

DOI:10.19651/j.cnki.emt.2519159

基于改进 RIME 算法与多特征融合的 VIT2M 股票预测模型*

秦肖阳^{1,2} 于文涛^{1,2} 李丽红^{1,2} 李志勋^{1,2} 陈彪^{1,2}

(1. 华北理工大学理学院 唐山 063210; 2. 华北理工大学河北省数据科学与应用重点实验室 唐山 063210)

摘要: 为应对股票价格预测中普遍存在的不稳定性与波动性,以及变分模态分解算法在预测过程中存在参数寻优的困难,本文提出了 CRIME-SE-VMD-VIT2M 二阶段组合预测框架。第 1 阶段,在原始霜冰优化算法的基础上引入 Chebyshev 混沌映射与透镜成像种群选择策略,并以 SE 作为适应度函数,构建改进的 CRIME-SE-VMD 寻优模型,从而提升参数寻优的全局搜索能力与分解质量。第 2 阶段,通过 PCC 筛选关键技术指标,将其与 VMD 分解所得的 IMFs 融合,形成多维度特征集,在此基础上结合第 1 阶段的寻优结果,设计并实现 VIT2M 并行双通道预测模型,对多尺度股票特征信息进行深度提取与建模。实验结果表明,CRIME-SE-VMD 在 4 个股票数据集上的适应度值较对比算法降低 0.000 318 9~0.000 703,显示出更优的分解性能;同时,VIT2M 模型在相同数据集的预测性能优于对比模型,验证了其在提升股票价格预测精度方面的有效性。

关键词: 股票预测;霜冰优化算法;组合预测;技术指标;变分模态分解

中图分类号: TN911.73;TP183 **文献标识码:** A **国家标准学科分类代码:** 510.1050

VIT2M stock prediction model based on improved RIME algorithm and multi-feature fusion

Qin Xiaoyang^{1,2} Yu Wentao^{1,2} Li Lihong^{1,2} Li Zhixun^{1,2} Chen Biao^{1,2}

(1. College of Science, North China University of Science and Technology, Tangshan 063210, China; 2. Key Laboratory of Data Science and Application of Hebei Province, North China University of Science and Technology, Tangshan 063210, China)

Abstract: To address the widespread instability and volatility in stock price forecasting, as well as the difficulty of parameter optimization in the variational mode decomposition (VMD) algorithm, this paper proposes a two-stage combined prediction framework, CRIME-SE-VMD-VIT2M. In the first stage, the Chebyshev chaos map and lens imaging population selection strategy are introduced on the basis of the original frost ice optimization algorithm. Using SE as the fitness function, an improved CRIME-SE-VMD optimization model is constructed to enhance the global search capability and decomposition quality of parameter optimization. In the second stage, key technical indicators are selected through PCC and fused with the IMFs obtained from VMD decomposition to form a multi-dimensional feature set. Based on this, combined with the optimization results of the first stage, a VIT2M parallel dual-channel prediction model is designed and implemented to deeply extract and model multi-scale stock feature information. Experimental results show that the fitness value of CRIME-SE-VMD on four stock datasets is 0.000 318 9~0.000 703 lower than that of the comparison algorithm, demonstrating better decomposition performance. At the same time, the prediction performance of the VIT2M model on the same datasets is better than that of the comparison model, verifying its effectiveness in improving the accuracy of stock price prediction.

Keywords: stock prediction; RIME optimization algorithm; combined forecast; technical indicators; variational mode decomposition

0 引言

在当今经济全球化快速发展的背景下,股票市场作为

现代金融体系的核心,其价格波动不仅直接影响投资者的决策行为和企业融资活动,更关系到宏观经济的稳定与健康发展。因此,如何有效预测股票价格走势,成为金融界与

收稿日期:2025-06-21

* 基金项目:河北省数据科学与应用重点实验室项目(10120201)、唐山市数据科学重点实验室项目(10120301)资助

学术界长期关注的研究热点^[1]。然而,由于股票价格受到宏观经济、公司基本面、市场情绪等多种复杂因素的共同作用,其波动具有显著的非线性、非平稳性及高噪声特性,使得精确建模与预测面临巨大挑战^[2]。

近年来,深度学习方法凭借自动特征提取和长序列建模能力,在股票预测中展现出优于传统方法的性能。常见的深度学习模型主要包括时间卷积神经网络(temporal convolutional network, TCN)、长短期记忆神经网络(long-short term memory, LSTM)以及 Transformer 类预测模型等,如 Gülmez^[3]使用 LSTM 模型对股票进行预测。Liu 等^[4]提出结合通道-时间注意力模块(CTAM)的 TCN 结构,用于提升股票预测性能。陈诗乐等^[5]提出 GA-Transformer 模型的股价预测方案。赵宗毅等^[6]在 Informer 基础引入多频段卷积增强模块替代原模型的常规卷积,也进一步提升了其预测性能。但对于长期预测任务而言,单一深度学习模型在结构上缺乏对多源特征交互和多尺度时序模式的有效建模能力,表现出预测精度有限和稳定性不足的缺陷,这也推动了多模型融合与分解预测方法的研究。

现有研究普遍采用“分解-组合”模型以提升股票价格预测的精度。Ren 等^[7]基于小波变换将原始序列分解为高、低频分量,并分别利用 Informer 和 Encoder Forest 进行预测,结果表明组合模型较单一模型更具优势。Gong 等^[8]采用改进的 ICEEMDAN 方法对股票序列进行分解,并构建 BiLSTM-SAM-TCN 模型以实现多子序列独立预测,实验结果显示优于对比模型。Yu 等^[9]则提出 CEEMDAN-Informer-LSTM 模型,通过 Informer 建模高频分量与 LSTM 捕捉低频分量的混合策略,有效提升了预测精度。尽管这些方法在一定程度上改善了信号分解的稳定性,但这些分解技术仍存在模态混叠与中心频率不可自适应的问题,从而导致部分分解子序列物理意义模糊,限制了后续建模的可靠性。相较之下,变分模态分解(variational mode decomposition, VMD)通过约束模态中心频率与带宽,有效缓解了模态混叠问题,并在股票预测研究中得到广泛应用。Liu 等^[10]提出 VMD-LSTMA-TCNA 双通道注意力模型,在多个实证研究中表现出更强的鲁棒性;Liu 等^[11]使用 VMD 将元学习与 LSTM 结合构建 VML 模型,有效提升预测准确性;黄后菊等^[12]通过 VMD-CSSA-LSTM 实现股票收盘价预测,同样获得了显著改进。

