



中国科技期刊卓越行动计划项目入选期刊

控制与决策

CONTROL AND DECISION



融合分层分块信息的轧制过程运行状态评估方法及应用

张永博, 张凯, 彭开香, 杨朋澄

引用本文:

张永博, 张凯, 彭开香, 杨朋澄. 融合分层分块信息的轧制过程运行状态评估方法及应用[J]. 控制与决策, 2024, 39(8): 2694–2702.

在线阅读 View online: <https://doi.org/10.13195/j.kzyjc.2023.0089>

您可能感兴趣的其他文章

Articles you may be interested in

工业信息物理系统安全风险动态表现分析量化评估模型

Quantitative evaluation model for dynamic performance analysis of security risk in industrial cyber physics systems

控制与决策. 2021, 36(8): 1939–1946 <https://doi.org/10.13195/j.kzyjc.2019.1479>

基于多块信息提取的AUV资源勘查系统故障检测

Fault detection of AUV resource exploration system based on multi-block information extraction

控制与决策. 2021, 36(4): 790–800 <https://doi.org/10.13195/j.kzyjc.2019.0732>

基于核典型相关性-熵成分分析的工业过程质量监测方法

A quality monitoring method for industrial process based on kernel canonical correlation-entropy component analysis

控制与决策. 2021, 36(12): 2999–3006 <https://doi.org/10.13195/j.kzyjc.2020.0548>

基于多层次特征的机械臂单阶段抓取位姿检测

Single-stage grasp pose detection of manipulator based on multi-level features

控制与决策. 2021, 36(8): 1815–1824 <https://doi.org/10.13195/j.kzyjc.2019.1840>

基于无标签、不均衡、初值不确定数据的设备健康评估方法

Equipment health risk assessment based on unlabeled, unbalanced data under uncertain initial condition

控制与决策. 2020, 35(11): 2687–2695 <https://doi.org/10.13195/j.kzyjc.2018.1493>

融合分层分块信息的轧制过程运行状态评估方法及应用

张永博¹, 张 凯^{1,2†}, 彭开香^{1,2}, 杨朋澄¹

(1. 北京科技大学 自动化学院, 北京 100083; 2. 工业过程知识自动化教育部重点实验室, 北京 100083)

摘 要: 工业过程运行状态评估方法是对过程当前运行状态进行合理的评价, 为工业过程的安全、高效运行提供有益的指导. 带钢轧制过程具有流程长且系统层级多等特点, 而传统的运行状态评估往往采用集中式的评估方法, 难以对轧制过程全流程的运行状态进行合理的评估. 针对此问题, 提出一种融合分层分块信息的轧制过程运行状态评估方法. 采用一种多层级分块的评估策略, 将全流程分为若干个层级和子块, 提高评估结果的可解释性. 提出一种联合核主成分分析 (kernel principal component analysis, KPCA) 和 t-分布随机邻域嵌入算法 (t-distributed stochastic neighbor embedding, t-SNE) 的特征提取方法, 并行地提取全局和局部的特征信息. 进而, 针对传统支持向量机 (support vector machine, SVM) 输出结果为硬判型输出, 将 SVM 的输出结果映射为后验概率, 并通过 D-S 证据理论融合多个层级的运行状态评估结果, 从而实现决策层面的信息融合, 提高评估结果的准确性. 最后, 将所提出方法应用于实际带钢轧制过程, 与各类传统的方法相比评估准确率提高近 18%.

关键词: 流程工业; 运行状态评估; 特征提取; D-S 证据理论; 热连轧过程

中图分类号: TP273

文献标志码: A

DOI: 10.13195/j.kzyjc.2023.0089

引用格式: 张永博, 张凯, 彭开香, 等. 融合分层分块信息的轧制过程运行状态评估方法及应用 [J]. 控制与决策, 2024, 39(8): 2694-2702.

An operating performance assessment method for tolling processes via integrating hierarchical and block information

ZHANG Yong-bo¹, ZHANG Kai^{1,2†}, PENG Kai-xiang^{1,2}, YANG Peng-cheng¹

(1. School of Automation, University of Science and Technology Beijing, Beijing 100083, China; 2. Key Laboratory of Knowledge Automation for Industrial Processes of Ministry of Education, Beijing 100083, China)

Abstract: Operating performance assessment methods are performed to evaluate the operating performance of industrial processes so as to provide helpful guidance for safe and high-efficient operations. The rolling processes are always featured with plant-wide characteristics and equipped with hierarchical information systems, while traditional assessment methods are centralized, which, thus, are less efficient. To solve this problem, this paper proposes a method that integrates hierarchical and block information. Firstly, a hierarchical strategy is adopted, and the entire process are partitioned into several levels and sub-blocks such that the interpretability of the assessment results can be improved. Then, in order to extract comprehensive operating state features, a combination of kernel principal component analysis and t-distributed stochastic neighbor embedding-based method is proposed, which can capture both global and local features in a parallel manner. In addition, to overcome the hard decision-making results caused by support vector machines (SVMs), this paper transfers the results of the SVM to calculate the posterior probability, and makes use of the D-S evidence theory to fuse the decision-making process. Finally, the proposed method is applied to a real hot rolling mill process, the results show that it improves at least 18% assessment rate compared with several traditional methods.

Keywords: process industry; operating performance assessment; feature extraction; D-S evidence theory; hot rolling mill process

0 引 言

随着流程工业不断自动化和复杂化, 为了确保产品质量和企业的经济效益, 需要对实际生产过程进行

实时监控. 同时, 在实际生产过程中, 过程的运行状态不一定总是处于正常状态, 由于外界的扰动以及内部参数的变化, 运行状态容易偏离正常状态进入故障状

收稿日期: 2023-01-28; 录用日期: 2023-06-16.

基金项目: 国家自然科学基金面上项目 (62073032); 国家自然科学基金区域联合重点项目 (U21A20483).

责任编委: 方华京.

[†]通讯作者. E-mail: kaizhang@ustb.edu.cn.

*本文附带电子附录文件, 可登录本刊官网该文“资源附件”区自行下载阅览.

