

基于因果思维树的电动汽车电池 SOC 预测模型

彭自然, 杨肖阳, 李雪勇, 周 玉

(湖南工业大学交通与电气工程学院, 湖南株洲 412007)

摘要: 针对现有电动汽车实时荷电状态(State-Of-Charge, SOC)预测模型在运行状态感知、动态校准与长时序预测精度方面存在的瓶颈问题, 本文提出了一种融合因果思维树推理机制与深度强化学习策略的时序预测框架。该框架通过引入动态进化与多分支因果推理机制, 使其在保持单模型高效计算结构的同时, 实现对复杂工况下电池状态变化的感知建模与自适应优化。首先, 设计了基于因果分级结构的多级近端策略优化(Proximal Policy Optimization, PPO)模型, 并提出了以时间序列网络为核心的 Actor 网络架构。该架构分层建模温度、电阻等关键变量对电池 SOC 的直接与间接因果影响, 并通过值函数迭代与长期回报优化策略, 实现了模型参数的持续进化与泛化能力提升, 增强了预测的可解释性与因果认知能力。其次, 引入了思维树结构构建多路径策略评估网络, 结合策略搜索、路径追踪与回溯修正机制, 在动态工况中实现策略的逐层优化与异常分支校正, 提升了模型的预测鲁棒性与泛化能力。实验结果表明: 在不同运行工况下, 本文算法在多个评价指标上均显著优于 Transformer、FED former、Mamba 和长短期记忆(Long Short-Term Memory, LSTM)网络模型。本文算法的平均绝对误差(Mean Absolute Error, MAE)低于 0.26%, 均方根误差(Root Mean Squared Error, RMSE)低于 0.35%, 决定系数(Coefficient of Determination, R^2)高于 99.5%, 在不同车型条件下均表现出卓越的鲁棒性与稳定性。

关键词: 电动汽车实时荷电状态(SOC)预测; 深度强化学习; 因果思维树; 时间序列网络

基金项目: 湖南省教育厅重点科研项目(No.22A0423); 湖南省自然科学基金(No.2023JJ60267, No.2022JJ50073)

中图分类号: TN919.5; TP183

文献标识码: A

文章编号: 0372-2112(2025)08-2970-15

电子学报 URL: <http://www.ejournal.org.cn>

DOI: 10.12263/DZXB.20250450

A Causal Tree-of-Thought-Based Model for Battery State-of-Charge Prediction in Electric Vehicles

PENG Zi-ran, YANG Xiao-yang, LI Xue-yong, ZHOU Yu

(School of Transportation and Electrical Engineering, Hunan University of Technology, Zhuzhou, Hunan 412007, China)

Abstract: To address the limitations of existing real-time state-of-charge (SOC) prediction models for electric vehicles in terms of operational state awareness, dynamic calibration, and long-sequence forecasting accuracy, this paper proposes a temporal prediction framework that integrates a causal tree-of-thought mechanism with a deep reinforcement learning strategy. By introducing dynamic evolution and multi-branch causal inference, the proposed framework maintains the computational efficiency of a single model while enabling adaptive modeling of battery state transitions under complex operating conditions. First, a multi-level proximal policy optimization (PPO) model based on a hierarchical causal structure is designed. A time-series network is constructed as the core of the Actor network to hierarchically model the direct and indirect causal influences of key variables such as temperature and internal resistance on SOC. Through value function iteration and long-term return optimization strategies, the model continuously evolves its parameters, enhancing its generalization capability, interpretability, and causal reasoning ability. Second, a tree-of-thought structure is introduced to build a multi-path policy evaluation network, which combines policy search, path tracking, and backtracking correction mechanisms to achieve layer-wise policy optimization and anomaly branch correction under dynamic conditions. This design significantly improves the robustness and generalization performance of the model. Experimental results show that under various operating conditions, the proposed algorithm significantly outperforms Transformer, FED former, Mamba, and long short-term memory (LSTM) models across multiple evaluation metrics, achieving a mean absolute error (MAE) below 0.26%, root

mean squared error (RMSE) below 0.35%, and coefficient of determination (R^2) above 99.5%, demonstrating outstanding robustness and stability across different vehicle types.

Key words: real-time state-of-charge (SOC) prediction for electric vehicles; deep reinforcement learning; causal tree-of-thought; time series network

Foundation Item(s): Key Scientific Research Project of Hunan Provincial Education Department (No. 22A0423); Natural Science Foundation of Hunan Province (No.2023JJ60267, No.2022JJ50073)

1 引言

动力电池组作为电动汽车的核心能源装置之一,通常由数百个电池单元以串联或并联的形式组成,从而满足其行驶所需的各种动力需求^[1,2]。目前,动力电池主要以铅酸电池、镍氢电池和锂离子电池这三类二次电池为主^[3]。然而,在使用过程中,动力电池的过度充放电或其他不当使用会对电池构成严重的安全威胁,导致电池性能下降和寿命缩短,进而影响动力电池的正常工作^[4-6]。为了确保电动汽车能够更加安全、稳定地运行,对动力电池荷电状态(State-Of-Charge, SOC)进行精确的估计和监测至关重要^[7]。

目前动力电池 SOC 估算的方法主要分为两类:基于模型的方法^[8]和基于数据驱动的方法^[9]。基于模型的方法是根据动力电池组的退化和失效机理,建立相应的数学和物理模型表征电池性能退化的过程,实现对 SOC 的估计^[10]。但这种方法十分依赖模型的准确性,且构建模型和求解过程较为困难^[11,12]。基于数据驱动的方法则不需要复杂的等效模型,可以直接从电池测试数据(如电压、电流、温度、阻抗等)中挖掘隐藏的电池健康信息及其演化规律,实现动力电池 SOC 的估计^[13,14]。这种方法不仅简化了预测流程,还提高了估计精度和效率,故而得到了广泛的应用。

数据驱动的方法主要通过神经网络(Neural Network, NN)算法模型学习数据的复杂关系^[15],实现对 SOC 和电池健康状态(State Of Health, SOH)的精准估计。文献[16]提出了一种差分进化灰狼算法与长短期记忆(Long Short-Term Memory, LSTM)网络联合估计电池 SOH 的新方法,从充放电过程中提取 15 个健康指标(Health Indicators, HIs)用于 SOH 估计,并用邻域成分分析(Neighborhood Component Analysis, NCA)消除 HIs 中的冗余信息,提高 SOH 的估计精度。但上述方法仅对 SOH 进行了单一估计,忽略了容量变化对 SOH 产生的影响^[17],导致估计结果准确性降低。文献[18]提出一种具有动态映射能力的改进动态递归神经网络(Dynamic Recurrent Neural Network, DRNN)估计 SOC 和 SOH 的方法,通过利用电池的温度、老化程度等特征,并用 SOH 估计值补偿 SOC 估计误差,提高 SOC 和 SOH 的估计精度。文献[19]提出一种融合卷积神经网络(Convolutional Neural Network, CNN)与改进门控循环单

元(Gated Recurrent Unit, GRU)结构的循环门控模型,实现 SOC 与 SOH 的同步高精度估算,有效缓解梯度消失并提升计算效率,预测误差控制在 0.5% 以内。但传统时间序列模型存在梯度消失和梯度爆炸等问题,导致模型性能下降和训练困难,进而影响 SOC 与 SOH 的估计精度^[20]。文献[21]提出了一种基于混合机器学习的电池状态估计框架,通过耦合 SOC 与 SOH 的关联特性显著提升了预测精度,但双模型联合架构导致计算复杂度较高,影响了模型的实用性。

为了进一步提升 SOC 的预测性能,文献[22]提出了一种基于时频域的 CNN 来估算 SOC,使用 CNN 的二维时频域频谱图分析来提取高级信息特征,以实现更准确的 SOC 估计,不存在梯度问题。但该模型对长时间依赖的捕捉能力有限并且难以处理复杂的时序特征。为解决这一问题,文献[23]利用 Transformer 的多头自注意力机制提出了一种专门用于预测电池 SOC 的模型,该模型能够识别和强调关键数据点,并允许关注不同的序列长度和依赖关系。进一步的,文献[24]则将自监督学习(Self-Supervised Learning, SSL)与 Transformer 相结合,不仅显著减少了模型的训练时间,还提升了 SOC 的估计精度。在电池 SOH 估计方面,文献[25]提出了一种基于 Informer 模型的电动汽车 SOH 估计方法,通过 Informer 模型捕捉和理解 HIs 和 SOH 之间的全面关联性和相互影响,提高 SOH 估计的准确性。此外,一些研究者为了提升输入特征的准确性,文献[26]提出了一种混合神经网络,结合卡尔曼滤波(Kalman Filter, KF)用于降噪,稀疏自编码器(Sparse AutoEncoder, SAE)用于特征提取和降维,进而提升输入特征准确性。在联合估计方面,文献[27]提出了一种融合扩展 KF 与改进 Informer 模型的 SOC 与 SOH 联合估计方法,通过滤波预处理提升复杂工况下的估计精度与鲁棒性。而在模型可扩展性方面,文献[28]开发了一种预测预训练 Transformer 模型,通过并行计算和迁移学习降低了计算成本,提升了可扩展性和适应性。尽管上述研究所提方法有所改进,但仍面临两大挑战:现有模型难以适应严重影响驾驶能耗的环境变化,缺乏主动进化调节机制^[29], Transformer 受限于计算效率低、位置编码精度不足^[30],这些瓶颈严重影响其在不同交通路况场景中的实时预测能力^[31,32]。

针对现有技术在 SOC 估计中的局限性,本文提出一种轻量化自适应实时估计框架.该框架融合时间序列网络、深度强化学习^[33]与因果思维树机制,显著增强模型对复杂工况的感知与自适应能力,实现无须 SOC-SOH 联合建模下的高精度 SOC 实时预测.本文的主要工作与创新点如下:

(1)将时间序列预测模型嵌入强化学习的策略与网络,结合 PPO 算法,通过策略迭代优化和长时序特征学习,感知电池状态变化,使模型具备持续的主动进化能力.

