

# 基于多智能体近端策略网络的数据中心双层优化调度

杨秀<sup>1</sup>, 张相寅<sup>1</sup>, 黄海涛<sup>1</sup>, 余文昶<sup>2</sup>, 陈永刚<sup>2</sup>, 曹俊波<sup>2</sup>

(1. 上海电力大学电气工程学院, 上海 200090; 2. 中国华电集团有限公司上海分公司, 上海 200120)

**摘要:** 随着新一代信息通信技术, 如5G、云计算和人工智能的不断演进, 世界正迅速迈入数字经济的快车道。针对数据中心中可再生能源和工作负载预测的不确定性, 提出了一种基于多智能体近端策略网络的数据中心双层优化调度方法。首先, 建立了数据中心双层时空优化调度框架, 对数据中心工作负载、IT设备、空调设备进行详细建模; 在此基础上, 提出数据中心的二层优化调度模型, 上层以互联网数据中心(Internet data center, IDC)运营管理商总运营成本最小为目标进行时间维度调度, 下层以各IDC运行成本最低为目标进行空间维度调度; 然后, 介绍多智能体近端策略网络算法原理, 设计数据中心双层优化调度模型的状态空间、动作空间和奖励函数。最后, 针对算例进行离线训练和在线调度决策, 仿真结果表明, 所提模型和方法能够有效降低系统成本和能耗, 实现工作负载的最佳分配, 具有较好的经济性和鲁棒性。

**关键词:** 多智能体; 数据中心; 近端策略优化; 时空调度; 工作负载分配

## Dual-Layer Optimization Scheduling of Data Center Based on Multi-Agent Proximal Strategy Network

YANG Xiu<sup>1</sup>, ZHANG Xiangyin<sup>1</sup>, HUANG Haitao<sup>1</sup>, YU Wenchang<sup>2</sup>, CHEN Yonggang<sup>2</sup>, CAO Junbo<sup>2</sup>

(1. School of Electrical Engineering, Shanghai University of Electric Power, Shanghai 200090, China; 2. Shanghai Branch of China Huadian Group Co., Ltd., Shanghai 200120, China)

**Abstract:** With the continuous evolution of new generation information and communication technologies such as 5G, cloud computing, and artificial intelligence, the world is rapidly entering the fast lane of the digital economy. A dual-layer optimization scheduling method for data centers based on multi-agent proximal strategy network is proposed to address the uncertainty of renewable energy and workload prediction in data centers. Firstly, a dual-layer spatiotemporal optimization scheduling framework for data centers is established, which provides detailed modeling of data center workloads, IT equipment, and air conditioning equipment; On this basis, a dual-layer optimization scheduling model for data centers is proposed. The upper layer schedules the time dimension with the goal of minimizing the total operating cost of IDC operators, while the lower layer schedules the space dimension with the goal of minimizing the operating cost of each IDC. Then, the principle of multi-agent proximal strategy network algorithm is introduced, and the state space, action space, and reward function of the dual-layer optimization scheduling model for data centers are designed. Finally, offline training and online scheduling decisions are conducted for the examples. Simulation results show that the proposed model and method can effectively reduce system costs and energy consumption, achieve optimal workload allocation, and have good economy and robustness.

**Key words:** multi-agent; data center; proximal strategy optimization; spatiotemporal scheduling; workload allocation

## 0 引言

随着信息时代的发展和全球互联网的普及,互联网数据中心(Internet data center, IDC)的用电需求迅速上升, IDC的需求响应是实现智能电网负荷管理的潜在资源<sup>[1-2]</sup>。最近的技术革命已经渗透到现代生活的大多数方面,不断增长的计算需求促使小型 IDC 向高度优化和高效的超大规模 IDC 转变,由亚马逊、Facebook、谷歌、微软和阿里巴巴等大型公司运营<sup>[3]</sup>。我国在 2021 年发布的《2030 年前碳达峰行动方案》中也提出进行“算力+电力”协同,加快以 IDC 为代表的新型基础设施节能降碳,从而实现碳达峰、碳中和的目标<sup>[4]</sup>。

