

基于分布式证据学习算法的 脱硫浆液品质监测模型研究

徐 侠¹, 朱万进¹, 薛钧赢¹, 苏志刚^{2*}, 郝勇生²

(1. 国能徐州发电有限公司, 江苏 徐州 221135; 2. 东南大学, 江苏 南京 210096)

摘要: 基于分布式计算框架与证据学习算法, 对脱硫浆液品质建立了健康品质监测模型, 突破了海量脱硫系统运行数据对基于传统机器学习的浆液品质监测方法所带来的计算瓶颈, 并利用该模型对江苏某 1 000 MW 电厂的浆液品质进行了在线监测。测试表明, 所建立的监测模型能够准确监测出脱硫浆液品质的恶化, 与其他 3 类先进监测方法对比结果说明了所建立模型能够达到最优的报警及时性。将分布式计算框架结合证据理论应用于脱硫浆液品质监测是可行的, 为脱硫浆液品质监测提供了一种新方法。

关键词: 脱硫浆液; 状态监测; 证据理论; 分布式计算框架

中图分类号: TH3

文献标志码: A

文章编号: 0253-4320(2024)S2-0348-07

DOI: 10.16606/j.cnki.issn0253-4320.2024.S2.061

An evidential condition monitoring model for desulfurization slurry

XU Xia¹, ZHU Wan-jin¹, XUE Jun-ying¹, SU Zhi-gang², HAO Yong-sheng²

(1. CHN Energy Xuzhou Power Co., Ltd., Xuzhou 221135, China; 2. Southeast University, Nanjing 210096, China)

Abstract: Based on distributed computing framework and evidence learning algorithm, a robust condition monitoring model is established for desulfurization slurry, which overcomes the computational bottlenecks brought by massive operational data of desulfurization systems to traditional machine learning-based slurry condition monitoring methods. This model is utilized to perform online monitoring of slurry condition in a 1 000 MW power plant in Jiangsu, China. Test results indicate that the monitoring model established is able to detect the deterioration of desulfurization slurry condition accurately. Through comparing with three other advanced monitoring methods, it is demonstrated that the model established can achieve the optimal alarm timeliness. It is feasible to apply the integration between distributed computing framework and evidence theory in desulfurization slurry condition monitoring, providing a new approach for similar monitoring.

Key words: desulfurization slurry; condition monitoring; evidence theory; distributed computing framework

我国燃煤电厂广泛采用石灰石-石膏湿法烟气脱硫(WFGD)工艺, 该工艺通过将石灰石转化成吸收浆液, 喷入脱硫塔内进行 SO₂ 的吸收, 最终转化为脱硫石膏^[1-3]。尽管该方法是治理 SO₂ 排放的有效方式, 但脱硫反应的多样性和复杂环境下的副反应使得实时监测塔内浆液品质及设备运行状况面临挑战^[4]。

随着分散控制系统(DCS)的应用, 脱硫系统累积了海量的运行数据。通过算法和机理公式化表达, 这些数据可以用于揭示状态评价^[5-6]、效率预测^[7-8]、故障诊断^[9-10]和优化运行^[11-12]等关键信息。然而, 这些研究多集中于脱硫系统的整体评价和优化, 而对于常规手段难以监测的浆液品质和喷淋层运行状况, 探讨相对较少。脱硫浆液的品质直接关

系到整个湿法脱硫工艺的效率。目前存在的挑战包括缺乏实时的定量和定性评价能力以及缺少统一的浆液品质评判标准。这些问题的存在导致难以及时调整和优化脱硫工艺^[13-14]。喷淋层的堵塞和漏液状况显著影响脱硫效率。但由于无法实时进入吸收塔内部进行检查, 仅能依赖于间歇性的停机检修, 这大大限制了对喷淋层运行状况的及时评估^[15-16]。

此外, 在大数据背景下, 脱硫系统的历史运行数据事实上成为了一种宝贵的资源。而现有研究大部分都基于抽样算法减少数据量从而减轻海量数据对传统机器学习算法的计算压力。事实上, 利用传统的抽样处理方法可能难以充分挖掘这些数据的潜在价值, 因为数据量庞大、结构复杂, 而且需要进行跨系统的综合性分析。而分布式计算框架的出现为该

收稿日期: 2024-04-08; 修回日期: 2024-08-29

基金项目: 国家自然科学基金项目(52076037)

作者简介: 徐侠(1989-), 男, 本科, 工程师, 主要从事火电机组节能优化运行工作, 16245101@ceic.com; 苏志刚(1979-), 男, 博士, 教授, 主要从事能源动力系统智能诊断与优化控制研究, 通讯联系人, zhigangsu@seu.edu.cn。