态。然而,传统的过程监测方法往往只考虑正常和故障两种状态,不能对实际生产过程的运行状态进行多等级评价,这种简单的方法已经不能满足现在绿色生产和高效生产的需求。所以,研究稳定而详细的运行状态评价方法对复杂流程工业具有十分重要的应用价值。带钢热轧生产过程是典型的流程工业,不同子系统之间相互耦合,关系复杂。因此,为了准确地评价带钢热轧生产过程的运行状态,需要根据其具体的工艺特点以及数据特性建立有针对性的运行状态评价方法。

目前,针对复杂流程工业的运行状态评估,相关学者已经进行了一定的研究。郑伟等^[1]考虑到各子系统之间的相互关系,提出了基于分层分块状态相关独立主成分分析的评估方法,并在火力发电过程中验证了该方法的可行性。姚远等^[2]针对工业过程动态性及非线性强等特点,提出一种基于动态局部保持主成分分析法的过程监测方法,通过构造扩展矩阵解决动态过程中各采样点间相关性强的问题,并将局部保持投影与主成分分析法相结合从而实现提取流形结构的最大方差信息。刘炎等^[3]开发了一种基于Fisher判别分析的状态评估方法,将其应用于湿法冶金过程并证明了其科学性。常玉清等^[4]利用互信息和加权投票机制改进了随机森林算法,将其应用于湿法冶金浸出过程并提高了评估的准确率。针对动态过程状态评估,Wang等^[5]将慢特征分析与高斯过程回归相结合实现了对动态过程的状态评估,将其应用于乙烯装置的脱硝器和裂解炉中并获得了较高的精度。为了获得更加全面的工况信息,褚菲等^[6]提出了基于关键性能指标的慢特征分析算法,实现了对重介质选煤过程的静态和动态特性协同感知。为了实现静态和动态的并行评估,Zou等^[7]将典型变量分析与慢特征分析结合在一起,提出了静态和动态并行评估方法,将其应用于三相流过程并获得了较好的评估准确率。针对烧结过程的碳效率评价问题,Zhou等^[8]提出基于神经网络和模糊综合评价的评估方法并获得了更高的准确性。针对钻井作业性能多模式的难评估问题,Fan等^[9]提出了一种基于最小二乘支持向量机的钻井作业性能多模式评估策略,通过与其他方法相比较证明了其优越性。针对部分工业过程故障样本不易获取以及样本数量不足的问题,Lu等^[10]提出了一种半监督聚类生成对抗网络,该方法有效利用未标记样本扩展标记数据集,应用于金浮选过程并获得了准确的评价结果。Zhong等^[11]将主成分分析与典型变量分析结合起来充分挖掘黄金湿

法冶炼过程中的数据信息,提高了检测非最优状态的能力。Wang等^[12]提出了基于实时学习和概率主成分回归的预测模型,并在工业加氢裂化过程中验证了该方法的有效性。针对工业过程普遍存在的非线性和不确定性,褚菲等^[13]将综合经济指标与稀疏降噪自编码器相结合,在误差项中引入综合经济指标,并成功地将其应用于重介质选煤过程中。针对现代流程工业具有动态、非线性以及大范围的特点,Zhang等^[14]提出了一种基于分布式改进最小冗余最大相关性和输出相关核共同趋势,分析关键性能指标相关运行状态的评估方法,并以热连轧过程为例验证了该方法的有效性。考虑到实际工业数据往往受到噪音污染以及非线性特点,刘炎等^[15]利用堆叠有监督降噪自编码器以及softmax模型实现了对湿法冶金过程的运行状态等级评价。在电凝净化过程中,存在反应器电极板钝化的问题,工艺运行状态不断变化,严重影响了电凝净化的效率和重金属离子的去除。Zhu等^[16]将长短期记忆网络与自回归综合移动平均模型相结合,得到了电凝净化过程中重金属离子去除率的预测值,并将预测值和影响净化效果的其他重要因素作为电凝净化工艺运行状态评价指标,根据模糊规则对运行状态评价等级进行划分。王印松等^[17]提出一种基于证据融合的诊断算法,首先利用基于信号分析的方法计算表征故障特征的指标,针对“一票否决”现象对指标结果加以改进;然后采用D-S证据理论融合基于最小二乘支持向量机的概率分类特征实现优势互补。针对批处理过程多个批次间存在的过渡区域识别与评估困难的问题,Li等^[18]提出了一种从数据中自动准确地检测过渡状态的方法,根据过渡态的特点系统地提出了若干特殊的指标,并以加氢裂化工业过程为例验证了方法的有效性。为了实现全生命周期的运行状态评估,Zhang等^[19]提出了一种基于鲁棒核典型变量分析的全生命周期运行性能评估框架,并在热连轧过程中实现了对全生命周期的运行状态评价。

综合以上研究进展,本文提出一种融合分层分块信息的轧制过程运行状态评估策略,具体贡献如下:1)考虑到热连轧过程具有长流程以及多系统多层级等特点,提出一种多层级分块的运行状态评估方法。从纵向来看,“多层级”是指将整个热连轧过程划分为全流程层、子系统层以及设备层;从横向来看,“多块”是指将子系统层划分为上游子块、中游子块和下游子块。2)针对过程数据具有非线性等特点,同时为了提取全面的运行状态信息,提出

一种联合核主成分分析(kernel principal component analysis, KPCA)和t-分布随机邻域嵌入算法(t-distributed stochastic neighbor embedding, t-SNE)的特征提取方法,并行地开展全局和局部信息的特征提取工作,提高特征提取模型的可靠性. 3)传统的支持向量机(support vector machine, SVM)的输出结果为硬判型输出,本文将SVM的输出结果映射为后验概率,并通过D-S证据理论融合多个角度的运行状态评估结果,从而实现决策层面的信息融合. 4)将所提出方法应用于实际热连轧生产过程数据,验证该方法的科学性和有效性.

1 基本原理与简介

本节主要介绍所提出运行状态评估方法使用的特征提取方法、分类评估方法以及信息融合方法.

1.1 核主成分分析算法

KPCA的目标是实现数据集整体方差的最大化^[20],即提取数据集的全局结构特征信息.通过非线性映射函数将过程数据集映射到高维空间中,再采用主成分分析算法在高维空间中提取主成分,选择贡献率较大的主成分作为全局结构特征信息.利用核函数的思想表示非线性映射函数,考虑到高斯核函数相较于其他核函数的复杂度低且效果佳,本文采用高斯核函数.

