

基于深度学习 - 强化学习联合算法的电力调度信息化择优决策

白桂君¹ 张 静¹

BAI Guijun ZHANG Jing

摘 要

针对现有决策方法在应用到电力调度中,存在消纳率过低,造成能源过度浪费的问题,文章开展了一种基于深度学习-强化学习联合算法的电力调度信息化择优决策研究。通过构建双向长短期记忆网络负荷预测模型精准捕捉源荷时序特征,结合改进近端策略优化算法动态优化调度策略,实现电力系统安全约束、经济运行与可再生能源消纳的多目标协同优化。通过专家策略预训练与强化学习在线调参的联合训练流程,提升算法收敛速度。通过对比实验证明,相较于现有决策方法,新的方法在不同可再生能源出力场景下均表现出更强的适应性,特别是在高波动工况中能有效平衡系统安全性与消纳率。

关键词

深度学习;联合算法;信息化;电力调度;强化学习

doi: 10.3969/j.issn.1672-9528.2025.12.022

0 引言

电网运营中传统“源随荷动”的单向运行模式逐步向“源荷互动”的双向协作模式过渡,电网耦合关系更为复杂,运行控制难度大幅提升。在此背景下,传统调度方式已很难应对这种动态、不确定环境下的动态、不确定因素,迫切需要引入智能技术来提高决策水平。

目前,电力系统中的人工智能技术已经得到了很大的发

展。在此基础上,李志华^[1]采用数据挖掘、机器学习等方法,对电网调度方案进行优化,以提高电网运行效率。但目前的研究主要集中在静态数据上,不能很好地反映电网的动态演变特征,也不能适应实时互动决策的需求。在此基础上,刘冉等^[2]研究新能源电力系统的可信计算与前瞻性决策,并在此基础上,引入可信评价机制,提升新能源电力系统的健壮性。然而,该方法过于依赖于离线优化,很难应对电网实时负荷变化及源-荷交互的随机特性,并且计算量大,难以用于在线调度。

1. 国网青海省电力公司信息通信公司 青海西宁 810000

参考文献:

- [1] 余传旗,郭海燕,王婷婷,等.基于图注意力网络和门控网络的轻量级单通道语音分离方法[J].信号处理,2025,41(4):706-717.
- [2] 褚俊佟,魏爽.基于深度学习的交叉残差连接网络应用于语音分离[J].上海师范大学学报(自然科学版中英文),2025,54(2):229-237.
- [3] 王树斌.NMF声场分离技术在语音识别中的优化研究[J].电声技术,2025,49(1):68-70.
- [4] 刘子寒,沈力,奚梦婷,等.基于支持向量机的复杂场景中多人对话语音智能识别方法研究[J].计算技术与自动化,2024,43(4):59-65.
- [5] 江楠,陈洁,肖潘,等.基于声纹识别的电力会议多角色语音的分离和识别研究[J].高电压技术,2023,49(S1):40-46.
- [6] 吴亮,王甲祥,施汉琴,等.基于多尺度自适应注意力机制的视听语音分离[J].人工智能,2024(3):1-14.
- [7] 屠彦辉,霍伟明,高建清,等.基于多模态波束方向特征的

多模语音分离及识别[J].人工智能,2024(3):36-44.

- [8] 孙林慧,王春艳,张蒙.基于全卷积神经网络多任务学习的时域语音分离[J].信号处理,2024,40(12):2228-2237.
- [9] 刘宏清,谢奇洲,赵宇,等.基于扩张卷积和Transformer的视听融合语音分离方法[J].信号处理,2024,40(7):1208-1217.
- [10] 李可.面向语言对话场景的智能语音交互关键技术研究[J].自动化与仪器仪表,2023(8):295-299.
- [11] 陈佳佳,张海剑,华光.基于Conformer的时域多通道语音分离方法[J].无线电工程,2023,53(9):2054-2060.