复杂问题提供了解决方案。通过分布式计算框架,可以有效地处理和分析大规模的历史运行数据,从中挖掘出隐藏的规律和趋势,为脱硫系统的优化和改进提供数据支持和决策依据。因此,将分布式计算框架与脱硫系统的历史运行数据结合起来,有助于提高工业生产的效率、降低能耗,同时实现环保和可持续发展。

本文以石灰石-石膏湿法脱硫系统为研究核心,运用证据学习方法^[17-18](证据聚类和证据KNN分类)与分布式计算框架Spark构建了浆液循环泵的故障诊断模型以及脱硫系统的运行优化模型。使用分布式计算框架Spark可以突破脱硫系统长时间运行大数据对故障诊断模型的计算瓶颈。这些模型初步实现了为WFGD系统运行提供维护建议、故障检测、推理解释以及自适应优化策略的功能,使得系统中的故障点可以被迅速且准确地诊断出来,并能够预先调整运行状态以提升操作效率。此外,研究成果也为脱硫装置的节能降耗运行目标提供了关键的技术支持。

1 证据理论基础与Spark计算框架

证据理论^[19]又称为Dempster-Shafer理论,是由Dempster于1967年首先提出,随后被Shafer在各方面完善。

1.1 证据表征

在证据理论中,对于某一变量 ω 取值的所有认知范围用一个识别框架(Frame of Discernment) $\Omega = \{\omega_1, \omega_2, \dots, \omega_c\}$ 表示。所关心变量 ω 的所有取值都在识别框架的子集中,对变量 ω 取值的信度被称为基本信度分配(Basic Belief Assignment, BBA)。BBA的基本定义如下:

记识别框架为 Ω ,若集函数 $m: 2^\Omega \rightarrow [0, 1]$ (2^Ω 为 Ω 的幂集),满足如下条件:

$$\sum_{A \subseteq \Omega} m(A) = 1 \quad (1)$$

则称该函数 m 为识别框架 Ω 上的一个基本信度分配,又称为一个质量函数(Mass function)。

在证据理论中,一个质量函数通常被称为一条证据。对于识别框架 Ω 中的任意命题 A ,其可信度 $m(A)$ 被定义为命题 A 为真的程度,也称为分配在命题 A 上的质量。具有 $m(A) > 0$ 的集合 A 被称为质量函数 m 的焦点。

除了使用质量函数以外,还可以使用信度函数(bel)来表示一条证据,其定义如下:

$$\text{bel}(A) = \sum_{\emptyset \neq B \subseteq A} m(B) \quad (2)$$

信度函数 $\text{bel}(A)$ 表达了人们对变量 $\omega \in A$ 的总的信任程度。

1.2 证据融合

当具有来自不同信息源的证据(质量函数)时,需要使用证据融合法则来融合不同的证据,从而获得描述某事物的最终证据。本小节介绍最常用的Dempster融合法则(Dempster's Rule)。

对于两个定义在同一识别框架下的质量函数 m_1 和 m_2 ,它们经过Dempster融合法则融合之后的证据记为 $m_1 \oplus m_2 = m_{1 \oplus 2}$,其中Dempster融合法则的定义如下:

$$m_{1 \oplus 2}(A) = \begin{cases} (\text{Kappa})^{-1} \sum_{B \cap C = A} m_1(B) m_2(C) & \forall A \subseteq \Omega, A \neq \emptyset \\ 0 & A = \emptyset \end{cases} \quad (3)$$

式中,Kappa为标准化因子,由如下公式计算:

$$\text{Kappa} = 1 - \sum_{B \cap C \neq \emptyset} m_1(B) m_2(C) \quad (4)$$

由此可知,经过Dempster法则融合之后的证据中分配给空集 \emptyset 的质量为0。

1.3 Spark计算框架

Spark是一种基于内存计算的分布式计算框架,其核心原理是将数据加载到内存中进行处理,从而避免了频繁的磁盘读写,极大地提高了计算效率。Spark采用了弹性分布式数据集(RDD)作为其基本抽象,通过RDD的惰性求值和容错机制实现了高效的数据处理。Spark包括了多个核心组件,包括Spark Core、Spark SQL、Spark Streaming、MLlib和GraphX等,其中Spark Core提供了RDD的抽象和基本操作,是Spark的最核心模块。

2 浆液品质监测模型建模

本节分别介绍了基于Spark计算框架和证据学习方法的监测建模思路、脱硫浆液品质监测模型输入参数以及建模具体步骤。

2.1 监测模型的建模思路

为了实现对浆液品质的在线监测,首先需要历史浆液数据中所包含的不同品质状态进行分析。这涉及到根据历史脱硫浆液数据来探索和描述其健康状态。在分析所有正常浆液数据的历史状态之后,待监测的浆液数据将与已探索的历史浆液品质数据进行比对,从而完成脱硫浆液品质的在线监测。需要注意的是,本论文着重于脱硫浆液品质的状态监测问题,而非故障诊断,尽管二者相关,但又存在

