

基于多智能体强化学习的源荷分离方法

华星源¹,段思婧^{1*},崔文朋²,陈玉哲²,杨清辰²,乔楠³,任炬¹

(1. 清华大学计算机科学与技术系,北京 100084;2. 北京智芯微电子科技有限公司,北京 102200;

3. 中南大学计算机学院,湖南长沙 410083)

摘要: 近年来,全球可持续能源发展的需求日益增长,推动了光伏发电系统在居民、商业及工业场景中的应用。为了保障电网运行安全和提升能源调度效率,如何准确预测用户侧的电力负荷变化尤为关键。然而现有方法忽略了实际用户侧接入电表所采集的数据中往往同时包含光伏发电(源)和用电负荷(荷)信息,二者未被有效分离,导致预测结果存在偏差。本文提出基于多智能体强化学习的源荷分离方法 SoLED,解决了用户之间复杂时空依赖关系建模能力不足,以及负荷与光伏发电混合数据未分离导致的预测偏差的问题。具体地,本文首先设计了一种特征提取模块,能够捕捉用户负荷短期与长期尺度上的变化规律,从而提升对电网时序特征的表达能力。然后,基于物理世界的电网结构构建仿真电网环境,通过在仿真环境中模拟增强时序特征对空间关系的敏感性。最后,提出了基于多智能体强化学习的源荷分离方法,利用仿真环境生成反馈来训练解耦模型,解决了数据分离困难导致的解耦不准确问题。通过仿真实验和实际系统验证,本文提出的方法在两个城镇电网真实数据集中的不同天气条件下相较现有方法分别提升 5.7%~24.6% 和 6.6%~24.2% 的准确率。该实验结果表明,提出的方法能够准确解耦功率,有效地解决了用户之间复杂时空依赖关系建模能力不足,以及负荷与光伏发电混合数据未分离导致的预测偏差的问题。

关键词: 光伏发电系统;时空特征建模;原型学习;强化学习;多智能体强化学习;源荷解耦

基金项目: 中国电子学会智芯科研专项(No.2024-02);国家自然科学基金(No.62432004, No.62402279);教育部“基础与交叉学科突破计划”(No.JYB2025XDXM122);国家资助博士后研究人员计划(No.BX20250375);中国博士后科学基金(No.2025M771503)

中图分类号: TP39;TM715

文献标识码: A

文章编号: 0372-2112(XXXX)XX-0001-13

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

DOI: 10.12263/DZXB.20260338

Source-Load Separation Method Based on Multi-Agent Reinforcement Learning

HUA Xingyuan¹, DUAN Sijing^{1*}, CUI Wenpeng², CHEN Yuzhe², YANG Qingchen², QIAO Nan³, REN Ju¹

(1. Department of Computer Science and Technology, Tsinghua University, Beijing 100084, China;

2. Beijing Smartchip Microelectronics Technology Co., Ltd., Beijing 102200, China;

3. School of Computer Science and Technology, Central South University, Changsha, Hunan 410083, China)

Abstract: In recent years, the growing demand for global sustainable energy development has accelerated the deployment of photovoltaic (PV) systems across residential, commercial, and industrial sectors. To ensure the safe operation of power grids and improve energy dispatch efficiency, accurately forecasting user-side power load variations has been increasingly critical. However, existing methods often overlook that the data collected from user-side smart meters typically contain both PV generation (source) and power consumption (load) components, which are not effectively separated, leading to biased prediction results. To address this issue, this paper proposes a multi-agent reinforcement learning-based source-load decoupling method, namely SoLED, that mitigates prediction bias caused by the unseparated mixed data of load and PV generation, as well as insufficient modeling of complex spatiotemporal dependencies among users. Specifically, we first design a feature extraction module to capture both short-term and long-term variations in user load, thereby enhancing the representation of spatiotemporal characteristics within the power grid. Then, based on the physical topology of the grid, we construct a simulation environment that models voltage responses under different load and PV power conditions. Finally, leveraging the feedback generated from this simulation, we train a decoupling model within a multi-agent reinforcement learning framework to achieve accurate source-load separation. Experimental results demonstrate that the proposed method improves prediction accuracy by 5.7%~24.6% and 6.6%~24.2% under different weather conditions on two real-world urban power grid datasets. These results confirm the effectiveness of the proposed approach in accurately decoupling user-side mixed

power data and enhancing the modeling capability of complex spatiotemporal dependencies, thereby reducing the prediction bias in user-side load forecasting.

Keywords: photovoltaic generation system; spatiotemporal feature modeling; prototype learning; reinforcement learning; multi-agent reinforcement learning; source-load decoupling

Foundation Item(s): Smartchip Research Project of the Chinese Institute of Electronics (No.2024-02); The National Natural Science Foundation of China (No.62432004, No.62402279); The Fundamental and Interdisciplinary Disciplines Breakthrough Plan of the Ministry of Education of China (No.JYB2025XDXM122); The Postdoctoral Fellowship Program (No.BX20250375); The China Postdoctoral Science Foundation (No.2025M771503)

0 引言

近年来,全球对于可持续能源发展的需求日益增长^[1],推动了以光伏为代表的清洁能源技术的快速发展^[2]。由于部署灵活,安装便捷且具备良好的可扩展性,光伏发电系统在居民、商业及工业场景中得到了广泛应用^[3-5]。在此背景下,分布式光伏的规模不断扩大,使得如何准确预测用户侧的电力负荷和发电功率变化变得尤为关键^[6]。这不仅是保障电网运行安全^[7],提升能源调度效率的基础^[8],也对于提升可再生能源利用率,促进源网荷储协同具有重要意义^[9]。

目前,学术界将电力负荷预测问题建模成时序预测问题并开展了大量研究,主要包括深度学习预测方法以及物理模型与数据驱动模型融合方法等。例如,长短期记忆网络^[10],Transformer^[11]等深度学习模型被用于捕捉电力负荷的时间特征并预测电网中用户侧负荷。为了进一步提升预测准确性,近期工作开始利用时空特征提取模块来联合建模时间演化规律与空间拓扑依赖关系,以充分挖掘电网时空数据中的关联^[12]。

虽然现有方法在电网负荷预测上取得了一定成果,然而其普遍面临电网用户之间复杂时空依赖关系的建模能力不足,以及负荷与光伏发电混合数据未分离导致预测存在偏差两个关键问题。首先,电网系统具有明显的时序性与拓扑相关性:时序性方面,用户负荷随时间呈现周期性与非平稳性^[13-15];电网拓扑性方面,不同节点之间通过电压、电流及功率流等物理耦合关系相互影响^[11]。现有深度学习模型多采用独立样本假设,或仅在时间维度上进行序列建模,难以同时捕捉用户之间的空间关联与动态交互机制。此外,电网的实际拓扑结构往往复杂且动态变化,节点间的功率传输路径随负荷变化而改变,使得模型在建模全局时空关系时面临维度高,非线性强的问题。这导致模型在负荷突变或异常波动情况下预测偏差较大,无法准确反映用户侧功率的空间传播特性。其次,在分布式能源背景下,用户侧接入的智能电表通常只能记录到总功率信息,即同时包含光伏发电功率与用电负荷功率的叠加信号。由于光伏发电受天气、辐照度等外部条件强烈影响,其功率曲线具有高波动

性与间歇性特征,而用户用电负荷通常具有较强的周期性和行为规律性。二者混合后,信号呈现复杂的非线性特征,使得基于传统负荷预测模型的时间序列假设不再成立。若未对源(光伏)与荷(负荷)进行有效解耦,模型学习到的特征将同时反映两种不同的物理机制,导致训练目标模糊,最终引发预测偏差和模型不稳定性。特别是在晴天与阴天、工作日与节假日等条件变化显著的情况下,这种混合效应会进一步放大误差,使预测结果偏离真实用电需求。

针对上述问题,本文提出一种基于多智能体强化学习的源荷分离方法(SoLED)。为解决现有负荷预测算法难以有效解决源荷耦合问题,本文设计了原型增强的多尺度特征提取网络,以及基于多智能体强化学习的源荷分离模块,其中特征提取网络融合了分钟级和季节级尺度的电网特征,并使用季节特征对其进一步增强天气对源荷功率的影响关系;源荷分离模块在仿真电网环境中获取仿真电压反馈并利用强化学习对负荷功率与光伏发电功率进行动态解耦。主要贡献包括以下三个方面:

(1)设计原型增强的多尺度特征提取模块,有效捕捉电力负荷和光伏发电的动态变化特性的同时利用负荷功率的季节特征进一步增强特征表达能力。相比传统单尺度方法,该模块提升了混合功率信号的时序特性的表征,为高精度源荷分离与负荷预测提供基础。

