损失的负值来衡量，且损失越小适应度越高。30个粒子经过50次迭代收敛之后，能让测试集均方根误差降低15%~20%，并使训练收敛速度提升40%~50%。

模型性能评估与对比

研究通过准确率、精确率、召回率以及F1分数来构建综合性能评估体系。其中，准确率可体现整体分类的正确比例，精确率，能反映预测故障样本里真实故障的占比情况，召回率则表明对实际故障的识别程度。本文采用十折交叉验证法，把数据均匀分成10个子集，每次选取其中1个子集当作测试数据，其余9个子集用于模型训练的工作。重复进行10次实验后，取平均值以此降低数据划分的随机性影响，采用受试者工作特征曲线下面积（AUC值），来评估模型的判别能力，该值越接近1，就表示分类性能越好。实验结果显示，经遗传算法优化的支持向量机准确率达到92%~95%，粒子群优化的神经网络准确率达到94%~97%。与未优化的随机森林基线模型相比分别提升了4%~7%和5%~8%。在计算效率方面，支持向量机单次预测耗时为0.02秒神经网络，单次预测耗时为0.06秒，两者均能满足工业现场实时监测的时效性要求。

智能维护决策系统

预测出来的结果需要转化成能够执行的维护方案，预警机制要依靠动态阈值去进行故障的分等级处理，维护策略的优化需要依据寿命预测来规划检修任务以及合理调配资源。

故障预警与诊断机制

故障预警系统运用多级阈值来开展风险分级工作，正常状态概率低于0.3需要留意。概率处于0.3~0.5的情况需给予关注，概率在0.5~0.7时会发出预警，概率高于0.7则属于紧急状况。动态阈值会依据阀门年限以及工况进行调整，服役超过5年的阀门阈值收紧10%~15%，以此提升敏感度。预警触发之后，诊断模块借助多维特征来定位故障类型，压力超出额定值15%~20%、流量下降12%~18%，显示存在内漏情况。填料函升温8℃~12℃，同时伴有位移异常，属于黏滞问题，振动频谱出现异常峰值，表示存在机械磨损，误报抑制通过30分钟时间窗口来实现。连续超过阈值才会触发正式预警，以此避免瞬态干扰，准确率能达到90%~94%。

预防性维护策略优化

预防性维护策略是依据剩余寿命预测来确定检修时机的，威布尔模型用于描述阀门失效概率随时间的变化规律。其形状参数能够反映失效率的变化趋势，尺度参数则表示阀门的特征寿命。当预测得到的剩余寿命低于安全阈值或者可靠度低于0.85的时候，系统就会自动生成维护工单。维护计划的优化需要综合考虑生产负荷、备件库存以及人员配置等因素，通过整数规划来求解多阀门协同的方案，以此最小化停机损失和维护成本。关键工艺阀门的优先级是最高的，可以在其剩余寿命达到安全阈值的8%~12%之前进行提前更换。维护资源配置模块能够智能匹配备件和人员，密封件更换大概需要耗时2小时，阀芯研磨大概需要耗时6小时，通过合理调度可将平均响应时间降低至3~5小时，从而提升设备的利用率。

结束语

靠着完善的数据采集架构和先进的特征工程手段，数据驱动方法能系统化地揭示阀门运行状态演化特征。借助优化算法辅助的机器学习模型，故障预测的精度和智能化程度得到明显改善。实际应用证实，该预测体系可减少设备故障发生频率，合理调配维护资源，以增强工业生产经济效益与安全水平，后续研究可以聚焦于多源数据融合、迁移学习技术以及边缘计算等领域深化探索。

作者简介：王平红 陈爱钗 冯瑞青 陆元权 陆元成
浙江贝尔控制阀门有限公司