1.2 t-分布随机邻域嵌入算法

t-SNE的目标是保持数据集的局部邻域结构^[21],即提取数据集的局部结构特征信息.它是一种非线性流形学习算法,是由SNE算法演变而来,在将高维数据映射到低维空间时改变了基于距离不变的思想,最大程度保证数据之间的分布概率不变,有效解决SNE算法中数据拥挤的问题.详细计算步骤见文献[22].

1.3 支持向量机分类评估

SVM是一种机器学习的方法^[23],基于VC维理论核结构化风险最小原则提出.针对非线性、小样本的数据,SVM具有良好的学习能力,其基本思想是在样本空间内找到最优超平面,该最优超平面可以使二分类样本的分隔间距最大化.

1.4 D-S证据理论

D-S证据理论是一套处理多证据融合问题的数学方法,主要用于描述对某一事件的信任程度,目前广泛应用于目标识别、故障诊断等领域.假设研究对象全体集合为辨识框架 Θ ,其中 2^Θ 为 Θ 的幂集, A 为 2^Θ 的任意子集.映射 $m: 2^\Theta \rightarrow [0, 1]$ 满足如下性质:

$$\begin{cases} m(\emptyset) = 0, \\ m(A) \geq 0, \\ \sum_{A \subset 2^\Theta} m(A) = 1. \end{cases} \quad (1)$$

其中 m 为 Θ 的基本概率分配函数(basic probability assignment, BPA), $m(A)$ 为相关证据对事件 A 的支持程度或者信任程度.D-S证据理论的基本思想是将证据集合划分为若干个独立证据,并分别利用它们对辨识框架进行独立判断,然后依据组合规则进行信息融合.假设 m_1 与 m_2 分别为同一辨识框架 Θ 上的2个独立证据,其相应的焦元分别为 $A_1, A_2, \dots, A_i$ 与 $B_1, B_2, \dots, B_j$,则D-S证据组合规则 $m(C)$ 可表述为

$$m(C) = \begin{cases} 0, & C = \emptyset; \\ \frac{\sum_{A_i \cap B_j = C} m_1(A_i)m_2(B_j)}{1 - K}, & C \neq \emptyset. \end{cases} \quad (2)$$

其中 K 为证据间的冲突程度,有

$$K = \sum_{A_i \cap B_j = \emptyset} m_1(A_i)m_2(B_j). \quad (3)$$

上述即为D-S证据组合规则,它满足交换律和结合律.

2 基于D-S证据理论的分层分块运行状态评估方法

考虑到复杂工业过程存在长流程以及多系统多层次等特点,提出一种融合分层分块信息的轧制过程运行状态评估方法.过多的运行状态等级会导致运行状态评估问题复杂化,且过于细腻的状态等级不利于运行状态信息的提取,同时过少的运行状态等级也不利于流程工业过程的精细化管理,所以对于复杂工业过程的运行状态等级的划分,需要以实际生产过程的运行情况以及工艺知识为参考,并结合实际的运行管理需求.本文以热轧过程为背景,考虑到实际热连轧过程的过程特性以及精细化管理需求,并结合生产工艺知识,将过程的运行状态分为4个等级,分别是“优”“良”“中”“差”,对应“1~4”四个运行状态等级.

2.1 联合KPCA和t-SNE的全局和局部信息特征提取

热连轧过程运行状态评估模型的离线建模过程包括:特征提取、各子块的运行状态评估、全流程运行状态评估3个部分.

KPCA是一种面向全局的特征提取方法,通过非线性映射使低维空间中线性不可分的原始数据在高维空间变得线性可分.这一过程是在最大化高维数

据全局特征方差的目标下,提取出原始数据的全局非线性特征. t-SNE 是一种非线性流形学习算法,它将高维空间数据点之间的相似性以概率形式表示,并假设数据点在高维空间服从高斯分布,在低维空间服从 t 分布. 从高维到低维的降维过程, t-SNE 能够充分挖掘嵌入在高维空间中的低维局部流形结构,获取数据的局部结构特征信息. 因此,综合二者的各自优势,本文充分考虑了样本数据的全局和局部结构特征,为了能够使 KPCA 和 t-SNE 算法在特征提取过程中实现优势互补,从全局和局部结构两个角度并行地进行运行状态信息特征提取工作.

2.2 各子块的运行状态评估

SVM 是近年来发展起来的基于统计学习理论和结构风险最小化原则的新型学习机,传统的 SVM 只能实现样本的二分类,且 SVM 的输出结果为硬判型输出. 为了与 D-S 证据理论相结合,需要在融合前将 SVM 输出映射为后验概率,利用逐对耦合法和一对一多分类 SVM 方法实现由二分类 SVM 到多分类概率型 SVM 的转换. 利用 2.1 节计算得到的全局结构运行状态特征和局部结构运行状态特征训练多分类概率型 SVM 评估器,最后通过求解下式中的优化问题计算出各类的后验概率 $p_i(y = i|x)$:

$$\begin{aligned} \min_p \quad & \frac{1}{2} \sum_{i=1}^h \sum_{j=1, j \neq i}^h (r_{ji}p_i - r_{ij}p_j)^2; \\ \text{s.t.} \quad & \sum_{i=1}^h p_i = 1, p_i \geq 0, \forall i. \end{aligned} \quad (4)$$

其中: $i, j = 1, 2, \dots, h$, h 为类别数; r_{ij} 为第 i 类和第 j 类匹配时 x 属于 i 类的后验概率.

D-S 证据理论的基本思想是通过合成规则将来自多个证据体的信任函数融合为一个新的信任函数,并以此作为决策依据. 构造客观合理的 BPA 是 D-S 证据理论应用中至关重要的一步,为了保证 BPA 赋值的客观性,需要对各个 $\text{SVM}_i (i = 1, 2, \dots, 2n)$ 评估结果进行评价. 混淆矩阵能够描述样本真实属性与评估结果之间的关系,可用于分类器性能评价. 对于 4 种运行状态评估问题, SVM_i 的混淆矩阵表示为

$$\text{CM}_i = \begin{bmatrix} c_{11} & c_{12} & \dots & c_{1h} \\ c_{21} & c_{22} & \dots & c_{2h} \\ \vdots & \vdots & c_{pq} & \vdots \\ c_{h1} & c_{h2} & \dots & c_{hh} \end{bmatrix}. \quad (5)$$

其中: c_{pq} 为真实属性为 p 的运行状态被分类器 SVM_i 判别为第 q 种运行状态的数目占 p 类运行状态样本总数的百分比; 元素的行下标表示运行状态的真实

类型,列下标表示分类器的识别类型.