(2)基于物理电网拓扑构建仿真环境,并通过多智能体强化学习实现节点间时空关联的协同建模与解耦优化。该方法充分利用节点响应电压,实现精确源荷分离,显著降低了负荷预测误差。

(3)在城镇电网真实数据集上开展验证实验,结果表明本文方法在不同天气条件下均优于现有方法,在负荷预测与功率解耦精度上取得显著提升,验证了方法的有效性,为智能电网的精准预测与能量管理提供可靠的数据基础方案。

1 相关工作

1.1 时空特征提取

电网负荷与光伏发电过程受到天气变化,用户行为及电网拓扑等多因素影响,具有明显的时空依赖特

性,因此时空特征提取通过提升模型对时序信息和环境的感知能力来增强负荷预测的准确性,并且被广泛应用于实际电网预测^[16-18]。然而在现有研究中,负荷预测方法多采用时间序列模型(如ARIMA、LSTM^[10]、GRU^[19]等)来捕捉时间维度上的变化规律,这类方法能够在单用户或局部时间范围内实现较好的短期预测性能,但难以刻画分布式电网中用户间的空间交互与关联结构。

随着对图结构化数据研究的发展,越来越多的研究开始引入空间特征提取模型,通过结合卷积神经网络(Convolutional Neural Networks, CNN)^[20]、图神经网络(Graph Neural Networks, GNN)或注意力机制(Transformer^[11]、ST-Transformer^[21]等)实现空间信息的特征建模。例如,基于图卷积网络(Graph Convolutional Network, GCN)^[22]的模型可以在一定程度上利用电网的拓扑结构捕捉邻近节点的相互影响,而基于注意力机制的模型则能够动态关注关键时刻与关键节点,从而提升全局建模能力。然而,这些方法通常假设电网节点的属性静态不变,或者仅考虑局部节点的特征交互,难以反映用户之间复杂且动态的时空依赖关系。此外,现有方法多针对单一特征(时间或空间),未能将时序特征融合到地理拓扑结构中,导致模型往往难以同时适应时间变化带来的动态性与电网拓扑带来的空间耦合性,从而在面对复杂的用户侧源荷混合场景时,预测结果容易出现偏差或不稳定。

1.2 电网负荷预测

电网负荷预测是智能电网运行与调控的核心前序任务之一,其目的是根据历史运行数据和外部环境因素,对未来的电力需求进行估计,从而迎合需求响应^[23]、电压稳定分析^[24]以及调度优化^[25]。早期研究主要采用传统统计建模方法,如自回归滑动平均模型(AutoRegressive Integrated Moving Average, ARIMA)、支持向量回归(Support Vector Regression, SVR)^[26]等,通过分析负荷的周期性与趋势性来进行短期预测。这类方法具有可解释性强、计算成本低的优点,但难以捕捉复杂的非线性关系与多变量交互特征。

随着深度学习技术的发展,研究者开始引入神经网络模型以提高负荷预测精度^[27]。循环神经网络(Recurrent Neural Networks, RNN)^[28]及其变体(如LSTM^[10]、GRU^[16])能够有效建模时间序列的长期依赖关系,在高频负荷预测中表现优异。卷积神经网络和基于注意力机制的模型(如Transformer^[11]、Informer^[29])进一步提升了模型对时序特征的提取与泛化能力。近期,很多工作设计了多种LSTM的变体来提升时序特征的表达能力^[30]。同时,为考虑用户之间的空间交互特性,近年来出现了结合图神经网络

(GNN)的时空预测方法^[31-32],通过电网拓扑关系建模节点间的能量流动,实现负荷的时空联合预测。

然而,尽管上述方法在捕捉电网负荷的时空动态方面取得了显著进展,但大多数研究仍建立在无光伏发电的场景中,即电表测得的功率数据仅来源于用电负荷,而未考虑分布式光伏发电的接入。随着实际场景中大量光伏设备的接入,用户侧电表采集的信号往往同时包含负荷功率与光伏发电功率,二者未被有效分离,导致模型在输入层即存在数据偏差,从而引发预测结果的不稳定与精度下降。

2 基于多智能体强化学习的源荷分离

为了解决时空表征不足和源荷耦合问题,本文提出了基于多智能体强化学习的源荷分离框架,主要包括三个模块:电网特征提取、仿真电网构建和源荷分离模块。整体框架如图1所示。

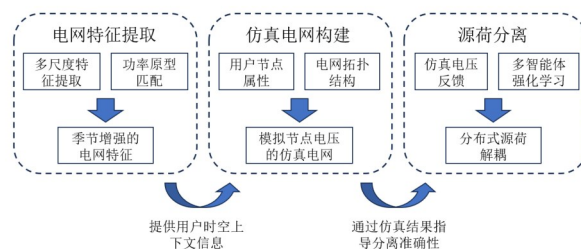


图1 基于多智能体强化学习的源荷分离的整体框架图

Figure 1 Overall framework diagram of source-load separation based on multi-agent reinforcement learning

2.1 电网特征提取

由于电网负荷在不同时间尺度上(如一天内的短时变化和全年范围内的季节性变化)通常表现出相似的周期性与波动模式,而这些相似特征在不同时间中的重复出现有利于模型捕捉负荷变化的潜在规律,从而提升预测的准确性。因此,可以设计多尺度编码模块,融合电网逆变器信息、历史天气数据和季节信息,对诸如太阳辐射、温度、湿度、风速等关键气象因素进行编码处理,提取其在短期与长期尺度上的变化特征,为后续负荷建模提供时空上下文信息支持。电网特征提取模块的框架如图2所示。

2.1.1 多尺度特征提取

首先是多尺度特征提取,目标是将电网状态、天气信息和季节信息作为输入,通过分钟级细粒度表征和季节级粗粒度表征对其进行增强,然后训练堆叠式Transformer网络^[18],以提取每个光伏逆变器的多尺度底层时间依赖关系。

在分钟级细粒度分支中,将表示电网状态的台区逆变器数据构成的电网静态特征 x_s 以及用此前 K 个

时刻的历史天气信息构成的天气动态特征 \mathbf{x}_m 拼接作为该用户的局部观测 $[\mathbf{x}_o, \mathbf{x}_m]$ 并输入 Transformer 网

络,旨在构建天气时序信息与电网状态变化的关系,生成维度为 $2h$ 的电网融合特征如下:

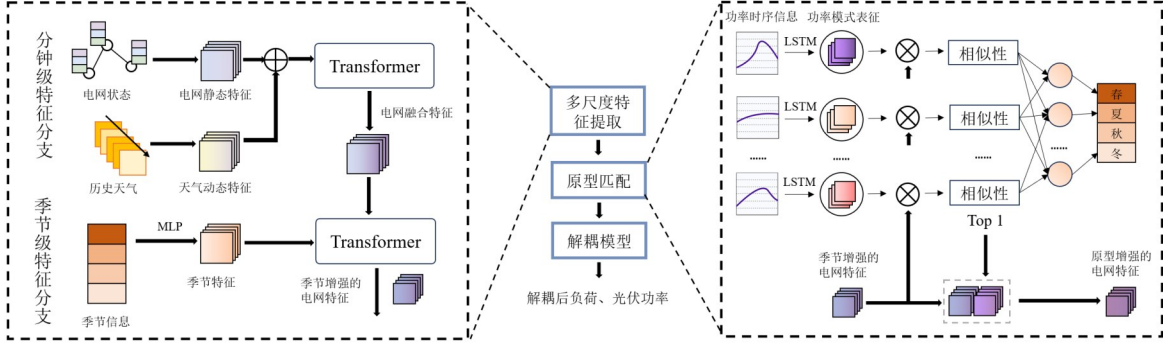


图2 电网特征提取模块的整体框架图

Figure 2 The overall framework diagram of the power grid feature extraction module

$$\mathcal{F}_m = f([\mathbf{x}_o, \mathbf{x}_m]; \psi_m) \in \mathbb{R}^{2h} \quad (1)$$