【作者简介】

赵成宝(1989—),通信作者(email:male202209@163.com),男,甘肃灵台人,本科,研究方向:会议通信。

张中海(1995—),男,甘肃天水人,本科,研究方向:会议通信。

(收稿日期:2025-06-17 修回日期:2025-11-28)

针对上述问题, 本文将开展基于深度学习 - 强化学习联合算法的电力调度信息化择优决策研究。

1 电力负荷预测

针对大规模新能源消纳导致的电力系统强随机、非线性等特点, 提出一种基于时间序列分析和回归模型的电力负荷预测新方法^[3]。在用电负荷预测时, 深度神经网络模型必须与电网的实际特点相结合, 对其进行优化。LSTM的门控机理可以很好地反映出负荷曲线上的周期变化, 例如, 在周末时, 用户的用电行为会发生变化。考虑到新能源接入电网所引起的波动性, 本项目将着重研究在光伏发电功率骤降、风力发电逆调峰等情景下的负载响应特征。针对传统RNN的梯度消失问题, 改为LSTM预测, 以此通过门控机制(遗忘门、输入门、输出门)控制信息流动^[4]。

遗忘门表示为:

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f) \quad (1)$$

式中: f_t 表示遗忘门; σ 表示 Sigmoid 激活函数; W_f 表示遗忘门权重矩阵; h_{t-1} 表示 $t-1$ 时刻隐藏状态; x_t 表示 t 时刻输入数据; b_f 表示遗忘门的偏置项。将历史负荷数据(如 96 点日负荷曲线)与气象数据(温度、湿度)作为输入矢量 x_t , 前一时段隐藏状态 h_{t-1} 初始化全零矩阵。权重矩阵 W_f 学习负荷周期特征。通过 sigmoid 输出 $[0,1]$ 值, 决定保留多少昨日同期负荷特征

输入门表示为:

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \quad (2)$$

式中: i_t 表示 t 时刻输入门输出结果; W_i 表示权重系数; b_i 表示偏置量; x_t 表示当前时刻光伏出力骤降事件编码为的特殊维度; h_{t-1} 表示矩阵学习逆调峰场景下的特征权重; σ 表示激活函数。

候选记忆表示为:

$$\tilde{C}_t = \tanh(W_C \cdot [h_{t-1}, x_t] + b_C) \quad (3)$$

式中: $\tilde{C}_t$ 表示 t 时刻候选细胞状态; h 表示状态系数; W_C 表示候选记忆计算的权重系数; b_C 表示候选状态系数; $\tanh$ 表示使用激活函数生成 $[-1,1]$ 的候选值, 刻画负荷突变方向。该公式表示学习气象 - 负荷非线性关系。

更新记忆表示为:

$$C_t = f_t \odot C_{t-1} + i_t \odot \tilde{C}_t \quad (4)$$

式中: C_t 表示 t 时刻细胞状态; C_{t-1} 表示 $t-1$ 时刻细胞状态; i_t 表示 t 时刻输入门输出; $\odot$ 表示哈达玛积。当 C_t 变化率超过阈值触发预警信号。

输出门表示为:

$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \quad (5)$$

式中: o_t 表示 t 时刻输出门输出结果; W_o 表示权重系数; b_o

表示输出门状态系数。

W_o 学习不同时间尺度特征, 在需求响应时段强制输出门值 ≥ 0.7 以保持策略一致性。

隐藏状态表示为:

$$h_t = o_t \odot \tanh(C_t) \quad (6)$$

式中: h_t 表示 t 时刻的隐藏状态。

将 h_t 接入全连接层输出预测负荷值, 采用 Pinball Loss 函数处理新能源场景下的非对称误差, 当预测误差连续 3 次超阈值时触发 LSTM 参数在线微调。结合单相 LSTM 时仅利用历史信息, 而采用双向 LSTM (BiLSTM), 则可以同时考虑历史信息和未来上下文信息, 以此可以有效提升预测结果的精度。BiLSTM 的隐藏状态可根据前向 $\vec{h}_t$ 和后向 $\overleftarrow{h}_t$ LSTM 隐藏状态拼接得出:

$$h_t = [\vec{h}_t, \overleftarrow{h}_t] \quad (7)$$

前向 LSTM 提取历史负荷趋势, 后向 LSTM 学习未来 24 h 天气预报特征。BiLSTM 在负荷突变场景下的预测误差较传统 LSTM 降低约 15%。气象、电价、节假日等都会影响到负荷, 因此需要构建更多变量的输入模型。同时, 通过注意力机制的运用, 实现动态加权, 进一步提高预测鲁棒性。在实际应用中, 需要对春节等重要节假日前、节假日前、后两个时间段的数据进行预测, 而双向 LSTM 由于可以有效地融合前后两个时间段的上下文信息, 其预报精度可提高 15% 以上。在此基础上, 本文提出基于视觉注意机制的建模方法, 实现对台风过程中降雨强度、风速等重要特征的动态提取, 提高对极端天气事件的预报稳健性。这种面向电力负荷特性的建模方法, 使得神经网络不仅能拟合常规负荷变化规律, 更能适应新型电力系统下的复杂运行场景。

2 建立电力调度信息化择优决策模型

在保证安全性的基础上, 通过对机组功率进行优化, 减少电网运营成本, 推动新能源消纳。传统的基于静态优化的调度模型很难有效地应对源 - 荷交互作用下的动态变化。在此基础上, 将深度学习和增强学习相结合, 建立基于信息的优选决策模型^[5]。在电力系统中, 包含火电机组、可再生能源机组, 则将最小化总运行成本作为调度目标, 得到电力调度信息化择优决策模型:

$$\min C = \sum_{t=1}^T \left(\sum_{i=1}^{N_g} c_i(P'_{g,i}) + \lambda \cdot \max(0, P'_{\text{load}} - P'_{\text{total}}) \right) \quad (8)$$

式中: $\min C$ 表示以最小化总运行成本为目标的电力调度信息化择优决策模型函数; $P'_{g,i}$ 表示火电机组 i 在时刻 t 的出力; c_i 表示成本函数; λ 表示惩罚系数; P'_{load} 表示负荷需求; P'_{total} 表示总出力。

在构建上述电力调度信息化择优决策模型时, 需要严格

遵循其物理和运行约束条件，从而确保电力系统运行的安全性和可靠性。首先，电力系统在任何时刻必须满足发电总量与负荷需求严格匹配，即所有火电机组、可再生能源机组的出力总和必须等于当前负荷需求，以此可以有效避免发电过剩或不足导致的频率波动或停电风险^[6]；其次，每台发电机组的实际出力需在其技术允许范围内，火电机组的出力不得低于最小技术出力，也不得超过额定容量，以此可以有效防止机组过载损坏或低效运行；最后，火电机组相邻时刻的出力调整幅度受限于最大爬坡能力，即当前出力与前一刻出力的差值绝对值不得超过机组最大爬坡速率，以此可以有效避免机组因出力突变导致机械应力损伤或燃烧不稳定。

3 基于深度学习 - 强化学习联合算法的决策模型求解

强化学习（RL）是一种通过与外界交互来获得优化策略的方法，但存在搜索效率低下、收敛缓慢等问题。本文将模仿学习（IL）和近端策略优化（PPO）相结合，研究改进的DRL算法。利用历史调度数据，对卷积神经网络进行训练，并根据训练结果生成初始策略网络。专家策略数据集表示为：

$$D = \{(s_1, a_1), (s_2, a_2), \dots, (s_N, a_N)\}$$

(9)

式中： D 表示专家策略数据集； $s_1, s_2, \dots, s_N$ 表示包含支路电流、负荷、可再生能源出力的参数； $a_1, a_2, \dots, a_N$ 表示机组出力调整量。

在对获取到的专家策略数据集进行训练时，主要获取电网历史运行数据，其中包含了机组出力参数、负荷曲线参数等。在此基础上，标注专家调度动作，并构建状态 - 动作配对数据集。可引入一维的卷积神经网络，实现从学习状态到动作状态的映射^[7]。将最小化预测动作与专家动作的均方误差作为训练目标，在训练结果满足目标后，生成初始策略网络参数，作为强化学习的起点输出。在强化学习在线优化的过程中，策略网络根据当前电网状态生成调度动作。在完成动作后，通过电网模拟器，返回新状态及奖励信号。将交互数据（状态、动作、奖励）存入经验回放缓冲区，每间隔一定步数，从缓冲区当中获取数据，并计算优势函数，更新策略网络。最后，将所建模型应用于实际电网调度，实现对源荷变化的实时响应。通过不断更新的数据，对模型进行调整，使其能够与电网结构和可再生能源的渗透率变化相适应。根据上述操作，完成对决策模型的求解，按照模型输出完成对电力调度的信息化择优决策。