明显的差异。本研究认为状态监测是故障诊断的先决问题,即只有在脱硫浆液品质状态监测中发现异常后,才能对具体的原因和位置进行诊断。

挖掘和描述脱硫浆液品质历史健康状态实际上包含了两个层次的内容。首先需要确定脱硫浆液品质的历史运行数据实际上包含了几种常见的健康品质状态,然后对每个历史运行时刻上脱硫浆液品质状态属于不同健康品质状态的隶属度进行划分。这实质上是对脱硫浆液品质数据进行“聚类分析”^[20]。在聚类分析中,脱硫浆液在不同传感器测得的数值被称为维度,每个历史时刻上不同维度所组成的向量称为一个样本数据点。因此,挖掘和描述脱硫浆液品质状态就是对脱硫浆液品质历史上所有样本数据点进行聚类分析的过程。

在完成脱硫浆液历史健康状态的挖掘和描述后,可以利用这些数据来训练证据 KNN 分类器。然后,该分类器可以对实时浆液品质数据进行分类,从而实现其健康状态的在线监测。

2.2 监测模型输入参数

湿法脱硫工艺是一种常用于减少工业排放中硫磺污染物的技术。脱硫效率取决于多种因素,包括脱硫浆液的品质和其他相关的设备运行参数。

脱硫浆液品质数据包括:入口烟气粉尘浓度(z_1)、浆液密度(z_2)、浆液 pH 值(z_3)、脱硫塔浆液液位(z_4)、氧化风量(z_5)、脱硫效率(z_6)、烟气排放量(z_7)。

与脱硫浆液有关的设备主要为浆液循环泵,其在石灰石-石膏湿法脱硫工艺中是关键设备之一,负责处理大量流动物质。由于工作环境的苛刻性,这些泵经常面临严重的磨损和腐蚀问题。具体来说,泵内未溶解的固体颗粒会对泵体造成磨损,而腐蚀作用则影响泵的整体结构。此外,浆液中的气泡在达到饱和后破裂,产生水力冲击,形成凹坑和缺口,这些损伤在长期运行中会扩大,最终导致泵的失效。这些问题不仅使得浆液循环泵的能耗比正常水平高出 6% 到 7%,造成显著的经济损失,还严重影响了泵的运行效率和整个脱硫系统的性能。事实上,浆液循环泵是烟气脱硫系统中耗电量最大的设备,一般占系统总电耗的 33% 至 35%。因此,本文另选择 4 台循环泵的平均出口压力(y_1)、总流量(y_2)、平均轴承温度(y_3)和总电机电流(y_4)作为监测模型的输入变量。这些参数能够提供关于泵运行状态的重要信息,有助于及时监测和预防潜在故障,从而提高浆液循环泵的运行效率和稳定性。

2.3 具体建模步骤

2.3.1 利用信度峰值聚类算法挖掘浆液品质状态

Rodriguez 和 Laio 提出了一种名为基于密度峰值的硬划分聚类算法 (Density peaks clustering, DPC)^[21]。在 DPC 算法中,每个类的中心具有比其他邻域内的其他数据点更高的密度,并且与其他密度较高的数据点具有较远的距离。其中某个数据点 x_i 的密度被定义为:

$$\rho_i = \sum_{j \neq i} \chi(d_{ij} - d_c) \quad (5)$$

$$\chi(a) = \begin{cases} 1, & \text{if } a < 0 \\ 0, & \text{otherwise} \end{cases} \quad (6)$$

式中: d_{ij} 为点 x_i 与点 x_j 之间的欧式距离; d_c 称为截断距离。数据点 x_i 的相对距离 δ_i 被定义为数据点 x_i 与大于其密度的数据点之间的最小距离:

$$\delta_i = \min_{j: \rho_j > \rho_i} (d_{ij}) \quad (7)$$

但是上述定义仅不适用于具有最大密度的数据点,因此,对于最大密度的数据点,其密度被定义为:

$$\delta_i = \max_{x_j} (d_{ij}) \quad (8)$$

在获取各个数据点的距离和密度后,可以绘制 x 轴表示距离、 y 轴表示密度的决策图。在该图中,被选为聚类中心点的是那些具有高密度和高距离的数据点。剩余的数据点则被分配到它们最近邻域内的聚类中心点所代表的集合中。DPC 算法的主要优势在于,它在聚类之前无需预先确定集合的数量。然而,无论是 DPC 算法还是改进的 DPC 算法,都属于硬划分聚类算法。相较于硬划分聚类算法,引入不精确性和不确定性在聚类问题中更具意义。考虑不精确与不确定性的聚类示意图如图 1 所示,强行将标有问号的点划分到虚线包含的聚类中显然不合适,因为某一特定时刻的浆液品质可能处于两种健康状态之间。实际上,DPC 算法在计算密度时仅利用了邻域内的对象的数值信息,即对象间的距离与截断距离的相对大小。然而,某个点的邻域内的点可以提供给它更多关于是否可以成为聚类中心的信息^[21-23]。