(2)针对 SOC 预测中因果建模与可解释性弱的问题,本文引入因果思维树机制,构建“信息积累-策略筛选-路径追踪-回溯修正”的推理流程,动态刻画变量间相互作用对 SOC 的综合影响.

2 解决方案及理论分析

2.1 问题描述

动力电池组作为电动汽车的能量核心,其 SOC 估计的准确性直接关系到整车的续航能力与运行安全.在实际运行中,SOC 与 SOH 高度耦合,主要表现为容量衰减和内阻变化对 SOC 估计结果的影响.通常依据当前有效容量 Q_{now} 与电流积分计算 SOC 值,计算式为

$$\text{SOC}(t) = \text{SOC}(t_0) - \frac{1}{Q_{\text{now}}} \int_{t_0}^t \eta I(\tau) d\tau \quad (1)$$

其中, $Q_{\text{now}} = Q_{\text{rated}} \cdot \text{SOH}$. 可见, SOC 的准确估计依赖于 SOH 状态的准确获取,尤其在电池老化后,若仍以额定容量估算 SOC,则容易引发系统性偏差,误导能量管理策略.

在复杂工况(如大电流加减速、高湿高温)及过充(OverCharge, OC)与过放(OverDischarge, OD)状态下,电池内部会发生一系列不可逆的物理变化,尤其内阻 R_{int} 的突变性增长最显著. OC 可能导致正极结构损伤、固体电解质界面(Solid Electrolyte Interphase, SEI)膜增厚,阻碍电子传输; OD 则可能引起锂沉积、电极钝化等现象,进一步抬升内阻.这一变化将直接影响基于电压建模的 SOC 估算,其等效模型可表述为

$$\begin{cases} \Delta R_{\text{int}} = a \cdot N_{\text{cycle}}^b + \zeta \cdot \text{Dover} \\ V_t = \text{OCV}(\text{SOC}_t) - I_t \cdot R_{\text{int}} \end{cases} \quad (2)$$

其中, Dover 为 OC/OD 深度; a 、 b 、 ζ 为退化系数. 当电流 I_t 较大或 R_{int} 异常上升时,终端电压 V_t 将偏离正常水平,进而导致误判 SOC 状态. 因此,考虑 OC/OD 状态对内阻的影响是提升模型精度与鲁棒性的关键.

此外,在动力电池在复杂交通场景下,常面临频繁启停、快速加减速、多源环境干扰等动态运行条件,导致 SOC 与多种物理变量(如温度、湿度、电压、电流等)之间呈现出明显的高度非线性、时变耦合关系,使得传

统等效电路模型在建模电池真实动态行为方面存在较大误差与结构局限.近年来,基于 Transformer 等架构的数据驱动方法因具备建模复杂时序特征和长距离依赖能力,逐渐成为 SOC 估计研究的主流方向.它们通过深层提取电压、电流等传感器信号中的潜在特征,实现对部分工况下 SOC 变化趋势的精准拟合.然而,在真实驾驶场景中,动力电池状态变化通常受到交通环境、用户操作行为及工况切换等多源扰动因素的影响,导致 SOC 呈现多模态演化特性,并伴随噪声扰动和长时间依赖.当应对以下几个关键问题时,现有方法仍存在以下不足:(1)缺乏状态感知能力,难以精准识别多变量动态变化对 SOC 的综合影响;(2)动态响应能力有限,在急加速、突放电等快速变化工况下,预测滞后明显;(3)长时序建模能力不足,难以捕捉跨周期的历史依赖性.此外,当前常见的 SOC/SOH 联合估计模型结构复杂,泛化能力有限,不利于部署和实际推广.为此,本文提出一种融合因果推断思维树机制与深度强化学习框架的汽车电池 SOC 预测模型,旨在实现对状态变量间因果关系的有效建模与分阶段策略优化.同时,提升模型在多变环境下的自适应性与长期建模能力.具体内容将在后续章节中展开介绍.

2.2 解决方案

为提升电动汽车电池在复杂动态环境下的 SOC 预测精度与适应性,本文将 SOC 估计建模为马尔可夫决策过程(Markov Decision Process, MDP),引入多级近端策略优化(Proximal Policy Optimization, PPO)强化学习算法,并结合具备长程依赖建模能力的时间序列预测网络(Time Series Network, TSN),构建全新网络结构: TSN-Actor 生成动作策略, Critic 评估状态价值.该设计旨在针对当前 SOC 估计面临的三类典型问题展开建模优化.首先,传统方法在状态感知方面存在不足,往往难以全面刻画电池状态与环境变量之间的复杂关联.本文通过将多维输入状态编码为统一策略状态向量,并引入策略-价值协同更新机制,使模型能够主动感知关键变量(如温度、SOH、电流等)在不同工况下的综合影响,从而提升状态辨识能力.其次,面向动态环境的响应能力有限,传统模型难以快速适应突发变化(如急加速/急减速). TSN 中引入的蒸馏式多尺度时序编码结构提升了模型对输入序列的局部变化与全局趋势的并行理解能力,在突变工况下,能保持预测稳定性.最后,多数方法在长时序建模方面能力有限,难以捕捉 SOC 演化过程中的长距离依赖特征. TSN 网络以自注意力与门控机制为核心,强化了历史状态的记忆建模能力,配合 PPO 策略优化的时序差分更新过程,提升了对 SOC 长期趋势的建模深度.整体框架如图 1 所示. TSN 作为主干网络,不仅具备电池状态长时序特征的提取

能力,还通过策略输出与 Critic 评估协同优化,构建起具备状态感知、自适应响应、长期推理能力的强化学习式 SOC 估计体系。

具体而言,系统在 t 时刻接收状态序列 S_t ,用于初始

化强化学习模型. 动作网络输出表示 SOC 影响因子的概率分布,价值网络则基于当前策略返回状态价值,反映在动作 a_t 下,状态 S_t 到达终点所需的预期累积电耗. 理论分析如下节所述。

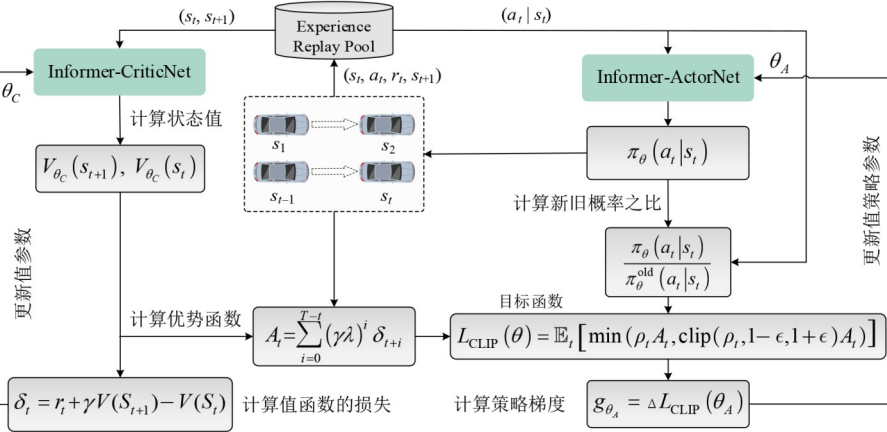


图1 网络整体框架

2.3 理论分析

为清晰地描述训练流程,本节依次介绍动作策略生成、状态价值估计、策略优化、参数更新机制以及模型收敛性分析. 首先,系统在每一时刻接收状态输入序列 S_t ,并输入至基于时间序列预测架构的 Actor 网络;其次, Critic 网络输出当前状态的价值估计,反映在采取动作策略因子 a_t 后达到终点所需的期望累积电耗;最后,将其产生的数据存储在经验池中. 不同步骤计算解释如下:

(1) 动作策略因子生成

状态 S_t 输入到共享编码器结构的 TSN-Actor 网络中生成动作分布的均值 μ_t 和方差 $\ln \sigma_t$,进而决定每一步控制策略 a_t 的取样范围,计算式为

$$\begin{cases} [\mu_t, \ln \sigma_t] = \text{TSN}_A(S_t) \cdot W_A + b_A \\ a_t \sim \mathcal{N}(\mu_t, \sigma_t^2) \end{cases} \quad (3)$$

其中, σ_t 为动作分布的标准差。

(2) 状态价值估计

依据上述所生成的策略,系统进一步利用 Critic 网络对当前状态 S_t 进行价值估计,计算其对应的状态价值函数 $V(S_t)$,计算式为

$$\begin{cases} V(S_t) = \text{Critic}_C(S_t) \cdot W_C + b_C \\ \text{SOC}_{t+1} = g(S_t, V(S_t)) \end{cases} \quad (4)$$

