

一种发电厂气体污染监测与预警模型设计

付康民¹, 邱建新¹, 解 标²

(1. 国能(绥中)发电有限责任公司 生产技术部, 辽宁 葫芦岛 125222;

2. 安徽德源环境科技有限公司, 安徽 蚌埠 233000)

摘要: 针对电厂气体污染监测的需要, 该文提出了一种融合多源无线传感器网络 (Wireless Sensor Networks, WSN) 与改进长短期记忆网络 (Long Short Term Memory, LSTM) 的气体污染监测与预警模型。通过设计多源 WSN 污染气体采集框架, 实现了对电厂多种污染气体的高效采集, 为预测模型提供高质量的数据输入。所设计的改进 LSTM 模型融合了模拟退火算法 (Simulated Annealing, SA) 和 SVM 模块, 其中 SA 用于优化 LSTM 的超参数, SVM 则作为分类器有效避免过拟合问题, 使改进后的 LSTM 能够充分提取多源传感器数据的特征并进行准确预测。实验结果表明, 结合 SA 和 SVM 模块有效提升了 LSTM 的预测性能, 与其他分类算法相比表现出了明显的优势, 准确率高达 97.83%, 相比于对比算法中表现最佳的 BiLSTM 提高了 9.64%。

关键词: 多源数据融合; WSN; 污染气体检测; LSTM; 模拟退火算法; SVM

中图分类号: TP391.9

文章编号: 1000-0682(2025)02-0021-05

文献标识码: A

DOI: 10.19950/j.cnki.CN61-1121/TH.2025.02.004

Design of a gas pollution monitoring and early warning model for power plants

FU Kangmin¹, QIU Jianxin¹, XIE Biao²

(1. The Production and Technology Department of Guoneng (Suizhong) Power Generation Co., Ltd., Liaoning Huludao 125222, China;

2. Anhui Deyuan Environmental Technology Co., Ltd., Anhui Bengbu 233000, China)

Abstract: In response to the need for gas pollution detection in power plants, this paper proposes a gas pollution monitoring and early warning model that integrates multi-source Wireless Sensor Networks (WSN) and improved Long Short Term Memory (LSTM) networks. By designing a multi-source WSN pollution gas collection framework, efficient collection of various pollution gases from power plants has been achieved, providing high-quality data input for prediction models. The improved LSTM model designed integrates Simulated Annealing (SA) algorithm and SVM module, where SA is used to optimize the hyperparameters of LSTM and SVM acts as a classifier to effectively avoid overfitting problems, enabling the improved LSTM to fully extract the features of multi-source sensor data and make accurate predictions. The experimental results show that the combination of SA and SVM modules effectively improves the predictive performance of LSTM, showing significant advantages compared to other classification algorithms, with an accuracy rate of up to 97.83%, which is 9.64% higher than the best performing BiLSTM among the comparison algorithms.

Keywords: multi-source data fusion; WSN; pollutant gas detection; LSTM; simulated annealing algorithm; SVM

0 引言

在现代电力行业中,随着环保法规的日益严格和公众对环境问题关注度的不断提升,发电厂的气体污染监测与预警已成为一个重要的研究领域。发电厂在运行过程中排放的二氧化硫、氮氧化物、二氧