在季节级粗粒度分支中,首先根据样本日期获取其所属季节,并将其转化为维度为4的独热编码向量 $\mathbf{z} = \{0, 1\}^4$ 。通过全连接神经网络 g_z 获得维度为 h 的季节特征 $\mathbf{x}_z = g_z(\mathbf{z}) \in \mathbb{R}^h$,该神经网络用于提取不同季节下各区域的季节尺度特征。为连接 \mathcal{F}_m 和 \mathbf{x}_z ,引入线性降维层 g_h (参数为 ψ_h) 将 \mathcal{F}_m 转为维度为 h 的特征 $\widehat{\mathcal{F}}_m = g_h(\mathcal{F}_m; \psi_h)$ 。然后将降维后的电网融合特征 $\widehat{\mathcal{F}}_m$ 与季节特征 \mathbf{x}_z 拼接形成季节级输入 $[\widehat{\mathcal{F}}_m, \mathbf{x}_z]$,并输入 Transformer 网络,在电网融合特征基础上进一步提取粗粒度季节尺度的依赖关系,生成维度为 $2h$ 的最终编码的季节增强的电网特征 \mathcal{F}_z :

$$\mathcal{F}_z = f([\widehat{\mathcal{F}}_m, \mathbf{x}_z]; \psi_z) \in \mathbb{R}^{2h} \quad (2)$$

2.1.2 功率原型匹配

为了进一步增强电网特征对不同季节功率特征的敏感性(例如光伏功率由于光照强度呈现出夏季高,冬季低的特性;负荷功率由于空调等电器的使用呈现夏季低,冬季高的特性),设计了基于相似度的功率原型匹配模块。首先根据光照强度、风速等天气数据人工选取一年里最具有明显季节特征的24 d^[33],并获取这24 d内三个主要的功率状态信息(有功功率、无功功率以及功率因数)来组成功率时序信息 $\mathbf{x}_p \in \mathbb{R}^{d \times 3}$ (其中 d 表示全天的实时控制周期数)。然后将该功率时序信息 \mathbf{x}_p 输入 LSTM 模型 g_p (参数为 ψ_p) 获得对应的24个维度为 h 的功率原型特征 $\mathcal{P} = \{\mathbf{p}_1, \mathbf{p}_2, \dots, \mathbf{p}_{24} | \mathbf{p}_i = g_p(\mathbf{x}_p; \psi_p) \in \mathbb{R}^h\}$ 。为了让多尺度表征模块获取的电网特征 \mathcal{F}_z 能够感知当前季节气候所具有的功率模式,利用式(3)计算功率原型与电网特征的相似度^[34]:

$$\text{sim}(\mathbf{p}_i, \mathcal{F}_z) = \log \left(\frac{|\mathbf{p}_i - \mathcal{F}_z|_2^2 + 1}{|\mathbf{p}_i - \mathcal{F}_z|_2^2 + \epsilon} \right) \quad (3)$$

其中, ϵ 被设置为一个较小的值,以防止被零除。

基于该相似度,可以匹配相似度最高的功率原型 \mathbf{p}^* 。最后利用全连接神经网络 g_c (参数为 ψ_c) 将电网特征 \mathcal{F}_z 和匹配的功率原型 \mathbf{p}^* 融合,生成功率原型增强的电网特征 $\mathcal{F}_p = g_c([\mathcal{F}_z, \mathbf{p}^*]; \psi_c) \in \mathbb{R}^h$,该特征将接下来输入负荷光伏解耦模块进行负荷和光伏的分离。

为了训练提取功率原型的 LSTM 模型 g_p ,定义损失函数 \mathcal{L}_{pl} 为

$$\mathcal{L}_{pl} = \mathcal{L}_{cc}(c \circ g_p \circ \mathcal{F}_z, \mathbf{z}) + \lambda_1 \mathcal{L}_{clst} + \lambda_2 \mathcal{L}_{sep} + \lambda_3 \mathcal{L}_{div} \quad (4)$$

其中, c 是预测季节概率的全连接层,以功率原型作为输入并输出对应于各季节的概率分布; \mathcal{L}_{cc} 表示季节分类的交叉熵损失; \mathbf{z} 表示季节标签; $\lambda_1, \lambda_2, \lambda_3$ 是控制损失权重的超参数。这么设计的原理是如果输入的功率原型经过该全连接层能够分类季节,那么代表该功率原型携带了充分的季节语义信息。接下来分别介绍式(4)的其余三个用于辅助训练的损失函数。

首先,聚类成本 \mathcal{L}_{clst} 促使编码特征和其对应季节的一个原型最接近:

$$\mathcal{L}_{clst} = \frac{1}{24} \sum_{k=1}^{24} \min_{j: \mathbf{p}_j \in \mathcal{P}_{s_k}} \|\mathcal{F} - \mathbf{p}_j\|_2^2 \quad (5)$$

其次,分离成本 \mathcal{L}_{sep} 促使编码特征与不属于其季节的原型之间的距离更远:

$$\mathcal{L}_{sep} = -\frac{1}{24} \sum_{k=1}^{24} \min_{j: \mathbf{p}_j \in \mathcal{P}_{s_k}} \|\mathcal{F} - \mathbf{p}_j\|_2^2 \quad (6)$$

最后,多样性损失 \mathcal{L}_{div} 通过惩罚过于接近的原型来促进学习到的原型的多样性:

$$\mathcal{L}_{div} = \sum_{i=1}^4 \sum_{i \neq j} \max(0, \cos(\mathbf{p}_i, \mathbf{p}_j) - \zeta) \quad (7)$$

$\mathbf{p}_i, \mathbf{p}_j \in \mathcal{P}_i$

其中, ζ 表示表示余弦相似度的阈值; i, j, k 均表示原型序号; t 表示季节序号; $\mathbf{p}_i, \mathbf{p}_j \in P_t$ 表示所有属于第 t 个季节的原型 $\mathbf{p}_i, \mathbf{p}_j$; y_k 表示序号为 k 的原型所处的季节; $\mathbf{p}_j \in P_{y_k}$ 表示所有属于季节 y_k 的原型 \mathbf{p}_j ; $\mathbf{p}_j \notin P_{y_k}$ 表示所有不属于季节 y_k 的原型 \mathbf{p}_j 。

2.2 仿真电网构建

仿真电网构建模块用于在用户侧光伏发电与用电负荷耦合场景下,基于电力系统的物理连接关系与用户的地理分布特征,构建一个具备拓扑结构的电网图环境。同时,为后续的多智能体强化学习过程提供精确和结构化的环境建模支撑,从而在时序特征的基础上增加空间关联。其具体实现过程包括以下步骤。

首先,模块采集与电网运行相关的静态结构信息,包括用户的空间位置坐标,配电变压器的接入点,电缆线路的连接关系,以及用户的用能类别(例如居民用电、工商业用电)。这些信息能够完整刻画电网的物理结构与用户的空间分布特性,为后续的图建模提供数据基础。

然后,将每一个接入用户抽象建模为电网图中的一个节点。由于用户节点可能包含不同类型的用能与发电设备,例如纯负荷用户,光伏发电用户或光伏与负荷的混合用户,因此节点属性中必须包含用户设备类型的标识信息。同时考虑到电表在实际运行中只能采集到光伏发电与负荷叠加后的净功率信号,无法直接区分二者的贡献,故在节点属性中进一步存储历史净负荷数据,以支持后续通过原荷分离模块进行解耦。除此之外,节点属性还包括天气特征(如温度、辐射、湿度等),地理位置编号以及用户类别等多维信息,以便全面表征用户节点的运行特征。然后,节点之间的边关系则根据是否存在电缆直接相连构建,并记录两个节点的地理距离和电缆的长度作为边属性。在完成节点和边的定义后,模块将节点属性与边属性结合,形成一个完整的电网图环境,从而保证了电网图能够反映多维度的交互关系。

通过在构建的仿真电网模拟电网在不同时刻的电压,可以让提取到的时序特征捕捉到电网拓扑结构的约束作用和空间耦合效应,从而更真实地反映光伏出力变化、负荷波动以及线路阻抗对局部电压的动态影响。最终,该仿真电网不仅能够多智能体提供交互环境,用于训练其在不同电压扰动下的调控策略,还能通过仿真结果反向指导原荷分离模块和负荷预测模块的输入特征优化,为整个系统的端到端性能提升提供基础支撑。

2.3 源荷分离模块

在实际用户侧接入场景中,电表通常仅能采集到光伏发电与用电负荷叠加后的净功率信号,而无法直

接区分两者的具体数值。这种“源-荷”耦合特性使得传统负荷预测模型在训练过程中受到显著干扰,难以准确刻画负荷的真实变化规律,从而导致电网调度优化与能量管理策略出现偏差。为了解决用户侧负荷与光伏发电耦合导致预测不准确的问题,本文提出一种基于多智能体强化学习的源荷分离方法。