其中, $g(\cdot)$ 为经过训练后从价值函数到 SOC 的映射网络。

该价值函数刻画了在当前状态下,遵循当前策略 a_t 获得的期望累积收益,从而为时间差分误差 δ_t 计算与优势函数 A_t 的评估提供基础,计算式为

$$\begin{cases} \delta_t = r_t + \gamma V(S_{t+1}) - V(S_t) \\ A_t = \sum_{i=0}^{T-t} (\gamma \lambda)^i \delta_{t+i} \end{cases} \quad (5)$$

其中, r_t 为基于单位时间电能消耗设计的正向奖励函数。

(3) 策略优化模块

为在策略更新过程中兼顾稳定性与优化效率,本文采用 PPO 裁剪目标函数构建策略损失项. 策略网络根据当前策略与旧策略的概率比值 ρ_t , 结合优势函数,构建裁剪目标函数 L_{CLIP} , 计算式为

$$\begin{cases} \rho_t = \frac{\pi_{\theta}(a_t | s_t)}{\pi_{\theta^{\text{old}}}(a_t | s_t)} \\ L_{\text{CLIP}}(\theta) = \mathbb{E}_t [\min(\rho_t A_t, \text{clip}(\rho_t, 1-\epsilon, 1+\epsilon) A_t)] \end{cases} \quad (6)$$

其中, π 代表先验概率分布。

为鼓励策略的多样性探索,引入熵正则项 L^H 对策略分布的熵 $H(\pi_{\theta})$ 进行最大化,提升模型对未见状态的泛化能力. 最终综合损失函数计算式为

$$L = L_{\text{CLIP}} + c_1 (V(S_t) - \tilde{V}_t)^2 - c_2 H(\pi_{\theta}) \quad (7)$$

其中, c_1, c_2 为损失加权系数,控制各子项对最终目标的贡献。

(4) 参数更新与收敛性分析

参数更新过程基于梯度下降法,交替优化 Actor 与 Critic 两个子网络参数,计算式为

$$\begin{cases} \theta_A \leftarrow \theta_A - \alpha_A \nabla_{\theta_A} L_{\text{CLIP}} \\ \theta_C \leftarrow \theta_C - \alpha_C \nabla_{\theta_C} L_{\text{CLIP}} \end{cases} \quad (8)$$

PPO 引入概率比进而限制策略更新的比值幅度,

避免策略振荡. 由式(6)可知,该模型通过裁剪损失进行优化,故而当学习率足够小、策略 π_θ 连续可导,且当满足 PPO 的 clip 更新约束时,策略将收敛至局部最优解 θ^* ,表达式为

$$\lim_{n \rightarrow \infty} \theta_n = \theta^* \quad (9)$$

综上所述,通过上述策略-值网络的协同优化,模型能够在维持 SOC 预测物理合理性的同时,实现对不同驾驶行为、工况波动的自适应控制,使得模型可以持续地主动进化.

3 SOC 预测思维树及策略、值网络的替换

传统 PPO 算法在 SOC 预测中存在以下缺陷:全连接网络难以捕捉电池数据的长期时序特征;单一策略输出难以建模多变量间的非线性耦合关系等缺陷. 为此,本文融合时间序列预测模型的长序列建模能力与强化学习的动态决策优势,通过稀疏注意力机制高效提取电池历史数据中的长期依赖特征. 同时,引入多路径思维树推理机制,将价值学习与因果路径建模相结合,构建状态分阶段的多路径估计框架,整合多策略路径及其对应的值函数,实现对复杂状态空间中策略行为的综合评估. 具体方法如下所述.

3.1 思维树的因果推断

为解决 SOC 预测中传统深度强化学习可解释性差、因果建模能力弱的问题,本节提出基于因果分级结构的 PPO 模型与多分支因果推理网络,通过多分支策略搜索与回溯修正机制,实现策略的动态优化与调整,避免陷入局部最优. 详细介绍如下所述.

3.1.1 因果推断

传统 SOC 预测方法多依赖历史数据的统计关联,忽略变量间的因果关系,考虑到因果推断可揭示温度、湿度、电池内阻等因素对 SOC 的直接与间接影响,并指导关键因子的调节,本文提出基于因果推断的多变量模型,系统分析八类特征对 SOC 的作用路径,并构建多

级 Actor-Critic 框架结合 PPO 算法优化预测策略. 原理结构见图 2,具体方法如下.

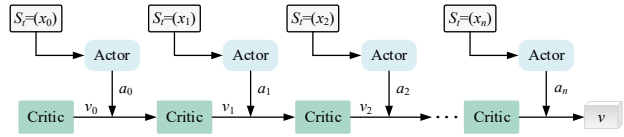


图 2 多级预测原理架构

在每一级 $n \in \{0, 2, \dots, 7\}$, 系统输入当前因果变量 x_0 与前一阶段 Critic 输出的状态值估计 V_{n-1} , 由 Actor 网络输出策略分布参数 $[\mu_n, \sigma_n]$, 以构造高斯策略采样动作 $a_n \sim N(\mu_n, \sigma_n^2)$. 随后, Critic 接收当前状态并输出阶段性的状态价值函数估计, 计算式为

$$V_n = \text{Critic}_c([x_n, V_{n-1}]) \cdot W_c + b_c \quad (10)$$

最终, 在第八级 Critic 的输出 V_7 即为整条输入路径下, 考虑因果链路后得到的最终 SOC 预测值, 计算式为

$$\text{SOC}_{t+1} = g(x_{t,7}, V_{t,7}) \quad (11)$$

3.1.2 思维树的应用和理论分析

为深入挖掘各输入变量在不同因果路径下对 SOC 预测的影响, 在前述基础上, 本文进一步构建了一种基于思维树结构的多分支推理策略网络. 该结构以 TSN 为核心, 保留 Actor-Critic 双网络框架, 融合强化学习中的策略搜索机制, 同时引入分支选择、路径追踪和回溯修正, 形成一棵因果推理思维树. 通过逐层输入特征、逐步生成策略与价值估计, 实现 SOC 预测过程中的信息累积与动态最优路径选择. 整体原理结构如图 3 所示, 具体如下展开说明.

在前节因果结构分析基础上, 本节进一步探讨融合思维树机制的动作网络 (ActorNet) 与价值网络 (CriticNet) 的运行机理. 两者协同构建自适应策略搜索树, 用于动态建模输入变量对 SOC 预测的影响. 为便于说明, 本文以其中一个阶段为例进行分析, 其协同结构如图 4 所示.

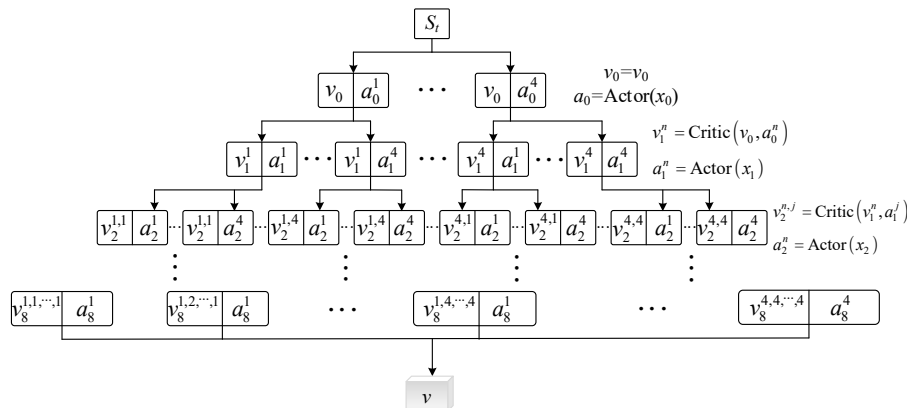


图 3 基于思维树的多分支思维结构

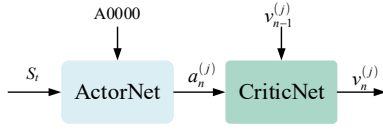


图4 协同结构原理图

该结构将八个因果输入变量定义为： $\{x_0, x_1, \dots, x_7\}$ ，每个 x_n 经由Actor网络产生1 000个动作策略因子 $\{a_n^{(j)}\}(j=1, 2, \dots, 1\ 000)$ 。为聚焦高价值信息并减少冗余路径，模型在每阶段仅保留价值最高的前4个策略，并计算对应的阶段性价值函数 $\{V_n^{(j)}\}(j=1, 2, 3, 4)$ ，以衡量当前输入对SOC预测的影响，总体计算式为

$$\begin{cases} a_n^{(j)} = \text{ActorNet}(x_n; \theta_A) \\ a_n^{(j_k)} = \text{Top-4}\left(\{a_n^{(j_k)}\}(j=1, 2, \dots, 1\ 000)\right) \\ V_i^{(j)} = \text{CriticNet}(V_{i-1}, a_i^{(j)}; \theta_C) \end{cases} \quad (12)$$

其中, $j=1, 2, \dots, 1\ 000$ 表示在该阶段下采样的策略候选动作策略因子的数目。

为进一步验证各特征的相对贡献,引入了归一化梯度加权的敏感性分析指标 ω_i ,计算式为

$$\omega_i = \frac{1}{n} \sum_{i=1}^n \left| \frac{\partial \tilde{\text{SOC}}_t}{\partial x_{i,t}} \right| \quad (13)$$