目前不少 IDC 运营者和专家对 IDC 能量管理开展了研究。文献[5]通过人工智能技术的应用, IDC 总能耗得到削减,提高能源利用率,运营更加经济高效。文献[6]为了降低超大规模 IDC 的购电成本,可延迟处理工作负荷,并研究其如何灵活与电网互动,利用电价的峰谷差以及时间上的批处理负荷,降低购电成本。文献[7]针对部分计算任务存在时间维度延迟迁移的特点,对能源消耗较大的 IDC 的灵活性潜力进行充分挖掘,促进了电力行业的低碳转型。文献[8]根据电价时间差异性提出一种最优的批处理负荷分配方法,该方法考虑将批处理任务分配到电价较低的时间段,以最大程度地降低能源成本。文献[9]建立了包含光能、风能等多种能源的互补发电端和 IDC 负荷端的调度模型,实现了电源侧与需求侧的协同调度。文献[10]通过对工作负荷的分配调控,提高新能源消纳率,优化 IDC 运行成本。文献[11]提出了 IDC 电、冷负荷的协同管理方式及模型,节约 IDC 购电成本。因此,工作负荷在很大程度上会影响 IDC 的能源消耗,可以看作一种广义的需求响应资源。然而,上述文献将 IDC 负荷进行了较为单一的分类,主要分为交互式负荷和批处理式负荷两种,忽略了 IDC 负荷可转移特性,即忽略了跨越不同地域进行负荷调度对系统运行的影响;同时如果参数选择不当或者迭代过程设计不合理,可能导致优化算法的收敛速度较慢。因此,有必要对传统优化方法进行改进,以加速其收敛速度。由于 IDC 内部大量的服务器构成了高维的状态空间和动作空间,大多数的优化方法难以求解。IDC 的运行需要处理大量的实时数据和复

杂系统的相互作用,深度强化学习模型通过学习适应这种繁杂环境的策略,使其能够更灵活地应对实际情况和环境变化。比起其他传统优化算法,深度强化学习算法更适用本文,因为其具有较强的环境意识和学习能力,并且不需要准确预测用户负荷和新能源出力。

近些年来,在电力系统优化调度中强化学习备受重视<sup>[12-13]</sup>。文献[14]提出了一种基于深度强化学习的微电网在线优化策略,降低了风光负荷的不确定性对微电网产生的影响。文献[15]基于深度 Q 网络,将源、荷不确定 IES 调度动态优化问题描述为 MDP 模型,得到最佳调度策略。文献[16]基于双深度 Q 网络的任务分配算法获得最优负荷分配策略,降低总成本和 IDC 能耗,实现多 IDC 的高效运行。但是上述深度强化学习方法一方面会降低 IDC 优化结果的准确性,动作离散不连续,另一方面离散动作的数量随着动作维度的增大而增大,这将会导致“维数破坏”问题。IDC 的运行环境时常面临变化,包括负载波动和能源成本的不断变化,本文使用多智能体近端策略优化算法(multi-agent proximal policy optimization, MAPPO)来求解多 IDC 的优化调度, MAPPO 算法可以求解连续和离散的动作空间问题,通过持续学习和适应,能够实时地调整任务调度策略,以最大程度地优化资源利用,具有良好的收敛性和对超参数的低灵敏度。

本文首先对 IDC 的 IT 设备、工作负荷、空调进行精细化建模,构建了包含 IDC 运营商与 IDC 的双层优化调度模型,其中上层模型最小化 IDC 与配电网的交互成本。下层模型则最小化各 IDC 的运行成本,确定目标函数和约束条件,对各 IDC 日内运行方案进行优化,并对 MAPPO 算法进行详细介绍,并将该问题表述为马尔科夫决策过程。最后通过算例分析验证,考虑负荷的时间平移特性,同时基于 IDC 的地理分布差异,通过传输 IDC 的数据负荷来实现电力负荷的空间传输,对比分析验证了本文所提策略和方法在 IDC 实际运行时的经济性和优越性,得到系统最佳调度策略。