4 对比实验

4.1 实验环境与参数配置

为验证本文提出的基于深度学习 - 强化学习联合算法的决策方法在电力调度当中应用的优越性，设计如下对比

实验：

将本文决策方法设置为实验组，将基于大数据的决策方法设置为对照A组，将基于可信计算和主动决策的决策方法设置为对照B组。实验采用IEEE 118节点系统作为测试基准，该实验环境中包含120个节点，85个负荷点，发电机数量为52台，其中火电机组为32台，可再生能源机组为16台，平衡机组为4台，输电线路为182条。在实验过程中，每分钟记录一次系统的运行状态，共采集18 000组历史数据，并按照8:2的比例，将其划分为训练数据集和测试数据集。在三组决策方法训练时，按照表1中内容完成对训练参数的设置。

表1 三组决策方法训练参数配置

方法类别	关键参数	训练方式
本文方法	LSTM 层数=2, PPO clip=0.2, 学习率=0.001	预训练（专家数据） + 在线强化学习
对照 A 组	树深度=10, 特征数=304	离线训练，静态策略
对照 B 组	可信阈值=0.85, 优化步长=5 min	在线优化， N-1 故障检测

在完成对3组决策方法的训练后，实验过程中分别设置3种不同的产出率条件作为实验变量，分别为低产出率、中产出率和高产出率，将3种产出率条件配置情况记录如表2所示。

表2 三种产出率条件配置表

产出率等级	出力范围（占额定容量）	典型场景描述	火电机组补偿策略	测试时长/h
低产出率	0%~40%	夜间 / 极端天气（如无风、阴天）	火电机组按最小技术出力运行，平衡机激活	8
中产出率	40%~60%	常态波动（如间歇性风速 / 部分日照）	火电机组动态调节，爬坡约束生效	12
高产出率	60%~100%	理想发电条件（强风 / 晴天）	火电机组深度调峰，优先消纳可再生能源	24