然而,现有基于信号分解与深度学习的研究仍存在多方面局限,首先,大多数方法采用不同模型对应特定频段或级联而非并行策略,难以在统一框架下同时建模长期依赖与短期动态,导致信息提取不完整;其次,这类方法通常仅依赖分解的 IMF_s 分量,未能充分整合技术指标等多源特征,因而难以揭示复杂的非线性关系;最后,VMD 分解的关键参数模态分解数 K 与惩罚因子 α 对分解质量影响显著^[13],但难以人工确定,如果不能选择合适参数,将会导致

后续模型学习到伪规律,进而影响最终的预测结果,针对高维且复杂的股票数据,传统优化算法常因搜索空间的非线性和高噪声特性而陷入局部最优,难以实现高效的全局寻优。

为了解决上述的问题,本文采用“分解-预测-集合”策略针对基于 VMD 分解和深度学习组合模型对未来股票价格进行预测研究。首先,针对霜冰优化算法(rime optimization algorithm, RIME)在处理复杂非线性及高维问题时的局限性,通过引入 Chebyshev 混沌映射和基于动态因子控制的透镜成像选择策略,提出了一种增强型霜冰优化算法(CRIME),并构建基于 CRIME 的 CRIME-SE-VMD 的寻优模型,以样本熵(sample entropy, SE)作为适应度函数,优化 VMD 的模态层数 K 和惩罚因子 α , 获得最优参数组合。其次,设计 VIT2M 并行非级联预测模型,在模型中使用最优参数组合,通过 Informer 通道建模股票序列的全局依赖与长期趋势,在 TCN 分支中,用 Mish 激活函数替代 ReLU 激活函数,并融入多头注意力机制(multi-head attention, MHA)提取短期依赖与局部动态变化,通过构建双通道结构关注不同的时序信息,实现长短期特征的协同建模。融合皮尔逊相关系数(pearson correlation coefficient, PCC)筛选的高相关性技术指标与 IMF_s, 形成多源增强特征。最后,重构各通道预测结果输出最终值,在多个真实数据集上的实验表明,所提方法提升了预测精度。

1 相关工作

1.1 CRIME-SE-VMD 寻优模型

模态分解数 K 与惩罚因子 α , 对 VMD 的精度起着决定性作用。鉴于实际中待分解信号呈现出高度复杂性与非平稳特性^[14], 通过人为方式设定分解数 K 及惩罚因子 α 极具挑战性, 不仅难以确保设置的科学性, 还极易致使 VMD 分解结果出现不可控的随机性偏差。 K 值如果过大将会导致股票序列过度分解, 过小则会欠分解; 如果 α 值过大, 会引起频带信息缺失, 过小则会引起信息冗余^[15]。因此基于收盘价子序列样本熵之和最小化的原则, 采用 CRIME 算法对模态分解层数及惩罚因子参数进行优选, 本寻优模型目的在于将原始股票价格序列解构为具备更强可解释性和高度结构化的 IMF_s。 CRIME-SE-VMD 寻优流程图如图 1 所示。

RIME 算法^[16]因其物理启发的搜索机制, 被广泛应用于超参数优化任务中, 杨博等^[17]通过 Logistic 混沌映射和莱维飞行策略增强 RIME 的探索能力, 应用于能源预测; 汪家伟等^[18]引入动态质心和差分变异, 优化路径规划; 谢灿坤等^[19]结合 Sine 混沌和随机因子搜索提升全局收敛。然而, 这些改进多针对工程优化, 其在复杂非线性、股票高维数据中的应用受限于初始种群多样性不足和易陷入局部最优。在基础 RIME 算法中, 初始种群是在指定搜索区域内随机生成的。这种初始化方式容易导致种群多样性和分布

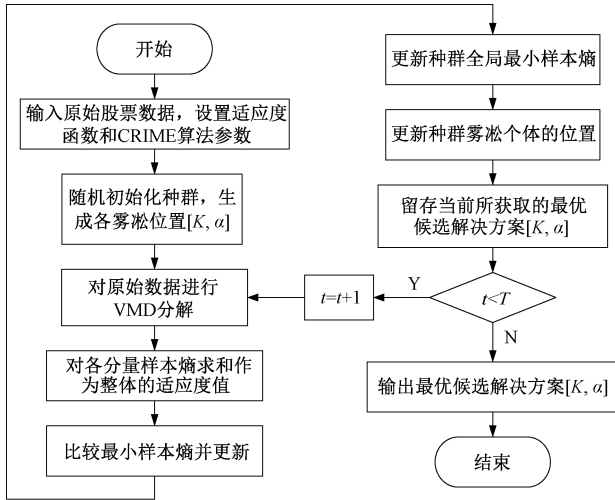


图 1 CRIME-SE-VMD 结构
Fig. 1 CRIME-SE-VMD structure

均匀性难以保障,进而造成算法在解空间探索过程中存在较大盲目性,直接影响到问题的寻优效率。

目前,研究人员常用的混沌映射包括 Tent 映射、Sine 映射、Logistic 映射和 Chebyshev 映射等。其中,Tent 映射属于分段线性映射,其更新机制导致粒子位置趋向于搜索空间的上下边界,限制了对解空间的全面探索,且粒子位置更新缺乏关联性,种群多样性不足,易使算法陷入局部最优。Sine 映射因其周期性特征,生成的序列可能出现重复或陷入周期循环,分布不均导致聚集效应,削弱种群多样性,进而可能造成算法收敛速度慢或停滞于局部最优解。Logistic 混沌映射在高维优化中会面临因分布不均、周期性陷阱及初始条件敏感性,导致种群多样性不足,从而陷入局部最优,限制了全局搜索效率。而 Chebyshev 映射在 $[-1, 1]$ 区间内具有均匀分布特性^[20],能够遍历整个搜索空间,提高种群多样性,并增强全局搜索能力,避免陷入局部最优,同时计算简单高效。为突破传统 RIME 算法在收敛速度与寻优精度方面的局限,引入 Chebyshev 混沌映射避免传统随机初始化可能导致的局部聚集问题。其数学公式为:

$$T_n(x) = \cos(n \cdot \arccos(x)) \quad (1)$$

当 $n \geq 2$ 时,其输出的函数值在 $[-1, 1]$ 上是呈现均匀分布的。种群初始化阶段,使用 4 阶 Chebyshev 多项式: $T_4(x) = 8x^4 - 8x^2 + 1$ 。输入 x 是 $[-1, 1]$ 上的均匀分布点,输出的 $T_4(x)$,在 $[-1, 1]$ 区间。通过 $T_4(x)$ 的混沌特性,初始解在解空间中分布更均匀且具有随机性,避免传统随机初始化的局部聚集。映射到解空间的公式为:

$$solution = lb + \frac{T_4(x) + 1}{2} \cdot (ub - lb) \quad (2)$$

$lb + \frac{T_4(x) + 1}{2} \cdot (ub - lb)$ 将混沌值线性映射到解空间 $[lb, ub]$ 。

作为霜冻机制的增强,进化扰动阶段能够增加跳出局部最优的能力,使用 2 阶 Chebyshev 多项式: $T_2(x) = 2x^2 - 1$,输入 x 是从 $[-1, 1]$ 区间内随机抽取的一个值,输出 $T_2(x)$ 仍在 $[-1, 1]$ 区间内。其扰动公式为:

$$perturbation = T_2(x) \cdot 0.5 \quad (3)$$

$T_2(x)$ 的值域为 $[-1, 1]$,乘以 0.5 以后,其扰动范围缩小到 $[-0.5, 0.5]$,其新解生成公式为:

$$new_solution = solution + T_2(x) \cdot 0.5 \cdot (ub - lb) \quad (4)$$

其中, $T_2(x) \cdot 0.5$ 是扰动幅度, $(ub - lb)$ 是解空间的宽度,扰动后的新解被限制在 $[lb, ub]$ 内。其中 $T_2(x)$ 引入非线性扰动,相较于简单的随机扰动,它能更好地跳出局部最优,同时扰动范围可控。

为了增强粒子的多样性,引入了基于动态因子控制的透镜成像策略,构建基于适应度偏移的保留机制,用于引导算法的搜索过程。相比自适应惯性权重策略的适配性有限、混沌扰动策略和莱维飞行策略的随机不稳定性、OBL 的边界聚集、多子群的资源需求及差分进化的参数复杂性,该策略提供了更稳定的多样性增强,这一策略将焦点集中在解空间中更为有利的区域,从而加快收敛速度,使算法能够更迅速地接近最优解。粒子的镜像解可以表示为:

$$\mathbf{R}_{reverse(i)} = \frac{(ub - lb)}{2} + \frac{(ub - lb)}{2 \times K} \frac{\mathbf{R}_i}{K} \quad (5)$$

K 为透镜成像的动态因子可表示为:

$$K = \left(1 + \left(\frac{T}{T_{max}}\right)^{0.5}\right)^{10} \quad (6)$$

其中, $\mathbf{R}_{reverse(i)}$ 为镜像解, \mathbf{R}_i 为原解。

1.2 VIT2M 预测模型

该模型遵循“分解-预测-集合”的原则,首先,基于创建的 CRIME-SE-VMD 寻优模型找出最优的 $[K, \alpha]$,将其作为后续 VMD 算法的输入参数。为了防止前视偏差,将原始股票收盘价序列通过时间滑动窗口机制划分为多个子窗口,使用 VMD 算法对子窗口进行分解,与此同时,通过 PCC 从多个技术指标中筛选出与收盘价相关性高的技术指标,将 IMFs 与技术指标构建一个多特征数据集输入到预测模型结构。模型结构方面,并行使用 Informer 和 TCN2M,通过两个分支并行学习,可以使模型在不同的时间尺度上进行信息提取和处理,使用 Informer 来捕捉长时间信号时序依赖特征长时间序列处理,并能够捕捉序列数据中的全局特征和长时间依赖关系,确保对时序特征进行全面建模;TCN2M 部分,能够有效捕捉短期依赖性和序列数据的动态变化,通过 Mish 激活函数和 MHA 对多源股票特征间非线性协同效应的自适应动态建模能力,协同提升了模型梯度流动稳定性、短期依赖建模能力,最后将 Informer 与 TCN2M 模型的输出特征进行拼接,通过全连接层进行特征融合,输出最终预测结果。整体结构如图 2 所示。

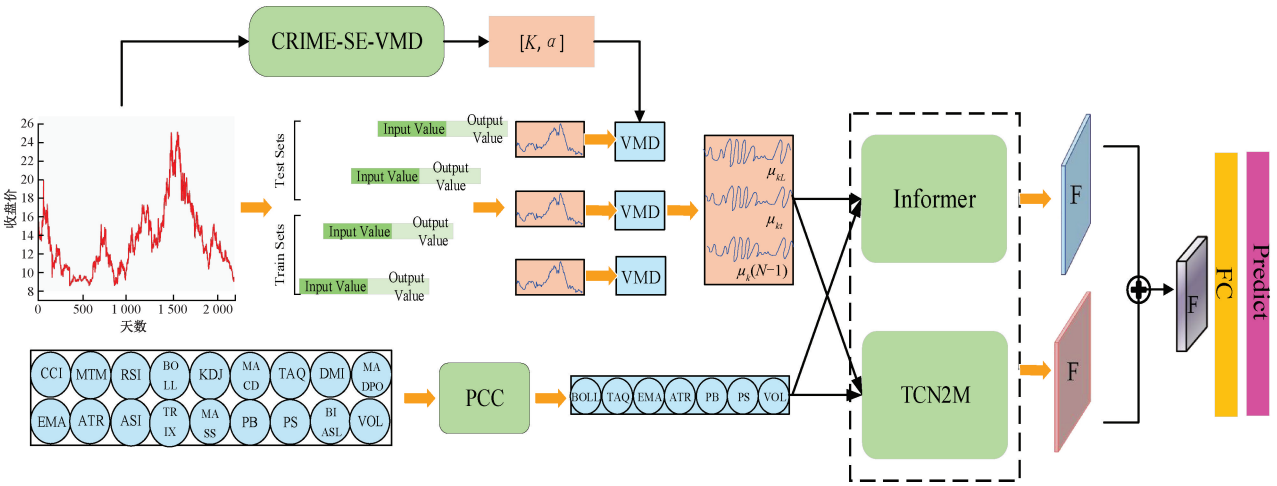


图 2 VIT2M 结构

Fig. 2 VIT2M structural

1.3 TCN2M 模型

为增强 TCN 模型捕捉股票数据中短期依赖关系的能力,本文在 TCN 模型的残差结构中,针对传统 ReLU 激活函数在深度网络中的神经元坏死问题,改进性地采用 Mish 激活函数进行替代。同时相较于 ReLU 函数, Mish 激活函数能在负数区间保留梯度以缓解深层网络的梯度消失问题、正数区间通过压缩特性抑制异常值干扰,同时以强非线性能力拟合复杂数据分布,相较于 ReLU 激活函数可有效提升模型训练稳定性与表征能力。如式(7)所示。

$$Mish(x) = x \cdot \tanh(\ln(1 + e^x)) \quad (7)$$

将其应用于 TCN 模型中,能通过其平滑非线性特性缓解梯度消失问题,提升网络训练效率,并且借助其平滑连续性增强长时序股价数据的历史信息传递效率,通过非线性映射捕捉量价关系等复杂市场模式,利用梯度保留与压缩特性提升噪声鲁棒性及多源特征的融合能力,从而优化模型对股票趋势、周期转换的预测精度与泛化性能。

