

DOI:10.3969/j.issn.2097-0706.2023.07.007

基于 MADDPG 算法的建筑群柔性负荷优化调控方法

Optimized control method for flexible load of a building complex based on MADDPG reinforcement learning

包义辛¹, 徐栲赟^{1,2}, 杨强^{1*}
BAO Yixin¹, XU Luoyun^{1,2}, YANG Qiang^{1*}

(1. 浙江大学 电气工程学院, 杭州 310027; 2. 浙江省白马湖实验室有限公司, 杭州 310056)

(1. College of Electrical Engineering, Zhejiang University, Hangzhou 310027 China;

2. Zhejiang Baima Lake Laboratory Company Limited, Hangzhou 310056, China)

摘要:随着电网调度环境和信息整理环境日趋复杂,电网调控的难度也随之增加。针对深度强化学习技术具有有效感知复杂系统运行状态、适应性强、可扩展性好等特点,提出了基于深度强化学习的配网优化调度方法。构建了考虑源-网-荷-储的模拟建筑体配网模型,从原理出发对多智能体深度确定性策略梯度(MADDPG)算法进行静态优化,将模型与真实数据导入适用于电网级目标的多智能体强化学习框架中,尝试用优化后的算法对配网系统进行电压调控。结果表明,所用算法基本消除了配网系统的违规峰值电压,降低了总体电压偏差;优化后的多目标导向算法在保持电压稳定的同时减小了负载-发电功率差,使负载功率损耗维持较低水平,表明基于深度强化学习的建筑群柔性负荷优化调控方法具有一定有效性。

关键词:微电网调控;能量管理;深度强化学习;确定性策略梯度;多目标优化;源网荷储;建筑群柔性负荷

中图分类号:TK 01*8;TU 18

文献标志码:A

文章编号:2097-0706(2023)07-0061-09

Abstract: The power grid dispatch environment and information organization environment have become more complex, and the difficulty of power grid regulation has gradually increased. Since deep reinforcement learning technology is of effective perception on complex system operation statuses, strong adaptability and good scalability, a distribution network optimization scheduling method based on deep reinforcement learning is proposed. Based on the simulated source-network-load-storage integrated distribution network model of a building complex, Multi-Agent Deep Deterministic Policy Gradient (MADDPG) algorithm was statically optimized from its principle. The model and real data were input into a multi-agent reinforcement learning framework suitable for grid-level objectives, and the optimized algorithm was tried to regulate the voltage of the distribution network system. The results show that the algorithm basically eliminates the abnormal peak voltages and reduces the overall voltage deviation. The optimized multi-objective oriented algorithm reduces the load-generated power difference while levelling the voltage off at a low level. The optimized control method for building complex flexible load based on reinforcement learning is proven to be effective.

Keywords: microgrid regulation; energy management; deep reinforcement learning; deterministic policy gradient; multi-objective optimization; source-grid-load-storage; flexible load of buildings

0 引言

电网调控系统贯穿于电力的发、输、配、用各个环节,集电网数据采集、存储和分析决策控制于一体^[1],具体调度过程还需要考虑和依照电力行业的

复杂执行经验、数据信息和行为准则^[2]。电网调控系统对电网的操作控制、调度管理等提供了不可忽视的重要支持^[3-4]。随着人工智能、特别是深度学习算法的快速发展,在电网调控领域应用深度学习算法准确预测电网的用电负荷以及平衡电网运行过程的各个环节^[5-6],能够在实现高精度、高效率决策的同时极大提高电网调控系统的自动化、智能化进程^[7-8]。

国内外根据各自的国情与电网分布特点,在研

基金项目:中国电机工程学会青年人才托举工程项目(CSEE-YESS-2021020)
The Young Elite Scientists Sponsorship Program by CSEE (CSEE-YESS-2021020)

究电网调控技术方向上也不尽一致。美国注重电力网络基础架构的升级更新,大力推进先进通信和控制技术以提高电网的智能性,并希望通过这些措施实现系统的智能化^[9];欧洲从1985年就开始了适合自己国情的智能电网研究开发计划,着重研究可再生能源的融入以及分布式电网的发展^[10-11];我国的智能电网调控技术起步较晚,自1990年以来开展了在电力系统故障诊断以及电网运行控制等方面的研究^[12]。通过与知识分析、数据挖掘相融合,电力系统中的人工智能应用技术探索了人机协同混合增强智能的实现机理。而在机器学习与深度学习技术研究不断深入的同时,两者在技术应用层面充分结合,在电网调控领域的各个方面^[13],包括在线安全评估、电力系统分析、配电网健康指数和负荷预测等发挥作用^[14]。

深度强化学习算法能够从大量的训练数据中获取最优权重参数以探索最优策略,有效感知复杂系统运行状态^[15]。而电网调控环节中由于时序性波动与运行状态、调控操作高度耦合导致策略动作过量,因此深度强化学习算法的这一特性能够更为有效的发挥作用^[16],为提高电网调控过程的智能化提供了一种更为合理、更为有效的途径^[17-18]。

本文在GridLearn框架已有的模型中迁移使用多智能体深度确定性策略梯度(Multi Agent Deep Deterministic Policy Gradient, MADDPG)算法进行电网调控并优化;调控32幢模拟建筑体模型组成的建筑群中保持柔性负荷总线电压稳定并降低其负载功率。通过应用纳什均衡原理协调多个智能体之间的策略选择、修改并优化奖励函数,借助具体算例验证多个调控目标的实现效果。

