

大数据驱动的管道泄漏特征提取 与定位模型：复杂工况下的适应性研究

卫彦超（中航油石化管道有限公司，北京 101300）

摘要：随着全球能源与资源管网规模的持续扩张，管道泄漏检测的精准性与鲁棒性成为保障公共安全与经济运行的关键。针对传统方法在复杂工况（如多相流扰动、压力突变、环境温湿度波动等）下适应性不足的问题，本文提出一种基于大数据驱动的管道泄漏特征提取与定位模型。通过构建多源异构数据融合平台，结合时序特征挖掘、关联模式提取与非线性降维技术，实现复杂场景下泄漏特征的动态捕捉；进一步设计自适应迁移学习框架，突破传统模型对固定工况的依赖，提升模型在不同运行状态下的泛化能力。实验验证表明，该模型在压力波动、介质组分变化等多工况场景下，泄漏定位误差较传统方法降低30%以上，误报率控制在2%以内，为复杂管网的智能监测提供了新思路。

关键词：管道泄漏；大数据驱动；特征提取；复杂工况；迁移学习

中图分类号：TE973.6 文献标识码：A 文章编号：1674-5167（2025）036-0157-03

Big Data-driven Pipeline Leakage Feature Extraction and Localization Model: Adaptive Research under Complex Conditions

Wei Yanchao (China National Aviation Fuel Petrochemical Pipeline Co., Ltd., Beijing 101300, China)

Abstract: With the continuous expansion of global energy and resource pipeline networks, the accuracy and robustness of pipeline leak detection have become critical for ensuring public safety and economic operations. To address the limitations of traditional methods in adapting to complex operating conditions (such as multiphase flow disturbances, pressure fluctuations, and environmental temperature/humidity variations), this paper proposes a big data-driven model for pipeline leak feature extraction and localization. By establishing a multi-source heterogeneous data fusion platform and integrating time-series feature mining, association pattern extraction, and nonlinear dimensionality reduction techniques, the model achieves dynamic capture of leak characteristics in complex scenarios. Furthermore, an adaptive transfer learning framework is designed to overcome the dependency of traditional models on fixed operating conditions, enhancing their generalization capabilities across different operational states. Experimental validation demonstrates that the model reduces leak localization errors by over 30% compared to conventional methods under pressure fluctuations and medium composition changes, while maintaining a false alarm rate below 2%. This provides innovative approaches for intelligent monitoring of complex pipeline networks.

Keywords: pipeline leakage; big data-driven; feature extraction; complex operating conditions; transfer learning

管道作为能源（如油气）、水资源及化工原料的核心运输载体，其安全性直接关系到国家经济命脉与社会稳定。据统计，全球每年因管道泄漏导致的直接经济损失超百亿美元，且泄漏引发的爆炸、污染等次生灾害后果严重。因此，高效、精准的泄漏检测与定位技术是管道运维领域的核心需求。传统泄漏检测方法（如负压波法、流量平衡法、声波检测法等）依赖物理模型与特定工况假设，当管道面临多相流切换、压力骤变、环境温度剧烈波动等复杂工况时，易受噪声干扰与非稳态流动影响，导致检测灵敏度下降甚至失效。针对上述问题，本文聚焦“复杂工况适应性”这一核心，从数据采集-特征工程-模型优化全流程设计大数据驱动的泄漏检测框架。重点解决多源数据的异质性融合、复杂工况下泄漏特征的动态提取，以

及模型跨工况迁移能力提升三大问题，为实际管网的智能监测提供理论支撑与技术参考。

1 复杂工况对泄漏检测的问题

1.1 多相流干扰

多相流是复杂工况下管道运行的常见状态，其内部介质的不均质性会导致压力、流量信号呈现强非线性与随机波动。传统泄漏检测方法依赖稳态流动假设，特征提取多基于均值、方差等静态统计量，或短时傅里叶变换等稳态时频分析，这些方法无法剥离多相流本身的剧烈波动与泄漏信号的差异。

1.2 动态边界条件

泵阀启停、压缩机切换等动态边界操作，会激发管道系统的瞬态压力/流量突变（如泵停机引发的回流冲击、阀门快关的水锤效应）。这类信号的时频特

征（短时间内压力骤降、高频振荡）与泄漏信号高度重叠，而传统模型多基于“泄漏 = 固定模式异常”的预设训练，缺乏对“非泄漏但具相似信号”事件的判别能力^[1]。例如，某长输管道曾因阀门切换的瞬态波动触发误报，而后续真实小泄漏却因信号弱于操作扰动未被检测。