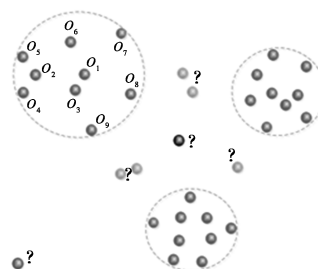


图 1 考虑不精确与不确定性的聚类示意图

受上述想法驱动,本文利用一种基于信度峰值的证据聚类算法(BPEC),旨在提供一个可合理响应数据内部结构的信度划分,更合理地挖掘历史浆液品质数据中所包含的质量信息。在信度划分中,当一个数据点不能明确地分类到某一类别时,可以使用一条证据来表示聚类的不确定性和模糊性。每个质量函数代表了对象属于该类别的程度。对于一个包含 n 个样本点的浆液质量数据集, n 维质量函数的集合是一个置信分区(表 1 给出了一个四维信度划分的示例,在该示例中,点 1 和点 3 具有较高的属于第一类别和第二类别的概率,而点 2 则属于完全无知的类别,表示对象属于的类别完全不清楚,而点 4 则代表一个噪声对象)。四维信度划分示例见表 1。

表 1 四维信度划分示例

m_i^Ω	\emptyset	$\{\omega_1\}$	$\{\omega_2\}$	Ω
m_1^Ω	0	0.7	0.2	0.1
m_2^Ω	0	0	0	1.0
m_3^Ω	0	0.2	0.8	0
m_4^Ω	1.0	0	0	0

对于给定的 n 个数据点,定义识别框架 $\Theta=(C, -C)$,其中 C 表示一个数据点被选为类中心, $-C$ 表示一个对象未被选为类中心。确定一个待决定的数据点是否为类中心的步骤总结如下:对于每个数据点,它的 K 个最近邻提供了 K 个证据,指示待决定的数据点是类中心的概率。通过 Dempster 法则融合 K 个证据后,可以得到一个质量函数 m_i 和对应的 Bel_i ,表示待决定的数据点是聚类中心的程度。邻居 x_j 对 x_i 提供的质量函数为:

$$\begin{cases} m_{ij}^C(A) = 1 - \phi(d_{ij}^2) & A = \{C\} \\ m_{ij}^C(A) = \phi(d_{ij}^2) & A = \{C, -C\} \\ m_{ij}^C(A) = 0 & \text{otherwise} \end{cases} \quad (9)$$

式中: d_{ij} 是 x_i 和 x_j 之间的欧氏距离,函数 ϕ 是一个递减函数。根据 Dempster 法则,通过融合 K 个证据可以得到最终的质量函数,并将其转换为置信函数 Bel_i^C 。在获得每个点的置信度后,可以根据公式(9)计算相互间的距离 δ 。然后,可以绘制一个决策图,其中 x 轴表示相对距离, y 轴表示置信度。在决策图中,具有较高置信度和较大相对距离的数据点将被选为聚类中心,决策图实例如图 2 所示。

在选择了聚类中心之后,剩余的数据点将被分配到与其自身最近的聚类中。本文中,第 2.2 节提到的浆液质量健康参数被用作每个数据点的特征,

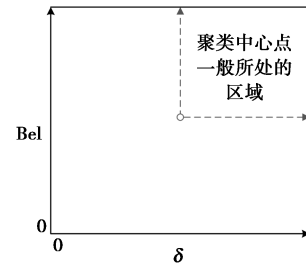


图 2 决策图实例

并且所有历史浆液质量数据都被输入到 BPEC 算法中,以获得聚类结果。

2.3.2 利用证据 KNN 算法在线诊断浆液品质状态

考虑在识别框架 $\Omega = \{\omega_1, \omega_2, \dots, \omega_c\}$ 中将一些测试样本 y_i 分类到某些类别的问题,在 EKNN 中,训练集 $T = \{(x_i, m_i) | i=1, 2, \dots, n\}$ 由 n 个 p 维数据点组成(本文中 p 维即为 2.2 节中提到的浆液参数),其中每个数据点的类别信息由证据函数 m_i 表示。与 KNN 分类器类似,当测试样本与训练样本之间的距离非常接近时,“它们具有相同类别属性”的命题的可信度越高;当测试样本与训练样本之间的距离非常远时,训练样本只能向测试样本提供一个“无知”的信息。下面的公式量化了训练样本和测试样本之间信息的传递:

$$m(A/x_i) = \begin{cases} \alpha m_i(\{\omega_q\}) \phi(d_i), & A = \{\omega_q\}, q = 1, 2, \dots, c \\ 1 - \alpha \phi(d_i) \sum_{q=1}^c m_i(\{\omega_q\}), & A = \Omega \end{cases} \quad (10)$$