1 智能体控制下的建筑群模型

1.1 建筑能量管理模型设计

在生活场景中,建筑群中各项柔性负荷的功耗波动情况受不同季节的气候条件、建筑体用途的影响较大^[19]。根据这一情况,将所使用的建筑模型依照其具体用途进行分类,见表1。基于美国能源部的建筑能源需求数据集,对9类共192项建筑能源系统分别在一年内每隔1h记录当前室内温度、平均未设定冷却设定值差值、各个系统(暖通空调、室内热水供应)耗电量情况等数据,在进行算法训练时随机调用某一建筑体作为原始数据^[20]。

1.2 设备模型设计

1.2.1 光伏系统

光伏系统代表了上述建筑群模型中的可再生能源发电部分,通过这一系统的持续发电从而在外

表1 建筑分类情况

Table 1 Classification of buildings

建筑类型	具体建筑用途	当前类型数量
1	中型写字楼	4
2	快餐厅	37
3	零售商店	5
4	购物商店	1
5,6,7,8,9	多户住宅	145

部电网供电以外对微电网系统进行出力补偿。

假设太阳能电池板发电容量在运行过程中保持稳定,而由于单位面积的光伏发电量在不同时段、不同季节有很强的间隙性,因此有必要在模型内引入时间与气候参数来模拟单位发电量的变化。光伏系统的具体输出功率与约束如下

$$-1 \leq S_{\text{curtail}} \leq 1, \quad (1)$$

$$c = 0.5 - 0.5 \times S_{\text{curtail}}, \quad (2)$$

$$P_{\text{solar}} = (1 - c) \times P_{\text{pv}} \times G_{\text{solar, weather, ts}}, \quad (3)$$

式中: S_{curtail} 为光伏发电过程中的弃电率; P_{solar} 为太阳能产生的功率; P_{pv} 为光伏电池的额定功率; $G_{\text{solar, weather, ts}}$ 为在当前气候条件下的时间步长内所接收的太阳辐射量;参数 c 将给定的弃电率从-1到1映射为0到1之间的数值得到太阳能功率。

1.2.2 暖通空调系统与热水供应系统

暖通空调(Heating, Ventilation and Air Conditioning, HVAC)系统与热水供应(Domestic Hot Water, DHW)系统仅代表上述模拟建筑体模型中的柔性负荷,不涉及冷热等能量协调管理。

HVAC系统与1.2.1章节中介绍的光伏系统一样,需要参照时间与气候条件运行负载。本文在假设制冷设备的额定功率与技术效率保持稳定的前提下,其具体的运行功率取决于当前平均冷却设定值与室内温度的差值大小^[21],为保证这一目标进而确定空调整冷系统在当前时间步骤下的可用功率 C 、空调热能储存能耗 C_c 。

DHW系统的使用则一般固定于一天中的某几个时间段内。在假设热泵或电加热器等热水器的额定功率与效率保持稳定的前提下,先从模拟结果中获取当前时间步骤下的热水需求量,进而确定热水加热设备在当前时间步长下的可用加热功率 P_{DHW} 、能量存储设备当前的热水存储能耗 C_{DHW} 以应对当前的热水需求。

1.2.3 储能系统

储能系统在各类建筑群柔性负荷中起到重要的调控作用。储能系统在发电高峰期通过电力调峰提高光伏系统的发电利用率,而对于暖通空调系统以及热水供应系统而言,每一次调度动作中与储

能相对应的控制动作会被发送到相应的设备对象(HVAC,DHW)并对其进行充电或放电^[22]。

储能电池在收到调度策略的控制动作后,通过计算将电池的当前电量转换为电池可承受的最大功率值,之后相应地调整额定功率并进行充电^[23]。这一方法使储能电池在参与并满足建筑群需求响应的同时将其充放电过程均约束在一定阈值内,保证储能系统的正常工作状态

$$P_{\max} = P_n \left[P_{b,i} + (P_{b,i+1} - P_{b,i}) \frac{S_{oc} - C_{b,i}}{C_{b,i+1} - C_{b,i}} \right], \quad (4)$$

$$\begin{cases} C_{b,\min} \leq C_{b,i} \leq C_{b,\max} \\ |P_{b,i}| \leq P_{b,\max} \end{cases}, \quad (5)$$

式中: $P_{\max}$ 为当前电池在不超过其最大功率的范围内可以输出的最大功率值; P_n 为电池的标称功率; $C_{b,i}$ 和 $P_{b,i}$ 分别为电池荷电状态和充放电功率曲线中第*i*个时间步长下的对应函数值; $C_{b,\min}$, $C_{b,\max}$, $P_{b,\max}$ 分别为电池荷电与充放电功率的最大、最小值; S_{oc} 为电池的实时储能百分比。

1.2.4 总线网络模型设计

为了对建筑群整体进行调控,本文采用PandaPower数据库^[24]和IEEE-33总线网络^[25]对建筑群进行建模。如图1所示,在32个配电节点下各分配一幢模拟建筑体,尝试在这一总线网络模型下对各个节点的电压偏差值进行调控,使其稳定在标准值左右。

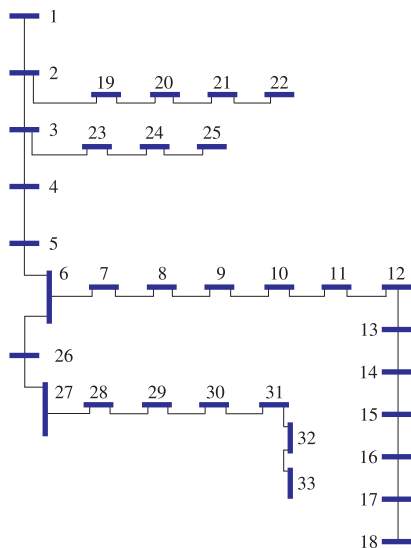


图1 IEEE-33总线系统单线示意

Fig. 1 Single line diagram of IEEE-33 bus system

2 基于深度强化学习算法的建筑微电网调控

2.1 MADDPG算法及其优化

多智能体强化学习算法可大致分为完全合作型、完全竞争型和竞争-合作型^[26],而本文中为调控

建筑群柔性负荷所使用的MADDPG算法为竞争-合作型。经典的强化学习(Deep Q Network,DQN)算法根据公式(6)中的值函数 $Q_{\pi}(s,a)$ 这一状态-动作对应的变化情况更新策略函数。但当单个智能体处于多智能体环境中时,独立使用函数 Q 从自身出发所感知的环境状态会因每个智能体的变化变的极不稳定,进而影响训练过程^[27]。

