

# 基于改进深度强化学习的电网电力智能调度分析模型研究

张君<sup>1</sup>, 林琳<sup>2</sup>, 郭芮<sup>1</sup>, 卢程程<sup>1</sup>

(1. 国网山东省电力公司滨州供电公司, 山东 滨州 256600;

2. 国网山东省电力公司, 山东 济南 250001)

**摘要:** 随着可再生能源的快速发展和广泛应用, 含可再生能源的绿色电网电力调度成为迫切需要解决的问题。研究设计并提出一个结合专家行为克隆方法的深度强化学习决策模型, 并基于此构建针对含绿色能源电网的电力调度智能决策模型。模拟实验结果表明, 该模型在绿色能源机组中值产出率条件下, 电网消纳率为99.1%, 显著优于其他方法。研究不仅为绿色电力网络调度提供了一种有效的决策工具, 同时为深度强化学习在复杂电网系统中的应用提供了思路。

**关键词:** 深度强化学习; 电网; 电力; 智能调度; 专家行为

中图分类号: TP181 文献标识码: A 文章编号: 1003-7241(2025)07-0139-05

## Research on Intelligent Dispatching Analysis Model for Power Grid Based on Improved Deep Reinforcement Learning

ZHANG Jun<sup>1</sup>, LIN Lin<sup>2</sup>, GUO Rui<sup>1</sup>, LU Chengcheng<sup>1</sup>

(1. Binzhou Power Supply Company of State Grid Shandong Electric Power Company, Binzhou 256600, China;

2. State Grid Shandong Electric Power Company, Jinan 250001, China)

**Abstract:** With the rapid development and widespread application of renewable energy, green grid power dispatch containing renewable energy becomes an urgent problem that needs to be solved. This study designs and proposes a deep reinforcement learning decision-making model that combines expert behavior cloning methods, and based on this, constructs an intelligent decision-making model for power dispatching in a grid containing green energy. The simulation experiment results show that the model has a grid consumption rate of 99.1% under the condition of median output rate of green energy units, which is significantly better than other methods. This research not only provides an effective decision-making tool for green power network scheduling, but also provides ideas for the application of deep reinforcement learning in complex power grid systems.

**Keywords:** deep reinforcement learning; power grid; electricity; intelligent dispatching; expert behavior

## 0 引言

随着全球能源结构的转型升级, 电力行业迎来了前所未有的发展机遇和挑战<sup>[1]</sup>。尤其是在可再生能源大规模接入与电网智能化水平日益提升的背景下, 对电力调度的自动化、智能化要求越来越高<sup>[2]</sup>。在这一过程中, 电力调度系统的效率和准确性直接影响着电力系统的安全稳定运行和经济效益<sup>[3]</sup>。然而, 现行的电力调度策略多基于传统的优化算法, 面临局限性及适应性不强等挑战, 难以应对复杂多变的电网运行环境。Maurer C等发现光伏等新能源电力的接入会给电网电能管理带来更多调整, 因此有必要对含新能源电网进行电力智能调度。作者参考欧洲创新电网技术供应商的研究结果以及多个欧洲国家的输电系统数据, 设计出三种能提升电网管理效率的

措施。测试结果显示, 所设计的管理方法能减少50%~60%的电网管理成本<sup>[4]</sup>。Ahsan S M等存储耦合太阳能光伏系统能起到优化区域电网电力供应的作用, 但也会提高电网的管理成本。设计了一项基于混合整数线性规划方法的电网电力调度模型。仿真计算结果显示, 该模型的调度方案能将电网管理成本缩减3%~16%<sup>[5]</sup>。而在国内, 赵晋泉等人的研究显示, 以深度学习、强化学习为代表的高新技术是当前电力系统, 特别是含有新能源机组的电网系统提升电力调度的关键, 还分析了传统优化算法技术在电网电力调度方面的应用缺点<sup>[6]</sup>。

深度强化学习(deep reinforcement learning, DRL)技术作为一种模仿人类学习机制的算法, 为电力智能调度提供了新的解决思路。该技术将深度学习与强化学习结合起来, 以求解决决策空间复杂。电力系统的调度与控制问题天然适合应用DRL技术。针对现有研究成果在算法效率、模型鲁棒性以及系统稳定性方面的不足, 研究专注

\*基金项目: 国家自然科学基金资助项目(62192753)