收稿日期: 2024-10-15

第一作者: 付康民(1983—),男,汉族,吉林通化人,硕士,高级工程师,研究方向为电力设备运行及检修技改。

化碳以及其他有害气体,直接影响着大气环境质量,进而对人类健康和生态环境产生不利影响^[1-3]。因此,如何有效监测和预测发电厂气体污染状况,及时采取措施防止污染加剧,已成为电力行业亟待解决的问题之一。传统的污染气体监测手段多依赖于固定监测点或人工采样分析,虽然能够提供污染状况数据,但由于监测范围有限、实时性较差等缺陷,难以满足现代污染防控的需求。尤其是在电厂复杂的运行环境中,气体污染具有空间分布广泛、时间变化快等特点,依靠单一数据源或监测技术通常无法获得全面、准确的污染信息。随着无线传感器网络(Wireless Sensor Networks, WSN)技术的发展,多源数据采集与融合技术为污染监测提供了新的解决方案。多源 WSN 可以在电厂内外部署多个传感器节点,实时收集多个污染物的浓度数据,形成广泛覆盖的监测网络。与此同时,借助大数据分析技术,可以对这些多源数据进行融合与处理,从而为污染状况的评估和预测提供更精确的依据。因此,该文提出了一种融合多源 WSN 与改进长短期记忆网络(Long Short Term Memory, LSTM)的发电厂气体污染监测与预警模型,用于发电厂气体污染的监测和预警。该模型利用多源 WSN 收集多种污染气体的相关信息,并通过改进后的 LSTM 挖掘复杂多源数据中的深层次特征以及捕捉时间序列数据中的长期依赖关系,从而实现了对发电厂气体污染的精确检测。

1 采集框架设计

在实际应用中,污染气体的排放具有多源性和动态性,监测系统需要覆盖广泛的空间区域,并具备实时响应能力。传统的监测手段如固定站点监测、人工采样等,虽然能够提供一定的监测数据,但在大范围、多维度的污染监测需求下显得效率低下。WSN 技术^[4-8]的引入,为污染气体的多源实时监测提供了技术路径。WSN 通过多个传感器节点,能够在不同空间位置和时间点持续采集污染气体的浓度信息,形成覆盖广泛的监测网络,有效提升监测系统的灵活性和准确性。基于多源 WSN 的污染气体数据采集框架中包含有数据的采集层、传输层和处理层 3 个主要部分,整体框架如图 1 所示。其中,数据采集层是整个框架的基础,负责通过多个传感器节点对发电厂区域内的污染气体进行实时采集。每个传感器节点内嵌有不同类型的传感器,能够检测多种污染物。为了保证数据的全面性和准确性,传感器节点分布在发电厂内外的不同区域,并根据风速、

排放源位置、地形等因素进行合理布置,确保覆盖到所有可能的污染区域。数据传输层负责将采集到的多源数据通过无线通信的方式传输到中心数据服务器。WSN 利用自组织网络机制,每个传感器节点不仅可以收集数据,还能作为中继节点,将其他节点的数据转发到下一节点或直接传输到中心服务器。这种机制提高了数据传输的灵活性和稳定性,尤其在大范围监控区域中,能够有效避免单点故障对整个系统带来的影响。数据传输层的设计还考虑了数据的实时性和传输效率,通过 ZigBee、LoRa 等通信协议^[9-10]实现了低功耗的高效传输。在最后的处理层中,将采集到的污染气体数据经过初步的预处理,如数据清洗、异常值剔除等,提升数据的质量,确保为后续预测模型提供高质量的输入。