在上述配置的3种实验条件下，分别记录3种决策方法应用下电力调度结果消纳率。消纳率越高说明相对应决策方法对能源的适应性越强，尤其是在波动性强的中高产出率场景当中。

4.2 实验结果

根据上述内容，将得到的实验结果绘制如图1~3所示。

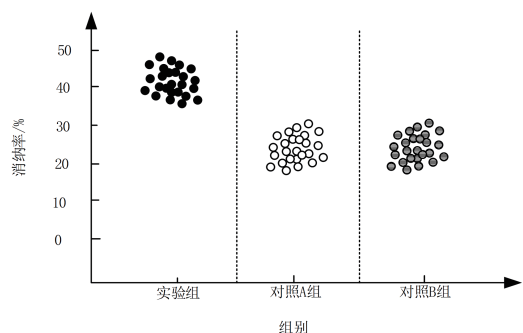


图1 低产出率条件下消纳率对比

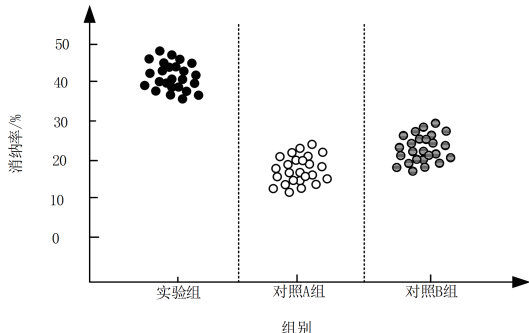


图2 中产出率条件下消纳率对比

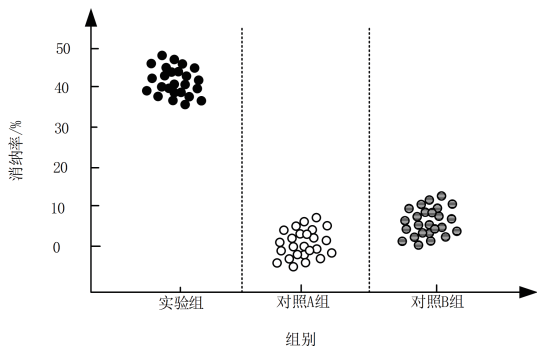


图3 高产出率条件下消纳率对比

从图1~3所示的实验结果可以看出，3种产出率条件下，实验组的消纳率始终高于对照A组和对照B组，且高产出率条件下，实验组与对照A组和对照B组的差距更明显。因此，通过这一实验结果可以看出，实验组决策方法具备对能源调度的更强适应性。同时，本文决策方法应用下，由于引入了深度学习-强化学习联合算法，因此可以实现实时学习优化策略，而对照A组和对照B组无法灵活匹配源荷变化。综合分析得出，本文决策方法具备更高的实际应用价值。

5 结语

本文提出了一种创新的电力调度信息化择优决策方法，该方法基于深度学习与强化学习的联合算法。通过精心设计的实验对比发现，在不同产出率条件下，该方法在消纳率方面均表现出色，显著高于现有的其他方法。这一结果充分证明了所提出方法的高应用价值。具体来说，本文所提出的电

力调度信息化择优决策方法，通过深度学习算法对大量历史数据进行分析和学习，能够准确预测电力系统的运行状态和需求变化。同时，结合强化学习算法，该方法能够在不断变化的电力市场环境中，实时调整调度策略，以达到最优的电力分配效果。这种联合算法的应用，不仅提高了电力系统的运行效率，还有效提升了可再生能源的利用率。

将本文提出的方法应用于实际电力调度中，可以显著提高电网对可再生能源的响应能力，减少对传统化石能源的依赖。这不仅有助于降低电力系统的碳排放，还能推动整个电力行业向绿色、低碳的方向发展。通过有效利用可再生能源，能够减少对环境的污染，促进可持续发展。然而，本文的研究也存在一些不足之处。主要问题在于，该方法在实际应用中需要依赖高质量的策略数据。然而，获取这些高质量数据的难度较大，这在一定程度上限制了方法的广泛应用。针对这一问题，将在未来的研究中进行更深入的探索。通过不断优化和改进，希望能够进一步提升该方法的实际应用效果，为电力系统的智能化和绿色化发展做出更大的贡献。

参考文献：

- [1] 李志华. 基于大数据的电力调度决策支持系统研究[J]. 通信电源技术, 2024, 41(14):55-57.
- [2] 刘冉, 汪凯威, 常星, 等. 基于可信计算和主动决策的新能源电力系统日前调度方法[J]. 河南科学, 2023, 41(2):157-163.
- [3] 仪忠凯, 梁寿愚, 王伟, 等. 电力系统调度决策：一种示教学习辅助加速的安全强化学习方法[J]. 中国电机工程学报, 2024, 44(13):5084-5096.
- [4] 崔杨, 张聪, 赵钰婷, 等. 计及碳捕集电厂调峰定价策略的多主体电力系统日前-实时灵活调度[J]. 中国电机工程学报, 2023, 43(18):7043-7060.
- [5] 吴任博, 刘淑琴. 考虑市场交易的新型电力系统调度辅助决策方法研究[J]. 电测与仪表, 2023, 60(11):81-89.
- [6] 张玉敏, 孙鹏凯, 吉兴全, 等. 基于并行多维近似动态规划的综合能源系统动态经济调度[J]. 电力系统自动化, 2023, 47(4): 60-68.
- [7] 蔡国伟, 姜雨晴, 黄南天, 等. 电力需求响应机制下基于多主体双层博弈的规模化电动汽车充放电优化调度[J]. 中国电机工程学报, 2023, 43(1): 85-99.

【作者简介】

白桂君（1993—），女，青海西宁人，本科，中级工程师，研究方向：电力信息通信运行维护、调度运行管理等。

张静（1993—），女，青海西宁人，本科，中级工程师，研究方向：电力信息通信运行维护、电力通信检修管理。

（收稿日期：2025-06-23 修回日期：2025-12-02）