如图 3 所示,右侧为更换为 Mish 函数的残差模块结构图,左侧为原始的残差模块结构图。

此外,为进一步提升多维金融特征间的关联建模能力,提出基于多头注意力机制的时序卷积网络架构。如图 4 所示。

该架构通过 Query-Key-Value 并行计算范式,赋能模型对 VMD 分解 IMFs、技术指标等异质特征的差异化关联建模能力,从根本上提升特征表达的丰富性与预测精度。传统 TCN 依赖预设扩张卷积核捕捉时序依赖,其特征交互机制存在局限性。卷积操作的感受野受核大小限制,难以建模跨尺度特征关联,卷积本质为线性加权求和,无法自适应刻画“技术指标+模态函数”等复合信号的非线性协同效应;而 MHA-TCN 通过学习的注意力权重动态构建特征关联网络,如式(8),根据市场状态灵活分配注意力,聚焦关键信号,抑制噪声干扰。这种差异使 MHA-TCN 更适配股票市场的高非线性、多尺度与动态变化

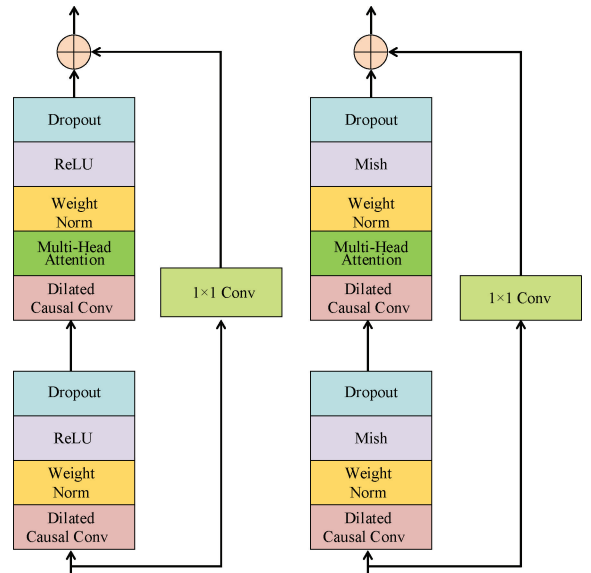


图 3 残差模块结构对比

Fig. 3 Comparison of residual module structures

特性。

$$Attention(Q, K, V) = softmax\left(\frac{QK^T}{\sqrt{d_k}}\right)V \quad (8)$$

相较于传统 TCN 模型,提出的改进方法在理论和实践层面均具有优势。首先, Mish 激活函数通过其独特的数学特性有效解决了 ReLU 激活函数的“神经元坏死”问题,其在负半轴保持的小幅梯度确保了股票数据标准化后负值信息的有效传递,而正半轴的软饱和特性则提供了天然的异常值抑制能力,这对于处理股票市场的厚尾分布和极值波动具有重要意义。其次,多头注意力机制从根本上突破了传统 TCN 依赖固定扩张卷积核的局限性,通过动态权重分配机制,实现了对多源股票特征间非线性协同效应的自适应建模,相比卷积操作的线性加权本质,能够更好地捕捉“技术指标+模态函数”等复合信号的复杂交互模

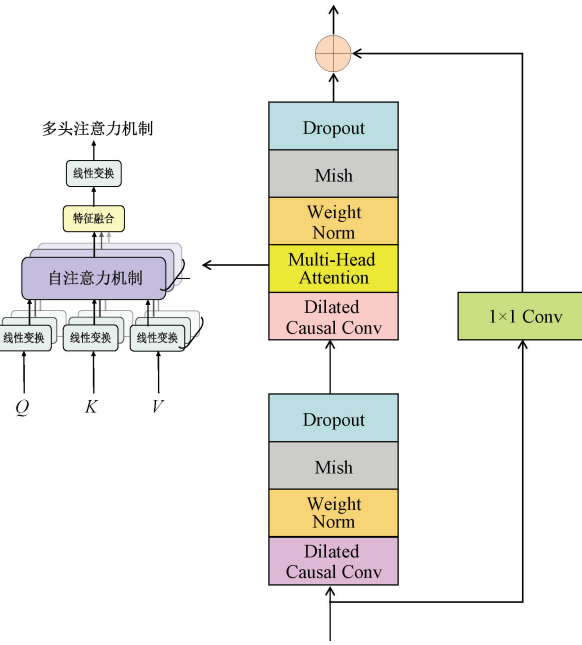


图 4 TCN2M 结构
Fig. 4 TCN2M structure

式。在这两项改进协同作用下，TCN2M 模型不仅在梯度流动稳定性和短期依赖建模能力方面优于传统 TCN，更重要的是其动态特征选择机制与股票市场的高非线性、多尺度和时变特性高度匹配，从而提高了预测精度和泛化性能。

2 实验分析

2.1 数据来源

本文研究的原始数据集为 Tushare 财经共享数据集

(<http://tushare.org/>)，选取 4 支不同行业的股票数据，旨在验证本文提出的模型在各行业中的预测效果，数据汇总如表 1 所示。选取 2015-01-05 日~2023-12-29 日近 9 年的基本指标和技术指标，共获取 2 190 个交易日历史数据，将 IMF_s 和筛选的关键技术指标作为样本数据集，选取样本数据集的前 70%，即 2015-01-05 日~2021-04-19 日约 1 533 个交易日的数据作为训练样本集，样本数据集后 30%，即 2021-04-20 日~2023-12-29 日约 728 个交易日的数据作为测试样本集。

表 1 数据对象
Table 1 Data object

股票代码	股票名称
000001.SZ	平安银行
000733.SZ	振华科技
SH. 600196	复星医药
SH. 600104	上汽集团

本文实验环境为 Windows 11 系统、采用 14th Gen Intel[®] Core i5-14600KF 处理器和 GPU RTX 4060 TI 显卡。在算法实现层面，系统整体架构基于 PyTorch 深度学习框架构建，功能单元基于 Python 3.8 编程语言开发完成。

2.2 特征选择分析

为了更准确的预测股票收盘价的走势，本文从技术指标中的超买超卖指标、趋势指标、均匀度指标、价格波动指标、人气指标、能量型指标以及成交量指标中选取 18 个股票技术指标，并使用 PCC 从中筛选与收盘价相关系数高的技术指标，表 2 和表 3 为 4 支股票相关系数值较大的技术指标。

表 2 技术指标相关性系数

Table 2 Correlation coefficient of technical indicators

股票名称	boll_lower_qfq	taq_min_qfq	ema_qfq_10	atr_qfq	pb	ps_ttm
平安银行	0.905 4	0.926 7	0.938 2	0.864 5	0.709 5	0.819 7
上汽集团	0.847 1	0.862 5	0.882 7	0.555 2	0.886 6	0.914 4
振华科技	0.970 7	0.994 9	0.898 2	0.909 0	0.839 4	0.941 5