$$Q_{\pi}(s,a) = E(R|s_t = s, a_t = a), \quad (6)$$

式中: E 为更新策略函数; R 为当前状态转移时的即时奖励; s, a_t 分别为决策时刻*t*时的环境状态与动作。

而深度确定性策略梯度(Deep Deterministic Policy Gradient,DDPG)算法的训练过程如图2所示,可以直接调整策略函数中的参量 θ 来优化策略函数,通过给出确定性策略 $\mu_{\theta}:S \rightarrow A$ 得到策略梯度 $\nabla J(\theta)$ (如式(8))^[28],直接沿 $\nabla J(\theta)$ 的方向移动来最大化策略函数 $J(\theta)$

$$J(\theta) = E_{s \sim p^*, a \sim \pi_{\theta}}[R], \quad (7)$$

$$\nabla J(\theta) = E_{s \sim D}[\nabla_{\theta} \mu_{\theta}(a|s) \nabla_a Q_{\mu}(s,a) \Big|_{a=\mu_{\theta}(s)}], \quad (8)$$

式中: p^* 为状态空间分布; π_{θ} 为一般动作策略; D 为离线经验池。

在多智能体环境中,考虑策略函数由 $\theta = \{\theta_1, \dots, \theta_N\}$ 参数化的具有*N*个智能体的确定性策略 μ_i ,可以得到智能体*i*的策略函数梯度

$$\begin{aligned} \nabla J(\theta_i) = & E_{x, a \sim D}[\nabla_{\theta_i} \mu_i(a_i|o_i) \times \\ & \nabla_{a_i} Q_i^{\mu}(x, a_i, \dots, a_N) \Big|_{a_i=\mu_i(o_i)}], \end{aligned} \quad (9)$$

式中: x 包含多智能体环境中的各个智能体的观测值, $x = \{o_1, \dots, o_N\}$; $D = \{x, x', a_1, \dots, a_N, r_1, \dots, r_N\}$,其中记录了整个训练过程中所有智能体的经验; o_i, a_i 为第*i*个智能体的观测空间与动作空间; $Q_i^{\mu}(x, a_1, \dots, a_N)$ 是一个集中的动作值函数,它将所有智能体的动作 $a_1, \dots, a_N$ 与多智能体环境中的观测值 x 作为输入,输出智能体*i*的*Q*值。 Q_i^{μ} 是通过分开学习得到的,使得每个智能体都能在满足自身协调的前提下通过竞争-合作机制进行训练。

式(9)与动作值函数 Q_i^{μ} 的获取均表明MADDPG算法是每个智能体在输入所有智能体的动作之后改变策略函数梯度。这就保证了在每个智能体的 $J(\theta)$ 变化的同时,单个智能体通过 Q_i^{μ} 和 $J(\theta)$ 所感知的环境保持稳定^[29]。

为更好地维持这一感知下的稳定,令每个智能体都使用单独actor-critic网络,其中actor网络产生动作,以最小限度减少进程损失进行训练,如式(10);critic网络进行*Q*值估计,以非政策方式最小

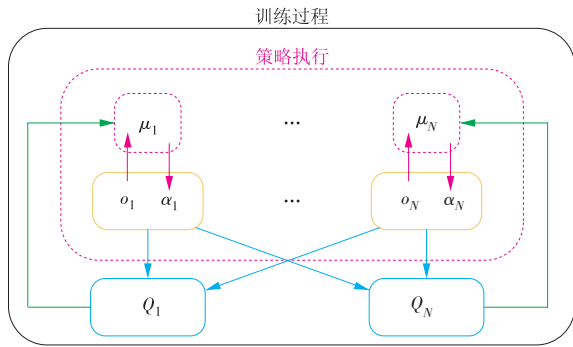


图2 MADDPG 算法的训练过程

Fig. 2 Training process of MADDPG algorithm

化均方贝尔曼误差 (MSBE)^[30], 如式(11)。在训练过程中, 每个智能体的 actor 网络都参考其他智能体的 critic 网络, 而 critic 网络只参考自身的奖励信号。

$$L_a^i(\phi^i) = -E_{o_i, a_i} \left\{ Q_i^{\mu}(s, [\dots a^{i-1}, \mu_{\phi^i}(s^i), a^{i+1}, \dots]) \right\}, \quad (10)$$

$$L_c^i(\theta^i) = E_{x, a, r, x'} \left\{ [Q_i^{\mu}(x, a_1, \dots, a_N) - (r_i + \gamma Q_i^{\mu'}(x', a_1', \dots, a_N'))]^2 \right\}, \quad (11)$$

式中: ϕ^i 为智能体 i 的近似值参数; γ 为下一奖励值的折扣系数。

2.2 深度强化学习中的博弈论应用

在平衡负荷、减少电网压力和提高能源利用率方面, 除了电压调控, 需求响应调控对于电网的稳定性和可靠性也具有重要意义^[31]。由于深度强化学习算法需要通过奖励函数实现环境与智能体之间的交互从而对目标进行调控, 本文考虑通过寻找多目标间的纳什均衡点合理安排电压与功率在奖励函数中的关系, 使系统目标获得较好的收敛性能。