具体地,源荷分离模块为每个用户节点分配一个独立的智能体,智能体首先将当前时刻的天气信息以及逆变器内可获取的参数(包括电压、电流、有功和无功比例)输入特征提取模块,然后将提取的功率原型增强的电网特征 \mathcal{F}_p 输入到全连接神经网络的解耦模型(参数为 θ)并输出负荷功率 P_{load} :

$$P_{\text{load}} = f(\mathcal{F}_p; \theta) \quad (8)$$

结合逆变器上的总净功率信息 P_{total} ,计算出光伏发电功率 P_{pv} :

$$P_{\text{pv}} = P_{\text{total}} - P_{\text{load}} \quad (9)$$

将负荷和光伏功率信息输入到仿真电网中,利用潮流计算方法计算出每个逆变器节点处的仿真电压 \widehat{V}_i ,并结合真实采集的电压 V_i ,采用平方误差的相反数计算出本次预测的奖励函数 R_i :

$$R_i = -(\widehat{V}_i - V_i)^2 \quad (10)$$

该奖励函数刻画了解耦后的负荷功率与光伏发电功率在物理层面的合理程度。当仿真电压与真实采集电压越接近时,说明源荷分离模型得到的负荷与光伏功率能够更准确地反映真实电网运行状态,此时平方误差较小,奖励值更高;反之,当两者偏差较大时,奖励值显著降低,驱动智能体在后续迭代中不断调整解耦策略。

然后使用多智能体强化学习来优化源荷解耦模型。首先构建经验回放池 \mathcal{D} ^[35]。在每一次交互周期内,按时间顺序收集若干条完整轨迹,并将每步的数据写入经验回放池以供后续批量计算与多轮优化使用。经验回放池 \mathcal{D} 中每条记录包含所有智能体的状态集合 $\{s_i^t\}_{i=0}^N$ (电网特征),动作集合 $\{a_i^t\}_{i=0}^N$ (解耦负荷值),奖励集合 $\{r_i^t\}_{i=0}^N$,以及动作对应的策略概率集合 $\{\pi_\theta(a_i^t|s_i^t)\}_{i=0}^N$:

$$\mathcal{D} = \left\{ \{s_i^t\}_{i=0}^N, \{a_i^t\}_{i=0}^N, \{r_i^t\}_{i=0}^N, \{\pi_\theta(a_i^t|s_i^t)\}_{i=0}^N \right\}_{t=0}^T \quad (11)$$

训练开始时从经验回放池中采样,利用优势函数衡量经验池中解耦负荷值(动作)相对于基线(价值函数)的增益。利用广义优势估计(Generalized Advantage Estimation, GAE)^[36]中的折扣因子 γ 与平滑参数 λ 在偏差与方差之间平衡,从而显著提升策略梯度方法的训练稳定性和样本效率。具体地,首先计算每一步的时序差分残差:

$$\delta_t = r_t + \gamma V_\phi(s_{t+1}) - V_\phi(s_t) \quad (12)$$

其中, r_t 为即时奖励; V_ϕ 为价值网络估计。然后将未来若干步的 TD 残差^[37]按 $\gamma\lambda$ 指数加权得到优势 \widehat{A}_t :

$$\widehat{A}_t = \sum_{l=0}^{T-1-t} (\gamma\lambda)^l \delta_{t+l} \quad (13)$$

接下来开始进行源荷分离模型的更新。主体算法采用近端策略优化 (Proximal Policy Optimization, PPO)^[38]来更新源荷解耦模型, 核心在于构造新旧策略概率比并用裁剪操作限制每步更新幅度, 从而避免策略发生剧烈变化, 具体计算如下:

$$r_t(\theta) = \frac{\pi_\theta(a_t|s_t)}{\pi_{\theta_{old}}(a_t|s_t)} \quad (14)$$

策略损失以 PPO 目标的负值作为要最小化的损失项:

$$\mathcal{L}^{\text{PPO}}(\theta) = \mathbb{E} \left[\min \left(r_t(\theta) \widehat{A}_t, \text{clip} \left(r_t(\theta), 1 - \epsilon, 1 + \epsilon \right) \widehat{A}_t \right) \right] \quad (15)$$

然后结合熵项鼓励探索, 并使用小批量多轮次优化来提升样本利用率, 从而得到解耦模型的损失函数 \mathcal{L} :

$$\mathcal{L} = -\mathcal{L}^{\text{PPO}}(\theta) - \beta \mathbb{E}_{t \sim D} \left[\mathcal{H} \left(\pi_\theta(\cdot | s_t) \right) \right] \quad (16)$$

其中, $\mathcal{H}(\pi_\theta(\cdot | s_t))$ 为策略熵; β 为熵权重, 用以平衡探索与利用。通过计算 $\nabla_\theta \mathcal{L}$ 并采用梯度下降和 Adam 优化器更新解耦模型参数 θ 。

近端策略优化还使用了价值网络来辅助训练。价值网络用于估计状态价值, 减小回报估计的方差并作为优势函数的基线。价值网络的目标是使预测的状态价值接近期望回报, 通常采用均方误差损失进行拟合, 并在同一批次数据上与策略更新并行进行若干次小步优化。首先计算价值函数的参考值, 通常采用优势辅助计算的方式:

$$\widehat{R}_t = \widehat{A}_t + V_\phi(s_t) \quad (17)$$

然后利用参考值与价值估计值计算损失函数, 计算 $\nabla_\phi \mathcal{L}_{\text{critic}}$ 并采用梯度下降法更新价值网络参数 ϕ :

$$\mathcal{L}_{\text{critic}}(\phi) = \mathbb{E}_{(s_t, a_t) \sim D} \left[\left(V_\phi(s_t) - \widehat{R}_t \right)^2 \right] \quad (18)$$

通过上述强化学习方法, 模型能够在无显式源荷真实标签的条件下, 借助电网物理约束与电压观测信息, 实现对用户侧负荷与光伏发电功率的优化, 从而逐步提升源荷解耦结果的准确性与稳定性。

2.4 模型训练

模型训练阶段, 利用电网的网络拓扑结构构建仿真电网, 将天气数据和逆变器采集的电网数据输入到模型中, 依次进行多尺度编码, 基于相似度的功率原型匹配的处理, 利用原型学习的损失函数进行训练,

充分捕捉天气对电网用户负荷的时间依赖关系。随后, 生成负荷光伏解耦结果, 结合逆变器当前时刻的功率, 利用模型解耦的负荷功率结果计算出光伏发电功率, 并作为参数输入到仿真电网进行潮流计算模拟每个逆变器的电压值。通过比较仿真电压和真实采集电压计算源荷解耦模型的奖励函数, 并将数据保存到经验回放池中。然后利用强化学习算法分别计算解耦模型和价值评估模型的损失, 从而得到总损失函数:

$$\mathcal{L} = \mathcal{L}^{\text{PPO}} + \lambda \mathcal{L}_{\text{critic}} \quad (19)$$

其中, \mathcal{L}^{PPO} 是源荷解耦模型损失; $\mathcal{L}_{\text{critic}}$ 是价值评估函数损失; λ 是权重系数。

通过反向传播算法计算总损失的梯度, 并利用梯度下降来更新模型参数。在每个训练周期结束后, 会将当前训练周期得到的模型在验证集上进行评估, 若验证集上的性能不再提升, 则停止训练, 从而得到性能最优的模型。完整的算法伪代码如算法 1 所示。

算法 1 基于多智能体强化学习的功率解耦

输入: 电网拓扑结构, 历史电网和天气数据

利用电网拓扑结构构建仿真电网

FOR 训练轮次 i :

FOR 用户 j :

提取用户 j 的电网特征

利用特征生成预测的负荷并计算光伏功率

将所有用户的负荷功率和光伏功率输入仿真电网

获取仿真电网的电压相应值并计算每个智能体的奖励

利用强化学习算法优化所有智能体

返回智能体负荷预解耦模型

3 实验结果与分析

3.1 数据集和评价标准

本文使用的两个电网数据集均来源于我国某城镇的真实数据。第一个数据集涵盖两个时间段, 分别为 2024 年 7 月至 9 月 (以下简称 2024Q3) 和 2025 年 5 月至 7 月 (以下简称 2025Q2)。所述数据集包含两种不同类型的数据, 即天气数据和逆变器状态数据。其中, 天气数据描述的是该用户位置的天气参数 (包括温度、风向、风速、总辐射、日累计辐射、湿度), 逆变器状态数据 (包含电压、电流、总功率、功率因子等)。该数据集包含 20 个用户节点, 其中电网拓扑结构的示意图如图 3 所示。第二个数据集涵盖两个时间段, 分别为 2024 年 1 月至 3 月 (以下简称 2024Q1) 和 2024 年 5 月至 7 月 (以下简称 2024Q2)。该数据集包含 67 个用户节点。

为了验证本实例提出的用户负荷功率解耦模型的性能, 比较了在仿真电网环境中的解耦电压误差。

选取 MAE 和 RMSE 作为评价指标:

$$MAE = \frac{1}{n} \left(\sum_{i=1}^n |y_i - \hat{y}_i| \right) \quad (20)$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (21)$$