表 3 技术指标相关性系数

Table 3 Correlation coefficient of technical indicators

股票名称	boll_lower_qfq	taq_min_qfq	ema_qfq_10	atr_qfq	pb	vol
复星医药	0.928 8	0.871 3	0.998 2	0.860 4	0.711 5	0.548 1

PCC 用于衡量两个变量之间线性相关程度的统计指标。通常用字母 r 表示，它的值介于 1 和 -1 之间，其系数数值越大，两个变量之间的相关性越强，其主要表达式为：

$$r = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2 \sum_{i=1}^n (y_i - \bar{y})^2}} \quad (9)$$

其中, \bar{x} 表示样本 X 的均值, 即 $\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$; \bar{y} 是样本 Y 的均值, 即 $\bar{y} = \frac{1}{n} \sum_{i=1}^n y_i$ 。

由于将股票收盘价分解后的 IMF_s 和技术指标同时作为输入的特征参数, 而两者数字范围上存在较大差异, 为了消除不同指标间的量纲影响, 提升模型精度和收敛速度, 本文对所有数据进行了归一化处理。归一化的公式为:

$$x_{i,s}^{norm} = \frac{x_{i,s} - x_s^{min}}{x_s^{max} - x_s^{min}} \quad (10)$$

其中, $x_{i,s}^{norm}$ 为归一化后的值, $x_{i,s}$ 为原始值, x_s^{max} , x_s^{min} 分别表示样本中的最大值以及最小值。

2.3 CRIME-SE-VMD 模型实验分析

为全面评估 CRIME 算法的性能, 选取了近年来在复

杂优化问题中表现优异且具有代表性的元启发式算法进行对比, 包括原始 RIME、改进的河马优化算法 IHO^[21]、性能优越人工兔优化算法 ARO^[22] 4 种算法的最大迭代次数均为 100, 且种群数量设置为 20, 其余算法参数均按标准设定。

从图 5 可以看出, 在对 4 支股票进行实验对比, 当达到最大迭代次数的时候, 4 支股票全部均是 CRIME 算法的适应度值最小, 说明其样本熵最小, 意味着分解后的子序列在整体上具备更高的有序性和更低的复杂性, 从而能够更有效地提取数据的内在特征和规律, 有力地证明了当前的优化方案在提升 VMD 分解效果方面具有的优越性, 并且在 4 种不同的数据集上面 CRIME 算法的收敛速度最快, 从适应度值可以得出, 本文提出的 CRIME-SE-VMD 分解结构相比于其他 3 种寻优模型迭代寻优效果最佳。

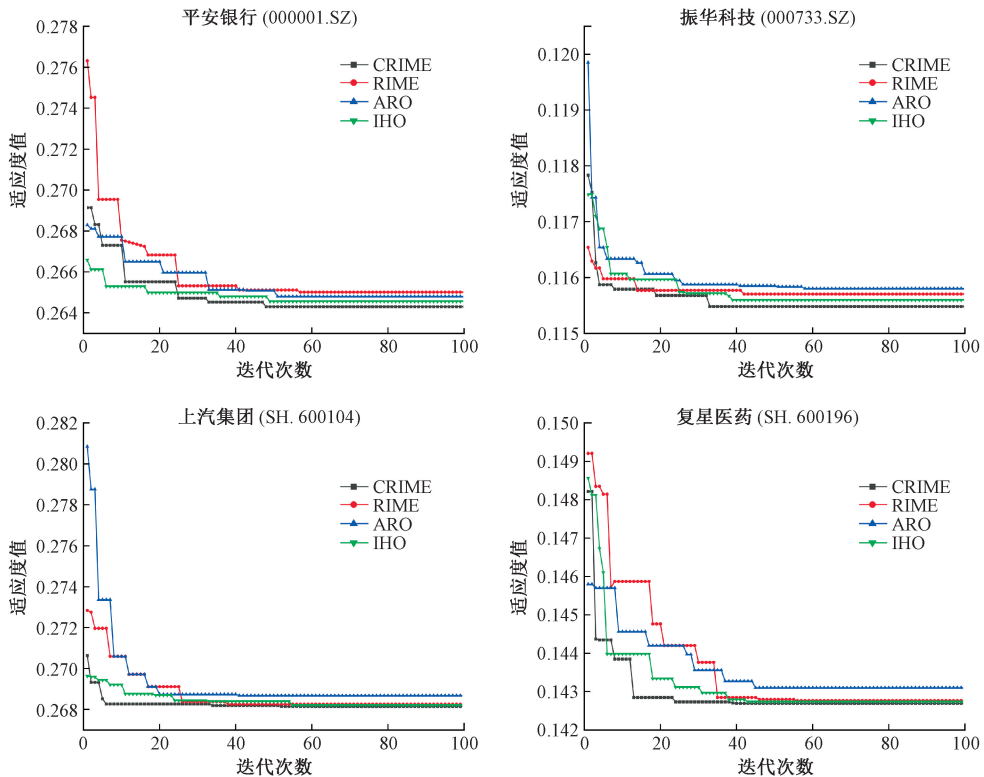


图 5 算法适应度收敛曲线

Fig. 5 Algorithm fitness convergence curve

表 4 展示了 4 支股票经 CRIME-SE-VMD 模型优化的参数结果, 包含最优模态分解层数及惩罚因子参数。以平安银行(000001.SZ)为例, 其 VMD 分解过程如图 6 所示, 原始收益率序列被有效分解为 6 个 IMF_s。IMF1 和 IMF2 分量具有相对平坦的趋势, 表明其是低频分量, 而其他分量表示高频分量, 这些分量与其他股票技术指标相结合, 以创建模型的多特征输入数据。

2.4 VIT2M 预测模型的对比分析

为保证对比实验的科学性与可靠性, 本实验运用多种评估指标, 对模型的预测性能展开了全方位的考量。具体涵盖决定系数(R^2)、平均绝对误差(MAE)、平均绝对百分比误差(MAPE)以及均方根误差(RMSE)。在这些指标体系中, 决定系数越高拟合效果越好; 误差值越低, 表明模型的预测值与实际值的契合度越高, 即模型的预测精准度越佳。

表 4 最佳参数及适应度值对比

Table 4 Comparison of optimal parameters and fitness values

股票名称	分解模态层数(K)	惩罚因子(α)	CRIME 最佳适应度值
平安银行	6	2 173	0.264 308
复星医药	3	2 065	0.142 695
上汽集团	6	1 591	0.268 148
振华科技	7	1 907	0.115 482