定义 SVM_i 的分类平均正确率为

$$\bar{\gamma}_i = \sum_{p=1}^h (c_{pp}N_p)/N. \quad (6)$$

其中: $N_p = n$, $N = hn$, 分别为第 p 类运行状态的样本总数和训练样本总数; γ_i 为对分类器 SVM_i 判决结果的一种全局可信度.

经过训练得到的混淆矩阵 CM_i 其行向量为 $[c_{11}, c_{12}, \dots, c_{1p}, \dots, c_{1h}]^T, p = 1, 2, \dots, h$, 其中 c_p 为第 p 类运行状态数据样本进行分离时对各运行状态模式的倾向性. 同样地, 混淆矩阵 CM_i 的列向量为 $[c_{11}, c_{21}, \dots, c_{q1}, \dots, c_{h1}]^T, q = 1, 2, \dots, h$, 其中 c_q 能够反映出分类器对第 q 类运行状态的局部可信度.

定义 SVM_i 的局部可信度为

$$\omega_{iq} = c_{qq} / \sum_{p=1}^h c_{pq}, \quad (7)$$

其中 $q = 1, 2, \dots, h$. 将其融入 SVM_i 的概率型判决输出 p_{iq} 并归一化可得

$$p'_{iq} = \omega_{iq}p_{iq} / \sum_{q=1}^h \omega_{iq}p_{iq}. \quad (8)$$

由此,可定义其 BPA 为

$$\begin{aligned} m_i(S_1, S_2, \dots, S_h, \Theta, \phi) = \\ (\bar{\gamma}_i p'_{i1}, \bar{\gamma}_i p'_{i2}, \dots, \bar{\gamma}_i p'_{ih}, 1 - \bar{\gamma}_i, 0). \end{aligned} \quad (9)$$

由式(9)可知,识别框架 $\Theta = (S_1, S_2, \dots, S_h, \Theta, \phi)$, 其中 S_h 为带钢热轧过程的第 h 种运行状态. 同时满足 $m_i(S_1) + m_i(S_2) + \dots + m_i(S_h) + m_i(\Theta) = 1, m_i(\phi) = 0$, 符合 BPA 的定义要求.

2.3 全流程运行状态评估

首先,从实际工业背景出发,现代流程工业本身具有大范围、多层级等特性,需要根据工艺特点将整个长流程分解为若干个短流程(子块)以达到监测局部行为的目的,从而实现精细化管控;然后,从模型的计算复杂度出发,分层分块的策略有利于降低模型尺度,减少计算的复杂程度,获得更好的评估性能,以上解释了分层分块的必要性. 在进行全流程运行状态评估时必须结合各子块的运行状态等级信息,所以要将所有子块的信息融合起来为全流程的评估提供重要的基础,但是每个子块对全流程影响的程度又不尽相同,需要各自赋予权重系数.

利用上一个阶段获得的各子系统的运行状态评估结果,根据实际带钢热轧生产过程的工艺知识进行全流程运行状态评估,全流程的运行状态评估可以通过对各子块的运行状态评估概率化结果进行加权计

算来确定,全流程的运行状态等级 S_T 表示为

$$S_T = \text{round}\left(\sum_B \alpha_B S_B\right). \quad (10)$$

其中: $\alpha_B \left(\sum_B \alpha_B = 1\right)$ 为每个子系统的权重系数; S_B 为每个子块的运行状态等级, $B = 1, 2, \dots, n$; $\text{round}(\cdot)$ 为一个四舍五入函数.

2.4 方法总结

本文所提出方法流程如图1所示. 首先,通过工业过程异常状态不同等级的数据来离线训练运行状态评估模型;然后,利用训练好的评估模型以及相应全局可信度和局部可信度进行在线运行状态等级评估. 所提出方法详细步骤如下.