其中, y_i 代表电网中第 i 个用户的真实电压值; \hat{y}_i 代表第 i 个用户源荷分离后的仿真电压值; n 代表用户总数。

3.2 基线方法

本文采用了 7 个基线方法用于对比试验: (1) 图卷积 (Graph Convolution Network, GCN) 电网特征提取^[18]; (2) 长短期记忆 (Long Short Term Memory, LSTM) 电网特征提取^[9]; (3) 时序图混合电网特征提取 (Temporal Graph-Mixer, TGM)^[39]; (4) 多样化时空

电网特征提取 (Diversified Spatio-Temporal Feature extraction, DSFT)^[40]; (5) 深度元坐标图 (Deep Meta Coordination Graphs, DMCG)^[41]; (6) 多智能体深度确定性策略梯度 (Multi-Agent Deep Deterministic Policy Gradient, MADDPG)^[42]; (7) 多智能体近端策略优化 (Multi-Agent Proximal Policy Optimization, MAPPO)^[43]。

3.3 实现细节和实验环境

本文中所有智能体的负荷预测网络均采用三层全连接神经网络结构, 每层包含 256 个神经元。模型训练过程中使用 Adam 优化器, 学习率设为 0.001, 批量大小 (batch size) 为 64。每轮训练采样 1 000 次, 共迭代 200 个训练轮次。实验代码基于 PyTorch 2.1.0 实现, 所有实验均运行在 Ubuntu 22.04.4 LTS 系统, 配备 4 块 RTX 4090 GPU 的服务器。

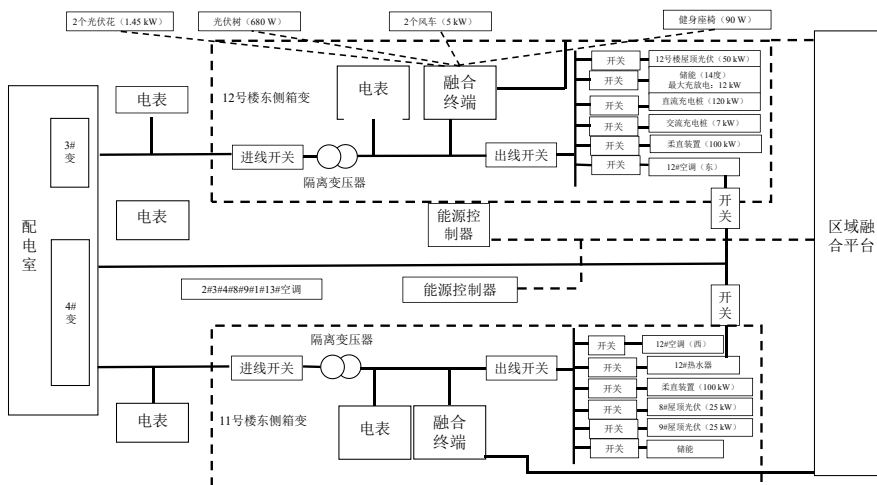


图3 我国某村镇电网拓扑结构的示意图

Figure 3 A schematic diagram of the topological structure of the power grid in a certain village in our country

3.4 实验结果

3.4.1 对比实验

从表 1 和图 4 可以看出, GCN^[18]、LSTM^[9]、TGM^[39]、MADDPG^[42]、DSFT^[39]、DMCG^[41] 和 MAPPO^[43] 等方法在不同程度上都存在性能不足。GCN 能够利用电网拓扑结构建模空间关系, 但缺乏时间特征提取能力; LSTM 能捕捉时间依赖, 却无法反映用户间的空间交互。TGM 侧重于时序建模, 能够捕获一定的时间动态特征, 但对电网的复杂空间关联刻画能力有限; DSFT 引入了时空特征融合机制, 在一定程度上增强了空间依赖建模能力, 但对源荷混合信号的解耦能力不足, 仍存在预测偏差。MADDPG、MAPPO 和 DMCG 通过多智能体机制实现了部分协同优化, 但由于未区分光伏与负荷的混合信号, 容易引入预测偏差。相比之下, 本文提出的源荷解耦模型 (SoLED) 在两个季度的数据集上均取得最优结果。这得益于模

型引入的多尺度特征提取模块与仿真电网环境, 能够同时建模时间动态与空间关联; 同时, 基于强化学习的解耦优化机制有效消除了光伏与负荷混合信号带来的干扰, 从而显著提升了预测精度和模型的稳定性。

3.4.2 消融实验

为了评估源荷解耦模型中不同组件的设计有效性, 进行了消融实验, 通过依次去除某些组件来观察对模型性能的影响。被消融的组件包括: (1) 多尺度编码; (2) 功率原型匹配; (3) 多智能体强化学习模块。结果如表 2 所示。

从表 2 中可以看出, 去除任何一个组件都会导致模型性能的下降, 尤其是去除电网特征提取模块中的多尺度编码和功率原型匹配, 显著降低了模型的预测精度, 这表明天气特征在刻画负荷的波动模式对提升预测性能至关重要。此外, 与基于多智能体强化学习

表1 2024Q3和2025Q2数据集对比实验结果

Table 1 Comparison experiment results of the 2024Q3 and 2025Q2

对比方法	2024Q3			2025Q2
	RMSE	MAE	RMSE	MAE
GCN ^[18]	7.853 2	7.142 1	7.634 3	6.949 8
LSTM ^[9]	7.473 5	6.864 2	7.251 6	6.652 7
MADDPG ^[42]	7.418 4	6.876 5	7.342 3	6.781 2
TGM ^[39]	8.663 7	7.948 4	8.519 2	7.887 2
DSFT ^[40]	8.346 3	7.709 3	8.097 7	7.472 4
DMCG ^[41]	8.361 0	7.696 6	8.162 5	7.676 5
MAPPO ^[43]	8.348 7	7.828 4	8.312 5	7.779 5
源荷解耦模型 (SoLED)	6.951 2	6.447 8	6.859 8	6.290 0

注:加粗数字表示最优结果。

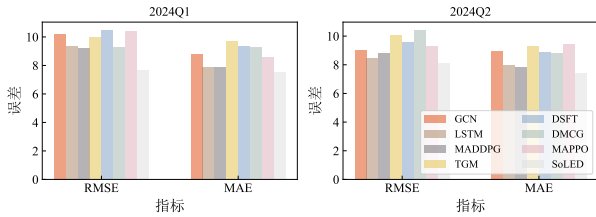


图4 2024Q1和2024Q2数据集对比实验结果

Figure 4 Comparison experiment results of the 2024Q3 and 2025Q2

表2 2024Q3和2025Q2数据集各组件的消融结果

Table 2 The ablation results of each component of the 2024Q3 and 2025Q2

消融组件	2024Q3		2025Q2	
	RMSE	MAE	RMSE	MAE
多尺度编码	7.453 2	6.847 5	7.248 9	6.652 7
功率原型匹配	7.856 4	7.152 1	7.654 3	6.948 8
多智能体强化学习	8.103 7	7.398 6	7.952 5	7.201 9
源荷解耦模型 (SoLED)	6.951 2	6.447 8	6.859 8	6.290 0

注:加粗数字表示最优结果。

的预测相比,单智能体学习由于不考虑拓扑结构的独立学习功率解耦反而降低了性能,证明了利用电网拓扑结构刻画用户之间的拓扑关系能够有效提升预测精度。

3.4.3 收敛性分析

为验证所提出的源荷解耦模型在训练过程中的稳定性与收敛特性,本节对训练阶段的平均奖励(reward)曲线进行分析。

如图5所示,所提出的多智能体强化学习源荷分离方法在训练过程中表现出良好的收敛性与稳定性。奖励曲线在初期快速上升,说明智能体能够有效利用电压物理反馈信息进行策略探索;随后奖励增长趋于平缓并最终稳定,表明策略逐步收敛至稳定区域。与对比方法相比,本方法不仅获得更高的最终平均奖励值,而且训练过程中的波动幅度很小,未出现明显震荡或策略退化现象。此外,由于奖励函数直接与电压

仿真误差相关联,奖励值的持续提升意味着解耦后的源荷功率逐渐满足电网物理约束,电压估计精度显著提高。

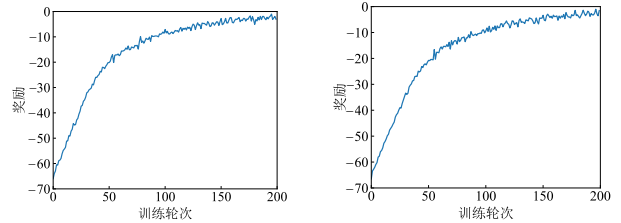


图5 训练阶段平均奖励随轮次变化

(a) The average reward variation during training in 2024Q3 (b) The average reward variation during training in 2025Q2