## 1 IDC 源荷储协同调度框架和数学模型

### 1.1 IDC 源荷储协同调度框架

本文研究的 IDC 源荷储协同调度框架如图 1 所示,在电源侧 IDC 以微电网的形式供电。目前 IDC

的主要供电设备包括风机、光伏、微型燃气轮机和储能设备等,与输电网进行双向电力传输<sup>[17]</sup>,同时并入大电网以保证随时可通过购售电来平衡内部功率。在负荷侧具有IDC用户和用于冷却IDC服务器散热的空调设备制冷系统。

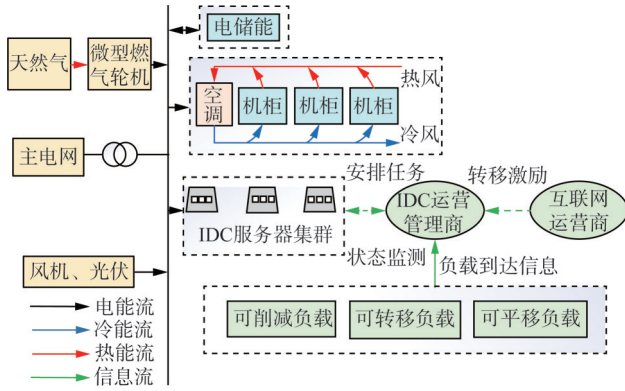


图1 IDC源荷储协同调度框架

Fig.1 Source load storage scheduling framework of IDC

## 1.2 IDC工作负荷及IT设备能耗建模

### 1.2.1 工作负荷模型

#### 1.2.1.1 可平移负荷

##### 1) IDC可平移负荷

IDC可平移负荷是指在一段时间范围内需要处理的总数据量不变,负荷可以从当前时段平移到其他时段内完成。通过调整各时间段对应计算量,利用负荷可平移性可以提升IDC可再生能源消纳水平,从而改变其时序用电、用冷需求。不同时段数据量具有灵活性,比如数据的挖掘备份、算法统计、图像处理等批处理负荷<sup>[18]</sup>。式(1)表示IDC各负荷到达率,式(2)表示可平移负荷存储总量约束。

$$\mu_t = \mu_{\text{batch},t} + \mu_{\text{tran},t} + \mu_{\text{cut},t} \quad (1)$$

$$L_{\text{batch},t} = L_{\text{batch},t-1} + \mu_{\text{batch},t} \Delta t - L_{\text{batch},\text{op},t} \quad \forall t \in \{T \mid t \neq t_0\} \quad (2)$$

$$L_{t,\text{max}} = 0 \quad (3)$$

$$0 \leq L_{\text{batch},t} \leq L_{\text{batch},\text{max}} \quad (4)$$

$$\mu_{\min} \leq \mu_t \leq \min(\varphi_{\text{ser},t}, B), \quad \forall t \in T \quad (5)$$

式中:  $\mu_t$  为  $t$  时段 IDC 数据负荷到达率;  $\mu_{\text{batch},t}$ 、 $\mu_{\text{tran},t}$ 、 $\mu_{\text{cut},t}$  分别为  $t$  时段 IDC 可平移负荷、可转移负荷、可削减负荷到达率;  $t_0$  和  $t_{\text{max}}$  分别为调度周期首、末时段编号;  $T$  为调度周期除  $t_0$  外的其他时段;  $L_{\text{batch},t}$  为  $t$  时段 IDC 存储的可平移负荷量;  $L_{\text{batch},\text{op},t}$  为安排在  $t$  时段处理的可平移负荷量;  $L_{t,\text{max}}$

为调度周期结束时 IDC 还未处理的可平移负荷数量;  $L_{\text{batch},\text{max}}$  为 IDC 存储可平移负荷的上限;  $\mu_{\min}$  为维持 IDC 正常工作的最小到达率;  $\varphi_{\text{ser},t}$  为  $t$  时段 IDC 服务率;  $B$  为数据网络带宽。