该指标衡量在全样本上输入特征 x_i 对预测结果 $\tilde{\text{SOC}}$ 的平均影响程度。通过分析可知,温度和电流和SOH的梯度响应较大,对预测贡献较显著,而湿度、OC等则更多体现为辅助调节变量。因此,本文在因果推理树的路径构建中引入优势估值机制,动态调整特征权重,计算如式(14)所示。策略网络基于当前特征 x_i 生成候选动作集,并通过价值网络评估各策略对状态值函数的影响权重 $A_n^{(k)}$,实现数据规律与物理特性的自适应建模。

$$A_n^{(k)} = Q(s_n, a_n^{(k)}) - V(s_n) \quad (14)$$

其中, $A_n^{(k)}$ 为第 k 个策略在输入特征 x_i 下的优势估计值。

这些最优分支作为下一阶段的输入继续传递,构建出逐层展开的策略搜索树,其中,第 i 层的输入为上层值函数 V_{i-1} 和当前策略 $a_i^{(j)}$,输出为当前阶段的价值函数 $V_i^{(j)}$ 。为规避把前一阶段的错误推断结论迭代到下一阶段的推理,逻辑推理过程按深度优先策略遍历逻辑树,结构原理如图5所示。

具体地,假设第 i 层当前策略路径下的最大状态价值为 V_i^{\max} ,计算式为

$$V_i^{\max} = \max_{k \in \{1, 2, \dots, 4\}} V(V_{i-1}, a_i^k) \quad (15)$$

其中, k 为当前候选策略数。

当上述最大值满足式(16)任一条件时,即判定当前路径预测的SOC结果超出物理约束边界,表明路径存在显著异常或非物理合理性,系统将触发回溯机制。

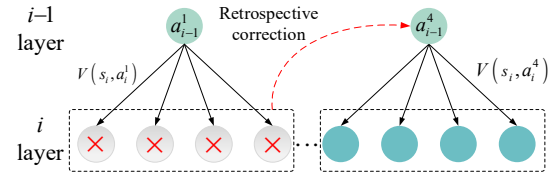


图5 回溯修正机制原理图

此时,系统将按照深度优先遍历策略回溯至第 $i-1$ 层,丢弃当前策略动作 a_{i-1}^k ,并从剩余策略候选集中重新选择新的动作 a_{i-1}^k ,构造新的路径。

$$\begin{cases} V_i^{\max} > P_{\text{SOC}} = \omega_i + \lambda \cdot \sigma_i \\ V_i^{\max} > 1 \end{cases} \quad (16)$$

其中, ω_i 为第 i 层所有候选策略对应状态值的均值; $\lambda \in (0.5, 2.0)$ 为调节因子,用于控制软边界的宽容度; σ_i 为其标准差。

该策略追踪与回溯逻辑构建了一个“逐层优化-动态纠偏”的决策闭环,使模型能够在保持整体高效搜索能力的同时,具备异常路径识别与自我修正能力,增强SOC预测路径的物理一致性与全局最优性。

最终的预测值定义为

$$\tilde{\text{SOC}} = g(V_7^{(j^*)}) \quad (17)$$

其中, j^* 为整棵树中路径代价最小且满足阈值的最优分支。

3.2 策略网络的替换

当处理电池SOC估算任务时,传统PPO算法中的策略网络存在一定局限,尤其是当面对长时间序列数据和复杂运行工况下的非线性动态特性时,难以有效建模并进行精确预测。为了解决这一问题,本文引入了具有长序列建模能力的时间序列网络TSN,结合强化学习的在线决策优化能力,利用稀疏注意力机制高效捕捉电池运行数据中的长期依赖关系,从而显著提升SOC估计在动态环境下的准确性与泛化能力。具体而言,传统的策略网络被改进的时间序列网络架构所替代,TSN结构能够更好地捕捉电池SOC变化中的长程依赖性和复杂动态行为。通过将TSN与PPO算法结合,模型能够在强化学习框架下,联合训练策略网络和价值网络,从而实现对电池SOC的精准估计和优化。该时间序列网络的编码器由多头概率稀疏自注意力机制(multi-head ProbsParse Self-Attention, PPSA)和自注意力蒸馏层构成,能够高效处理长序列数据中的信息流,并保留对重要特征的关注。整体架构如图6所示。

在该模型中,PPSA通过筛选并保留对结果影响较大的注意力分数,实现稀疏化,从而减轻计算负荷。其计算方式如下:首先,通过式(18)定义第 i 个向量查询的稀疏度量。

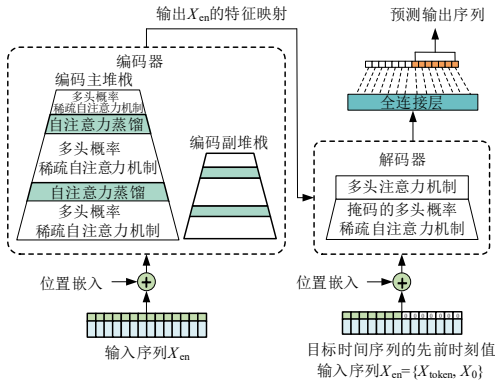


图6 时间序列网络结构图

$$\bar{M}(\mathbf{q}_i, \mathbf{K}) = \max_j \left(\frac{\mathbf{q}_i \mathbf{k}_j^T}{\sqrt{d}} \right) - \frac{1}{L_K} \sum_{j=1}^{L_K} \frac{\mathbf{q}_i \mathbf{k}_j^T}{\sqrt{d}} \quad (18)$$

$$A(\mathbf{Q}, \mathbf{K}, \mathbf{V}) = \text{softmax} \left(\frac{\bar{\mathbf{Q}} \mathbf{K}^T}{\sqrt{d}} \right) \mathbf{V} \quad (19)$$

通过式(18)计算出得分最高的 m 个主要查询向量。然后通过式(19)计算 m 个主要查询向量 \mathbf{q}_i 的注意力得分。

4 基于思维树的 SOC 预测

4.1 SOC 预测强化学习建模

为实现电动汽车 SOC 的高精度预测,本文将 SOC 建模任务视为一个 MDP 过程,形式化定义为

$$M = \langle S, A, P, R, \gamma \rangle \quad (20)$$

其中, S 为状态空间; A 为动作空间; P 为状态转移概率; R 为奖励函数; γ 为折扣因子。

相较传统方法依赖固定输入集合,本文在状态建模中引入电池系统核心特征变量,构建状态向量 \mathbf{S} 为

$$\mathbf{S}_t = [T_t, H_t, R_t, I_t, V_t, \text{SOH}_t, \text{OC}_t, \text{OD}_t] \quad (21)$$

其中, T_t 为当前温度,影响电池内部反应速率和电池内阻; H_t 为湿度,高湿可能加速老化; R_t 为内阻,决定功率损耗 ($P_{\text{loss}} = I_t^2 R_t$); I_t 和 V_t 分别为电流与电压,构成能量转移核心参数; SOH_t 表示电池健康状态,影响有效容量; OC_t 与 OD_t 为 OC、OD 的二值指标,用于刻画极端工况下对 SOC 估算的非线性影响。

将 SOC 看作上述变量的非线性函数,计算式为

$$\text{SOC}_t = f(T_t, H_t, R_t, I_t, V_t, \text{SOH}_t, \text{OC}_t, \text{OD}_t) + \varepsilon_t \quad (22)$$

其中, ε_t 为在 t 时刻的模型预测误差或不可观测扰动项。

进一步通过能量守恒模型构建变量间的定量关系。瞬时放电导致的 SOC 下降以及有效容量受 SOH 影响而动态变化的公式化表述式为

$$\begin{cases} \Delta \text{SOC} = -\frac{I_t \cdot \Delta t}{C_{\text{nominal}}} \\ C_{\text{effective}} = \text{SOH}_t \cdot C_{\text{nominal}} \\ \text{SOH}_t = \frac{Q_{\text{remain}}}{C_{\text{effective}}} \end{cases} \quad (23)$$

其中, C_{nominal} 为电池标称容量; Q_{remain} 为剩余电量。

奖励函数设计式为

$$r_t = \int_t^{t+\Delta t} \left(V_{\text{bat}} I_{\text{bat}} + \frac{F_{\text{motor}} V_t}{\eta_{\text{drivetrain}}} \right) dt \quad (24)$$

其中, V_{bat} 为电池组端电压; I_{bat} 为从电池流出的电流; $\eta_{\text{drivetrain}}$ 为传动系统效率; F_{motor} 为电动机作用于车轮的驱动力, F_{motor} 越大,说明电机输出功率越高、电流越大,电池能耗越高,从而引起 SOC 下降更快。

本研究的动作空间 a_t 不再采用传统控制量(如加速度、转矩),而是定义为一种策略因子,即在当前状态下对后续 SOC 路径选择的偏好,计算式为

$$a_t = \pi_{\theta}(S_t) \quad (25)$$

该策略由基于时间序列网络结构构建的 Actor 网络输出,用以引导多阶段价值传播。考虑 SOC 估计的多因果路径,本文引入逐级策略推理结构,每一输入维度在结构中作为一层推进节点,定义为

$$\begin{aligned} V_0 &= \text{Init}(\text{history}), V_1 = V_{\theta_c}(V_0, a_0), \\ V_2 &= V_{\theta_c}(V_1, a_1), \dots, V_7 = V_{\theta_c}(V_6, a_6) \end{aligned} \quad (26)$$

其中, V_i 为第 i 阶段由 Critic 网络估计的状态值; θ_c 为对应价值网络参数;动作策略因子 a_i 分别由第 i 维输入(如温度、湿度等)对应的 Actor 网络生成。