首先如式(12), 计算微电网系统中电力系统总损耗 $Total_Loss$, 当总供应 p_{gen} 与总需求 p_{dem} 间存在偏差 $\Delta = p_{gen} - p_{dem}$ 时, 式(13)调整损失权重 ω 改变奖励损失的负偏差 (惩罚收益)。

$$Total_Loss = \sum_{i=2}^N p_{m,i} - p_{gen}, \quad (12)$$

$$\omega = 1 - \frac{1}{\sqrt{N}|\Delta| + 1}。 \quad (13)$$

总线电压偏差与负荷需求偏差的权值系数 ω_v , ω_p 同样可通过损失权重 ω (式(14)、式(15)) 进行更迭, 最终得到奖励函数 $reward$ 。

$$\omega_v = l_a \omega + (1 - l_a) \omega_{v,old}, \quad (14)$$

$$\omega_p = 1 - \omega_v, \quad (15)$$

$$reward = -1[(\omega_v V_n)^2 + (\omega_p P_n)^2 + (\omega Total_Loss)^2], \quad (16)$$

式中: $\omega_{v,old}$ 为上一轮迭代中采用的电压权重系数; l_a 为动态学习率调整系数。同时, 为防止微电网系统中电力系统总损耗 $Total_Loss$ 过大、电压超限或输电线路负载百分比过高等现象, 可以在式(16)的基础

上, 在以上参数超限时为奖励 $reward$ 施加惩罚进行软调控。如当总线电压 $v \geq 1.05$ 或 ≤ 0.95 时, $reward_{final} = reward - 50$ 。

3 算例分析

3.1 算例概况

为评估第1章节中提出的模型和第2章节中使用的方法在电压调节和能耗管理方面的性能, 本文指定 IEEE-33 总线网络为测试配电网, 选择建筑类型 5~9 的多户住宅模型连入节点, 指定 32 个基于 MADDPG 算法的电网调控智能体进行训练, 要求节点电压的最大偏差不超过 5% 并尽量降低负载功率损耗。在训练过程中每小时时间步长设置为 4, 即每 15 min 读取当前微电网的各项状态参数并进行调控, 在 1 天之内有 96 次调度, 对 1 个月的调度情况进行仿真。

研究过程中, 模拟建筑体中各类设备如负载、储能系统与光伏发电的相关数据选用美国能源部提供的美国建筑能源需求真实数据进行建模。以 1.1 章节中提到的第 4 类建筑, 即购物商店建筑体为例, 展示模型中微电网系统可控单元的相关参数见表 2。其中电热水器、热泵的具体功率由热水需求与制热性能系数决定, 暖通空调系统具体功率由温度需求与空调性能系数决定, 最高均不大于 20 kW; 储能系统的功效功率曲线表示对电力输出比例下电池功率转换效率, 例如当输出比例为 0.3 时, 电池功率转换效率为 0.85。容量功率曲线同理, 表示对于不同的电池容量水平电池可以提供的最大功率百分比。训练过程中设置的具体训练参数见表 3。

表2 可控单元基本参数

Table 2 Basic parameters of controllable units

可控单元	参数	值
光伏系统	太阳能板容量/kW	40
	热泵性能系数	0.22
	冷/热储水罐容量(标准单位)	1.5/0
	冷/热储水罐热损失系数	0.006/0.008
热水供应系统	加热/降温目标度/°C	45/9
	电热水器效率	0.92
	电池容量/(kW·h)	75
	电池效率	0.9
储能系统	额定功率/kW	30
	功率效率曲线	[0.0, 1] [0.3, 0.85]
	容量功率曲线	[0.0, 1] [0.8, 0.95]

表3 相关训练参数

Table 3 Related training parameters

参数	默认值
经验回放缓冲区容量	750
经验抽取数	32
折扣因子	0.95
目标网络更新权值	0.01
actor网络学习率	0.001
critic网络学习率	0.001
训练前数据收集轮次	3
动态学习率调整系数	0.3

3.2 结果与分析

由于模型训练过程中智能体数量多,导致电网系统调控过程中联合动作空间过大、训练速度慢,在训练轮次 $\text{eposide} = 100$ 轮后曲线始终在一个较低的水平上来回波动,如图3所示。

MADDPG算法中的 Q 值估计依赖于多智能体环境中所有智能体的联合状态空间与动作空间而非单一智能体的状态,这使得环境无法保持平稳,训练结果呈现出发散的特点。

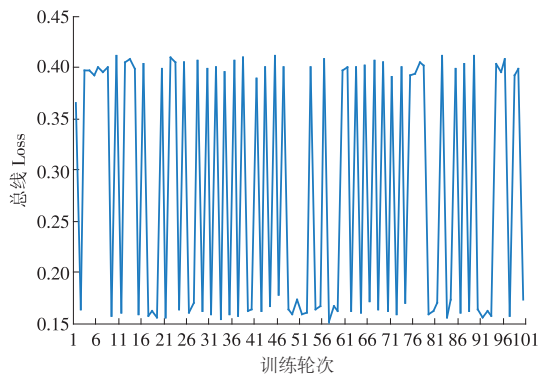


图3 MADDPG算法配电网调控的 Loss 曲线

Fig. 3 Line loss of the distribution network regulation made by MADDPG

电压调控过程中将标准总线电压 $u < 0.95$ (标么值) 或 $u > 1.05$ (标么值) 的情况定义为违规电压。为说明 MADDPG 算法在降低电压偏差、消除失控电压上起到的效果,绘制 IEEE-33 总线模型某节点上以标准电压标么值为纵坐标、以时间为横坐标的节点总线电压图,对比调控前后的电压在时间轴上的波动情况,如图4所示。图4选择了靠近母线末端节点 22、25 节点分别绘制了调控前与调控后从 10—14 d 处的电网波动图像。其中蓝色曲线为调控前节点总线电压波动,红色曲线为调控后节点总线电压波动。