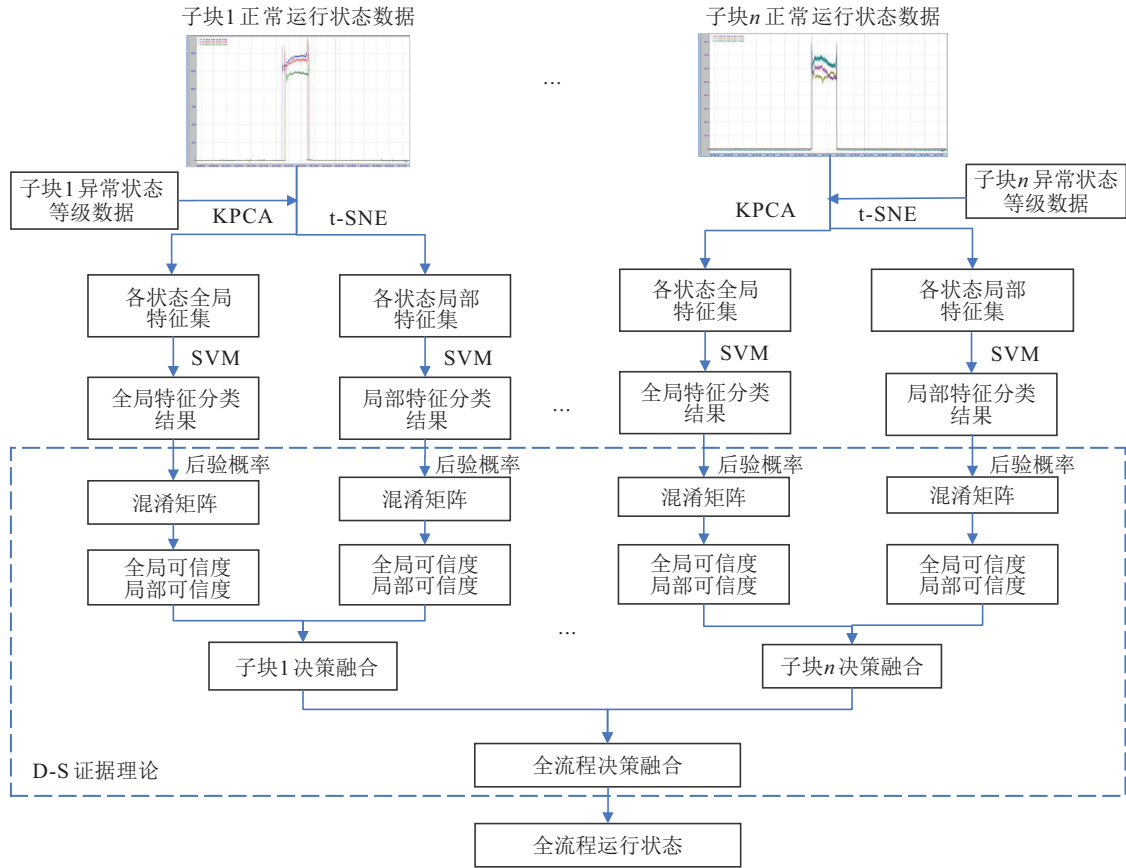


图1 融合分层分块信息的轧制过程运行状态评估方法流程

1) 离线建模.

① 对异常状态不同等级的数据 X 进行标准化处理;

② 结合工艺特点和专家知识,对全流程进行分层分块,得到 $x_1, x_2, \dots, x_n$ (n 为子块总数);

③ 通过 KPCA 算法和 t-SNE 算法分别提取 $x_1, x_2, \dots, x_n$ 的全局特征和局部特征;

④ 将③提取到的全局和局部特征作为 SVM_i ($i = 1, 2, \dots, 2n$) 的训练输入向量,输出为相应的混淆矩阵 CM_i ($i = 1, 2, \dots, 2n$);

⑤ 由式(5)~(7)计算得到各混淆矩阵的全局可信度 γ_i ($i = 1, 2, \dots, 4n$) 和局部可信度 ω_i ($i = 1, 2, \dots, 4n$).

2) 在线评估.

① 采集得到待评估的数据 $\tilde{X}$ 并进行标准化处理;

② 利用已知的专家和工艺知识对标准化后的数据进行分层分块,分块后得到 $\tilde{x}_1, \tilde{x}_2, \dots, \tilde{x}_n$ (n 为子块的总数);

③ 联合 KPCA 算法和 t-SNE 算法分别对 $\tilde{x}_1, \tilde{x}_2, \dots, \tilde{x}_n$ 进行全局和局部特征的并行提取;

④ 利用已经训练好的 SVM_i ($i = 1, 2, \dots, 2n$) 模型分别对相应的全局和局部特征进行预测评估,并结合式(4)得到对应的概率型判决输出 p_i ($i = 1, 2, \dots, 2n$);

⑤ 通过式(8)和(9)分别将后验概率 p_i ($i = 1, 2, \dots, 2n$)、全局可信度 γ_i ($i = 1, 2, \dots, 4n$) 与局部可信度 ω_i ($i = 1, 2, \dots, 4n$) 结合起来得到其 BPA_i ($i = 1, 2, \dots, 2n$);

⑥ 基于式(2)和(3)分别对各 BPA_i ($i = 1, 2, \dots, 2n$) 进行 D-S 证据理论融合,得到各子块的运行状态的概率化评估结果 S_B ($B = 1, 2, \dots, n$),从而实

现对各子块的运行状态评估;

⑦ 根据各子块对全流程重要性的不同对每个子块分别赋予不同的权重系数,并通过式(10)计算得到全流程运行状态的概率化评估结果 S_T ,从而实现对全流程的运行状态评估.

3 实际带钢轧制过程应用验证

3.1 运行状态评估模型的建立

采用某钢铁公司1 700 mm带钢热连轧生产线现场采集的大量反映实际生产过程的数据验证所提出方法的有效性和实用性. 实际生产过程的数据包括过程变量和带钢产品的综合经济指标数据,其中综合经济指标数据主要包括出口厚度、平整度、凸度、力学性能以及设备耗能指标等,此类指标数据均可作为

评判运行状态优劣的参考依据. 结合实际带钢热轧生产过程对参数获取的难易程度,选择最具有代表性的出口厚度以及受影响变量的个数作为评判带钢热轧生产过程运行状态等级的参考指标.

针对实际带钢热轧生产过程,为了验证所提出方法在实际生产过程中的有效性和科学性,针对由钢厂采集的带钢热轧生产过程数据,共选取1 200组过程数据. 随机抽取600组数据作为训练数据,剩余600组数据作为测试数据,其中训练数据集“优、良、中、差”4个状态等级的数据均为150组. 考虑工厂的实际生产需求并结合专家知识,按照偏离正常厚度值的范围(即厚度差,TD)以及受影响变量的数目(NUM)将运行状态划分为4个状态等级,具体如表1所示.

表1 运行状态等级划分及故障原因

厚度差	受影响变量个数	运行状态等级	故障原因
$0\text{ mm} \leq \text{TD} \leq 0.005\text{ mm}$	$\text{NUM} = 0$	优	无故障
$0.005\text{ mm} < \text{TD} \leq 0.01\text{ mm}$	$3 < \text{NUM} \leq 6$	良	第3机架发生弯辊力故障
$0.01\text{ mm} < \text{TD} \leq 0.015\text{ mm}$	$6 < \text{NUM} \leq 9$	中	第3机架发生轧制力故障
$0.015\text{ mm} < \text{TD} \leq 0.02\text{ mm}$	$\text{NUM} > 9$	差	第3机架冷却水阀发生堵塞

利用前600组数据以及KPCA和t-SNE模型构造 SVM_i 的训练向量. 以 SVM_1 为例,其混淆矩阵为

$$\text{CM}_1 = \begin{bmatrix} 150 & 0 & 0 & 0 \\ 0 & 150 & 72 & 0 \\ 0 & 0 & 78 & 0 \\ 0 & 0 & 0 & 150 \end{bmatrix}.$$

按照式(6)和(7)的可信度计算方法,可以计算出 SVM_1 的全局可信度为0.880,局部可信度为1.000、0.676、1.000、1.000. 其余 SVM_i 的全局可信度和局部可信度在表2中给出.