式中, $\alpha \in (0, 1)$ 是一个人为设定的超参数,一般为 0.95。完成了对所有 K 个最近邻的证据折现过程后,利用式(3)中定义的 Dempster 法则融合得到的 K 个证据。关于 x_i 的分类结果是具有最大质量函数值的类别。在使用 EKNN 进行分类时,最终的分类结果以质量函数的形式获得。这意味着在分类结果中存在一种可能性,描述了测试样本属于“无知”类别,即 $m_i(\Omega)$ 。在设备状态监测的背景下,这种“无知”类别的概率可以理解为当前状态点偏离历史上所有正常操作状态的程度。当检测到历史运行数据中的所有正常状态时,偏离所有正常状态的可能性可以解释为“异常状态发生的可能性”。利用这一特性,本文完成了设备状态监测模型的建立。

2.3.3 模型分布式构建

由 2.3.2 节可知,BPEC 算法总共包括 4 个步骤:找到每个样本点的 K 个最近邻、计算置信度和相对距离、选择类中心以及计算置信分区。计算置

信度和相对距离的复杂度为 $O[p^2n(n-1)/2] + O(2^2Kn)$ 。BPEC 算法的时间复杂度主要取决于数据点数量 n 。其中两个耗时最多的项目是探索 KNN 和计算每个数据点的 δ_i ，其复杂度达到 n 的平方。当处理数据量较小的数据集时，计算 δ_i 的复杂度可以降低到 $O(n \log n + n)$ 。这是因为将不相似矩阵存储在内存中，并且可以直接在不相似矩阵中搜索最小的成对距离，而无需重新计算。然而，由于对于大型数据集来说，不相似矩阵无法存储在内存中，因此，在为具有大量运行数据的设备构建证据库时，无法省略此重新计算过程。

分布式 BPEC 算法中搜索 K 近邻与信度计算流程如图 3 所示。每个数据点 x_i 的 K 近邻的分布式搜索可分为 2 个阶段：映射 (Map) 和归约 (Reduce)。可以观察到，在映射阶段 1 中采用了“分而治之”的方式。首先将浆液品质数据集 O 加载为 RDD 形式，并将其切分为分布在 L 个不同计算核心上的 L 个相似数据点数量的片段 O_1, O_2, \dots, O_L 。每个数据点 x_i 也作为映射阶段 1 的输入，其候选的 K 个最近邻在每个片段上进行搜索。在接收到来自 L 个片段的所有候选 K 近邻之后，通过 Spark 函数 `reduceByKey (Combine)` 得到真正的最近邻，然后在映射阶段 2 根据公式 (2) 计算每个数据点的信度。通过上述分布式计算框架的搭建，可以将本文所提出的证据学习监测模型扩展到十万级甚至上百万级的脱硫浆液品质数据上。

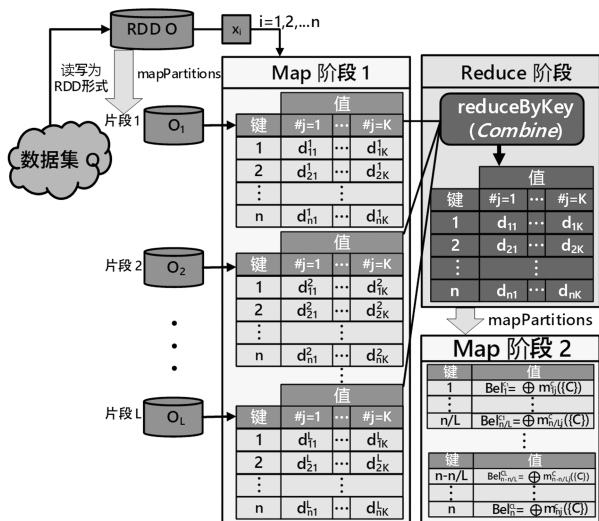


图 3 分布式 BPEC 算法中搜索 K 近邻与信度计算

3 模型测试与分析

本节利用工业实例验证基于证据学习算法所构

建的浆液品质在线监测模型的有效性。本文所使用的计算机具有 40 个型号是 Intel(R) Xeon(R) Gold 6230 CPU@2.10 GHz 580 的 CPU 核心。本节为了验证所构建的监测模型的性能，提取江苏某 1 000 W 热电厂从 2021 年 1 月 30 日至 2021 年 4 月 30 日共 133 200 条脱硫浆液品质数据。首先将脱硫效率、烟气排放量、入口烟气粉尘浓度、浆液密度、浆液 pH 值、脱硫塔浆液液位、氧化风量以及循环泵的出口压力、流量、轴承温度和电机电流作为信度峰值证据聚类算法的输入，确定这些历史品质数据中包含了多少个健康品质状态。其算法运行结果的决策图如图 4 所示。