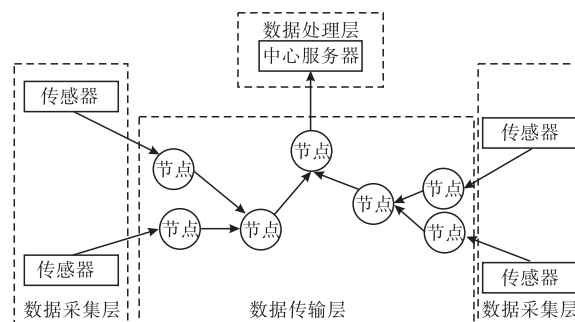


图 1 多源 WSN 污染气体采集框架图

2 预测模型设计

2.1 气体污染检测算法整体框架

为了更智能、准确地对电厂排放的污染气体进行监测和预测,该文提出了一种结合模拟退火算法(Simulated Annealing, SA)、LSTM 与 SVM 的污染气体检测算法,其整体框架设计如图 2 所示。首先,由多源无线传感器网络收集多种污染物的相关数据,并对其进行清洗与预处理,去除异常值并填补缺失数据,同时选取影响气体浓度的关键特征信息。接下来,将数据输入到该文所提出的基于 SA-LSTM-SVM 的污染气体检测模型中。该模型利用 LSTM 获取多源数据中的深层次特征,并捕捉时间序列数据中的长期依赖关系。在 LSTM 的训练过程中,利用 SA 算法优化 LSTM 的超参数,包括隐藏层单元数、学习率和批次大小等,通过在每次迭代中随机选择新参数并计算验证集上的损失,优化 LSTM 的参数设置,从而进一步提高模型的预测精度。其次将 SVM 作为最后的分类器,并将 LSTM 所获取到的特征信息作为 SVM 的输入,最终输出预测结果。

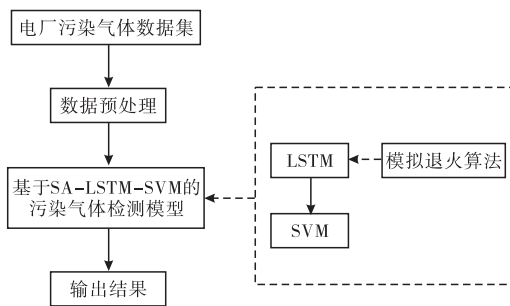


图2 污染气体检测算法整体框架设计图

2.2 模拟退火算法

在训练过程中,随机初始化权重,使其不会偏向于某个特定的输入,且模型超参数的变化会显著影响训练过程中的损失程度。学习率决定了训练过程中权重更新的幅度,较高的学习率可以使模型在较短的时间内收敛到最优权重,但其过高则会导致模型跳跃,无法准确到达最优点。而较低的学习率可以有效控制训练过程,使其到达最优点。因此,学习率的选择对模型的准确率具有较大影响。而衰减率决定了学习率随着每次更新衰减的速度,学习率会随着迭代次数的增加而逐渐减小,减小学习率可以加快训练过程,因此需要为优化后的学习率确定一个最优的衰减率。批大小则决定了每次迭代中用于训练的数据样本数量,最优的批大小决定了训练权重值是否趋向于准确值。为了在训练过程中使损失最小化并在一个区间内达到全局最优,此次使用SA算法^[11-13]对学习率、衰减率、批量大小这3个超参数进行优化。SA具有可避免局部优化的优势,适合在复杂和多模态优化问题中寻找全局解决方案,与梯度下降(Gradient Descent)和粒子群优化算法(Particle Swarm Optimization, PSO)相比不会陷入局部最小值,且SA概率探索机制使其能够探索广泛的解决方案并避免过早收敛。因此该文引入了SA对LSTM进行改进,利用SA优化LSTM的超参数,使其更适用于电厂污染气体的多源数据特征提取,从而提高污染检测的准确率。

SA算法由两部分组成:退火过程和Metropolis算法,分别对应算法的外循环和内循环。首先外循环即退火过程,其会将整个优化过程理解为固态退火过程,固体初始被加热到高温 T_0 ,然后按照选定的冷却因子 α 进行冷却,当温度达到最终温度 T_{end} 时,整个退火过程结束。内循环是Metropolis算法^[12],在每个温度下进行 L 次迭代,寻找该温度下的最小能量或最优值。