表2 SVM_i 全局可信度和局部可信度

评估器	$\bar{\gamma}_0$	ω_{i0}	ω_{i1}	ω_{i2}	ω_{i3}
SVM_1	0.880	1.000	0.676	1.000	1.000
SVM_2	0.860	1.000	0.920	0.816	0.704
SVM_3	0.927	0.898	0.831	1.000	1.000
SVM_4	0.878	1.000	1.000	0.673	1.000
SVM_5	0.812	0.746	1.000	0.708	1.000
SVM_6	0.918	1.000	0.801	0.849	1.000

3.2 在线评估

为了验证所提出方法的实用性,在实际热连轧生产过程数据上进行实验. 其中测试数据集为600组,“优、良、中、差”4个状态等级的数据均为150组. 将各子块不同状态等级的数据送入对应的 SVM 中进行评估,获得概率化的评估结果结合表2中各 SVM 的全局可信度和局部可信度,进一步得到各采样点的BPA赋值,并结合D-S证据理论实现对各子系统的运行状态评估. 热连轧过程通常由7个机架串联

而成,一旦其中一个机架处于异常运行状态,便会影响后续机架,甚至影响最终的产品质量,给企业带来经济损失. 由于采用了自动控制系统,上游机架的异常可以通过后续工序的调整进行补偿. 然而,当中游或下游的机架发生异常时很难进行补偿,同时从中游机架和下游机架分布的空间位置看,下游机架位是整个流程工业的最后一个工序,所以下游机架的运行状态等级对于最终的产品质量影响是最大的. 换言之,在全流程中,下游机架是最重要的,所以给下游机架赋予的权重是最大的,中游机架的权重次之,上游机架权重最小. 综上所述,在进行全流程运行状态评估时,对上游子块、中游子块和下游子块分别赋予0.2、0.3、0.5的权重系数,最后利用各子系统的评估结果并结合权重值实现对全流程的运行状态评估.

采用融合全局和局部信息的运行状态评估方法对600组测试集样本进行评估,先根据带钢热轧生产过程的工艺知识将整个机架组分为上游、中游、下游3个部分,即3个子块,同时也形成了3个层级,分别为

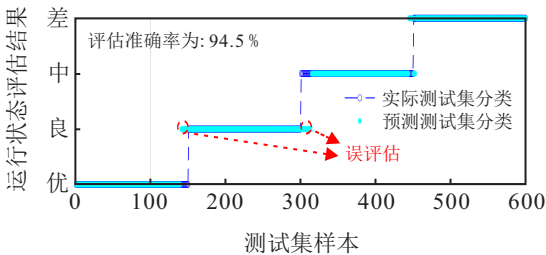


图2 基于D-S证据理论融合各个子块运行状态信息的全流程运行状态评估结果

设备层、子系统层和全流程层,这样便实现了对带钢热轧生产过程的分层分块运行状态评估,详细的评估结果如图2所示。

3.3 对比分析

3.3.1 其他方法对比分析

分别验证基于KPCA+SVM、基于t-SNE+SVM和基于AdaBoost的分层分块运行状态评估方法。通过比较不同方法之间的评估准确率可以发现:上述3种方法的评估准确率均低于第3.2节基于D-S证据理论的分层分块运行状态评估方法的评估准确率。实验结果表明:相较于基于单一特征的评估方法,采用特征融合的评估方法确实能够提高评估的准确率。上述3种方法对全流程详细的运行状态评估结果如图3~图5所示。

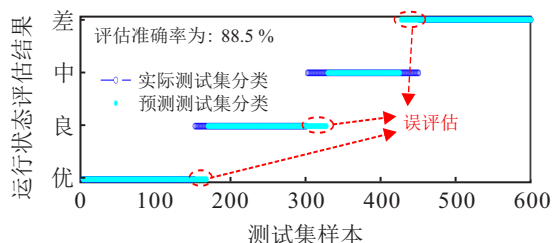


图3 基于KPCA+SVM融合各个子块运行状态信息的全流程运行状态评估结果

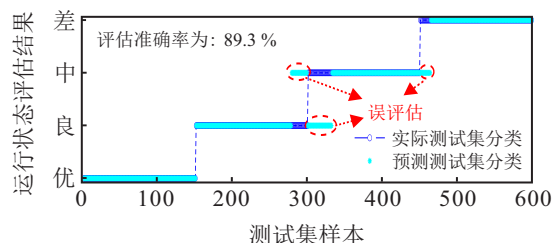


图4 基于t-SNE+SVM融合各个子块运行状态信息的全流程运行状态评估结果

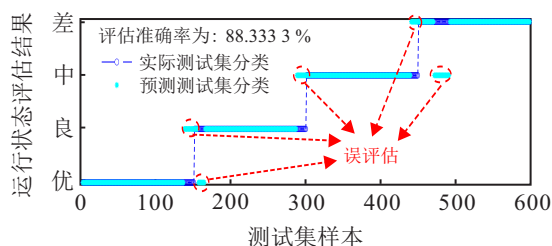


图5 基于AdaBoost融合各个子块运行状态信息的全流程运行状态评估结果

相较于第3.2节的基于D-S证据理论融合全局和局部信息的分层分块运行状态评估方法,本节所对比方法只融合了数据的全局信息和局部信息,并没有采用分层分块的策略。通过对比评估准确率不难发现:所对比方法的评估准确率低于基于D-S证据理论融合全局和局部信息的分层分块运行状态评估方法的评估准确率,实验结果表明了对全流程进行分层分块

的必要性。所对比方法对全流程详细的运行状态评估图如图6所示。

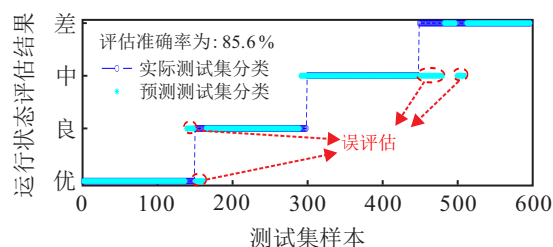


图6 基于D-S证据理论融合全局和局部信息的未分层未分块全流程运行状态评估结果

最后,验证基于KPCA+SVM、基于t-SNE+SVM以及基于AdaBoost的未分层未分块运行状态评估方法。将3种方法与上面已经验证过的评估方法进行比较,可以得出由于这3种方法既没有采用特征融合策略又没有采用分层分块策略,对全流程的评估准确率都是最低的,该实验结果进一步阐明了针对工业过程的运行状态评估问题采用特征融合方法以及分层分块策略的必要性。上述3种方法对全流程详细的运行状态评估图如图7~图9所示。