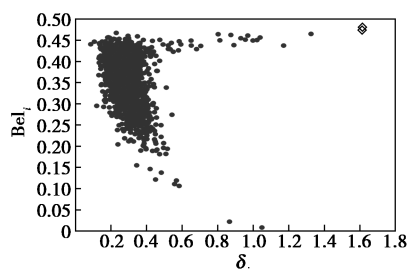
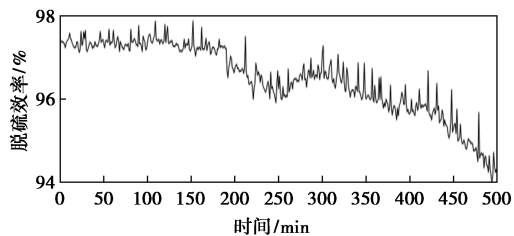
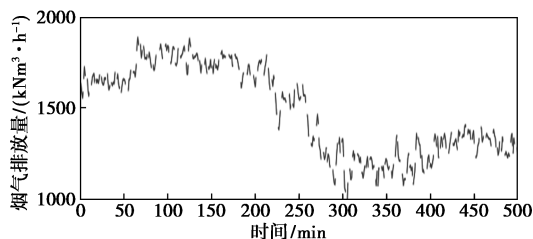


图 4 浆液品质决策图

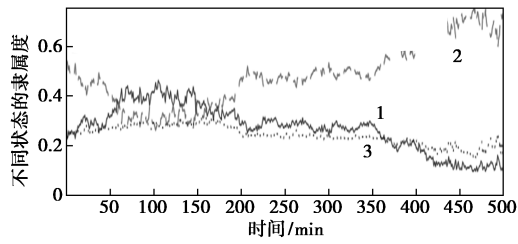
由图 4 可见，在 2021 年 1 月 30 日至 2021 年 4 月 30 日浆液品质可以被分为 2 个（由图右上角的 2 个菱形）状态。经过观察数据得知，发现历史数据被聚为两个健康类。其中一类的脱硫效率较高（大于 96%），烟气排放量较大；另一类的脱硫效率次高（小于 96% 大于 94%），烟气排放量较小。将剩余的样本点完成聚类后，用来训练证据 KNN 分类器，构建状态监测模型。将 2021 年 4 月 2 日 00:00 起的 500 条数据作为测试数据，得到了如图 5 所示的监



(a) 脱硫效率



(b) 烟气排放量

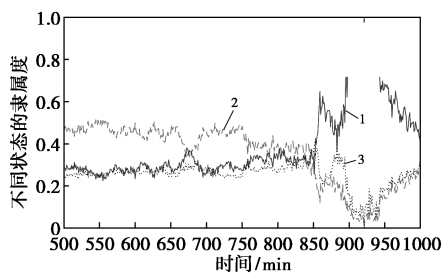


1—属于 ω_1 状态的隶属度;2—属于 ω_2 状态的隶属度;
3—属于 Ω 状态的隶属度
(c)不同类别隶属度

图5 4月2日前500 min浆液品质监测结果图

由图5可知,在4月2日的前220 min浆液属于第二类的隶属度略高于第一类,而属于无知类的隶属度(即浆液健康品质劣化的可能性)则一直较低。在220 min后,可以看到烟气排放量的减少,脱硫效率下降。经观察数据得知,发现循环泵的电流量减小,功率下降。此时,脱硫浆液的属于第二类的隶属度继续上升。

随后将4月2日500~1 000 min的数据作为测试数据放入模型,结果见图6。由图6可见,属于浆液品质发生少许劣化的可能性在600 min后有轻微上升。在排查对应数据的参数后发现,石灰石浆液密度在500 min后不断下降,这需要提升其浓度防止继续劣化。然而在实际操作中,浆液的浓度也不可过高,这是因为过高的浓度会造成浆液循环泵的管道出现堵塞。这种堵塞情况通常与吸收塔内的浆液浓度有很大关系,特别是当石膏浆液的浓度超出了石膏的过饱和度,就会导致石膏结晶在吸收塔的内壁、底部和循环泵的入口滤网等地方形成沉积物。为了确保脱硫效果能够达到排放标准,在实际操作过程中,通常会将石膏浆液的浓度保持在一个相对较高的水平,这一水平是高于设计值的。我们将属于无知类别隶属度0.2作为阈值(在所有对比方法