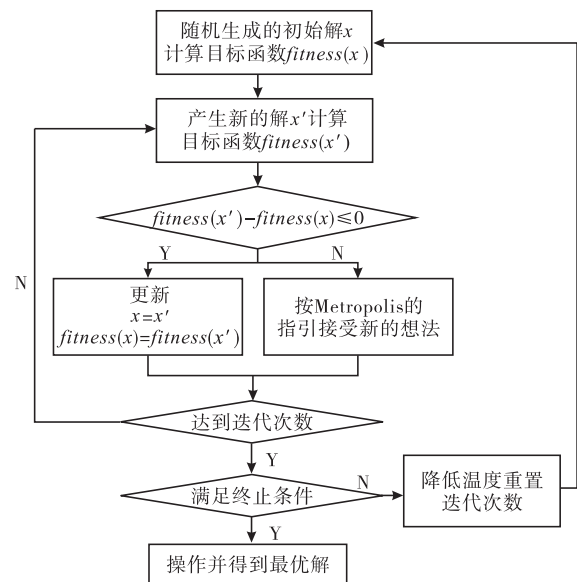


图3 SA算法优化流程

SA基本流程如图3所示,计算步骤如下:当给定温度状态 T 时,当前状态为超参数值 x ,在 x 附近设定一个邻域范围,使得 $x' = x$,此时 x' 为新状态。两个状态中的能量,即对应的目标模型损失值分别为 $fitness(x)$ 和 $fitness(x')$ 。当 $fitness(x') \leq fitness(x)$ 时,接受 x' 为当前状态后继续退火计算。当 $fitness(x') > fitness(x)$ 时,则根据式(1)定义,判断概率 p 是否大于 $(0,1)$ 之间的随机数 r ,再据此决定是否接受新状态为当前状态。确定当前状态后,回到原路继续降温,迭代直至在规定的范围内搜索到平衡状态的最优值。

$$p = \frac{1}{1 + e^{-(fitness(x') - fitness(x))/T}} \quad (1)$$

2.3 基于SA的改进LSTM

为了建立LSTM模型^[15-16]并将其应用于真实世界数据的预测和训练,需要定义LSTM内部神经网络模型的基本参数。初始参数如表1所示。建立LSTM模型后,为了找到最优的超参数使模型的预测效果达到最佳,采用SA算法对模型中的学习率、衰减率以及批大小3个超参数进行优化。根据优化标准,为了加快模型的收敛速度并提高训练精度,防止过拟合问题以及参数过大造成的梯度爆炸和训练失衡,该文将学习率在 $1 \times 10^{-6} \sim 1 \times 10^{-3}$ 范围内进行优化。在优化过程中选取训练和预测过程中的损失值作为评价标准,确定超参数对模型的反馈。

表 1 LSTM 模型初始参数

参数名称	初始化参数值
LSTM 层数/层	2
第一层 LSTM 神经元数量/个	32
第二层 LSTM 神经元数量/个	64
全连接层数量/个	2
第一层连接层神经元数量/个	32
第二层连接层神经元数量/个	1
衰减率	0.4
迭代次数/次	50
初始化学率	0.001
Adam 优化器的初始化衰减值	0.9
初始化批大小/个	16

该文所提出的基于 SA 的 LSTM 模型整体结构,如图 4 所示。首先,算法开始时随机选择一组超参数作为当前解,并训练 LSTM 模型,计算在验证集上的损失。在每次迭代中,根据当前温度生成新解,随机微调学习率、衰减率或批大小,计算新解的损失。并通过比较当前解与新解的损失值决定是否接受新解:若新解的损失降低,则接受该解;如果新解的损失上升,则以概率 p 接受次优解。随着迭代的进行,逐步降低温度,直至达到设定的终止条件。通过这种方式,SA 有效探索超参数空间,最终得到最佳的学习参数,进而提升 LSTM 模型的性能。