为提升决策效率与预测精度,每阶段从生成的 1 000 个动作中筛选策略值最高的前 4 个,避免低价值路径干扰。当预测值异常或超出 SOC 边界时,触发回溯机制调整路径。通过“策略筛选—路径追踪—回溯修正”的推理结构,结合 PPO 迭代优化,实现基于因果结构的 SOC 高精度建模。

4.2 基于思维树的 SOC 预测网络总体结构与应用

本研究针对当前 SOC 预测方法在因果关系建模、策略最优性及模型泛化能力方面的不足,提出了一种融合因果推断与强化学习的创新预测框架,如图 7 所示。

该框架以“状态—动作—价值”三位一体的建模理念为基础,结合电动汽车运行环境与电池系统的多维特征,构建逐层因果信息积累与动态路径优化的推理结构。模型采用多阶段 Actor-Critic 架构,每阶段引入一个输入特征(如温度、湿度、电池内阻、电流、电压、SOH、OC/OD 标记),由 Actor 生成候选策略, Critic 评估其对 SOC 预测的影响。从每阶段生成的 1 000 个策略中筛选前 4 个价值最高者,形成多分支策略搜索树,在路

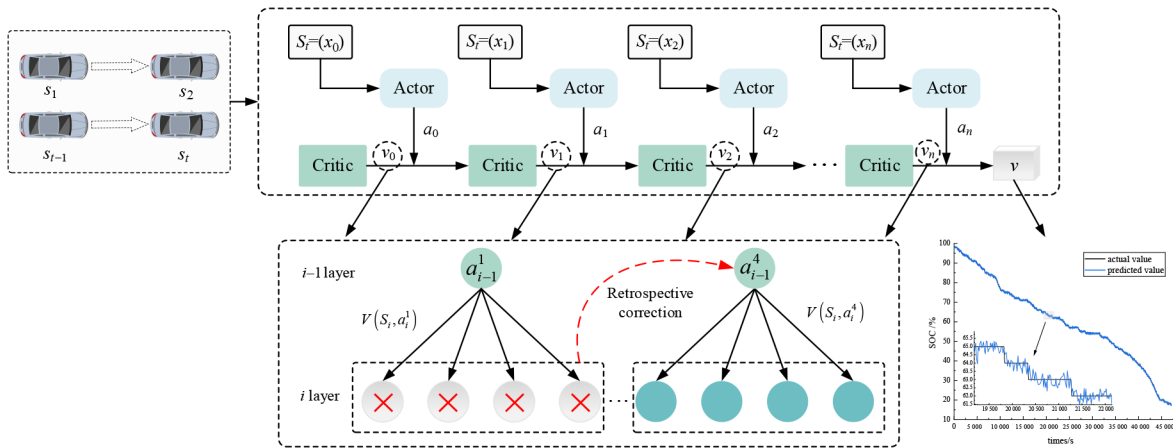


图7 总体框架

径上逐步累积因果信息与策略价值. 若路径中任一节点的SOC估值超出阈值,系统将触发回溯机制,返回上一阶段探索替代路径,以规避低质量推理分支. 最终,根据全路径估值选取最优分支,输出SOC预测结果. 该方法实现了“信息积累—策略筛选—路径追踪—回溯修正”的因果建模过程,能够动态捕捉输入变量间的联动关系及其对SOC的综合影响,提升了模型的准确性与泛化能力.

5 实验验证分析

5.1 数据处理及参数设置

5.1.1 数据预处理

为全面评估模型性能,本文采用A、B两类共10辆纯电动汽车5a的实车运行数据,涵盖城市、高速、乡村等全路况及-20~40℃环境,包含电池温度、环境湿度等多维参数,全面记录不同季节下的典型工况,具有良好的代表性与多样性,输入特征详细信息见表1.

表1 模型输入特征说明

特征名称	数据范围	说明
温度	-20~40℃	覆盖寒冷冬季至高温夏季,影响内阻与SOC反应速度
湿度	10%~90% RH	包含干燥(<30%)至潮湿(>80%)环境,体现环境对SOH演化的调节作用
电流	-1 000~1 000 A	实测最大放电/充电电流范围,包含急加速(大放电)与强回馈制动(大充电)等状态
电压	0~1 000 V	覆盖从低SOC至高SOC的全过程,受温度、电流影响显著
电阻	0.1~5 mΩ	实时估算的电芯内阻,反映衰老、温度与负载状态变化
SOH	0%~100%	不同老化阶段的电池状态,受循环次数与环境因素影响
OC标记	0/1	1表示在当前时间窗口内存在OC行为,指示异常应力状态
OD标记	0/1	1表示在当前时间窗口内存在OD行为,提示潜在容量损伤

A车型中低速工况稳定,电流波动小;B车型高速响应强,SOC波动大,二者在动力系统、电池容量与控制策略等方面存在差异. 考虑到数据来源于自然运行状态,未按照国标循环工况进行采集,如新欧洲驾驶循环(New European Driving Cycle, NEDC)、中国轻型汽车测试循环(China Light-duty vehicle Test Cycle, CLTC),本文在预处理阶段基于车辆速度、加速度与电流波动特征构建了经验规则,划分出怠速($v \approx 0$ km/h)、匀速巡航($a \approx 0$ m/s²)、加速($a > 0$)、减速($a < 0$)及停车等典型运行状态,作为弱标注工况标签,已用于后续模型训练与误差分析中,提升状态识别能力与解释性. 数据按7:2:1划分为训练集、验证集和测试集,确保模型训练、调优与评估在不同数据上独立进行,更能够准确反映实际预测能力.

为了避免特征数值范围的差异对模型训练产生倚倚,本文采用Z-Score标准化方法,将数据转化为均值为0、标准差为1的分布,计算式为

$$z_i = \frac{x_i - \mu}{\sigma} \tag{27}$$

其中, x_i 为原始数据点; μ 为均值; σ 为标准差.

5.1.2 模型参数和评价标准

本实验平台为CPU(Intel Core i5-9300H CPU@2.40 Hz)、GPU (NVIDIA GeForce GTX 1650),编译平台为Pycharm,编译器为Pytorch. 其中,学习率为0.000 1,折扣因子为0.93,初始探索率为0.65,PPSA机制头数为4,本文选取平均绝对误差(Mean Absolute Error, MAE)和均方根误差(Root Mean Square Error, RMSE)作为模型的评价指标,具体计算式为

$$\begin{cases} \text{MAE} = \frac{1}{n} \sum_{i=1}^n (|y_i - \hat{y}_i|) \\ \text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \end{cases} \quad (28)$$

其中, y_i 为真实值; \hat{y}_i 为预测值; n 为数据点的数量.

5.2 预测效果及其对比实验

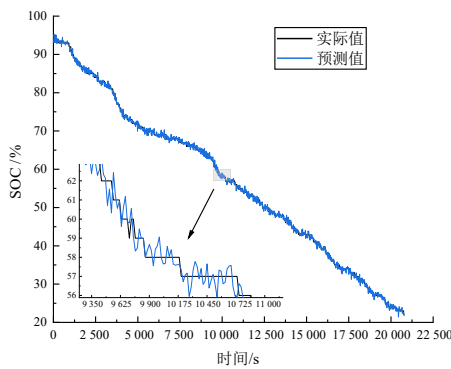
5.2.1 预测效果

为全面评估融合强化学学习与因果思维树推理预测网络实时SOC估计的性能及在不同运行条件下的泛化能力. 本研究分别选取A、B两种典型车型进行实验验证, 通过预测所得能量的消耗进而计算其对应的SOC.

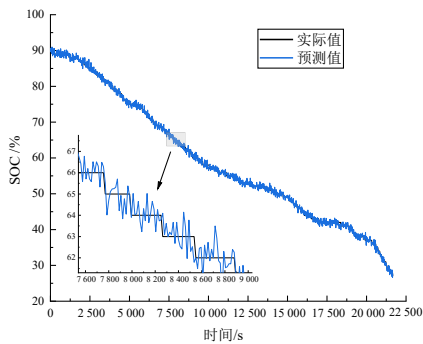
图8和图9展示了当电池SOH变化时, 该网络在两种车

型上的估计效果.

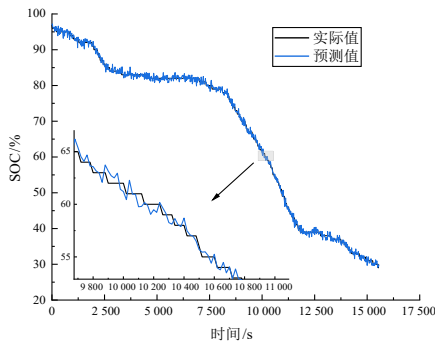
通过图8(a)和图9(a)可以直观地看出, 本文所提出的融合强化学学习与因果思维树推理的SOC预测模型在不同车型下均表现出卓越的估计能力. 模型输出的估计曲线与实际SOC曲线高度吻合, 误差极小, 充分验证了所提模型在准确性和稳定性方面的优异性能. 此外, 将图8和图9中在不同SOH条件下的估计结果对比可以发现, 即使在电池SOH发生变化的情况下, 模型依然能够保持出色的预测效果, 进一步体现了其良好的泛化能力与鲁棒性.