图 6 VMD 分解子序列图

Fig. 6 VMD decomposition subsequence diagram

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (11)$$

R^2 是评估回归模型拟合优度的核心指标。它衡量了自变量对因变量变化的解释程度,反映了模型的可靠性。

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (12)$$

RMSE 是衡量预测值与真实值之间差异程度的常用指标,其核心作用是量化模型预测结果的平均误差大小。

$$MAPE = \frac{1}{n} \sum_{i=1}^n \frac{|y_i - \hat{y}_i|}{y_i} \times 100\% \quad (13)$$

MAPE 是衡量预测模型精度的核心指标之一,其核心特点是通过“百分比”形式量化预测值与真实值的偏差,能直观反映误差占真实值的比例。

$$MAE = \frac{1}{n} \sum_{i=1}^n |\hat{y}_i - y_i| \quad (14)$$

MAE 是衡量预测模型误差的基础且常用的指标,核心是通过计算预测值与真实值之间绝对偏差的平均值,反映模型预测结果与实际情况的平均偏离程度。

为了验证所提出的 VIT2M 预测模型性能的优越性,将本模型与在股票预测领域常用的两种基本深度学习 LSTM 模型、TCN 模型以及其他论文中两种强化组合模型 VMD-LSTMA-TCNA 和 LSTM-GRU-SA-AM 进行精度对比分析,预测结果如图 7 所示。

由表 5 可知,VIT2M 模型在 4 支股票数据集上实验得到的 RMSE、MAE 和 MAPE 均小于其他模型, R^2 更接近

1。实验结果表明,与其他模型相比,VIT2M 模型在股票价格预测中表现出更高的拟合度,预测结果更贴近真实值。验证结果进一步说明,在处理股票价格预测任务时,对原始价格序列进行分解再预测,能够有效提升模型预测性能。尤其是 VIT2M 模型在预测误差方面小于其他 4 种模型,主要归因于其能够从多源特征中提取更具代表性的有效信息,从而预测误差更小。

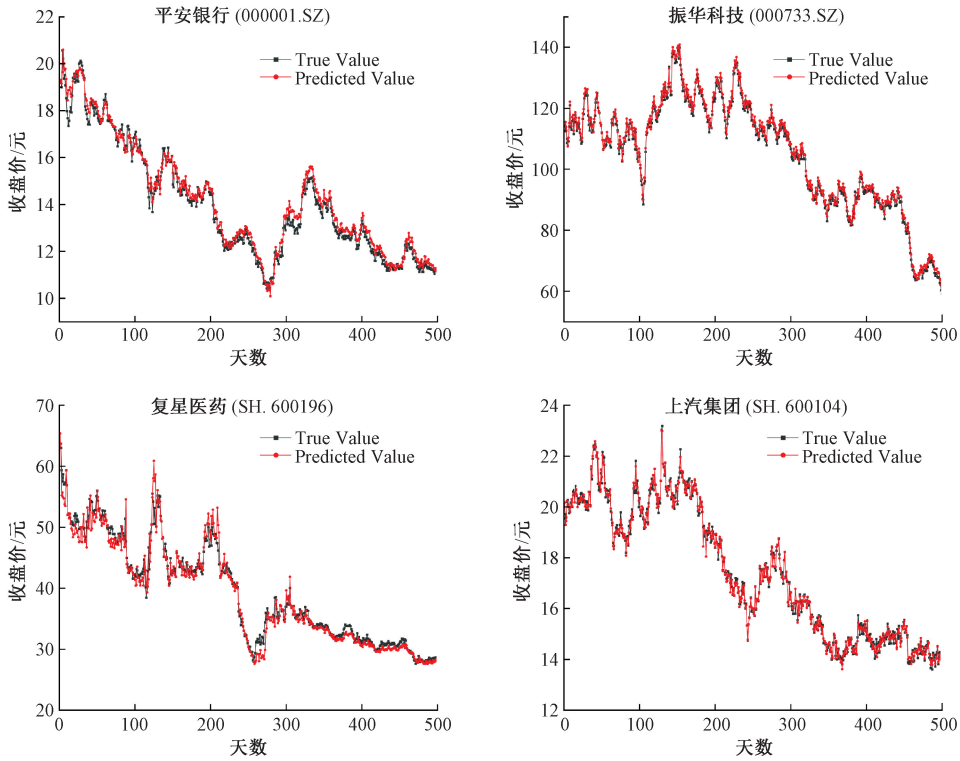


图 7 VIT2M 模型对 4 支股票的预测结果

Fig. 7 The prediction results of VIT2M model for four stocks

表 5 对比实验评价指标值

Table 5 Evaluation index values of comparative experiments

数据集	模型	RMSE	MAE	MAPE/%	R^2
平安银行(000001.SZ)	LSTM	0.701	0.489	3.415	0.937
	TCN	0.834	0.575	3.918	0.905
	VMD-LSTMA-TCNA	0.431	0.359	2.564	0.976
	LSTM-GRU-SA-AM	0.551	0.411	2.942	0.966
	VIT2M	0.385	0.301	2.156	0.984
复星医药(SH. 600196)	LSTM	0.663	0.441	2.873	0.931
	TCN	0.737	0.553	3.187	0.918
	VMD-LSTMA-TCNA	0.485	0.365	1.101	0.964
	LSTM-GRU-SA-AM	0.419	0.333	1.915	0.977
	VIT2M	0.353	0.326	1.536	0.982
上汽集团(SH. 600104)	LSTM	0.663	0.441	2.873	0.931
	TCN	0.737	0.553	3.187	0.918
	VMD-LSTMA-TCNA	0.485	0.365	1.101	0.964

表 5(续)

Table 5 (continued)

数据集	模型	RMSE	MAE	MAPE/%	R ²
振华科技(000733.SZ)	LSTM-GRU-SA-AM	0.419	0.333	1.915	0.977
	VIT2M	0.353	0.326	1.536	0.982
	LSTM	0.408	0.309	1.951	0.911
	TCN	0.339	0.270	1.664	0.937
	VMD-LSTMA-TCNA	0.301	0.229	1.353	0.957
平安银行	LSTM-GRU-SA-AM	0.337	0.263	1.571	0.945
	VIT2M	0.243	0.179	1.097	0.978

为系统评估 CRIME 算法在 VMD 关键参数优化任务中的有效性,本文基于 4 支不同股票数据集设计了对比实验,实验以平均绝对百分比误差(MAPE)作为统一评估指标,为确保公平性与科学性,在每组数据实验中,所有对比方法均采用该数据集最优 K 值所对应的同一 α 值下进行测试,从而最大程度排除其他变量对结果的干扰。