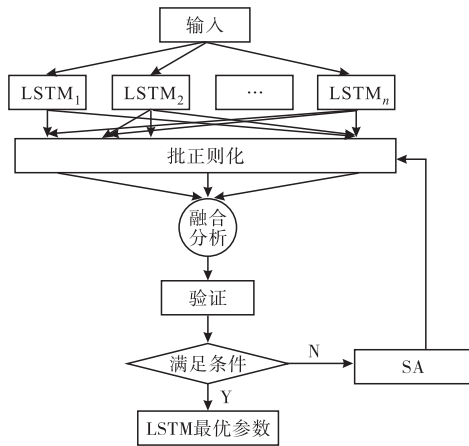


图 4 基于 SA 的改进 LSTM 模型结构图

2.4 基于 SA - LSTM - SVM 的污染检测模型

LSTM 能够有效捕捉时间序列中的动态变化和长短期依赖关系,在电厂污染气体检测中,气体浓度随时间变化,LSTM 能够提取出这些时序特征。然而,LSTM 的输出通常是高维特征向量,直接进行分类可能会导致复杂性增加,尤其是在样本量较小的情况下,容易出现过拟合现象。而 SVM^[17-18]能够在小样本和高维数据环境中通过寻找最佳超平面来

划分不同类别的数据,最大化类间距离,从而提高分类准确性。将 LSTM 提取的特征输入到 SVM 中,可以利用 SVM 的强大分类能力,提供更稳健的分类结果。因此文中将 SVM 作为分类器加入到污染气体检测模型中,提出了基于 SA - LSTM - SVM 的气体污染检测模型,具体结构如图 5 所示。该模型通过将 SA 与 LSTM 模型结合,充分发挥了 LSTM 在特征提取上的优势,并且利用 SVM 的高效分类性能,从而提升了整个模型的准确性和鲁棒性。

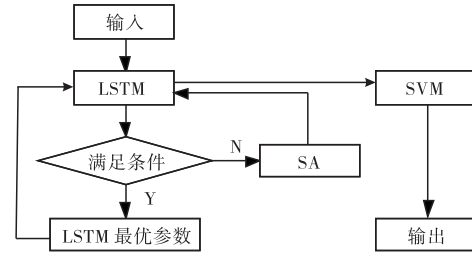


图 5 基于 SA - LSTM - SVM 的气体污染检测模型

3 实验结果与分析

3.1 实验数据

实验数据样本来自于某电厂的多源传感器设备,收集了为期 3 个月的污染气体浓度数据。数据集包含主要污染物的实时监测记录,如二氧化硫、氮氧化物和一氧化碳等污染气体的浓度值。此外,环境变量如温度、湿度和风速也被同步记录,以便对气体浓度的变化进行分析。数据集共包含 5000 条记录,每条记录均包含气体浓度值、时间戳和环境变量的特征。为确保数据质量,经过数据清洗与预处理后,还需剔除异常值并填补缺失数据。此次随机抽取了 80% 的数据用于模型训练,20% 的数据用于验证和测试模型性能。

3.2 实验结果分析

为验证所提出气体污染检测模型的有效性,文中进行了多组消融实验对比,具体结果如表 2 所示。所提模型的准确率为 97.83%,表明 SA 和 SVM 同时改进的 LSTM 具有良好的识别性能。

表 2 消融实验结果

算法	准确率/%
LSTM	84.47
SA - LSTM	94.21
SVM - LSTM	93.86
SA - SVM - LSVM	97.83

与其他分类算法进行比较实验的结果,如表 3 所示。由表 3 可知,SVM 和 ELM 的预测效果较差,

准确率分别为 75.49% 和 79.27%，而 GRU 与前两个相比有一定的提升，准确率为 83.82%。接着使用 BiLSTM 和 BiGRU 进行预测，准确率为 86.37% 和 88.19%，相比于上述算法均有优势，但是仍比该文模型低了 11.46% 及 9.64%。由对比结果可以证明，结合 SA 算法优化的 LSTM 与 SVM 在处理复杂电厂气体污染检测任务时，具有更强的识别能力和更高的准确性。

表 3 对比实验结果

算法	准确率/%
SVM	75.49
ELM	79.27
GRU	83.82
BiLSTM	86.37
BiGRU	88.19
SA-SVM-LSVM	97.83