收稿日期: 2024-02-27

通过改进深度强化学习框架,以加强电网电力智能调度分析的准确性及效率,这也是本研究的创新性所在。

## 1 面向含绿色能源的电网电力调度决策的改进DRL模型

### 1.1 含绿色能源的电力调度模型与专家调度策略

本研究设计的电力调度决策模型的应用对象为含有大比例绿色能源的电力系统。这种电力系统的调度模型优化目标函数通过系统运行安全性、运行支出、绿色能源消耗效率构成<sup>[7-9]</sup>。调度模型的约束条件由以下部分组成:机组出力存在限制、任何时刻系统均满足功率平衡、潮流计算后的最大平衡机出力需要处于下限的90%与上限的110%范围之内、电网火电机组需满足爬坡约束。再对于以上电力调度模型构建对应的调度决策模型,决策模型由状态空间、动作空间、回报函数三者组成,电力系统的状态信息输入给调度模型与决策模型,决策模型再计算出调整动作,电力系统根据调整动作修改状态并再次

输入给决策模型,从而达到动态调整电力系统电力的目的。近端策略优化(proximal policy optimization, PPO)算法具备较好的稳定性与运算性能,本研究选择使用DRL中的PPO算法求解含绿色能源的电力系统调度模型问题。

但是考虑到PPO算法训练前往往采用随机生成方式获得初始参数和策略网络,可能会生成过大的搜索空间,导致盲目探索,从而影响算法性能<sup>[10]</sup>。另外DRL算法由于监督信号的来源过于单一,存在样本训练效率低下问题。因此现在对DRL算法进行改进<sup>[11-12]</sup>。具体来说,改进后的算法会按照真实电网数据计算选择专家策略,并按照行为克隆思想生成初始策略网络。生成的策略网络将作为DRL的策略网络。

首先设计改进DRL中的专家调度策略生成模块。根据电力系统的历史运行数据,可以生成专家调度策略,算法即可基于此建立电网状态和专家调度策略之间的映射关系,形成初始策略网络,从而赋予深度强化学习算法一定的先验能力。基于模型驱动的电网专家策略的决策流程见图1。

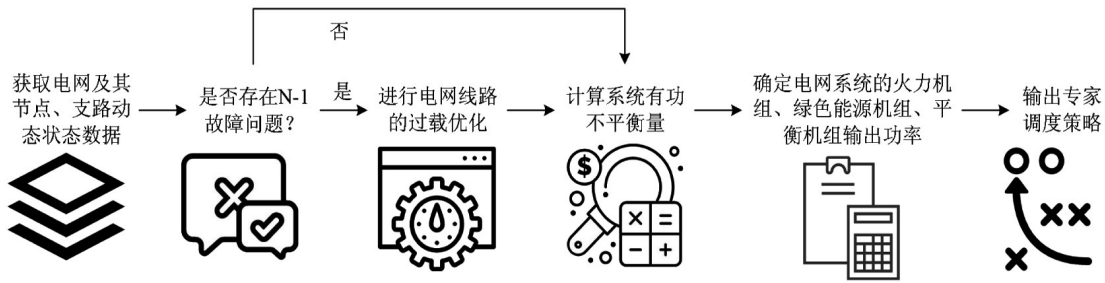


图1 DRL中专家决策的计算流程

决策流程第一步是确定电网实时状态,也即根据电网实时状态,确定各节点和支路的状态信息。第二步是进行N-1故障分析和线路过载优化,N为线路或设备数量。当任意一条输电线或设备发生故障时,算法会快速定位故障位置,以避免故障蔓延和电力系统的意外中断。

在故障分析后,对过载或有过载风险的线路进行优化。通过调整过载线路上下游机组的有功出力来进行线路过载优化。假设过载支路*l*的上游机组*x*和下游机组*y*有功出力调整值分别为 $\Delta P_{x,up}$ 和 $\Delta P_{y,down}$ 。 $\alpha$ 是设定的机组出力调整系数。 $P_{x,max}$ 和 $P_{y,max}$ 分别是*x*、*y*机组的最大和最小有功出力。*X*和*Y*分别是上游和下游的机组集合。则具体的优化方法见式(1)。

$$\begin{cases} \Delta P_{x,up} = -\alpha P_{x,max} \cdot I_{x,l} > 0, x \in X \\ \Delta P_{y,down} = \alpha P_{y,max} \cdot I_{y,l} > 0, y \in Y \end{cases} \quad (1)$$

线路过载优化产生的机组有功不平衡量  $P_{ext}$  按式(2)计算。

$$P_{ext} = \sum_{x \in X} \Delta P_{x,up} + \sum_{y \in Y} \Delta P_{y,down} \quad (2)$$