图5 训练阶段平均奖励随轮次变化

Figure 5 Graph of average reward during the training phase varies with the number of rounds

3.4.4 超参数敏感性分析

为验证所提出的源荷解耦模型对关键超参数的鲁棒性,分别对学习率、原型数量以及折扣因子进行敏感性分析。

从表3可以观察到,当学习率设置为0.001时,模型在两个数据集上均取得最优性能。当学习率增大至0.01或0.005时,模型误差指标明显上升,说明过大的学习率会导致参数更新步长过大,从而影响模型稳定优化;而当学习率减小至0.0005时,性能亦未进一步提升,表明过小的学习率限制了模型收敛效率,难以充分挖掘数据特征。此外,可以发现不同学习率下的性能变化趋势在两个数据集上保持一致,说明所提出方法在不同时间段数据上具有较好的稳定性和泛化能力。综合来看,学习率设置为0.001时,模型在收敛速度与优化稳定性之间取得了较优平衡,因此本文所有实验均采用该参数设置。

表3 2024Q3和2025Q2数据集学习率对算法性能的影响
Table 3 The influence of learning rates in the 2024Q3 and 2025Q2

学习率	2024Q3		2025Q2	
	RMSE	MAE	RMSE	MAE
0.01	7.482 1	6.931 5	7.365 4	6.802 7
0.005	7.103 6	6.612 8	7.021 9	6.543 3
0.001	6.951 2	6.447 8	6.859 8	6.290 0
0.000 5	7.028 4	6.501 6	6.947 3	6.362 5

注:加粗数字表示最优结果。

从表4可以观察到,随着原型数量的增加,模型性能(以电压预测误差为指标)呈现先明显提升、随后逐渐趋于平稳的趋势。这是因为当原型数量较少时,难以充分覆盖不同季节与运行工况下的功率变化模式,导致模型对多样化源荷特征的表达力不足,从而影响解耦精度;随着原型数量增加,模型能够更细粒度地刻画季节性与日内变化特征,因此性能得到提升。然而,当原型数量进一步增大时,新增原型与已有原型之间的差异逐渐减小,提供的有效信息有限,同时模型已能够较充分地覆盖主要功率模式,因此性能提升趋于饱和,不再带来显著增益。此外,过多原型还会增加模型复杂度并引入冗余特征,可能降低泛化能力并增加推理时间。

从表5可以观察到,在保持其他参数不变的情况下,我们分别设置不同折扣因子进行对比。实验结果表明,当折扣因子较小时,模型更加关注即时奖励,策略倾向于优化短期电压误差,容易忽略长期时序关联信息,导致整体分离精度较低;随着折扣因子增大,模型能够更充分地考虑长期累计奖励,从而更好地利用时序依赖关系,性能逐渐提升。然而,当折扣因子进一步接近1时,长期回报占比过高,会增加价

表4 2024Q3和2025Q2数据集原型数量对算法性能的影响
Table 4 The impact of number of proto in the 2024Q3 and 2025Q2

原型数量	2024Q3		2025Q2	
	RMSE	MAE	RMSE	MAE
6	8.950 6	8.160 2	9.734 8	8.475 6
12	7.959 8	6.745 6	7.335 3	6.908 7
24	6.951 2	6.447 8	6.859 8	6.290 0
36	6.630 0	6.241 4	6.622 2	6.289 9
48	6.432 9	5.925 6	6.502 2	5.802 6

值估计方差,使训练稳定性下降,同时性能提升趋于饱和。

表5 2024Q3和2025Q2数据集折扣因子对算法性能的影响
Table 5 The impact of discount in the 2024Q3 and 2025Q2

折扣因子	2024Q3		2025Q2	
	RMSE	MAE	RMSE	MAE
0.9	8.520 5	6.826 9	7.601 0	8.093 1
0.95	7.137 7	6.676 3	6.980 1	6.413 4
0.99	6.951 2	6.447 8	6.859 8	6.290 0
0.995	7.235 3	7.275 6	7.745 6	7.174 9

注:加粗数字表示最优结果。

3.4.5 鲁棒性分析

为评估所提方法的鲁棒性,我们分别从三个角度进行验证:(1)对节点功率数据叠加随机噪声,以反映实际测量误差。(2)在原始数据中随机移除部分时间步的功率与电压观测值,以模拟通信丢包或采样不同步情况;同时随机遮蔽部分节点或测点的数据,以模拟实际系统中测量设备缺失或局部采集失败。(3)在原始数据上添加不同强度的随机噪声以模拟真实场景中的测量误差与数据波动。实验结果如图6所示。

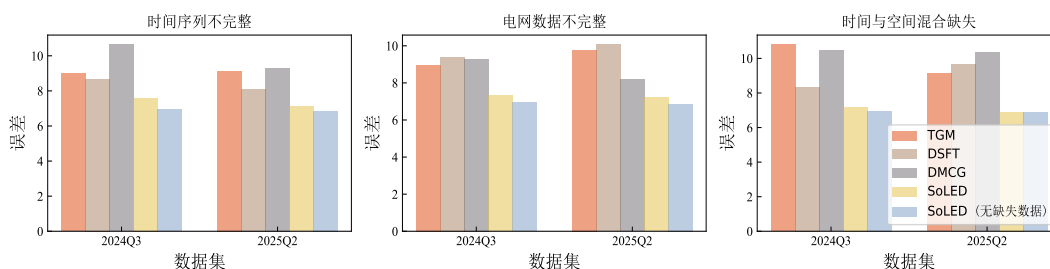


图6 2024Q3和2025Q2数据集数据缺失对算法性能的影响

Figure 6 The impact of missing data in the 2024Q3 and 2025Q2 datasets

实验结果表明,本文所提出的方法在三种实际场景中均能保持稳定性能,说明其在存在数据扰动和环境不确定性的情况下仍具有良好的鲁棒性。首先,模型并非完全依赖单一仿真结果进行决策,而是基于多时序样本以及不同源荷模式进行联合学习,通过从多

样化运行状态中提取共性特征来构建策略,从而有效降低单次仿真偏差或偶然噪声对最终决策的影响。其次,本文方法引入原型学习机制,通过聚合类别或任务的中心表示进行决策,使模型更加关注整体分布特征而非个别异常样本。这种基于中心表示的建模

方式能够在一定程度上缓解噪声样本和离群点带来的干扰,从而提升模型对数据扰动的适应能力,并赋予方法天然的鲁棒性。

3.4.6 训练与推理开销分析

为了验证模型在实际场景中的工程适用性,我们进行了模型复杂度与计算开销分析,并从训练与推理两个阶段进行评估。具体而言,我们分别统计了训练和推理流程中各模块的计算时间占比,结果如表6所示。

表6 推理和训练过程各模块时间

Table 6 Time for each module of the inference and training process

模块	时间/ms	模块	时间/ms
分钟级特征提取	368	多尺度Transformer编码	231
季节级特征提取	346	原型匹配	1 152
原型匹配	231	潮流计算	288
		多智能体强化学习	3 816

实验结果表明,训练阶段的主要计算开销来源于多智能体强化学习与潮流仿真过程。这是由于多智能体强化学习在策略优化过程中需要进行多轮环境交互与策略更新,每一轮迭代均依赖于环境反馈来评估当前策略的优劣;同时,在电网场景中,每次交互通常需要调用潮流仿真以获取系统运行状态与约束信息,该过程计算复杂度较高,从而进一步增加了整体训练时间,此外。多智能体之间的协同学习还涉及联合状态空间建模与策略收敛过程,这也在一定程度上提高了训练阶段的计算成本。

相比之下,在推理阶段强化学习策略参数已固定,不再需要进行多轮交互与策略更新,计算流程显著简化。此时模型仅需执行一次前向传播,其中主要计算开销集中在多尺度Transformer编码模块,用于提取不同时间尺度下的源荷特征并完成策略映射。然而,该部分计算复杂度相对可控,且不涉及迭代仿真过程,因此整体推理效率较高。实验结果显示,模型单次推理时间小于1 s,能够满足电网调度中分钟级甚至更短时间尺度的在线决策需求,说明所提出方法在保证性能的同时具备良好的工程应用可行性。