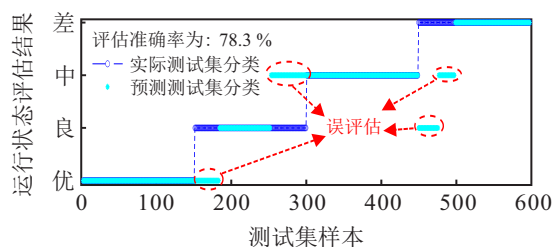


图7 基于KPCA+SVM未分层未分块运行状态评估结果

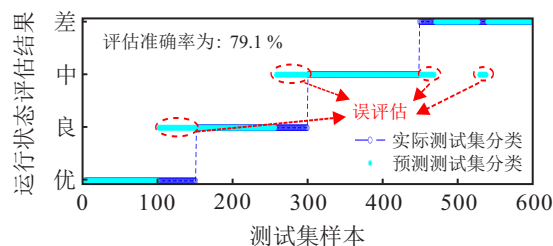


图8 基于t-SNE+SVM未分层未分块运行状态评估结果

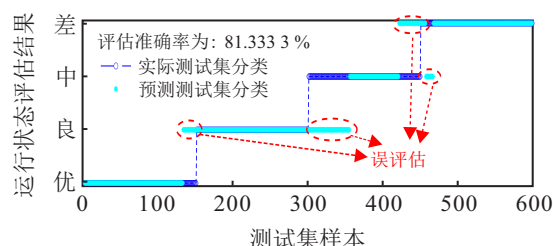


图9 基于AdaBoost未分层未分块运行状态评估结果

3.3.2 统计对比分析

评估准确率等于预测正确的样本点数目与全部待预测样本点数目之商。为了保证评估准确率的科学性和可信性,另外选取3组测试集样本,每组600个样本,测试集样本通过热连轧过程故障注入系统获

得.热连轧过程故障注入系统是通过正常工况下的过程数据结合工艺机理模型注入不同严重程度的故障,从而获得不同运行状态等级的数据.整个故障注入系统提供了故障的开始时间和结束时间、故障类型、故障大小、故障位置等功能区.重复第3.2节和第3.3.1节的实验步骤,计算各种运行状态评估方法准确率的平均值.表3为6种方法对全流程层运行状态评估准确率的平均值.

表3 全流程层运行状态评估准确率的平均值 %

	序号	评估方法	评估准确率
未分层未分块	1	KPCA+SVM	0.773
	2	t-SNE+SVM	0.782
	3	AdaBoost	0.829
	4	KPCA+t-SNE+SVM	0.872
分层分块	5	KPCA+SVM	0.869
	6	t-SNE+SVM	0.834
	7	AdaBoost	0.893
	8	本文提出的方法	0.951

在子系统层中,针对各个子块运行状态评估,融合全局和局部信息后相较于从单一角度进行运行状态评估具有更高的准确率.在全流程层中,针对全流程的运行状态评估,无论是对比表3中的方法1、方法2、方法4,还是对比方法4、方法5、方法8,融合全局和局部信息后相较于从单一角度进行运行状态评估都具有更高的准确率.并且,对比方法1与方法5、方法2与方法6、方法3与方法7、方法4与方法8,采用分层分块的全流程评估策略相较于未分层未分块的全流程评估策略具有更高的准确率.所提出基于D-S证据理论融合全局和局部信息的分层分块运行状态评估方法的评估准确率为95.1%,而传统的基于KPCA+SVM的运行状态评估方法和基于t-SNE+SVM的运行状态评估方法的评估准确率仅为77.3%和78.2%,相比之下,所提出方法评估准确率提高了近18%.

对比以上提到的运行状态评估方法在子系统层和全流程层的评估准确率可以发现,无论在子系统层还是在全流程层,融合全局和局部信息后的评估准确率均高于从单一角度的评估准确率,表明了进行多维特征信息融合必要性.同时,采用分层分块的评估方法相较于未进行分层分块的评估方法具有更高的评估准确率,进一步验证了采用分层分块评估策略的合理性.所以,通过热连轧生产过程实际数据的实验研究表明了融合多维特征信息分层分块的热连轧过程运行状态评估方法的科学性和有效性.同时,由于本文将热连轧生产过程分为“3个子块、3个层级”,可以对各个子块进行有针对性地监测和维护,提高了

企业运维的效率.

4 结 论

本文提出了一种融合分层分块信息的轧制过程运行状态评估方法.首先将整个流程分为“3个层级和3个子块”,依次为全流程层、子系统层、设备层以及上游子块、中游子块、下游子块;然后,在每个子块分别建立运行状态评估模型,为了获得完整的运行状态信息,分别采用KPCA算法和t-SNE算法从全局和局部角度进行子块运行状态信息的提取;最后,将获得的基于全局和局部的信息输入到SVM中,利用D-S证据理论融合SVM的后验概率,进而评估各个子块的运行状态等级,通过组合权重理论融合各子块的运行状态等级,实现全流程的运行状态等级评估.在热连轧过程中的实验结果表明,相较于从单一角度进行特征提取的算法以及未分层分块的评估策略,所提出算法不仅从两个角度对运行状态信息进行特征提取,利用决策融合的方法将评估结果进行融合,而且采用分层分块的评估策略,获得了更高的评估准确率.

参考文献(References)

- [1] 郑伟,姚远,刘炎,等.基于分层分块状态相关独立成分分析的火力发电过程运行状态评价[J].控制与决策,2022,37(4):1081-1088.
(Zheng W, Yao Y, Liu Y, et al. Operation performance assessment of thermal power generation process based on HMPRICA[J]. Control and Decision, 2022, 37(4): 1081-1088.)
- [2] 姚远,佟佳蓉,高军,等.基于分层分块DLPPCA-SVM的复杂工业过程监测与故障诊断方法[J].控制与决策,2022,37(5):1402-1408.
(Yao Y, Tong J R, Gao J, et al. Monitoring and fault diagnosis method of complex industrial process based on DLPPCA-SVM[J]. Control and Decision, 2022, 37(5): 1402-1408.)
- [3] 刘炎,常玉清,王福利.基于Fisher判别分析的过程运行状态在线评价[J].控制与决策,2014,29(9):1655-1660.
(Liu Y, Chang Y Q, Wang F L. Online assessment of process operating performance based on Fisher discriminant analysis[J]. Control and Decision, 2014, 29(9): 1655-1660.)
- [4] 常玉清,孙雪婷,钟林生,等.基于改进随机森林算法的工业过程运行状态评价[J].自动化学报,2021,47(9):2214-2225.
(Chang Y Q, Sun X T, Zhong L S, et al. Industrial operation performance evaluation of industrial processes based on modified random forest[J]. Acta Automatica Sinica, 2021, 47(9): 2214-2225.)
- [5] Wang H D, Wang X, Wang Z L. Performance assessment method of dynamic process based on SFA-GPR[J]. Journal of Process Control, 2022, 111: 27-34.