对于未调控前的蓝色基线电压,22 节点相对稳定、25 节点则波动性较强,但整体上都表现出围绕着标准总线电压上下波动的趋势。这 2 个节点处的

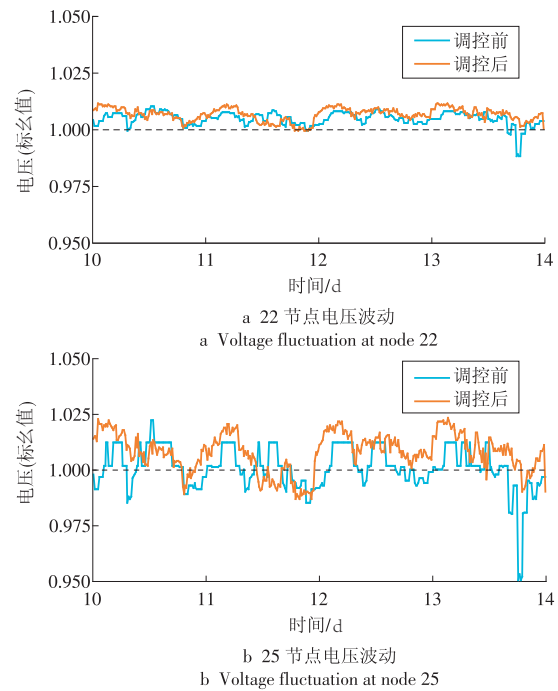


图4 配网总线电压调控前后对比

Fig. 4 Distribution network bus voltage before and after the regulation

电压都在第13天结束时出现峰值。

而 MADDPG 算法调控后的节点电压具有同样的波动趋势,但大部分时间内,调控后的电压具有靠近 1.00 (标么值) 的趋势,准确的说是具有降低总体电压偏差的能力。可以做出总线电压调控前后的概率分布直方图对比如图5所示。由图5明显可以看出,经过 MADDPG 算法调控的总线电压频率分布更接近于中心位置,在 $u = 1.00$ (标么值) 周围总体概率高,调控后在 0.99 (标么值) $\leq u \leq 1.01$ (标么值) 区间内的电压值数量增加了 31.302%,说明电压在调控后更趋近于稳定,具体参数见表4。

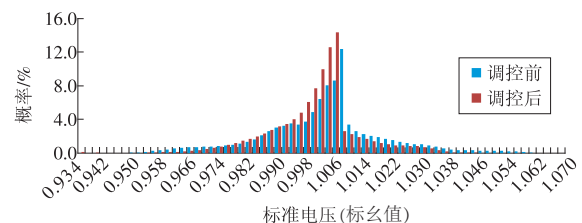


图5 配网总线电压直方图对比

Fig. 5 Distribution network bus voltage histograms

但 MADDPG 算法调控后的节点电压在某些时刻不太稳定,如从第 11 天中段开始时 25 节点电压略高于基线电压以及标准电压。从原理分析,这是由于深度强化学习算法能够在经验池中对策略进行大量且无止境的探索,重复已获取的经验与使用过的策略导致其调控效果往往具有滞后性;模拟建

表 4 MADDPG 算法调控前后总体电压偏差情况

Table 4 Overall voltage deviation before and after

MADDPG regulation		
项目	调控前	调控后
0.99~1.01 电压比例/%	35.458	46.557
平均值	1.003 14	1.003 03
样本方差	0.000 360 08	0.000 337 17

筑体中光伏系统渗透率较低、容量较小也使得基于储能系统充放电的电网调控效果并不显著。而这种维持配网总线电压稳定的特性在基线电压有突然峰值的情况下,如第9天和第13天结束时尤为显著,能够快速地将电压峰值“移平”,控制总线电网偏移量维持稳定。对比在同一模型28节点中使用邻近策略优化(Proximal Policy Optimization, PPO)算法^[20]的调控效果,如图6所示。

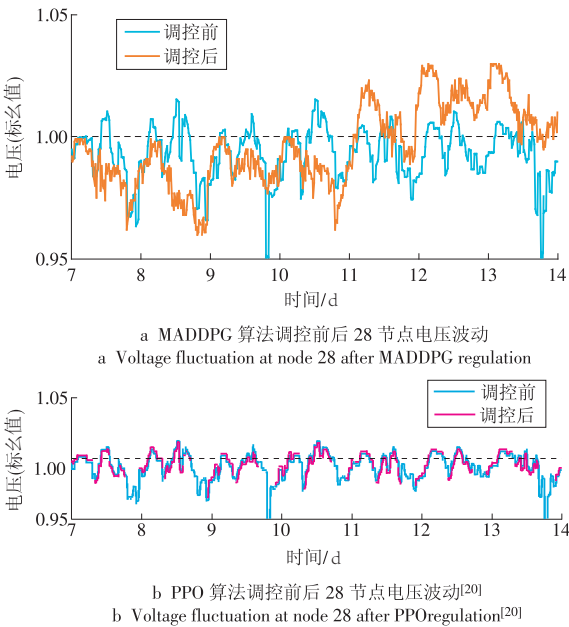


图 6 配网总线电压调控 MADDPGPPO 算法与 PPO 算法对比

Fig. 6 Distribution network bus voltage regulation made by MADDPG algorithm or PPO algorithm

图6中,28节点处出现了低于0.95(标么值)的失控电压,占比为0.326%。PPO算法对违规电压几乎没有起到调控效果,而MADDPG算法对违规电压峰值有良好的控制作用。相比于PPO算法,MADDPG算法在策略上的更新步骤梯度更大,在确定的目标下能够更好地收敛到最优解或局部最优解。