(a) SOH=100%时的估计曲线

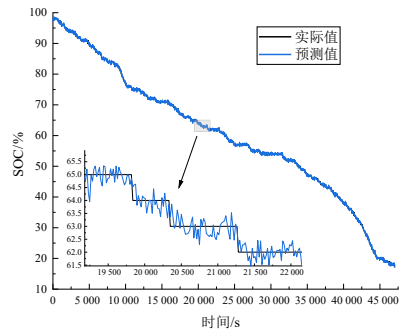


(b) SOH=90%时的估计曲线

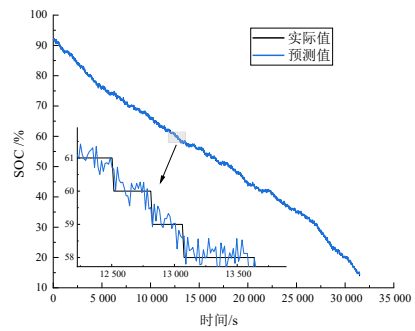


(c) SOH=80%时的估计曲线

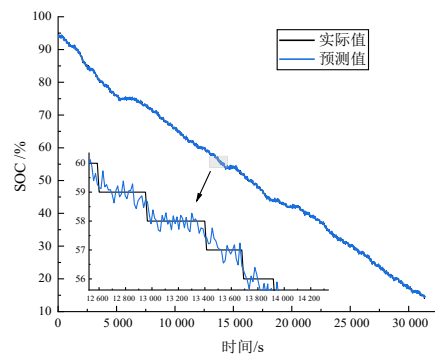
图8 车型A实时SOC程估计曲线



(a) SOH=100%时的估计曲线



(b) SOH=90%时的估计曲线



(c) SOH=80%时的估计曲线

图9 车型B实时SOC程估计曲线

为全面评估模型在不同应用场景下的预测性能, 本文在多种车型上不同SOH水平下统计分析了其关键

性能指标 MAE、RMSE 及 R^2 , 详见表 2. 由于实验室资源有限, 我们曾尝试联系多家车企获取包含动态 SOH 变化的联合数据集, 但因数据保密和对接困难未能获得支持, 故采用不同 SOH 条件下的 SOC 预测作为替代性验证方式.

表 2 不同车型估计汽车 SOC 的评价指标 单位: %

评价指标	A			B		
	MAE	RMSE	R^2	MAE	RMSE	R^2
SOH						
100	0.256	0.309	99.65	0.255	0.312	99.66
90	0.258	0.313	99.63	0.251	0.301	99.76
80	0.254	0.303	99.71	0.259	0.315	99.59

5.2.2 状态感知能力验证实验

为深入评估模型在状态感知方面的能力, 本文在全变量输入(八类特征)基础上, 设计了逐步移除关键变量的消融实验, 包括 SOH、温度与湿度、OC/OD 标志, 以及仅保留电流与电压的极简输入组合. 以 A 车 SOH=90% 的实验结果为例, 进行分析说明, 各组合在测试集上的 MAE、RMSE 与 R^2 指标如表 3 所示.

由实验结果可知, 不同输入组合对 SOC 预测性能

表 3 不同输入关键性能指标 单位: %

输入组合	MAE	RMSE	R^2
全变量	0.258	0.313	99.63
移除 SOH	0.314	0.387	97.91
移除温度与湿度	0.294	0.348	98.92
移除 OC/OD 信号	0.289	0.348	98.97
仅保留电流与电压	0.362	0.421	96.41

影响明显. 相比全变量输入 MAE 为 0.258%, R^2 为 99.63%, 移除 SOH 后的模型 MAE 上升至 0.314%, 误差增加约 21.71%, R^2 下降约 1.73%, 说明 SOH 是影响预测精度的关键变量. 移除温度与湿度使 MAE 增加了约 13.95%, R^2 下降约 0.71%, 表明环境因素对电池状态也有一定影响. 移除 OC/OD 信号的影响相对较小, MAE 增加约 12.02%, 说明该变量主要在边界状态下提供辅助决策信息. 当仅保留电流与电压作为输入时, MAE 升至 0.362%, 较全变量组合上升约 40.31%, R^2 大幅下降至 96.41%, 相比减少了约 3.23%, 显示出当缺乏其他状态变量时模型预测能力显著削弱.

综上所述, SOH 与环境变量(温度、湿度)对模型性能影响较大, 而电流电压虽为基础输入, 但难以支撑复杂状态建模. 全变量输入能最大程度地提升模型准确性与稳定性, 体现了所提方法在状态感知方面的有效性与多变量协同建模能力.

5.2.3 不同算法估计效果对比

为全面评估本文所提出的融合深度强化学习与因果思维树机制的多模态时序预测框架的综合性能, 本文选取了 Transformer、文献[18]、Mamba 以及文献[27], 这些近三年提出的代表性联合模型及其基准结构作为对比对象, 从多个维度对比验证所提方法在 SOC 预测任务中的相对优越性. 对比实验涵盖不同车型与工况变化的场景, 各算法的关键性能指标统计如表 4 所示, 系统反映了其在估计精度、鲁棒性与泛化能力方面的差异, 加粗字体表示各评价指标中的最优结果.

表 4 不同算法关键性能指标

单位: %

算法	评价指标 SOH	A			B		
		MAE	RMSE	R^2	MAE	RMSE	R^2
Transformer	100	0.432	0.515	96.11	0.434	0.523	96.31
	90	0.498	0.575	87.84	0.487	0.571	88.72
	80	0.549	0.675	81.56	0.543	0.669	82.17
文献[18]	100	0.279	0.366	99.05	0.275	0.363	99.16
	90	0.288	0.371	98.75	0.281	0.369	98.77
	80	0.337	0.401	97.79	0.331	0.398	97.88
文献[27]	100	0.296	0.341	97.64	0.293	0.337	97.77
	90	0.317	0.382	97.53	0.313	0.375	97.59
	80	0.361	0.426	96.22	0.365	0.433	96.05
Mamba	100	0.297	0.333	98.51	0.287	0.327	98.44
	90	0.398	0.457	90.07	0.381	0.451	90.61
	80	0.494	0.551	87.12	0.501	0.586	86.62
CoT-RL (本文模型)	100	0.256	0.309	99.65	0.255	0.312	99.66
	90	0.258	0.313	99.63	0.251	0.301	99.76
	80	0.254	0.303	99.71	0.259	0.315	99.59

为了更直观地展示本文模型在实际运行状态下的动态适应能力与长时序预测表现, 进一步选取 B 车型在

SOH 为 100% 与 80% 两种典型状态下的预测结果进行重点分析. 图 10 展示了各算法在不同 SOH 条件下的

SOC 估计曲线,图 11 则对比了各方法在上述工况中的关键性能指标,从定性与定量两个层面验证了本文所提出框架在复杂状态中的推理能力、估计稳定性与策略演化优势.

观察图 10(a)和图 10(b)可知,当 SOH 为 100% 时,各算法均能较好地估计实际值,但 SOH 变化后,估计精度差异显著. Transformer 的估计结果波动较大,偏离实际值较远; Mamba 虽能跟随趋势,但误差随时间累积;文献[18]和文献[27]所提模型整体稳定,能较准确地拟合衰减曲线. 本文网络始终表现最佳,预测曲线与实际值几乎重合,且波动最小、拟合度最高,细节放大图

进一步证明,其在不同工况下依然保持出色的鲁棒性和精度.

观察图 11(a)和图 11(b)可知,当 SOH 从 100% 降至 80% 时,各算法性能差异显著. Transformer 在 SOH 为 80% 时相对于 SOH 为 100% 时的 MAE 和 RMSE 分别上升约 25.12% 和 27.92%, R^2 下降约 14.68%, 适应性较差; 文献[18]和文献[27]的 MAE、RMSE 几乎无变化,但 R^2 仍有所降低,适应性有限. 相比之下,本文所提网络的实时能耗估计性能表现最优, MAE 和 RMSE 仅小幅度变化, R^2 基本保持稳定,展现出卓越的估计精度和工况适应能力.