如图 8 所示,经 CRIME 算法优化所得的 VMD 参数组合 $[K, \alpha]$ 在 4 组数据集上均展现出优势。实验结果表明,除复星医药数据集外,当 $K=3$ 时评价指标结果均为最差,验证了 K 值过小则出现欠分解问题;除振华科技数据集外,当 $K=7$ 时,另外 3 种数据集评价指标均高于其最优 K 值对应的评价指标,验证了模态分解数 K 取值过大,会造成股票序列分解过度。相比其他优化方法,该算法增强了种群多样性与全局寻优能力,有效克服了陷入局部最优的弊端。实验结果表明,CRIME 能够稳定获得更优的模态分解参数组合 $[K, \alpha]$,在提高分解质量的同时增强了模型对股票价格序列中潜在规律的捕捉能力。所获得的高质量 IMFs 作为深度预测模型的输入,在多个真实股票数据集上均提升了预测精度。

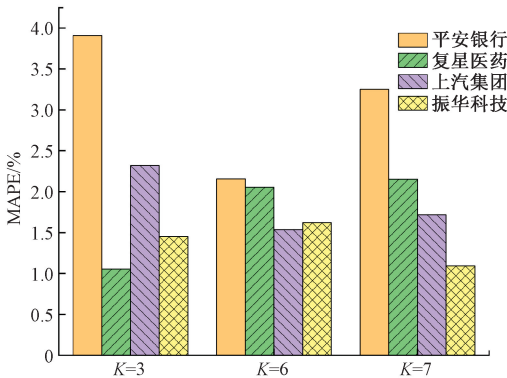


图 8 不同 K 值评价指标

Fig. 8 Different K value evaluation indicators

2.5 CRIME-SE-VMD-VIT2M 模型对比

为证明提出的 CRIME-SE-VMD-VIT2M 二阶段预测模型的整体性能,本节将以 RIME-SE-VMD-VIT2M、ARO-SE-VMD-VIT2M、IHO-SE-VMD-VIT2M 进行实验

对比。选取上汽集团(SH. 600104)为实验数据,分别采用 4 种不同寻优模型得出的 $[K, \alpha]$,作为 VMD 的参数进行实验对比,其中在上汽集团数据集下 4 种不同寻优模型参数如表 6 所示。

表 6 不同寻优模型参数

Table 6 Different optimization model parameters

模型	K	α	适应度值
CRIME-SE-VMD	6	1 591	0.268 148
RIME-SE-VMD	3	2 107	0.268 259
ARO-SE-VMD	7	2 003	0.268 682
IHO-SE-VMD	5	1 695	0.268 210

实验对比结果如图 9 所示。在所有评价指标下,本文提出的 CRIME-SE-VMD-VIT2M 二阶段组合预测模型均表现最优,其具体指标值如下:RMSE 为 0.353,MAE 为 0.326,MAPE 为 1.536%,R² 为 0.982。均显著优于实验中所选取的另外 3 种对比模型,充分验证了本文所提二阶段组合预测模型的有效性与优越性。

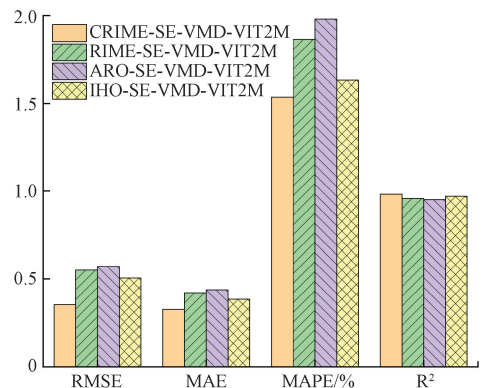


图 9 CRIME-SE-VMD-VIT2M 模型对比结果

Fig. 9 Comparison results of the CRIME-SE-VMD-VIT2M model

2.6 消融实验

为深入验证所提出改进方法中各组件的有效性贡献,本节以平安银行(000001.SZ)为实验数据设计了系统性的

消融实验。本实验通过消融研究评估 Informer+TCN 模型在不同组件配置下的性能表现,具体分析了 MHA,Mish

激活函数以及两者的组合(MHA+Mish)对模型预测精度的影响。实验数据如表 7 所示。

表 7 消融实验结果

Table 7 Results of ablation experiment

Informer+TCN			RMSE	MAE	MAPE/%	R ²
MHA	Mish	MHA+Mish				
✓			0.566	0.388	2.718	0.965
	✓		0.764	0.522	3.620	0.928
		✓	0.385	0.301	2.156	0.984
—	—	—	1.005	0.684	4.717	0.883

根据实验结果显示,相比无组件,MHA 降低了误差, RMSE 降低约 43.7%,MAE 降低约 43.3%,MAPE 降低约 42.4%,R² 提升至 0.965,表明有效增强了模型对股票市场数据中高度非线性、多尺度动态特征的自适应捕捉能力。单独替换 Mish 激活函数同样优于基础模型, RMSE 降低约 24.0%,MAE 降低约 23.7%,MAPE 降低约 23.3%,这验证了 Mish 函数通过其平滑的非线性特性和梯度保留能力,能够更有效地强化特征的非线性映射,提升模型表征能力,MHA 与 Mish 的组合展现了最佳性能,各项误差指标均降至最低, RMSE 降低约 61.7%,MAE 降低约 56.0%,MAPE 降低约 54.3%,R² 达到 0.984。该结果有力说明,MAH 的动态特征权重分配机制与 Mish 的梯度保留特性形成了有效互补。二者的协同作用优化了模型的特征提取和融合过程,从而实现了预测误差的进一步降低。

3 结 论

本文提出的 CRIME-SE-VMD-VIT2 二阶段预测框架在股票价格预测领域展现出较好效果。通过引入 CRIME 优化算法自动搜索 VMD 的最优模式组合,有效克服了传统 VMD 方法依赖人工调参的局限。同时,VIT2M 模型通过 Informer 与 TCN2M 并行双通道设计,实现了多时间尺度下时序特征的高效捕捉与精准预测,融合关键技术指标进一步丰富了特征维度,增强了模型对复杂股票价格波动的建模能力。

本研究为非线性、多尺度时间序列预测提供了较好的解决办法,不仅提升了股票价格预测的精度,也为复杂时序数据的建模提供了新的思路。然而,受限于股票价格受投资者主观情绪等不可量化因素的影响,本框架在捕捉市场非理性波动方面的能力仍有不足。未来研究可通过引入投资者情绪因子、社交媒体数据等外部变量,进一步提升模型对市场动态的适应性。此外,本框架目前主要聚焦于股票价格预测场景,其通用性尚需进一步验证。后续研究可将该框架拓展至能源负荷预测等其他复杂时间序列任务,深入探索其在多领域中的泛化性。