考虑居民区易出现集中式用电高峰,为保证总线电压安全,防止峰值电压对配电系统的危害,MADDPG算法在电网调控中具有独特的优势,如图7—8所示。

图7a,7b纵坐标为系统瞬时功率,图7c,7d纵坐

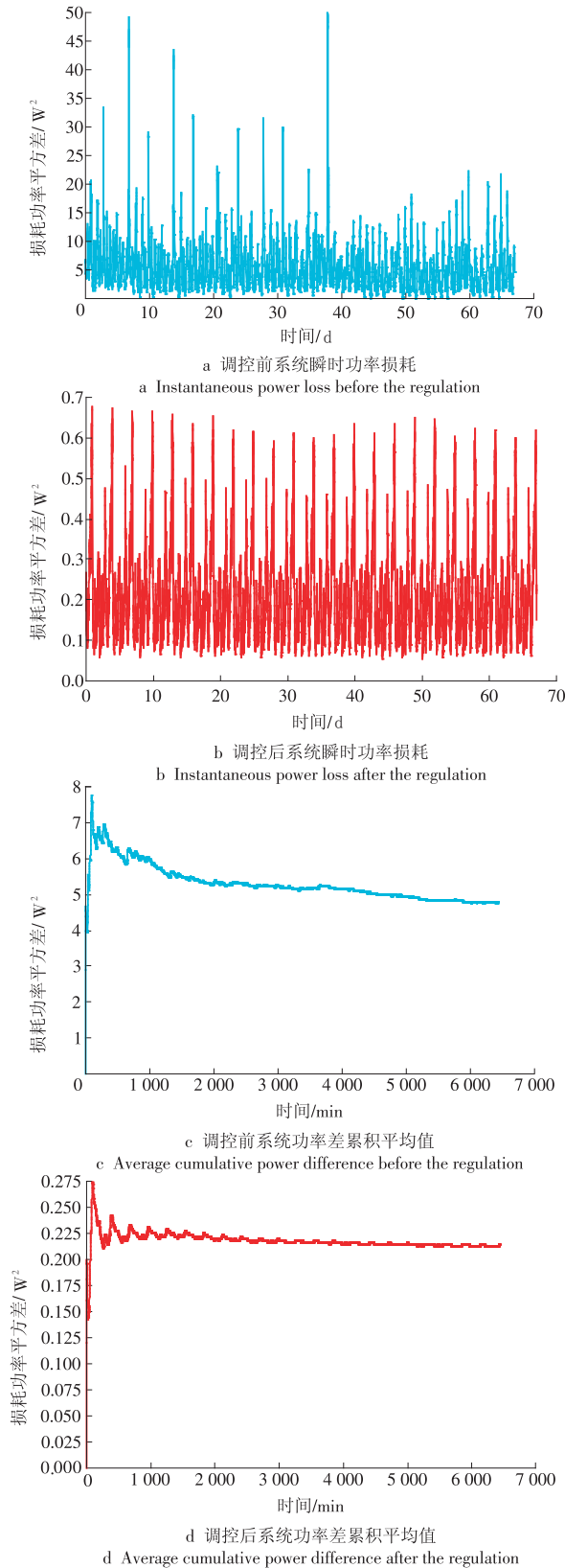


图 7 调控前后负载功耗水平对比

Fig. 7 Power consumption before and after the optimization

标为累积负载-发电功率差平方的平均值,均用于反映系统调度能力瞬态表现与负载功耗水平。

对比后可以看出,经过多目标优化后 MADDPG 算法调控后配网的负载-发电功率差降低到原有基

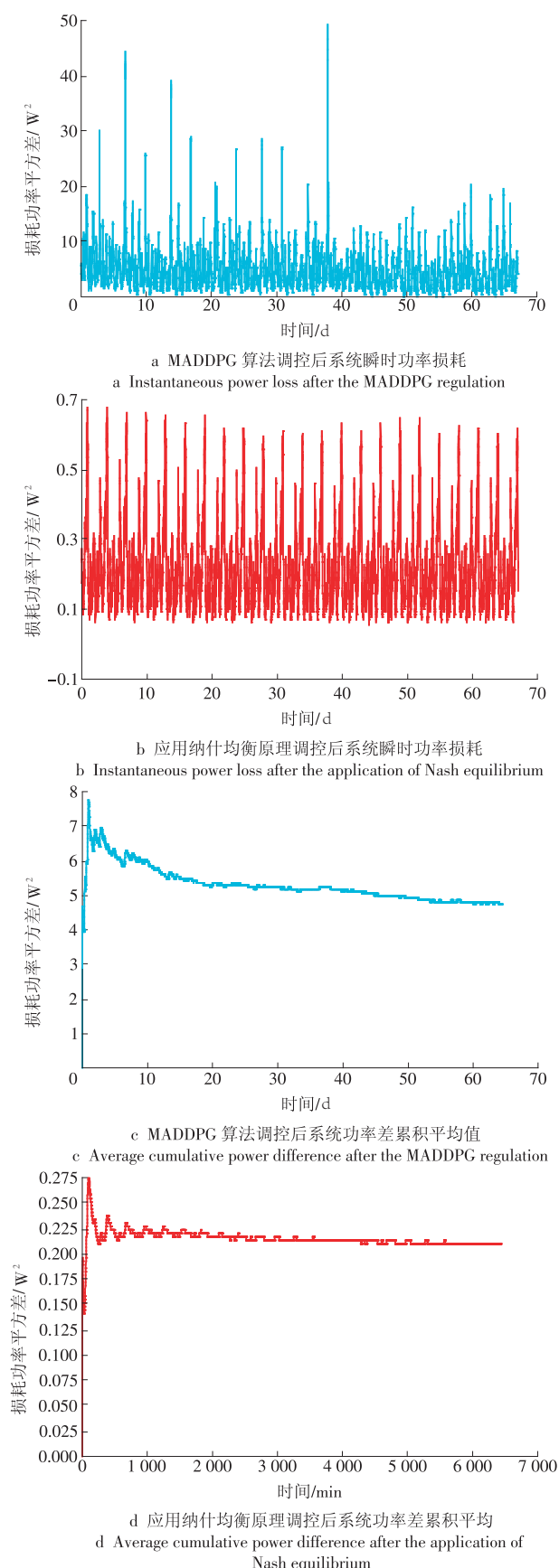


图8 纳什均衡原理应用前后配网系统负载
功耗水平对比

Fig. 8 Load and power consumptions of the distribution
network system before and after the application of
Nash equilibrium