4 结束语

为了应对电厂气体污染检测中准确性和实时性的挑战，该文提出了一种结合 SA 与 SVM 优化 LSTM 的电厂气体污染检测模型。该模型通过 SA 优化 LSTM 的超参数，提升了特征提取的效率，并利用 SVM 增强了分类能力。实验结果证明，所提模型能够有效地处理多源传感器数据，实现对污染气体浓度的精准监测与预测，并有效克服了传统方法在复杂数据环境中的局限性。

参考文献：

- [1] 张瑞山, 柏建华, 陈军, 等. 基于布谷鸟算法的火力发电厂经济调度与碳排放研究[J]. 制造业自动化, 2023, 45(9): 94-98.
- [2] 张彦军, 谢志成, 董强, 等. 替代燃料应用于燃煤发电厂的碳排放研究[J]. 能源与环境, 2024(2): 70-72+88.
- [3] 付金杯, 李梦南, 徐炜达, 等. 基于卫星遥感等多源数据的发电厂 CO₂ 排放分析[J]. 中国环境科学, 2024, 44(4): 1805-1815.
- [4] 张旺, 陈丹, 管维亚, 等. 一种基于层次信任的智能电网 WSN 安全检测[J]. 信息技术, 2022(7): 75-80.
- [5] 聂文梅, 宋晓霞. 基于自适应粒子群优化算法的无线传感器网络覆盖控制[J]. 沈阳工业大学学报, 2023, 45(4): 459-464.
- [6] 艾雪瑞, 许放, 曹瑾, 等. 智慧园区无线传感网络能量均衡覆盖优化系统[J]. 电子设计工程, 2024, 32(11): 110-114.
- [7] 郑晓, 汪豪, 梁伟昕, 等. 基于 WSN 与异常数据识别的拉线动态监测方法研究[J]. 电子设计工程, 2023, 31(15): 79-83.
- [8] 朱本科, 高丙朋, 蔡鑫. 基于多因素均衡动态分簇的 WSN 路由协议算法[J]. 科学技术与工程, 2024, 24(16): 6799-6808.
- [9] 吴海欣, 肖蕾, 李冠希, 等. 基于 ZigBee 的无线物联网传感系统[J]. 自动化与仪表, 2024, 39(6): 142-146+152.
- [10] 巩潇, 崔登祺, 万彬彬, 等. 基于 LoRa 无线网络的网络安全问题和实验验证研究[J]. 工业控制计算机, 2024, 37(7): 67-68+82.
- [11] 徐亚军, 吴红洪, 赵一阳, 等. 基于模拟退火粒子群算法的甚高频台站补盲[J]. 科学技术与工程, 2024, 24(21): 8996-9001.
- [12] 张学友, 董翔宇, 葛健, 等. 基于模拟退火算法的特高压直流 VDCOL 控制环节优化方法[J]. 电力科学与技术学报, 2024, 39(4): 53-60.
- [13] 徐博, 韩晓霞, 董颖超, 等. 自适应状态转移模拟退火算法及其应用[J]. 计算机应用研究, 2024, 41(1): 150-158.
- [14] 叶梓萌, 张大斌. 融合反向学习与 Metropolis 准则求解 TSP 的遗传算法[J]. 软件导刊, 2023, 22(8): 104-110.
- [15] 柳行青, 赵国帅, 韩素敏. 基于 CSSA-LSTM 的 IGBT 模块退化趋势预测[J]. 电子科技, 2024, 37(8): 60-67.
- [16] 翟兴哲, 李鹏竹, 王会民, 等. 基于 LSTM 算法的火电厂智能辅助脱硝系统开发与工程应用[J]. 工业仪表与自动化装置, 2024(5): 3-8+23.
- [17] 郭金玉, 李涛, 李元. 基于多模型 SVM 的多模态过程故障检测[J]. 沈阳化工大学学报, 2023, 37(6): 533-541.
- [18] 梁盛楠, 刘文博, 李雅芝. 基于改进局部极化准则的多核 SVM 模型[J]. 东北师大学报(自然科学版), 2023, 55(4): 32-38.