1.3 环境时变性

环境温度、湿度的持续变化会破坏信号传播的稳定性：一方面，管道材料因热胀冷缩发生形变，改变泄漏声波/压力信号的传播路径与反射特性；另一方面，介质物性（如油气黏度、水密度）随温度波动，导致信号传输速度、衰减系数发生变化。传统静态特征模型基于固定环境参数训练（如夏季的声速阈值、常温的流量梯度），当环境变化后，其特征权重无法匹配新信号规律^[2]。

2 大数据驱动的泄漏特征提取与定位模型设计

2.1 多源异构数据融合平台构建

针对复杂工况下数据维度高、类型杂的特点，首先构建多源数据融合平台，如下图1所示：

数据采集层：部署压力、流量、温度、加速度（振动）等传感器，同步采集管道全线时序数据，采样频率覆盖高频（振动信号，10kHz）至低频（压力/流量，1Hz）；

数据预处理层：通过滑动窗口去噪（如中值滤波）、缺失值插补（基于KNN的时空邻近填充）、量纲归一化（Min-Max标准化）等方法，消除传感器误差与传输丢包影响；

数据关联层：基于时间戳与空间位置（GPS定位）建立多参数时空关联矩阵，刻画压力-流量-温度的协同变化规律，为后续特征提取提供全局上下文。

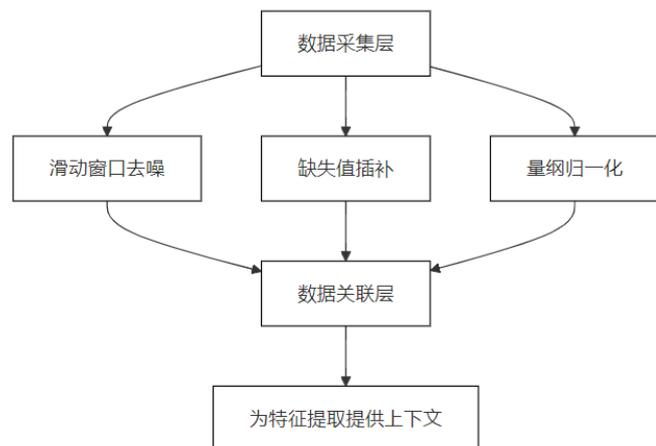


图1 多源异构数据融合平台构建流程

2.2 复杂工况下的泄漏特征提取

复杂工况下，传统泄漏特征提取依赖人工设计的

均方根、峰值因子等经验特征，难以适配多源扰动与非稳态耦合的多样性场景，常因无法捕捉泄漏与正常波动的本质差异导致特征失效。对此，上文提出“分层递进式”特征提取策略，通过多维度特征逐层深化实现对复杂工况的全面分析。

聚焦基础时序特征，从压力、流量、温度等多传感器数据中提取统计特征（均值、方差、偏度）与时频特征（短时傅里叶变换的能量分布、小波包分解的频带能量占比），前者刻画信号整体分布异常，后者捕捉泄漏引发的局部突变，在小相流或稳态波动中可初步定位疑似异常区间^[3]。

拓展关联模式特征，通过互信息量化压力-流量、流量-温度等参数间的耦合强度，构建联合分布直方图，识别泄漏特有的跨参数异常关联。例如，泄漏时压力下降常伴随流量突增但温度无明显变化，此类关联模式能有效区分泄漏与泵阀切换等操作引发的单一参数波动。

第三，引入工况标识特征，基于“稳定输送”“泵阀切换”“多相流”等工况标签训练自编编码器，提取工况判别特征，使模型同步学习不同工况下的信号基线规律，为后续自适应迁移学习提供先验知识。三层特征逐级递进，既覆盖了局部异常的细节，又关联了全局参数的协同变化，更嵌入了工况背景信息，显著提升了复杂工况下泄漏特征的判别力与泛化性。

2.3 自适应迁移学习的泄漏定位模型

为解决跨工况泛化问题，设计基于迁移学习的定位模型：

在典型工况（如稳定输送）下，使用标注的泄漏样本训练基础模型（如LSTM-Attention网络），学习泄漏信号的长程依赖与局部突变模式；当面临新工况（如压力波动）时，冻结基础模型的底层特征提取层（保留通用时序模式），仅微调顶层分类器与定位头，利用少量目标域数据快速适应^[4]；建立工况分类器，根据实时数据判断当前工况类型，动态调用对应工况的子模型或融合多个子模型的输出，进一步提升定位鲁棒性。

3 实验验证与结果分析

3.1 实验数据与场景设置

选取某油气长输管道的实际运行数据（时长6个月），包含3类典型工况：①稳定输送（占60%）；②压力波动（泵阀频繁启停，占25%）；③多相流（气液比1:3~1:8交替，占15%）。人工标注泄漏事件127例，涵盖小泄漏（<1%流量损失）、中泄漏（1%~5%）、大泄漏（>5%）三种级别。下表1统计了油气长输管道的运行数据，包括不同工况下的泄漏事件及其级别。