参考文献

- [1] NOOR M H N, IGE A O. A survey on state-of-the-art deep learning applications and challenges [J]. ArXiv preprint arXiv:2403.17561,2024.
- [2] 徐凤如,李博涵,胥帅. 基于深度学习与大语言模型的序列推荐研究进展[J]. 计算机科学与探索,2025,19(2):344-366.
XU F R, LI B H, XU SH. Research progress on sequence recommendation based on deep learning and large language models [J]. Journal of Frontiers of Computer Science & Technology, 2025, 19 (2): 344-366.
- [3] GÜLMEZ B. Stock price prediction with optimized deep LSTM network with artificial rabbits optimization algorithm [J]. Expert Systems with Applications, 2023, 227: 120346.
- [4] LIU H J, ZHAO T L, WANG S W. A stock rank prediction method combining industry attributes and price data of stocks [J]. Information Processing & Management, 2023,60(4): 103358.
- [5] 陈诗乐,王笑,周昌军. 基于 GA-Transformer 模型的多因子股票预测[J]. 广州大学学报(自然科学版), 2021, 20(1):44-55.
CHEN SH L, WANG X, ZHOU CH J. Multi-factor stock prediction based on GA-Transformer model[J]. Journal of Guangzhou University (Natural Science Edition), 2021,20(1):44-55.
- [6] 赵宗毅,张振宇,徐清振. 融合双频域增强模块与 Informer 的股价预测研究[J/OL]. 计算机工程与应用, 1-14 [2025-07-07]. <https://doi.org/10.3778/j.issn.1002-8331.2502-0124>.
ZHAO Z Y, ZHANG ZH Y, XU Q ZH. Research on stock price prediction based on dual frequency domain enhancement module and Informer[J/OL]. Computer Engineering and Applications, 1-14 [2025-07-07]. <https://doi.org/10.3778/j.issn.1002-8331.2502-0124>.

- [7] REN SH SH, WANG X, ZHOU X, et al. A novel hybrid model for stock price forecasting integrating Encoder Forest and Informer[J]. Expert Systems with Applications, 2023, 234: 121080.
- [8] GONG H, XING H Y. Predicting the highest and lowest stock price indices: A combined BiLSTM-SAM-TCN deep learning model based on re-decomposition[J]. Applied Soft Computing, 2024, 167: 112393.
- [9] YU Y Y, DAI D SH, YANG Q, et al. An intelligent framework based on optimized variational mode decomposition and temporal convolutional network: Applications to stock index multi-step forecasting[J]. Expert Systems with Applications, 2025, 268: 126222.
- [10] LIU Y P, HUANG S Y, TIAN X Y, et al. A stock series prediction model based on variational mode decomposition and dual-channel attention network[J]. Expert Systems with Applications, 2024, 238: 121708.
- [11] LIU T T, MA X, LI SH, et al. A stock price prediction method based on meta-learning and variational mode decomposition[J]. Knowledge-Based Systems, 2022, 252: 109324.
- [12] 黄后菊,李波. 基于VMD-CSSA-LSTM组合模型的股票价格预测[J]. 南京信息工程大学学报, 2024, 16(3): 332-340.
HUANG H J, LI B. Stock price prediction based on VMD-CSSA-LSTM combination model[J]. Journal of Nanjing University of Information Science and Technology, 2024, 16(3): 332-340.
- [13] 周星,行鸿彦,叶如,等. 基于优化变分模态分解的海杂波去噪方法[J]. 电子测量与仪器学报, 2023, 37(11): 81-90.
ZHOU X, XING H Y, YE R, et al. Sea clutter denoising method based on optimized variational mode decomposition[J]. Journal of Electronic Measurement and Instrumentation, 2023, 37(11): 81-90.
- [14] 马亮,安腾飞,刘文黎,等. 基于声信号的给水管微小泄漏检测技术研究[J]. 电子测量与仪器学报, 2024, 38(12):113-123.
MA L, AN T F, LIU W L, et al. Research on water pipe micro-leakage detection technology based on acoustic signals [J]. Journal of Electronic Measurement and Instrumentation, 2024, 38(12):113-123.
- [15] 王强,刘宏伟,聂子凡. 基于VMDT-POA-DELM-GPR的两阶段短期负荷预测[J]. 国外电子测量技术, 2024, 43(1):101-109.
WANG Q, LIU H W, NIE Z F. Two-stage short-term load forecasting based on VMDT-POA-DELM-GPR [J]. Foreign Electronic Measurement Technology, 2024, 43(1):101-109.
- [16] SU H, ZHAO D, HEIDARI A A, et al. RIME: A physics-based optimization [J]. Neurocomputing, 2023, 532: 183-214.
- [17] 杨博,张芮,胡袁炜骥,等. 基于改进霜冰优化算法的混合风-光-波一体化系统阵列优化[J]. 南方电网技术, 2025, 19(7):50-61.
YANG B, ZHANG R, HU Y W J, et al. Array optimization of hybrid wind-solar-wave integrated system based on improved frost-ice optimization algorithm[J]. China Southern Power Grid Technology, 2025, 19(7): 50-61.
- [18] 汪家伟,付盛伟,黄海松. 改进霜冰优化算法用于无人机三维路径规划[J]. 电子测量技术, 2025, 48(9): 44-55.
WANG J W, FU SH W, HUANG H S. Improved frost optimization algorithm for UAV three-dimensional path planning[J]. Electronic Measurement Technology, 2025, 48(9):44-55.
- [19] 谢灿坤,于丽娅,张涛,等. 增强型霜冰优化算法的复杂环境下机器人路径规划[J]. 计算机应用研究, 2025, 42(1):185-195.
XIE C K, YU L Y, ZHANG T, et al. Enhanced frost optimization algorithm for robot path planning in complex environments [J]. Computer Application Research, 2025, 42(1):185-195.
- [20] 刘华伟,石润华,张红杉,等. 基于切比雪夫混沌映射的V2G跨境双向认证方案[J]. 计算机工程与科学, 2025, 47(3):472-484.
LIU H W, SHI R H, ZHANG H SH, et al. V2G cross-domain bidirectional authentication scheme based on Chebyshev chaotic map[J]. Computer Engineering and Science, 2025, 47(3):472-484.
- [21] PEI SH Y, SUN G, TONG L. An improved hippopotamus optimization algorithm based on adaptive development and solution diversity enhancement[J]. PeerJ Computer Science, 2025, 11: e2901.
- [22] WANG L, CAO Q, ZHANG ZH X, et al. Artificial rabbits optimization: A new bio-inspired meta-heuristic algorithm for solving engineering optimization problems[J]. Engineering Applications of Artificial Intelligence, 2022, 114: 105082.

作者简介

李丽红(通信作者),教授,硕士生导师,主要研究方向为数据挖掘、三支决策。

E-mail:22687426@qq.com

秦肖阳,硕士研究生,主要研究方向为深度学习、股票预测。

E-mail:qinxiaoyang828@stu.ncst.edu.cn

于文涛,硕士研究生,主要研究方向为入侵检测、Web安全。