流程第三步为确定系统的有功不平衡量。系统有功

不平衡量是根据下一个时刻的负荷预测值  $P_{l,t+1}$  和当前时刻的负荷值  $P_{l,t}$  确定的负荷变化量  $\Delta P$ ,它是系统有功不平衡量的一部分,按式(3)计算:

$$\Delta P = P_{l,t+1} - P_{l,t} \quad (3)$$

如果需要对线路过载优化,则需从系统有功不平衡量中去除用于过载优化的部分。因此,最终的系统有功不平衡量即为机组的出力调整值  $\Delta P_{adj}$  按式(4)计算。

$$\Delta P_{adj} = \Delta P - P_{ext} \quad (4)$$

专家决策流程的第四步为确定各类机组的出力。在构建专家策略时,需要设定以调度时刻最大化可再生能源消纳为前提<sup>[13-14]</sup>。当下一调度时刻可满足可再生能源机组全功率输出时,火电机组的出力调整将按节点的连接支路数进行“同比例”出力调整。在确定各类机组的出力时,需按照约束条件设置,并判定可再生能源机组是否全功率输出。可再生能源机组能否全功率运转需要通过式(5)判定。

$$P_d = \Delta P_{con,t+1}^{up} + \Delta P_{th,t+1}^{down} \quad (5)$$

式中, $P_d$ 是可再生能源机组全功率运转的判据; $\Delta P_{con,t+1}^{up}$ 为下一

时刻所有可再生能源机组的出力调整上限总和。 $\Delta P_{th,t+1}^{down}$ 是下一时刻所有火电机组的出力调整下限总和。而火电机组的有功出力调整值 $\Delta P_{th}$ 可按式(6)计算。

$$\Delta P_{th} = r \left( \Delta P_{th,max} - \Delta P_{th,min} \right) \frac{D_{gi}}{\bar{D}} + \Delta P_{th,min} \quad (6)$$

式中, $\Delta P_{th,max}$ 、 $\Delta P_{th,min}$ 分别代表 $\Delta P_{th}$ 的上限与下限; $D_{gi}$ 为机组所处节点的连接支路数量; $\bar{D}$ 是机组所在节点平均连接支路数; $r$ 为机组的“同比例”调整系数。计算方法见式(7)。

$$r = \begin{cases} \frac{\left( \Delta P_{adj} - \Delta P_{con,t+1}^{up} \right) - \Delta P_{th,t+1}^{down}}{\Delta P_{con,t+1}^{up} - \Delta P_{th,t+1}^{down}}, P_d \leq \Delta P_{adj} \\ \frac{\left( \Delta P_{adj} - \Delta P_{th,t+1}^{down} \right) - \Delta P_{con,t+1}^{down}}{\Delta P_{con,t+1}^{up} - \Delta P_{con,t+1}^{down}}, P_d \leq DP_{adj} \end{cases} \quad (7)$$

式中, $\Delta P_{adj}$ 是系统的有功不平衡量。 $\Delta P_{con,t+1}^{down}$ 是下一调度时刻所有可再生能源机组的出力调整下限总和。 $\Delta P_{con,t+1}^{up}$ 是下一调度时刻所有火电机组的出力调整上限总和。专家决策流程的第五步是按照各机组的有功出力组成当前电网状态的专家调度策略,支持绿色能源电网电力调度决策模型中的专家调度策略部分设计完成。

## 1.2 基于改进深度强化学习的电网决策模型求解方法

模仿学习与深度强化学习同样致力于通过策略网络训练来控制智能体行动,但二者的方法有所差别:模仿学习是向专家策略看齐,而强化学习依赖于回报函数的反馈来优化策略,追求累计回报的最大化。行为克隆是主要的模仿学习方法,也是一种机器学习模型预训练的有效方法,这里在DRL模型中添加行为克隆模块,用于克服DRL模型在训练初期面临的探索-利用困境。但是行为克隆法也存在缺点,主要是该方法难以适应复杂和动态变化的环境,而且无法生成新的知识,所以这里将其与DRL结合,使得整体模型在优化性能的同时,保留生成新知识的能力。具体来说,在决策模型中,行为克隆法通过记录专家行为并将行为数据转化为训练数据的方式来提供DRL算法的训练数据,这即为两者的结合方式。通过行为克隆法的数据处理步骤,DRL模型具备了模仿专家行为的能力,也能完成自动化任务。现在假设专家策略生成步骤输出的信息为 $(s, a)$ 状态-动作对,令所有状态动作对构成 $\mathbf{B}$ ,也即存在式(8)。