#### 2) 服务质量调节

服务质量调整的本质是在最大允许延迟内调整服务器和IT设备的带宽,从而间接地降低IDC的能量消耗。根据M/M/1排队原理,提出了一种新的数据负荷平均队列时延估计方法。

$$T_{\text{que},t} = \frac{1}{\varphi_{\text{ser},t} - \mu_t}, \quad \forall t \in T \quad (6)$$

$$T_{\text{han},t} = \frac{N_{\text{batch},\text{allot},t}}{\varphi_{\text{ser},t}}, \quad \forall t \in T \quad (7)$$

$$\varphi_{\text{ser},t} = \sum_{n \in M} n_{\text{batch},\text{allot},t} \times \mu_{\text{ser}} \quad (8)$$

$$T_{\text{que},t} + T_{\text{han},t} \leq T_{\text{max},t} \quad (9)$$

式中:  $T_{\text{que},t}$ 、 $T_{\text{han},t}$ 、 $T_{\text{max},t}$  分别为  $t$  时段可平移负荷平均排队延迟、平均处理延迟、用户可容忍的最大延迟时间;  $n_{\text{batch},\text{allot},t}$  为  $t$  时段安排处理可平移负荷的服务器数量;  $M$  为 IDC 服务器数量;  $\mu_{\text{ser}}$  为 IDC 数据负荷服务率,即单位服务器单位时间内可处理的负荷量。

#### 1.2.1.2 可转移负荷

1) IDC具有用能空间可调特性,将负荷以数据流形式在不同IDC间进行转移,从而改变内部电、热、冷负荷的空间分布,通常用于处理必须马上执行并可以迁移的负荷。建立  $X$  个前端服务器与  $N$  个 IDC 的数据负荷空间均衡模型如式(10)~(13)所示,其中式(13)为多IDC负荷均衡约束。

$$\mu_{\text{tran},x,t} = \sum_{i=1}^N \mu_{\text{tran},x,i,t}, \quad n = \{1, 2, \dots, N\}, \quad \forall t \in T \quad (10)$$

$$0 \leq \mu_{\text{tran},x,i,t} \leq \min(\varphi_{\text{ser},i,t}, B_{x,i}), \quad \forall t \in T \quad (11)$$

$$\mu_{i,t} = \sum_{x=1}^X \mu_{x,i,t}, \quad x = \{1, 2, \dots, X\}, \quad \forall t \in T \quad (12)$$

$$\sum_{i=1}^N L_{\text{tran},i,t} = \sum_{i=1}^N \phi_{\text{tran},i,t}, \quad n = \{1, 2, \dots, N\}, \quad \forall t \in T \quad (13)$$

式中:  $\mu_{\text{tran},x,t}$  为  $t$  时段第  $x$  个前端服务器的可转移负荷到达率;  $\mu_{\text{tran},x,i,t}$  为  $t$  时段第  $x$  个前端服务器分配第  $i$  个 IDC 的可转移负荷到达率;  $\varphi_{\text{ser},i,t}$  为  $t$  时段第  $i$  个 IDC 服务率;  $B_{x,i}$  为第  $x$  个前端服务器分配第  $i$  个 IDC 的数据网络带宽;  $L_{\text{tran},i,t}$  为  $t$  时段需要处理的可转移负荷量;  $\mu_{x,i,t}$  为第  $x$  个前端服务器分配第  $i$  个 IDC 的数据负荷到达率;  $\mu_{i,t}$  为  $t$  时段第  $i$  个 IDC 数据负荷



到达率;  $\phi_{\text{tran}, i, t}$  为到达该 IDC 运营商分配至 IDC $i$  的可转移负荷量。

图 2 为 IDC 负荷时空调度模式, 基于历史数据来预测未来的负荷情况, 提前做好资源分配和调度。通过时空转移负荷, 能避免机房内能耗局部集中导致服务器过热, 还可以最大程度实现调峰填谷、为电网提供辅助调频服务、促进可再生能源消费和 IDC 的安全运营。