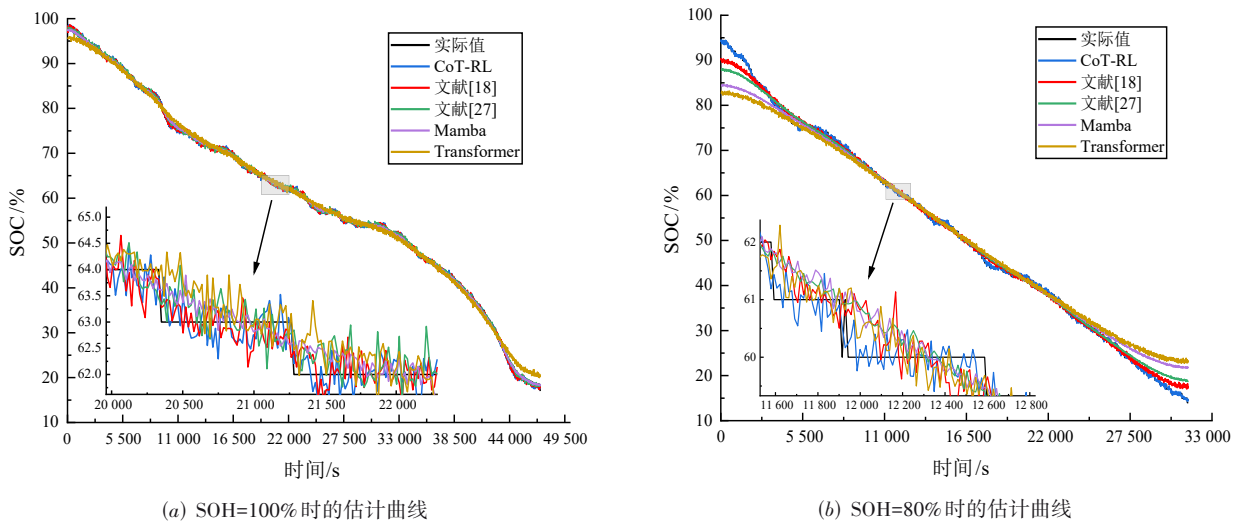


图 10 不同算法 SOC 预测所得曲线

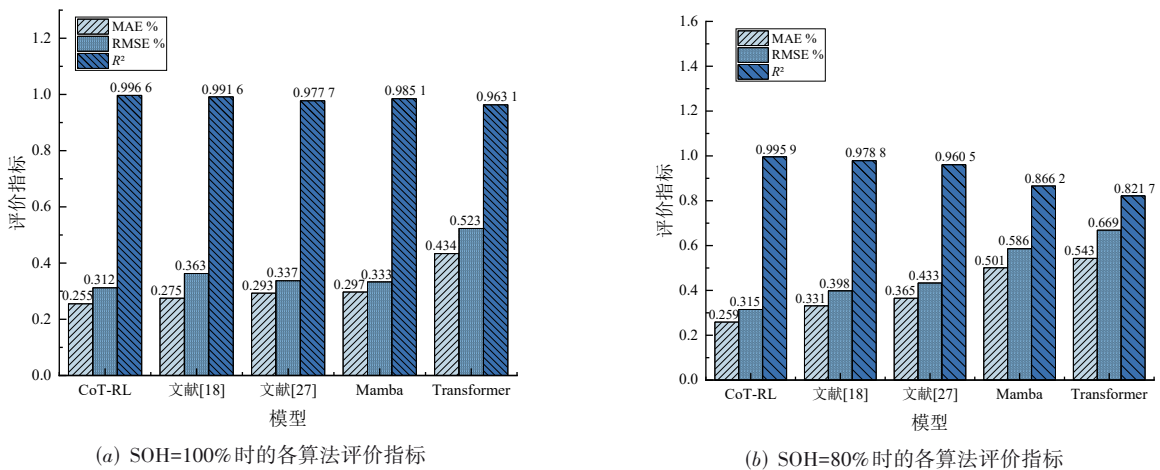


图 11 不同算法关键性能指标对比图

5.2.4 模型计算效率评估与实时性分析

为确保所提出因果思维树强化学习框架的实际部署能力,本文从推理时间、模型参数规模和内存占用三个方面对比近五年的优秀模型,结果如表 5 所示.

尽管本文模型引入了多阶段因果策略建模与回溯机制,略增加了模型复杂度,但其平均单次推理时间为 78.3 ms, 仍然满足工业应用中“软实时”响应(<100 ms)的要求. 同时,参数规模保持在 3.2 M 以内,内存占用也

表 5 模型计算效率与实时性对比评估结果

模型	平均推理时间/(ms·次 ⁻¹)	参数规模/M	1 000 次总耗时/s	平均内存/MB
CoT-RL(本文模型)	78.3	3.2	78.7	214
文献[18]	80.4	4.5	80.9	281
文献[27]	67.7	3.4	68.1	230
Mamba	29.7	1.7	29.8	177

未超过 250 MB,具备良好的可部署性.综合考虑精度提升与效率控制,本文模型在准确性与实时性之间达成了有效平衡,适用于新能源汽车能量管理等实际应用场景.

5.3 消融实验

为验证融合强化学习与因果思维树推理的动力电池 SOC 预测模型(Chain-of-Thought Reinforcement Learning, CoT-RL)在进行实时 SOC 估计时关键模块的有效性,本研究设计了消融实验.以 TSN 网络为基础,通过替换或增加创新模块,测试不同车型在不同工况下的性能指标.表 6 结果显示,各模块对 SOC 预测的贡献度存在差异,其中,加粗字体表示各评价指标中的最优

结果.

以车型 B 的实验数据为例进行具体分析.首先,为验证深度强化学习中的策略更新机制对 TSN 网络性能的优化作用,通过对比表 6 中 RL-TSN 与原始 TSN 的实验结果可知,RL-TSN 在不同 SOH 条件下始终保持稳定的 MAE、RMSE 和 R^2 表现.例如,当 SOH 从 100% 下降至 80% 时,RL-TSN 的 MAE 和 RMSE 略有下降, R^2 提升了约 0.19%;而 TSN 的预测误差均明显增加,其中 R^2 下降了约 12.16%.这一结果表明:基于 PPO 算法的值函数迭代与长期回报优化策略,通过其特有的信赖域约束与策略梯度稳定性机制,有效驱动模型参数的持续自适应进化,显著增强了 TSN 网络在动态环境中的泛化能力与鲁棒性.

表 6 消融实验结果

单位:%

改进网络	RL	CoT	评价指标 SOH	A			B		
				MAE	RMSE	R^2	MAE	RMSE	R^2
TSN	×	×	100	0.319	0.373	96.08	0.327	0.385	95.77
			90	0.428	0.497	89.75	0.425	0.323	89.83
			80	0.496	0.562	83.77	0.487	0.554	84.12
RL-TSN	√	×	100	0.311	0.369	96.14	0.326	0.376	96.17
			90	0.308	0.349	96.33	0.319	0.373	96.77
			80	0.313	0.355	95.93	0.324	0.373	96.35
CoT-TSN	×	√	100	0.287	0.348	97.11	0.291	0.353	97.09
			90	0.294	0.351	96.97	0.288	0.347	97.13
			80	0.275	0.341	97.27	0.277	0.343	97.19
CoT-RL(本文模型)	√	√	100	0.256	0.309	99.65	0.255	0.312	99.66
			90	0.258	0.313	99.63	0.251	0.301	99.76
			80	0.254	0.303	99.71	0.259	0.315	99.59

最后,为验证因果思维树机制对预测网络性能的提升作用,并进一步说明各模块联合应用的必要性,分析表 6 中的实验数据可知,与 RL-TSN 和 CoT-TSN 相比,本文提出的模型 CoT-RL 当 SOH=100% 时,MAE 分别降低了约 21.78% 和 12.37%,RMSE 分别降低了约 17.02% 和 11.61%, R^2 分别提升了约 3.63% 和 2.65%;当 SOH=80% 时,MAE 分别降低了约 20.06% 和 6.5%,RMSE 分别降低了约 15.6% 和 8.16%, R^2 分别提升了约 3.36% 和 2.47%.随着 SOH 的变化,CoT-RL 在各个 SOH 条件下均表现出最优的预测精度,其 MAE 和 RMSE 进一步下降, R^2 达到最高水平.

上述结果充分验证了因果思维树机制通过逐层引

入关键输入变量并动态评估其在因果路径中的影响力,实现了从输入特征到 SOC 输出的多路径选择与信息累积,同时也进一步印证了各优化模块联合使用对于提升模型稳健性与泛化能力的必要性和可行性.

此外,为验证模型长时序建模能力,评估其对长期依赖关系的建模效果,本文在所提出框架下设置 5 s、30 s 和 60 s 三个预测时间步任务,分别对应短期响应、中期趋势与长期预测场景.实验中采用所提出的 TSN 结构与 FED former 作为策略网络进行对比,其他网络结构保持一致,通过不同时间尺度下的预测误差变化,分析 TSN 在长时序建模方面的性能优势与稳定性.以 A 车 SOH=90% 的实验结果为例进行分析说明,各组合在

测试集上的 MAE、RMSE 与 R^2 指标如表 7 所示。

表 7 不同时间尺度下的实验结果

预测时长/s	策略网络	MAE/%	RMSE/%	R^2 /%
5	TSN	0.261	0.322	99.57
	FED former	0.281	0.325	98.92
30	TSN	0.256	0.309	99.65
	FED former	0.315	0.366	98.13
60	TSN	0.266	0.325	99.51
	FED former	0.412	0.469	94.87

由实验结果可知,随着预测时长从 5 s 延长至 60 s, TSN 结构在各时间尺度下均保持稳定的预测性能, MAE 仅从 0.261% 微降至 0.256% 再上升至 0.266%, 整体波动不足 0.01%, R^2 始终维持在 99.5% 以上. 而对比模型 FED former 的 MAE 从 0.281% 升至 0.412%, 误差增加约 46.62%, R^2 从 98.92% 下降至 94.87%, 表现出明显的性能退化. 在 60 s 预测任务中, TSN 相比 FED former 的 MAE 降低约 35.44%, RMSE 降低约 30.70%, R^2 提升 4.89%, 表明 TSN 在处理长序列输入时具有更强的稳定性和长期依赖建模能力.

综上, TSN 在不同时间尺度下均展现出优越的预测精度和鲁棒性, 验证了其作为策略网络在长时序建模任务中的显著优势, 为实现高精度 SOC 估计提供了结构支撑.

6 结论与展望

本文针对电动汽车 SOC 预测中存在的状态感知不足、动态响应能力弱及长时序建模效果差等问题, 提出了一种融合因果思维树机制与深度强化学习的预测框架. 该方法通过构建多阶段 Actor-Critic 结构, 结合 PPO 策略优化与因果思维树推理, 实现对关键变量因果路径的动态建模与策略选择. 实验结果表明: 该方法在不同 SOH 与车型工况下均表现出更高的预测精度与泛化能力, 验证了所提模型的有效性与适应性. 未来, 本文方法具备良好扩展潜力, 有望在车载边缘计算平台中应用, 构建端云协同的能量管理体系. 需要说明的是, 文中所提边缘部署设想尚处探索阶段, 未开展具体实现. 后续将重点推动模型轻量化与部署可行性研究, 并进一步引入动态复杂工况与对抗扰动实验, 提升模型在非平稳环境下的适应性与泛化能力.