$$\mathbf{B} = \{ (s_1, a_1), (s_2, a_2), \dots, (s_n, a_n) \} \quad (8)$$

式中, $s_n$ 、 $a_n$ 分别为专家策略在时刻 $n$ 的状态的、专家策略基于 $s_n$ 输出的动作,两者分别为DRL学习部分的输入与输出。

本研究中的专家策略状态含有支路电流负载 $\rho_{ho}$ ,当前时刻有功负荷 $P_{load,t}$ ,当前时刻的绿色能源机组最大有

功输出功率 $P_{th,max,t}$ ,三者共304维。动作数据则包含49维的发电机组、平衡机组、绿色能源机组有功出力。为了拟合专家策略,本研究采用一维卷积神经网络训练初始策略网络,目的是拟合专家策略的状态和动作数据。考虑到这里一维卷积神经网络的用途,现在设计出如表1所示的结构。另外表1中全部层结构的填充方式和激活函数分别为Same与Relu。

表1 一维卷积神经网络的结构参数

层编号	层名称	层类型	移动步长	卷积神经元数量	尺寸
#1	Conv1	卷积	2	32	5×5
#2	Pool1	平均池化	1	/	2×2
#3	Conv2	卷积	2	32	3×3
#4	Pool2	平均池化	1	/	2×2
#5	Conv3	卷积	2	64	3×3
#6	Pool3	平均池化	1	/	2×2

按照上述的行为克隆与专家策略进行运算,可以得到DRL算法的初始策略网络,有助于提高后者的学习效率和算法收敛速度。基于改进DRL算法的含绿色能源电网电力调度决策模型运行流程如图2所示。

## 2 绿色电网电力调度模型的仿真测试

为测试本研究设计的调度决策模型应用效果,现在设计并开展一项基于多节点含绿色能源的电力系统的仿真电力调度测试实验。该电力系统共含130个节点、88个负载、49台发电机、172条负载线。发电机中有33台属于火力机组,15台为绿色可再生能源机组,以及一台平衡机组。每一分钟记录一次电网系统潮流断面的历史数据,共获取12488条数据,数据集按照3:7比例切分为测试集和训练集。仿真实验中的实时数据通过安装在电力系统中的检测设备自动收集,并汇集于计算机上,再通过Python语言实现的决策模型进行处理计算,得到电力调度动作数据。电力调度动作数据将被作为面临再次输入到电力系统执行,从而达到动态调节电网的作用。实验中选择深度确定性策略梯度(deep deterministic policy gradient, DDPG)、SAC(soft actor-critic)、PPO构建对比的电力系统调度决策模型。所有模型的参数均按照网格搜索法确定。此次实验选择初始策略网络训练过程中的损失函数、准确率(Accuracy)以及电网绿色可再生能源消纳率作为评价指标。

本研究设计模型的初始策略网络在训练阶段的损失函数与准确率变化情况见图3。图3中横轴代表网络的迭代计算次数,左纵轴代表损失函数值,右纵轴代表准确率数值。观察图3可知,随着迭代次数的增长,初始策略网络的损失函数值在波动中快速下降并稳定于1.0附近,准确率在波动中上升,并在迭代到224次左右时完成收敛,收敛后的准确率数值处于99.6%左右。