1—属于 ω_1 状态的隶属度;2—属于 ω_2 状态的隶属度;
3—属于 Ω 状态的隶属度

图6 4月2日500~1 000 min浆液品质监测结果

中此值均作为阈值来设定),对比不同状态监测算法(PCA,FD-KNN^[24],CMEW^[25])的报警性能,对比结果如表2所示。

表2 不同状态监测方法对比结果

	EVM	FD-KNN	PCA	CMEW
报警时间/min	574	585	591	579

注:EVM为本文所提出的方法。

由于本文所提出的方法(EVM)中的基于证据KNN分类的方法考虑的是数据的 K 近邻局部分布,因此具有非线性的性质,并且利用证据理论优化了分类监测诊断过程中的信息融合过程,所以,本文所提出的方法在预警及时性方面要优于剩余的其他3种方法。

我们也测试了使用并行计算框架所节省的运行时间与达到的加速效果。根据文献[21]中所提出的加速比(speed up)来定义本文中的加速指数SP:

$$SP = \text{Time}_1 / \text{Time}_U \quad (11)$$

式中, Time_1 和 Time_U 分别代表了使用一个核心程序所消耗的时间和使用 U 个核心所消耗的时间。不同数量核心的运行时间和加速指数见表3。

表3 不同数量核心的运行时间和加速指数

核心数	1	10	20	40
运行时间/s	4324	443	229	125
SP指数	—	9.76	18.88	34.59

由表3结果可知,在使用越多的计算核心后,对于脱硫系统历史数据的聚类时间越少,相应的加速指数也会增加。更具体来说,基本能够达到线性加速(如使用10个核心数所消耗的时间大约仅仅是使用1个核心数所消耗的时间的十分之一)。

4 结论

本文基于证据学习算法,对脱硫浆液品质建立了健康品质监测模型,利用该模型对江苏某电厂的浆液品质进行了在线监测。经测试表明:本文所建立的监测模型能够准确监测出脱硫浆液品质的恶化,与其他3类先进监测方法对比结果说明了本文所建立模型仍能达到最优的报警及时性。说明将证据理论应用于脱硫浆液品质监测是可行的,为脱硫浆液品质监测提供了一种新方法。

参考文献

- [1] 张志勇,莫华,王猛,等.600 MW燃煤机组烟气污染物控制研究[J].中国电力,2022,55(5):204-210.

- [2] 禾志强,周鹏,张铭,等.有机酸添加剂在石灰石-石膏法脱硫中应用的试验研究[J].中国电机工程学报,2011,31(23):34-40.
- [3] Lim J H, Cho H T, Lim J. Optimization of wet flue gas desulfurization system using recycled waste oyster shell as highgrade limestone substitutes[J].Journal of Cleaner Production,2021,318:1-15.
- [4] 曲江源,齐娜娜,关彦军,等.湿法烟气脱硫塔内传递与化学反应过程 CFD 模拟[J].化工学报,2019,70(6):2117-2128.
- [5] 侯建勇,严芳,王浩,等.基于因果-模糊层次分析的湿法脱硫系统运行稳定性综合评价[J].化工进展,2022,41(2):569-583.
- [6] 曹国庆,邢金城,涂光备.基于灰色层次分析理论的烟气脱硫技术评价方法[J].中国电机工程学报,2006,26(4):51-55.
- [7] 马双忱,周权,曹建宗,等.湿法脱硫系统动态过程建模与仿真[J].化工学报,2020,71(8):3741-3751.
- [8] 吴磊,康英伟.基于改进粒子群优化长短时记忆神经网络的脱硫系 SO₂ 预测模型[J].热力发电,2021,50(12):66-73.
- [9] 许丹,沈凯,张亚平,等.基于模糊理论的湿法脱硫故障诊断和优化模型研究[J].环境工程,2018,36(2):92-97.
- [10] 张锦航,乔宗良,司凤琪.基于模糊理论的脱硫系统故障诊断[J].工业控制计算机,2021,34(9):92-96.
- [11] 徐遵义,刘文慧,张旭冉,等.基于数据驱动的浆液循环泵运行优化研究[J].中国电力,2021,54(3):197-204.
- [12] 范昊鹏,夏凤毅,包军宇,等.基于 PFC 预测的燃煤机组脱硫优化控制[J].动力工程学报,2020,40(10):808-814.
- [13] 贾西部,金万元,李兴华,等.石灰石-石膏湿法烟气脱硫系统浆液起泡原因分析[J].中国电力,2015,48(9):157-160.
- [14] Gu H, Cui Y F, Zhu H X, *et al.* A new approach for clustering in desulfurization system based on modified framework for gypsum slurry quality monitoring[J].Energy,2018,148:789-801.
- [15] 薛建明,王小明,刘建民,等.湿法烟气脱硫设计及设备选型手册[M].北京:中国电力出版社,2011:133-139.
- [16] 董锐锋,陈浩军,王锋涛,等.单塔双循环脱硫系统结垢问题研究[J].中国电力,2018,51(4):149-154.
- [17] Su Z, Denooux T. BPEC: Belief-peaks evidential clustering [J]. IEEE Transactions on Fuzzy Systems,2018,27(1):111-123.
- [18] Denooux T. A k-nearest neighbor classification rule based on Dempster-Shafer theory[J].IEEE Transactions on Systems, Man, and Cybernetics,1995,25(5):804-813.
- [19] Shafer G. Dempster-Shafer theory[J].Encyclopedia of Artificial Intelligence,1992,1:330-331.
- [20] Gong C, Su Z, Wang P, *et al.* Cumulative belief peaks evidential K-nearest neighbor clustering [J]. Knowledge-Based Systems,2020,200:105982.
- [21] Gong C, Su Z, Wang P, *et al.* Distributed evidential clustering toward time series with big data issue[J].Expert Systems with Applications,2022,191:116279.
- [22] Gong C, Su Z, Wang P, *et al.* An evidential clustering algorithm by finding belief-peaks and disjoint neighborhoods [J]. Pattern Recognition,2021,113:107751.
- [23] Gong C, Li Y, Fu D, *et al.* Self-reconstructive evidential clustering for high-dimensional data [C]. 2022 IEEE 38th International Conference on Data Engineering (ICDE), Beijing:IEEE,2022.
- [24] He Q P, Wang J. Fault detection using the K-nearest neighbor rule for semiconductor manufacturing processes [J]. IEEE Transactions on Semiconductor Manufacturing,2007,20(4):345-354.
- [25] Chen X, Wang P, Hao Y, *et al.* Evidential KNN-based condition monitoring and early warning method with applications in power plant [J]. Neurocomputing,2018,315:18-32. ■