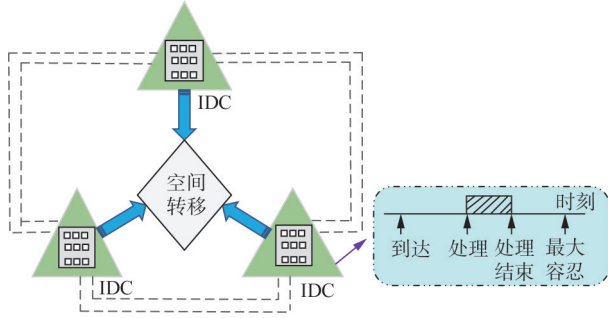


图 2 IDC 负荷时空调度模式

Fig. 2 Spatiotemporal scheduling mode of IDC load

## 2) 服务质量调节

当负荷以数据流的形式在空间上传输时, 会出现通信延迟和传输延迟, 需要满足用户可容忍的最大延迟时间。

$$T_{\text{cor}, t} = \sum_{\rho \in P} s_{\rho, t} t_{\text{cor}, t}, \forall t \in T \quad (14)$$

$$T_{\text{tran}, t} = \sum_{\rho \in P} s_{\rho, t} \frac{L_{\rho, t}}{B_{\rho}}, \forall t \in T \quad (15)$$

$$T_{\text{que}, t} + T_{\text{han}, t} + T_{\text{cor}, t} + T_{\text{tran}, t} \leq T_{\text{max}} \quad (16)$$

式中:  $T_{\text{cor}, t}$ 、 $T_{\text{tran}, t}$  分别为  $t$  时段通信、传送延迟的时间;  $\rho$ 、 $P$  分别为网络分支编号和相应的集合;  $s_{\rho, t}$  为  $t$  时段标志负荷是否通过分支  $\rho$  通信、传输,  $\rho$  为 1 时表示通过分支, 否则为 0;  $t_{\text{cor}, t}$ 、 $B_{\rho}$  分别为分支  $\rho$  对应的通信延迟和带宽;  $L_{\rho, t}$  为  $t$  时段由分支  $\rho$  传输的负荷;  $T_{\text{max}}$  为用户可容忍的最大延迟时间。

### 1.2.1.3 可削减负荷

IDC 可削减负荷是指对处理的数据任务进行裁剪和削减, 确定不重要任务的数据备份和业务恢复的数量, 减少任务同时传输的频率<sup>[19]</sup>。

$$0 \leq L_{\text{cut}, t} \leq L_{\text{cut}, t} \quad (17)$$

式中  $L_{\text{cut}, t}$ 、 $L_{\text{cut}, t}$  分别为需求响应前、后  $t$  时段的可削减负荷量。

综上, IDC 在  $t$  时段数据负荷总量需求  $L_t$  如式 (18) 所示, 由可平移、可转移、可削减负荷构成。

$$L_t = L_{\text{batch}, t} + L_{\text{tran}, t} + (L_{\text{cut}, t} - L_{\text{cut}, t}) \quad (18)$$

### 1.2.2 IT 设备能耗模型

服务器集群是 IDC 的主要用电需求来源, 也是实现 IDC 功能的基本单元, 集计算、储存为一体。通常每个服务器的使用水平以数据负载到达率与服务率为特征<sup>[20]</sup>, 并通过集成诸如服务器数量和额定功率之类的参数来建模。

$$P_{\text{IT}, i, t} = (P_{\text{peak}} - P_{\text{idle}, t}) \varpi_t + P_{\text{idle}, t} M_t \quad (19)$$

$$\varpi_t = \frac{L_t}{N_{\text{up}, \text{op}, t} \mu_t} \quad (20)$$

式中:  $P_{\text{IT}, i, t}$  为  $t$  时段第  $i$  个 IDC 的 IT 设备总能耗;  $P_{\text{peak}}$  为单台服务器满载能耗;  $P_{\text{idle}, t}$  为  $t$  时段单台服务器空载能耗;  $\varpi_t$  为单台服务器  $t$  时段利用率;  $M_t$  为  $t$  时段开启的服务器数量;  $N_{\text{up}, \text{op}, t}$  为安排开启的服务器数量;  $\mu_t$  单台服务器  $t$  时段数据处理量。