础上的4.361%,负载功耗水平显著降低,系统的负载-发电功率基本实现平衡。

由于模型本身功耗水平、系统光伏渗透率均较低,柔性负载功耗优化相对于建筑群系统的功耗总量而言并不显著,差别不明显。但在数据分析并对比图8c与8d后可知,在整体时间上运用纳什均衡原理能够降低瞬时负载与实际发电能力之间的差异,负载-发电功率差平方的平均值在基本稳定后在MADDPG算法调控的基础上降低了2.159%,说明考虑纳什均衡原理的多目标优化方案对于配网系统的功率调度瞬态表现起到了一定的优化效果,相比于未调控前负载功率损耗收敛于较低的范围内。

4 结论

本文将由各类柔性负荷组成的建筑群作为具体的研究场景,尝试将深度强化学习算法应用于电网的优化调控。算例结果表明:MADDPG算法通过保证感知环境的相对稳定与离线经验池D的使用提高了训练效果,能够在多智能体调控柔性负荷时使得庞大的联合动作空间以更快的速度实现收敛;通过深度确定性梯度策略与纳什均衡来优化电网调控效果的多目标优化方案具有一定的效果,能够显著平衡电压峰值并实现负载损耗的降低。

参考文献:

- [1] 王晓燕,宋方宇轩,卢珊.电网调控中人工智能应用的关键技术研究[J].科技与创新,2022,(23):4-6,11
WANG Xiaoyan, SONG Fangyuxuan, LU Shan. Research on key technologies of artificial intelligence application in power grid regulation [J]. Science and Technology & Innovation, 2022, (23): 4-6, 11.
- [2] 崔文虎.电力系统故障演化建模与分析[D].成都:电子科技大学,2018.
CUI Wenhui. Power system fault evolution modeling and analysis [D]. Chengdu: University of Electronic Science and Technology of China, 2018.
- [3] 姜丽珍,董淑杰,闫振伟,等.电网调控技术在电力系统中的应用[J].电子制作,2019(12):69-70.
JIANG Lizhen, DONG Shujie, YAN Zhenwei, et al. Application of power grid regulation technology in power system [J]. Practical Electronics, 2019(12): 69-70.
- [4] 吉斌,孙绘,昌力,等.黏性电力用户参与需求侧响应的行为决策建模与分析[J].综合智慧能源,2022,44(2):80-88.
JI Bing, SUN Hui, CHANG Li, et al. Modeling and analysis on decision making behavior of loyal users participating in demand-side response [J]. Integrated Intelligent Energy,