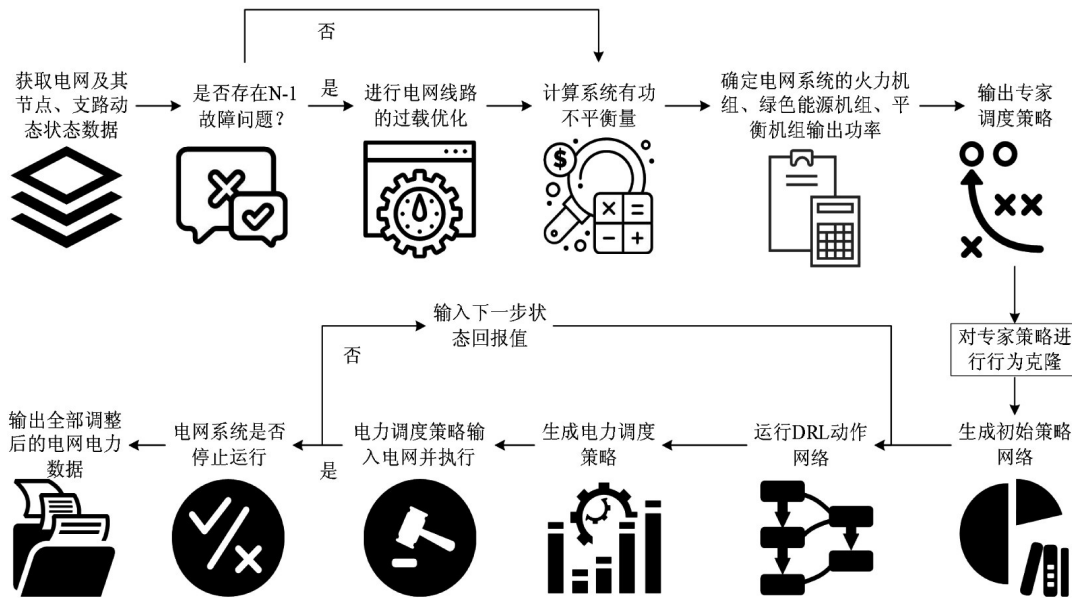


图2 基于改进DRL的含绿色能源电网电力调度决策模型流程

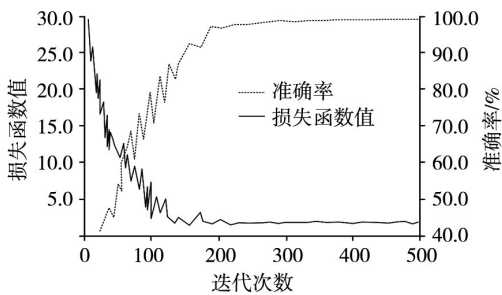


图3 初始策略网络训练时的损失函数与准确率

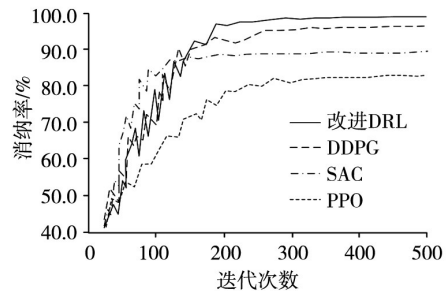


图4 中值产出率下的各模型训练阶段消纳率

绿色能源机组按照中值产出率的各决策模型训练阶段所对应消纳率变化情况见图4。图4中横轴代表迭代次数,纵轴代表对应的电网系统消纳率。不同样式曲线代表不同的决策模型。观察图4可知,随着迭代次数增长。各模型干预后的电力系统绿色能源消纳率逐渐提升,当迭代次数达到500次时,本研究设计的改进DRL、DDPG、SAC、PPO模型的消纳率分别为99.1%、96.4%、89.6%、83.5%。

最后分析各模型在不同绿色能源机组产出率条件下的收纳率,统计结果见图5。图5(a)、图5(b)、图5(c)分别展示的是绿色能源机组在低产出率(产出率0%~40%)、中等产出率(产出率40%~60%)、高产出率(产出率60%~100%)条件下的各模型消纳率数据。观察图5可知,在低产出率条件下,各模型的绿色能源消纳率没有明显差别,均处于20%~40%范围内。在中等产出率和高产出率条件下,此次设计的改进DRL决策模型消纳率明显高于其余各模型。在实际应用价值较高的绿色能源高产

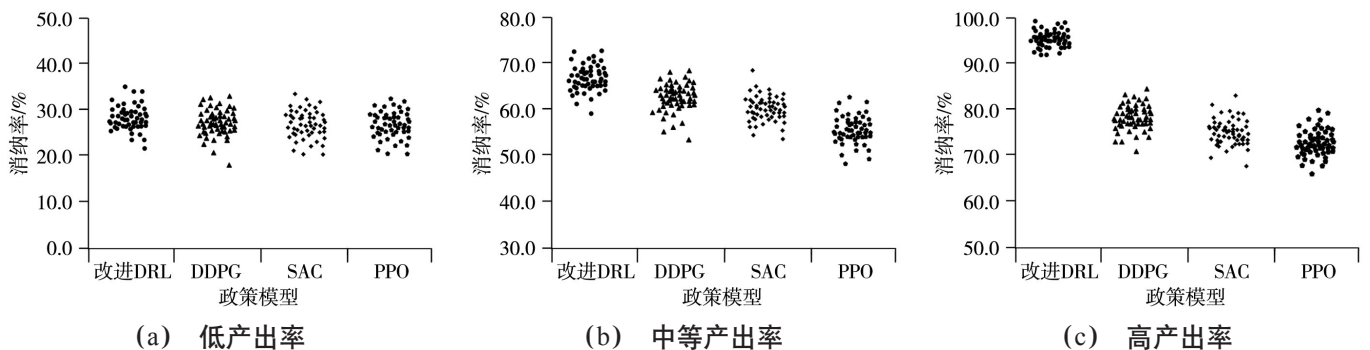


图5 不同绿色能源机组产出率的各模型消纳率

(下转第177页)