表 1 油气长输管道的运行数据

工况类型	占比	总泄漏事件数量	泄漏级别	流量损失范围	泄漏事件详情 (例数)
稳定输送	60%	76	小泄漏	<1%	62
			中泄漏	1%-5%	12
			大泄漏	>5%	2
压力波动	25%	32	小泄漏	<1%	25
			中泄漏	1%-5%	5
			大泄漏	>5%	2
多相流	15%	19	小泄漏	<1%	15
			中泄漏	1%-5%	3
			大泄漏	>5%	1
合计	100%	127	-	-	127

3.2 性能评估指标

采用定位绝对误差 (m)、误报率 (%)、工况切换适应时间 (s) 作为核心指标, 对比传统方法 (如基于流量平衡的阈值法) 与单一深度学习模型 (如 LSTM) [5-6]。

3.3 实验结果

3.3.1 定位精度

本文模型在稳定工况下误差为 8.2m, 压力波动工况下误差 9.5m, 多相流工况下误差 10.1m; 传统方法误差分别为 15.7m、28.3m、35.6m, 单一 LSTM 模型误差为 12.4m、22.1m、29.8m。不同工况下定位精度误差对比如下图 2 所示。

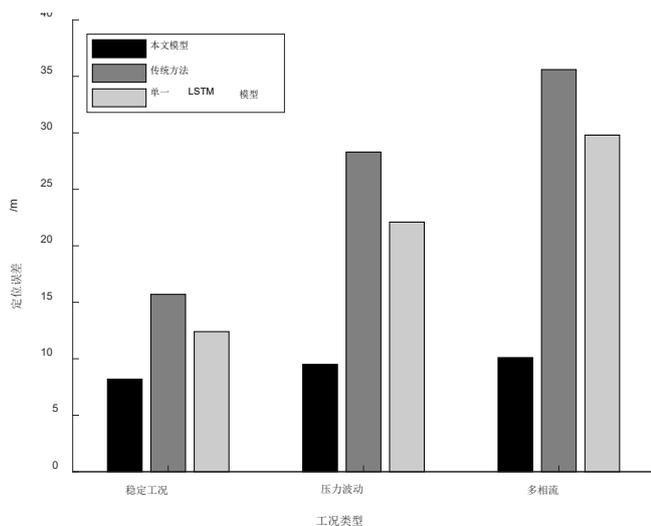


图 2 不同工况下定位精度误差对比

3.3.2 误报率

复杂工况下 (同时存在压力波动与多相流), 本文模型误报率为 1.8%, 传统方法升至 12.3%, 单一模型升至 9.7%。

3.3.3 适应时间

当工况切换时 (如从稳定输送转为压力波动),

模型仅需 3.2s 完成参数微调, 而重新训练单一模型需 2.1h。

4 结束语

综上所述, 本文所提出的大数据驱动模型通过多源数据融合、分层特征提取与自适应迁移学习, 显著提升了复杂工况下的泄漏检测适应性。然而, 实际应用中仍面临以下挑战: ①极端罕见工况 (如地震引发的管道断裂) 数据稀缺, 模型泛化能力受限; ②实时性要求高的场景下, 边缘端部署的轻量化设计需进一步优化。未来可结合数字孪生技术, 构建虚拟仿真环境生成稀缺工况数据, 并探索联邦学习框架, 在保护数据隐私的前提下提升模型全局适应性。针对复杂工况下管道泄漏检测的适应性问题, 本文构建了大数据驱动的特征提取与定位模型。通过多源数据融合、分层特征挖掘与自适应迁移学习, 模型在压力波动、多相流等场景下表现出更强的鲁棒性, 定位精度与误报率指标显著优于传统方法。研究成果为长输管道、市政管网等复杂系统的智能安全监测提供了可行方案, 具有重要的工程应用价值。

参考文献:

- [1] 谢欣彤, 等. 基于大数据分析的天然气管道泄漏预测与风险评估模型 [J]. 化工管理, 2025,(26):68-70+74.
- [2] 程可心, 等. 数据驱动背景下的油气管道风险评价挑战 [J]. 管道保护, 2025,2(01):88-96.
- [3] 左忠霖. 多工况下无泄漏数据驱动的在役天然气管道泄漏检测方法 [D]. 重庆大学, 2023.
- [4] 马大中, 等. 基于数据驱动的管道云边协同泄漏检测方法 [J]. 控制与决策, 2023,38(08):2415-2424.
- [5] 郭颖. 基于数据驱动的管道泄漏检测与定位技术研究 [D]. 沈阳工业大学, 2021.
- [6] 任程泽. 基于数据驱动的管网缺失数据补全及泄漏故障诊断方法研究 [D]. 东北大学, 2020.