(上接第 347 页)

参考文献

- [1] Hussin N E, Johari A, Kidam K, *et al.* Major hazards of process equipment failures in the chemical process industry [J]. Applied Mechanics and Materials,2015,735:75-79.
- [2] Datta S, Sarkar S. A review on different pipeline fault detection methods [J]. Journal of Loss Prevention in the Process Industries,2016,41:97-106.
- [3] Abid A, Khan M T, Iqbal J. A review on fault detection and diagnosis techniques: Basics and beyond [J]. Artificial Intelligence Review,2021,54:3639-3664.
- [4] Chen K Y, Chen L S, Chen M C, *et al.* Using SVM based method for equipment fault detection in a thermal power plant [J]. Computers in Industry,2011,62(1):42-50.
- [5] Li Y, Lu N Y, Jiang B, *et al.* An islanding fault detection method with CFDF-SVM based RPV approach under pseudo islanding phenomenon [J]. IFAC-Papers on Line,2018,51(24):1349-1355.
- [6] 周莉,鲍志伟.基于灰狼算法优化 SVM 的变压器故障诊断 [J]. 长江信息通信,2022,35(9):27-29.
- [7] 李英顺,张国莹,张杨,等.基于 SSA 算法优化 SVM 的发动机润滑油信息状态评估 [J]. 润滑与密封,2023,48(2):129-134.
- [8] Abualigah L, Diabat A, Mirjalili S, *et al.* The arithmetic optimization algorithm [J]. Computer Methods in Applied Mechanics and Engineering,2021,376:113609.
- [9] Hearst M A, Dumais S T, Osuna E, *et al.* Support vector machines [J]. IEEE Intelligent Systems and their Applications,1998,13(4):18-28.
- [10] Patle A, Chouhan D S. SVM kernel functions for classification [C]//2013 International Conference on Advances in Technology and Engineering (ICATE), IEEE,2013:1-9.
- [11] 李炳宇,萧蕴诗.新的进化计算算法——粒子群优化算法 [J]. 计算机科学,2003,(6):19-22.
- [12] Sohail A. Genetic algorithms in the fields of artificial intelligence and data sciences [J]. Annals of Data Science,2023,10(4):1007-1018.
- [13] Devarapalli R, Bhattacharyya B, Kumari A. Enhancing oscillation damping in a power network using EWOA technique [M]//Advances in Smart Grid Automation and Industry 4.0; Select Proceedings of ICETSGAI4.0. Singapore: Springer Singapore,2021:27-36.
- [14] Francesca C, Erik V, Murat K, *et al.* The revised Tennessee Eastman process simulator as testbed for SPC and DoE methods [J]. Quality Engineering,2019,31(2):212-229.
- [15] Larsson T, Hestetun K, Hovland E, *et al.* Self-optimizing control of a large-scale plant: The Tennessee Eastman process [J]. Industrial & Engineering Chemistry Research,2001,40(22):4889-4901. ■