参考文献

[1] 张浩, 胡昌华, 杜党波, 等. 多状态影响下基于 Bi-LSTM 网络的锂电池剩余寿命预测方法[J]. 电子学报, 2022, 50(3): 619-624.
ZHANG H, HU C H, DU D B, et al. Remaining useful life prediction method of lithium-ion battery based on Bi-

LSTM network under multi-state influence[J]. Acta Electronica Sinica, 2022, 50(3): 619-624. (in Chinese)

- [2] ZHAO L H, QIN P L. Accurate SOC prediction and monitoring of each cell in a battery pack considering various influencing factors[J]. IEEE Transactions on Industrial Electronics, 2023, 70(1): 1025-1035.
- [3] 赵轩, 李美莹, 余强, 等. 电动汽车动力锂电池状态估计综述[J]. 中国公路学报, 2023, 36(6): 254-283.
ZHAO X, LIM Y, YU Q, et al. State estimation of power lithium batteries for electric vehicles: A review[J]. China Journal of Highway and Transport, 2023, 36(6): 254-283. (in Chinese)
- [4] VICHARD L, RAVEY A, VENET P, et al. A method to estimate battery SOH indicators based on vehicle operating data only[J]. Energy, 2021, 225: 120235.
- [5] KUMAR R R, BHARATIRAJA C, UDHAYAKUMAR K, et al. Advances in batteries, battery modeling, battery management system, battery thermal management, SOC, SOH, and charge/discharge characteristics in EV applications[J]. IEEE Access, 2023, 11: 105761-105809.
- [6] ZENG X Y, SUN Y K, XIA X Y, et al. A framework for joint SOC and SOH estimation of lithium-ion battery: Eliminating the dependency on initial states[J]. Applied Energy, 2025, 377: 124624.
- [7] NG M F, ZHAO J, YAN Q Y, et al. Predicting the state of charge and health of batteries using data-driven machine learning[J]. Nature Machine Intelligence, 2020, 2(3): 161-170.
- [8] MENG J H, RICCO M, LUO G Z, et al. An overview and comparison of online implementable SOC estimation methods for lithium-ion battery[J]. IEEE Transactions on Industry Applications, 2018, 54(2): 1583-1591.
- [9] HU X S, JIANG J C, CAO D P, et al. Battery health prognosis for electric vehicles using sample entropy and sparse Bayesian predictive modeling[J]. IEEE Transactions on Industrial Electronics, 2016, 63(4): 2645-2656.
- [10] RESHMA P, MANOHAR V J. Collaborative evaluation of SoC, SoP and SoH of lithium-ion battery in an electric bus through improved remora optimization algorithm and dual adaptive Kalman filtering algorithm[J]. Journal of Energy Storage, 2023, 68: 107573.
- [11] HU L P, TANG J J, XU F Q, et al. SOC prediction for electric buses based on interpretable transformer model: Impact of traffic conditions and feature importance[J]. Energy, 2025, 324: 135935.
- [12] DEMIRCI O, TASKIN S, SCHALTZ E, et al. Review of battery state estimation methods for electric vehicles - Part I:

- SOC estimation[J]. *Journal of Energy Storage*, 2024, 87: 111435.
- [13] WANG S L, TAKYI-ANINAKWA P, JIN S Y, et al. An improved feedforward-long short-term memory modeling method for the whole-life-cycle state of charge prediction of lithium-ion batteries considering current-voltage-temperature variation[J]. *Energy*, 2022, 254: 124224.
- [14] LIN X Y, WU J Y, WEI Y M. An ensemble learning velocity prediction-based energy management strategy for a plug-in hybrid electric vehicle considering driving pattern adaptive reference SOC[J]. *Energy*, 2021, 234: 121308.
- [15] 彭自然, 许怀顺, 肖伸平. 一种基于 CatBoost 优化的光伏阵列故障诊断模型[J]. *电子学报*, 2024, 52(7): 2418-2428. PENG Z R, XU H S, XIAO S P. A CatBoost optimization-based fault diagnosis model for photovoltaic arrays[J]. *Acta Electronica Sinica*, 2024, 52(7): 2418-2428. (in Chinese)
- [16] MA Y, SHAN C, GAO J W, et al. A novel method for state of health estimation of lithium-ion batteries based on improved LSTM and health indicators extraction[J]. *Energy*, 2022, 251: 123973.
- [17] REN H B, ZHAO Y Z, CHEN S Z, et al. Design and implementation of a battery management system with active charge balance based on the SOC and SOH online estimation[J]. *Energy*, 2019, 166: 908-917.
- [18] CHE Y B, LIU Y S, CHENG Z, et al. SOC and SOH identification method of Li-ion battery based on SWPSO-DRNN[J]. *IEEE Journal of Emerging and Selected Topics in Power Electronics*, 2021, 9(4): 4050-4061.
- [19] 彭自然, 王顺豪, 肖伸平, 等. 一种精确估算电动汽车动力电池 SOC&SOH 的循环门控模型[J]. *电子测量与仪器学报*, 2024, 38(9): 11-23. PENG Z R, WANG S H, XIAO S P, et al. Cycle gating model for accurate estimation of SOC & SOH of power battery in electric vehicles[J]. *Journal of Electronic Measurement and Instrumentation*, 2024, 38(9): 11-23. (in Chinese)
- [20] PENG Z R, YANG X Y. Short- and medium-term power load forecasting model based on a hybrid attention mechanism in the time and frequency domains[J]. *Expert Systems with Applications*, 2025, 278: 127329.
- [21] LI Y H, LI K, LIU X, et al. A hybrid machine learning framework for joint SOC and SOH estimation of lithium-ion batteries assisted with fiber sensor measurements[J]. *Applied Energy*, 2022, 325: 119787.
- [22] KIM K H, OH K H, AHN H S, et al. Time-frequency domain deep convolutional neural network for Li-ion battery SoC estimation[J]. *IEEE Transactions on Power Electronics*, 2024, 39(1): 125-134.
- [23] XIONG R, CAO J Y, YU Q Q, et al. Critical review on the battery state of charge estimation methods for electric vehicles[J]. *IEEE Access*, 2017, 6: 1832-1843.
- [24] HANNAN M A, HOW D N T, HOSSAIN LIPU M S, et al. Deep learning approach towards accurate state of charge estimation for lithium-ion batteries using self-supervised transformer model[J]. *Scientific Reports*, 2021, 11: 19541.
- [25] HE Z G, NI X G, PAN C F, et al. Full-process electric vehicles battery state of health estimation based on Informer novel model[J]. *Journal of Energy Storage*, 2023, 72: 108626.
- [26] XIONG Y F, SHI Q L, SHEN L X, et al. A hybrid neural network based on KF-SA-Transformer for SOC prediction of lithium-ion battery energy storage systems[J]. *Frontiers in Energy Research*, 2024, 12: 1424204.
- [27] 彭自然, 杨肖阳, 肖伸平. 基于 EKF-HInformer 模型估计汽车动力电池的 SOC&SOH[J]. *电子测量与仪器学报*, 2025, 39(3): 21-33. PENG Z R, YANG X Y, XIAO S P. SOC and SOH of the battery are estimated based on the EKF-HInformer model[J]. *Journal of Electronic Measurement and Instrumentation*, 2025, 39(3): 21-33. (in Chinese)
- [28] ZHAO J Y, WANG Z H, WU Y Y, et al. Predictive pre-trained transformer (PPT) for real-time battery health diagnostics[J]. *Applied Energy*, 2025, 377: 124746.
- [29] ZHAO T T, LI G X, ZHAO T, et al. Learning explainable task-relevant state representation for model-free deep reinforcement learning[J]. *Neural Networks*, 2024, 180: 106741.
- [30] MA J H, DAN J P. Long-term structural state trend forecasting based on an FFT-informer model[J]. *Applied Sciences*, 2023, 13(4): 2553.
- [31] 严莉, 张凯, 徐浩, 等. 基于图注意力机制和 Transformer 的异常检测[J]. *电子学报*, 2022, 50(4): 900-908. YAN L, ZHANG K, XU H, et al. Abnormal detection based on graph attention mechanisms and transformer[J]. *Acta Electronica Sinica*, 2022, 50(4): 900-908. (in Chinese)
- [32] EDDINE M D, SHEN Y M. A deep learning based approach for predicting the demand of electric vehicle charge[J]. *The Journal of Supercomputing*, 2022, 78(12): 14072-14095.
- [33] PENG Z R, HE Z Y. Optimization of regenerative braking control strategy for dual-motor electric vehicles based on deep reinforcement learning[J]. *IEEE Transactions on Intelligent Transportation Systems*, 2025, 26(7): 10954-10967.

作者简介



彭自然 男,1969年10月出生于湖南省益阳市.现为湖南工业大学交通与电气工程学院副教授、硕士生导师.研究方向为人工智能、信号处理、智能检测仪表、智能移动终端等方面.
E-mail: pengziran@hut.edu.cn



杨肖阳 男,2000年12月出生于河南省周口市.现为湖南工业大学交通与电气工程学院硕士研究生.研究方向为人工智能、为电动汽车动力电池荷电状态和健康状态估算.
E-mail: 2873633605@qq.com