- 2022,44(2):80-88.
- [5] 闪鑫, 陆晓, 瞿明玉, 等. 人工智能应用于电网调控的关键技术分析[J]. 电力系统自动化, 2019, 43(1): 49-57.
SHAN Xin, LU Xiao, ZHAI Mingyu, et al. Analysis of key technologies for artificial intelligence applied to power grid dispatch and control [J]. Automation of Electric Power Systems, 2019, 43(1): 49-57.
- [6] LIU Y, HUANG X, LI S, et al. A construction method of power grid monitoring knowledge graph [J]. Journal of Physics: Conference Series, 2022, 2166(1): 12-14.
- [7] 赵俊峰, 庄哲寅, 承轶青, 等. 资源描述框架语义网视角下的智能电网模型[J]. 华电技术, 2014, 36(4): 19-21.
ZHAO Junfeng, ZHUANG Zheyin, CHENG Yiqing, et al. Intelligent power grid model under visual angle of RDF Semantic Web [J]. Huadian Technology, 2014, 36(4): 19-21.
- [8] CHEN Z, MI W, LIN J, et al. Discussion on intelligence assistant scheme of dispatching and control operation in power grid[J]. Automation of Electric Power Systems, 2019, 43(22): 173-178.
- [9] 吴俊勇. 国内外智能电网的发展战略[J]. 变频器世界, 2011, 34(9): 36-37.
WU Junyong. Development strategy of smart grid at China and abroad[J]. The World of Inverters, 2011, 34(9): 36-37.
- [10] 巫飞新. 国内外智能电网技术发展现状[J]. 电气开关, 2012, 50(2): 3-6.
WU Feixin. Development status of smart grid technology home and abroad [J]. Electric Switchgear, 2012, 50(2): 3-6.
- [11] 史梦, 张志生, 罗学礼, 曹敏. 智能电网建设思路及国内外发展战略[C]//2010年云南电力技术论坛论文集(文摘部分), 2010: 1571-1575.
- [12] FANG J, WANG Y, LEI Z, et al. Control Strategy and Performance Analysis of Electrochemical Energy Storage Station Participating in Power System Frequency Regulation: A case study of the jiangsu power grid [J]. Sustainability, 2022, 14(15): 9189.
- [13] HUANG Y, LI H, WANG Z, et al. Research on the mechanism and forecast opower grid regulation policy under the background of new electricity reform in China [J]. Journal of Physics: Conference Series, 2019, 1176(4): 2030.
- [14] 李明节, 陶洪铸, 许洪强, 等. 电网调控领域人工智能技术框架与应用展[J]. 电网技术, 2020, 44(2): 393-400.
LI Mingjie, TAO Hongzhu, XU Hongqiang, et al. The technical framework and application prospect of artificial intelligence application in the field of power grid dispatching and control [J]. Power System Technology, 2020, 44(2): 393-400.
- [15] ELFAKI, ABDELRAHMAN O, SIM L, et al. Designing learning object repository using first order logic [C]//2nd IIAI International Conference on Advanced Applied Informatics, IIAI-AAI 2013, 413-414, 2013.
- [16] 范士雄, 李立新, 王松岩, 等. 人工智能技术在电网调控中的应用研究[J]. 电网技术, 2020, 44(2): 401-411.
FAN Shixiong, LI Lixin, WANG Songyan, et al. Bowen. application analysis and exploration of artificial intelligence technology in power grid dispatch and control [J]. Power System Technology, 2020, 44(2): 401-411.
- [17] CLAESSENS B, VRANCX P, RUELENS F. Convolutional neural networks for automatic state-time feature extraction in reinforcement learning applied to residential load control[J], IEEE Transactions on Smart Grid, 2016, 9(4): 3259 - 3269.
- [18] 李志. 智能电网调度控制系统的运行安全策略[J]. 集成电路应用, 2022, 39(11): 132-133.
LI Zhi. Strategy of operation security of smart grid dispatching control system [J]. Application of IC, 2022, 39(11): 132-133.
- [19] LIU S, GAO Y, YANG H, et al. Optimal guidance strategy for flexible load based on hybrid direct load control and time of use[J]. Global Energy Interconnection, 2023, 6(3): 297-307.
- [20] AISLING P, CONSTANCE C, KYRI B, et al. GridLearn: Multiagent reinforcement learning for grid-aware building energy management [J]. Electric Power Systems Research, 2022, 213.
- [21] 华咏竹, 谢强强, 秦会斌, 等. 计及用户端调节容量的变频空调自适应电压调控策略[J]. 综合智慧能源, 2022, 44(2): 21-28.
HUA Yongzhu, XIE Qiangqiang, QING Huibin, et al. Adaptive voltage regulation strategy for inverter air conditioners considering the regulation capacity on user side [J]. Intrgrated Intelligent Energy, 2022, 44(2): 21-28.
- [22] 童家麟, 洪庆, 吕洪坤, 等. 电源侧储能技术发展现状及应用前景综述[J]. 华电技术, 2021, 43(7): 17-23.
TONG Jialin, HONG Qing, LYU Hongkun, et al. Development status and application prospect of power side energy storage technology [J]. Huadian Technology, 2021, 43(7): 17-23.
- [23] 赵建立, 汤卓凡, 王桂林, 等. 具有储能作用的用户侧资源运行特性[J]. 综合智慧能源, 2022, 44(2): 8-14.
ZHAO Jianli, TANG Zhuofan, WANG Guilin, et al. Operation characteristics of user-side resources with energy storage function [J]. Intrgrated Intelligent Energy, 2022, 44(2): 8-14.
- [24] LEON T, ALEXANDER S, FLORIAN S, et al. Pandapower an open-source python tool for convenient modeling, analysis, and optimization of electric power systems [J]. IEEE Transactions on Power Systems, 2018, 33(6): 6510-

- 6521.
- [25] DOLATABADI S, MAEDEH G, PIERLUIGI S, et al. An enhanced IEEE 33 bus benchmark test system for distribution system studies [J]. IEEE Transactions on Power Systems, 2021, 36(3): 2565–2572.
- [26] BUŞONIU L, BABUŠKA R, BART D. A comprehensive survey of multiagent reinforcement learning [J]. IEEE Transactions on Systems, Man, and Cybernetics, 2008, 310(2): 156–172.
- [27] TAN M. Multi-agent reinforcement learning: Independent vs. cooperative agents [C]//Proceedings of the Tenth International Conference on Machine Learning, ICML 1993, 330–337.
- [28] WILLIAMS J. Simple statistical gradient-following algorithms for connectionist reinforcement learning [J]. Machine Learning, 1992, 8(3): 229–256.
- [29] 谢昕怡, 应黎明, 田书圣, 等. 基于MADDPG和智能合约的微电网交易决策优化[J]. 电力建设, 2022, 43(11): 142–150.
- XIE Xingyi, YING Liming, TIAN Shushen, et al. Optimization of microgrid trading strategy based on maddpg and smart contracts [J]. Electric Power Construction, 2022, 43(11): 142–150.
- [30] DAVID B, ZHANG X, DYLAN W, et al. Power gridworld: A framework for multi-agent reinforcement learning in power systems [C]//Proceedings of the 2022 13th ACM International Conference on Future Energy Systems, Energy, 2022, 565–570.
- [31] 孟安波, 王鹏, 丁伟锋, 等. 基于强化学习及纵横交叉粒子群算法的电网最优潮流计算[J]. 华电技术, 2021, 43(8): 74–82.
- MENG Anbo, WANG Peng, DING Weifeng, et al. Optimal power flow calculation of power grid based on reinforcement learning and crisscross PSO algorithm particle swarm optimization [J]. Huadian Technology, 2021, 43(8): 74–82.
- (本文责编: 齐琳)
- 收稿日期: 2023–06–05; 修回日期: 2023–07–03
上网日期: 2023–07–25; 附录网址: www.ienergy.cn
- 作者简介:
包义辛(2001), 男, 在读硕士研究生, 从事电力系统优化规划与运行调控等方面的研究, 3190103975@zju.edu.cn;
徐楞赞(1993), 男, 博士, 从事新能源微电网高效运行与稳定性等方面的研究, luoyun.xu@outlook.com;
杨强*(1979), 男, 教授, 博士, 博士生导师, 从事从事综合能源系统规划与运行控制研究, qyang@zju.edu.cn。
*为通信作者。