- [6] 褚菲, 许杨, 尚超, 等. 基于静-动态特性协同感知的复杂工业过程运行状态评价[J]. 自动化学报, 2023, 49(8): 1-14.
(Chu F, Xu Y, Shang C, et al. Evaluation of complex industrial process operating state based on static and dynamic cooperative perception[J]. *Acta Automatica Sinica*, 2023, 49(8): 1-14.)
- [7] Zou X Y, Zhao C H. Concurrent assessment of process operating performance with joint static and dynamic analysis[J]. *IEEE Transactions on Industrial Informatics*, 2020, 16(4): 2776-2786.
- [8] Zhou K L, Chen X, Wu M, et al. A novel performance assessment method of the carbon efficiency for iron ore sintering process[J]. *Journal of Process Control*, 2021, 106: 44-53.
- [9] Fan H P, Wu M, Cao W H, et al. An operating performance assessment strategy with multiple modes based on least squares support vector machines for drilling process[J]. *Computers & Industrial Engineering*, 2021, 159: 107492.
- [10] Lu D, Wang F L, Wang S, et al. Operating performance assessment based on semi-supervised cluster generative adversarial networks for gold flotation process[J]. *IEEE Transactions on Instrumentation and Measurement*, 2023, 72: 1-13.
- [11] Zhong L S, Chang Y Q, Wang F L, et al. Distributed operating performance assessment of the plant-wide process based on data-driven hybrid characteristics decomposition[J]. *Industrial & Engineering Chemistry Research*, 2020, 59(35): 15682-15696.
- [12] Wang Y L, Li L, Wang K. An online operating performance evaluation approach using probabilistic fuzzy theory for chemical processes with uncertainties[J]. *Computers & Chemical Engineering*, 2021, 144: 107156.
- [13] 褚菲, 傅逸灵, 赵旭, 等. 基于ISDAE模型的复杂工业过程运行状态评价方法及应用[J]. 自动化学报, 2021, 47(4): 849-863.
(Chu F, Fu Y L, Zhao X, et al. Operating performance assessment method and application for complex industrial process based on ISDAE model[J]. *Acta Automatica Sinica*, 2021, 47(4): 849-863.)
- [14] Zhang C F, Peng K X, Dong J, et al. KPI-related operating performance assessment based on distributed ImRMR-KOCTA for hot strip mill process[J]. *Expert Systems with Applications*, 2022, 209: 118273.
- [15] 刘炎, 龚思哲, 王福利, 等. 基于堆叠有监督降噪自编码的运行状态评价[J]. 仪器仪表学报, 2022, 43(4): 271-281.
(Liu Y, Gong S Z, Wang F L, et al. Process operating performance assessment based on stacked supervised denoising auto-encoders[J]. *Chinese Journal of Scientific Instrument*, 2022, 43(4): 271-281.)
- [16] Zhu H Q, Wang Q L, Zhang F X, et al. Fuzzy comprehensive evaluation strategy for operating state of electrocoagulation purification process based on sliding window[J]. *Process Safety and Environmental Protection*, 2022, 165: 217-229.
- [17] 王印松, 孙天舒. 一种基于证据融合的执行器故障诊断方法[J]. 控制与决策, 2022, 37(8): 2026-2032.
(Wang Y S, Sun T S. A method of actuator fault diagnosis based on evidence fusion[J]. *Control and Decision*, 2022, 37(8): 2026-2032.)
- [18] Li L, Wang Y L, Sun B, et al. Operating performance assessment for transition state of industrial processes[J]. *Chemical Engineering & Technology*, 2020, 43(12): 2567-2575.
- [19] Zhang C F, Peng K X, Dong J. A lifecycle operating performance assessment framework for hot strip mill process based on robust kernel canonical variable analysis[J]. *Control Engineering Practice*, 2021, 107: 104698.
- [20] 王玲, 朱慧. 基于KPCA和G-G聚类的多元时间序列模糊分段[J]. 控制与决策, 2021, 36(1): 115-124.
(Wang L, Zhu H. Fuzzy segmentation of multivariate time series with KPCA and G-G clustering[J]. *Control and Decision*, 2021, 36(1): 115-124.)
- [21] Retsinas G, Stamatopoulos N, Louloudis G, et al. Nonlinear manifold embedding on keyword spotting using t-SNE[C]. *The 14th IAPR International Conference on Document Analysis and Recognition*. Kyoto, 2018: 487-492.
- [22] van der Maaten L, Hinton G. Visualizing data using t-SNE[J]. *Journal of Machine Learning Research*, 2008, 9(11): 2008: 2579-2605.
- [23] 谢国民, 蔺晓雨. 基于改进SSA优化MDS-SVM的变压器故障诊断方法[J]. 控制与决策, 2023, 38(2): 459-467.
(Xie G M, Lin X Y. Transformer fault diagnosis method based on improved SSA optimized MDS-SVM[J]. *Control and Decision*, 2023, 38(2): 459-467.)

作者简介

张永博(1998—), 男, 硕士生, 从事工业过程监控与故障诊断、机器学习及大数据分析等研究, E-mail: 18846093367@163.com;

张凯(1987—), 男, 副教授, 硕士生导师, 从事复杂工业过程先进控制、运行状态评估及故障诊断等研究, E-mail: kaizhang@ustb.edu.cn;

彭开香(1971—), 男, 教授, 博士生导师, 从事工业大数据分析、制造过程数字主线及闭环优化等研究, E-mail: kaixiang@ustb.edu.cn;

杨朋澄(1998—), 女, 硕士生, 从事工业过程监控与故障诊断、机器学习及大数据分析等研究, E-mail: ypc199898@163.com.