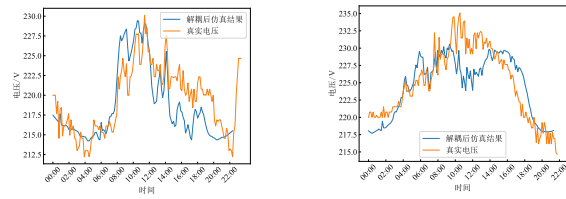
3.4.7 解耦准确性可视化

为了进一步显示本文提出方法在源荷分离上的优越性,可视化了一天內真实电压和仿真电压的变化曲线,如图7和图8所示。从图中可以看出,基于多智能体强化学习的源荷解耦模型能够很好仿真出真实电压,验证了本模型在用户侧光伏发电与用电负荷耦合场景下的源荷分离任务中的优越性。

3.4.8 真实系统应用

为了验证源荷解耦作为光伏能源调控前序步骤

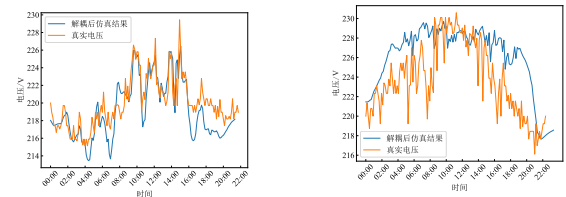
的必要性与实际价值,即通过将分布式光伏与负荷功率进行结构化分离,构建物理可解释、因果关系清晰的数据基础,从而提升后续电网能源调控模型的训练效率与控制性能。在实验中,我们选取实际运行台区的历史量测数据,对总功率进行源荷解耦处理,并基于解耦后的有功扰动特征与电压响应构建训练数据集;随后利用该数据训练光伏调控模型,并与直接基于原始耦合数据训练的模型进行对比分析,最终将模型部署至实际运行环境(如图9所示)进行在线验证。



(a) 夏季某日仿真电压与真实电压随时间变化图
(b) 冬季某日仿真电压与真实电压随时间变化图
(a) The variation of simulated voltage and actual voltage over time on a certain summer day
(b) The variation of simulated voltage and actual voltage over time on a certain winter day

图7 模型在2024Q3数据集仿真电压与真实电压随时间变化图

Figure 7 The variation of simulated voltage and actual voltage in 2024Q3



(a) 夏季某日仿真电压与真实电压随时间变化图
(b) 冬季某日仿真电压与真实电压随时间变化图
(a) The variation of simulated voltage and actual voltage over time on a certain summer day
(b) The variation of simulated voltage and actual voltage over time on a certain winter day

图8 模型在2025Q2数据集仿真电压与真实电压随时间变化图

Figure 8 The variation of simulated voltage and actual voltage in 2025Q2

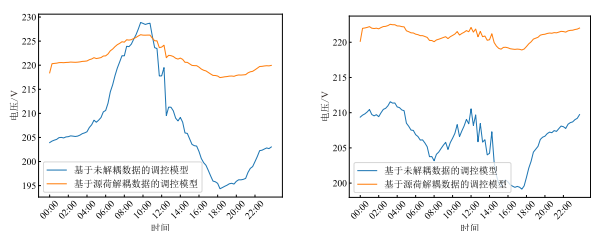
如图10所示,实验结果表明,基于源荷解耦数据训练的调控模型在实际运行中表现出更高的稳定性与控制精度。与未解耦方案相比,系统电压越限次数明显下降,光伏消纳率得到提升,同时调节动作更加平滑,避免了频繁波动带来的设备冲击。在高光伏出力及快速波动场景下,模型能够更准确识别源侧扰动趋势并提前调整控制策略,实现电压安全与新能源消



图9 调控模型在智能电表终端的实际部署示意图

Figure 9 The schematic diagram of the actual deployment of the regulation model in the smart electricity meter terminal

纳之间的有效协调。上述结果验证了源荷解耦作为光伏调控前序步骤在工程应用中的有效性与必要性。



(a) 9月某日调控后电压随时间变化图
(a) The variation of voltage over time after the regulation on a certain date in September
(b) 10月某日调控后电压随时间变化图
(b) The variation of voltage over time after the regulation on a certain date in October

图10 使用源荷解耦与未解耦数据训练的调控模型电压随时间变化
Figure 10 The voltage of the regulation model trained using source-load decoupling and non-decoupling data changes over time

4 结论

本文针对用户侧光伏发电与用电负荷耦合导致的源荷解耦困难问题,提出了一种基于多智能体强化学习的源荷分离方法。本方法利用多尺度特征提取以及功率原型匹配,有效捕捉了负荷与光伏功率的复杂时空依赖关系。通过引入多智能体强化学习框架,本方法实现了节点级源荷分离策略的自适应优化,使各用户智能体能够在电网整体约束下协同学习最优分离策略。实验结果表明,该方法在不同天气条件下均表现出较高的负荷与光伏功率解耦精度,显著优于传统方法。同时,利用解耦后数据训练的调控策略展现出更好的稳定性与鲁棒性,说明源荷解耦能够为电网能源调控提供更加清晰可靠的数据基础,为构建高效稳定的调控模型提供了一种新的思路。

参考文献

[1] Detchon R, Van Leeuwen R. Policy: Bring sustainable energy to the developing world[J]. Nature, 2014, 508(7496):

309-311.
[2] Batzelis E, Rather Z H, Barton J, et al. Solar integration in the UK and India: technical barriers and future directions[PP/OL]. V1. arXiv(2021-04-01)[2023-06-15]. <https://doi.org/10.17028/rd.lboro.14453133.v1>.
[3] 杨峰, 刘心松, 邱元杰, 等. 电网调度自动化的任务服务模式 and 调度算法研究[J]. 电子学报, 2002, 30(12): 1811-1815.
Yang Feng, Liu Xinsong, Qiu Yuanjie, et al. Study on task service model and scheduling algorithm in electric power automatization[J]. Acta Electronica Sinica, 2002, 30(12): 1811-1815. (in Chinese)
[4] 于昕妍, 沈艳霞, 陈杰, 等. 考虑概率区间的微电网短期负荷多目标预测方法[J]. 电子学报, 2017, 45(4): 930-936.
Yu Xinyan, Shen Yanxia, Chen Jie, et al. A multi-objective prediction method for short-term microgrid load considering interval probability[J]. Acta Electronica Sinica, 2017, 45(4): 930-936. (in Chinese)
[5] 王后能, 殷劲骁, 廖小兵, 等. 微电网并网切换系统混合樽海鞘优化控制策略[J]. 电子学报, 2025, 53(11): 4091-4103.
Wang Houneng, Yin Jingxiao, Liao Xiaobing, et al. Hybrid salp swarm optimization control strategy for microgrid grid-connected and islanded mode switching system[J]. Acta Electronica Sinica, 2025, 53(11): 4091-4103. (in Chinese)
[6] Chen X, Qu G N, Tang Y J, et al. Reinforcement learning for selective key applications in power systems: Recent advances and future challenges[J]. IEEE Transactions on Smart Grid, 2022, 13(4): 2935-2958.
[7] 吴忠强, 程洪强. 网络攻击下考虑状态受限的微电网安全运行与控制[J]. 电子学报, 2024, 52(9): 3240-3250.
Wu Zhongqiang, Cheng Hongqiang. Safe operation and control of microgrid considering state constraints under network attacks[J]. Acta Electronica Sinica, 2024, 52(9): 3240-3250. (in Chinese)
[8] Xu F Y, Liu S Y, Qing Y P, et al. Temporal prototype-aware learning for active voltage control on power distribution networks[C]//Proceedings of the 30th ACM SIGKDD Conference on Knowledge Discovery and Data Mining. New York: ACM, 2024: 3598-3609.
[9] Liu S Y, Luo W, Zhou Y Z, et al. Transmission interface power flow adjustment: A deep reinforcement learning approach based on multi-task attribution map[J]. IEEE Transactions on Power Systems, 2024, 39(2): 3324-3335.
[10] Shi X J, Chen Z R, Wang H, et al. Convolutional LSTM Network: A machine learning approach for precipitation

- nowcasting[C]//Proceedings of the 29th International Conference on Neural Information Processing Systems - Volume 1. New York: ACM, 2015: 802-810.
- [11] Vaswani A, Shazeer N, Parmar N, et al. Attention is all you need[C]//Proceedings of the 31st International Conference on Neural Information Processing Systems. New York: ACM, 2017: 6000-6010.
- [12] Wu Z H, Pan S R, Chen F W, et al. A comprehensive survey on graph neural networks[J]. *IEEE Transactions on Neural Networks and Learning Systems*, 2021, 32(1): 4-24.
- [13] Masters C L. Voltage rise: The big issue when connecting embedded generation to long 11 kV overhead lines[J]. *Power Engineering Journal*, 2002, 16(1): 5-12.
- [14] Yang G Y, Marra F, Juamperez M, et al. Voltage rise mitigation for solar PV integration at LV grids: Studies from PVNET. dk[J]. *Journal of Modern Power Systems and Clean Energy*, 2015, 3(3): 411-421.
- [15] Agalgaonkar Y P, Pal B C, Jabr R A. Distribution voltage control considering the impact of PV generation on tap changers and autonomous regulators[J]. *IEEE Transactions on Power Systems*, 2014, 29(1): 182-192.
- [16] Huang J, Wei Q, Wang T K, et al. Spatio-temporal feature fusion-based hybrid GAT-CNN-LSTM model for enhanced short-term power load forecasting[J]. *Energies*, 2025, 18(21): 5686.
- [17] Su Z E, Zheng G Q, Hu M S, et al. Short-term load forecasting of regional integrated energy system based on spatio-temporal convolutional graph neural network[J]. *Electric Power Systems Research*, 2024, 232: 110427.
- [18] Cheng M, Yu J W, Wu M K, et al. A spatiotemporal deep learning framework for joint load and renewable energy forecasting in stability-constrained power systems[J]. *Information*, 2025, 16(8): 662.
- [19] Chung J, Gulcehre C, Cho K, et al. Empirical evaluation of gated recurrent neural networks on sequence modeling[PP/OL]. V1. arXiv (2014-12-11) [2023-06-15]. <https://doi.org/10.48550/arXiv.1412.3555>.
- [20] LeCun Y, Bottou L, Bengio Y, et al. Gradient-based learning applied to document recognition[J]. *Proceedings of the IEEE*, 1998, 86(11): 2278-2324.
- [21] Liu Y C, Zhang J H, Fang L J, et al. Multimodal motion prediction with stacked transformers[C]//2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2021: 7573-7582.
- [22] Kipf T N, Welling M. Semi-Supervised Classification with Graph Convolutional Networks[C]//International Conference on Learning Representation. Washington DC: ICLR, 2017.
- [23] Javed F, Arshad N, Wallin F, et al. Forecasting for demand response in smart grids: An analysis on use of anthropologic and structural data and short term multiple loads forecasting[J]. *Applied Energy*, 2012, 96: 150-160.
- [24] Yuan Y, Yuan X P, Wang H Y, et al. Net load forecasting method in distribution grid planning based on LSTM network[J]. *Science and Technology for Energy Transition*, 2024, 79: 57.
- [25] Kumar N. Optimized power distribution using machine learning for load forecasting, fault detection, and voltage regulation[J]. *International Journal of Research and Review in Applied Science, Humanities, and Technology*, 2024: 94-102.
- [26] Awad M, Khanna R. Support vector regression[M]//Efficient Learning Machines: Theories, Concepts, and Applications for Engineers and System Designers. Berkeley, CA: Apress, 2015: 67-80.
- [27] Wang G, Xie X Y, Zhang F, et al. bCharge: Data-driven real-time charging scheduling for large-scale electric bus fleets[C]//2018 IEEE Real-Time Systems Symposium. Piscataway: IEEE, 2018: 45-55.
- [28] Zaremba W, Sutskever I, Vinyals O. Recurrent neural network regularization[PP/OL]. V5. arXiv (2015-02-19)[2023-06-15]. <https://doi.org/10.48550/arXiv.1409.2329>.
- [29] Zhou H Y, Zhang S H, Peng J Q, et al. Informer: Beyond efficient transformer for long sequence time-series forecasting[J]. *Proceedings of the AAAI Conference on Artificial Intelligence*, 2021, 35(12): 11106-11115.
- [30] Lu S, Bao T T. Short-term electricity load forecasting based on NeuralProphet and CNN-LSTM[J]. *IEEE Access*, 2024, 12: 76870-76879.
- [31] Jia Y L, Wang J, Hosseini M R, et al. Temporal graph attention network for building thermal load prediction[J]. *Energy and Buildings*, 2024, 321: 113507.
- [32] Orhi U, Guven C, Stowell D. Enhanced load forecasting with GAT-LSTM: leveraging grid and temporal features[PP/OL]. V1. arXiv (2025-02-12)[2023-06-15]. <https://doi.org/10.48550/arXiv.2502.08376>.
- [33] Yarats D, Fergus R, Lazaric A, et al. Reinforcement learning with prototypical representations[C]//Proceedings of the 38th International Conference on Machine Learning. New York: PMLR, 2021:11920-11931.
- [34] Zhang Z X, Liu Q, Wang H, et al. ProtGNN: Towards self-explaining graph neural networks[J]. *Proceedings*

of the AAAI Conference on Artificial Intelligence, 2022, 36(8): 9127-9135.

- [35] Mnih V, Kavukcuoglu K, Silver D, et al. Human-level control through deep reinforcement learning[J]. Nature, 2015, 518(7540): 529-533.
- [36] Schulman J, Moritz P, Levine S, et al. High-dimensional continuous control using generalized advantage estimation[PP/OL]. V6. arXiv (2018-10-20) [2023-06-15]. <https://doi.org/10.48550/arXiv.1506.02438>.
- [37] Fujimoto S, Van Hoof H, Meger D. Addressing function approximation error in actor-critic methods[C]//Proceedings of the 35th International Conference on Machine Learning. New York: PMLR, 2018: 1587-1596.
- [38] Schulman J, Wolski F, Dhariwal P, et al. Proximal policy optimization algorithms[PP/OL]. V2. arXiv (2017-08-28)[2023-06-15]. <https://doi.org/10.48550/arXiv.1707.06347>.
- [39] Bilal M, Lopez L C. Temporal Graph MLP Mixer for Spatio-Temporal Forecasting[PP/OL]. V1. arXiv (2025-01-17)[2023-

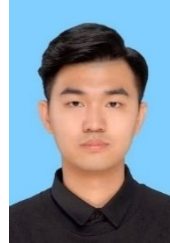
06-15]. <https://doi.org/10.48550/arXiv.2501.10214>.

- [40] Wang X, Wang X J, Huang F L, et al. DSTF: A Diversified Spatio-Temporal Feature Extraction Model for traffic flow prediction[J]. Neurocomputing, 2025, 621: 129280.
- [41] Gupta N, Hare J Z, Kannan R, et al. Deep Meta Coordination Graphs for Multi-agent Reinforcement Learning[PP/OL]. V1. arXiv (2026-02-10) [2023-06-15]. <https://doi.org/10.48550/arXiv.2502.04028>.
- [42] Wu J, Huang Z X. Promoting diversity in mixed complex cooperative and competitive multi-agent environment[C]//Proceedings of the 32nd ACM International Conference on Information and Knowledge Management. New York: ACM, 2023: 4355-4359.
- [43] Bayen A, Gao J X, Velu A, et al. The surprising effectiveness of PPO in cooperative multi-agent games[C]//Advances in Neural Information Processing Systems 35. Neural Information Processing Systems Foundation, Inc. (NeurIPS), 2022: 24611-24624.

作者简介



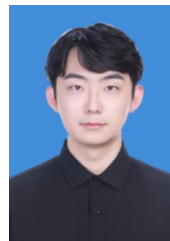
华星源 男,2002年5月出生于北京市。现为清华大学计算机科学与技术系博士生。主要研究方向为强化学习、智能体后训练。
E-mail: huaxy24@mails.tsinghua.edu.cn



杨清辰 男,1998年5月出生于山东省东营市。现为国家电网北京智芯微电子科技有限公司工程师。主要研究方向为人工智能芯片及其应用。
E-mail: qingchen.yang98@gmail.com



段思婧 女,1994年7月出生于湖南省娄底市。现为清华大学计算机系助理研究员。主要研究方向为移动计算、边缘智能、具身智能。
E-mail: duansj@tsinghua.edu.cn



乔楠 男,1999年11月出生于山西省阳泉市。现为中南大学计算机学院博士研究生。主要研究方向为强化学习、大模型训练和端侧智能决策。
E-mail: nanqiao.ai@gmail.com



崔文朋 男,1987年5月出生于山东省济宁市。现为国家电网北京智芯微电子科技有限公司人工智能事业部总经理助理。主要研究方向为人工智能芯片及其应用。
E-mail: cuiwenpeng@sgchip.sgcc.com.cn



任炬 男,1987年12月出生于湖南省汨罗市。现为清华大学计算机系副教授。主要研究方向为端侧智能与边缘智能。
E-mail: renju@tsinghua.edu.cn



陈玉哲 男,1999年3月出生于山东省青岛市。现为国家电网北京智芯微电子科技有限公司大模型应用开发组副组长。主要研究方向为人工智能芯片及其应用。
E-mail: chenyzhe1@sgchip.sgcc.com.cn