### 1.3 IDC 空调能耗模型

服务器运行会产生非常大的热量, IDC 冷却设备可以调节机房的温度, 来限制机房内各服务器的最高运行温度, 使得环境温度控制在安全阈值以内。IDC 的大部分冷负荷来源于吸收 IT 设备产生的热能和环境热量, 约占 IDC 空调设备冷负荷的 90%。

#### 1) 空调冷负荷来源

$$Q_{\text{air}} = Q_{\text{IT}} + Q_{\text{env}} \quad (21)$$

$$Q_{\text{IT}} = 0.97 \times P_{\text{IT}} \quad (22)$$

$$Q_{\text{env}} = C_T \times S_{\text{ser}} \quad (23)$$

式中:  $Q_{\text{air}}$  为空调的冷负荷;  $Q_{\text{IT}}$  为 IT 设备产生的热能 (IT 设备制造商为 IT 设备提供功耗和散热数值, 0.97 表示散热比<sup>[21]</sup>);  $Q_{\text{env}}$  为环境热量;  $C_T$  为室温系数, 与室温成正比, 华南地区取 0.18, 北方取 0.10, 其他地区取 0.10—0.18<sup>[21]</sup>;  $S_{\text{ser}}$  为 IT 设备总投影面积。

#### 2) 空调设备能耗

IDC 通常用冷负荷  $Q_{\text{air}}$  和空调能耗  $P_{\text{air}}$  的比值性能系数  $\theta_{\text{cop}}$  来描述空调的制冷效率,  $\theta_{\text{cop}}$  越高说明空调性能越好。

$$\theta_{\text{cop}} = 0.0068 K_{\text{out}}^2 + 0.0008 K_{\text{out}} + 0.458 \quad (24)$$

$$P_{\text{air}} = \frac{Q_{\text{air}}}{\theta_{\text{cop}}} = \frac{0.97 P_{\text{IT}} + C_T S_{\text{ser}}}{\theta_{\text{cop}}} \quad (25)$$

式中  $K_{\text{out}}$  为空调设备送风温度<sup>[9]</sup>。

#### 3) IT 设备与空调换热

IT 设备在工作时 97% 的电力消耗转化为热量,

并由流经其中的冷空气输送。在热力学中,这种现象被称为对流换热。通常用等效热阻模型表示为:

$$K_{in} = 0.97\eta_h P_{IT} + K_{out} \quad (26)$$

$$K_{IT} = P_{IT} R_{in} + K_{in} \quad (27)$$

式中:  $K_{in}$  为IT设备入口气流温度;  $K_{IT}$  为服务器芯片温度;  $R_{in}$  为IT设备输入端对流换热的等效热阻<sup>[22]</sup>, 在对流换热过程中, 视为换热面气体相对速度的函数, 通常流经服务器的气流设定为  $0.0147\text{ }^\circ\text{C/W}$  的恒定值;  $\eta_h$  为传热系数, 在设备均匀布置时取常数。

因此, IDC能耗  $P_{tol,i,t}$  视为服务器能耗加上空调能耗, 如式(28)所示。

$$P_{tol,i,t} = P_{IT,tol,i,t} + P_{air,tol,i,t} = \sum_{i=1}^M P_{IT,i,t} + \sum_{j=1}^L P_{air,j,t} \quad (28)$$

式中:  $M$  为服务器数量;  $L$  为空调设备数量;  $P_{IT,tol,i,t}$  为  $t$  时段服务器能耗;  $P_{air,tol,i,t}$  为  $t$  时段空调设备总能耗;  $P_{IT,i,t}$ 、 $P_{air,j,t}$  分别为  $t$  时段第  $i$  台IT设备和第  $j$  台空调设备的能耗。

## 2 IDC源荷储双层优化调度模型

基于负荷的时空转移特性, 上层以IDC与配电网交互成本最小进行时间维度调度; 下层以各IDC运行成本最低为目标进行空间维度调度。考虑工作负荷服务质量、空调约束、功率平衡、燃气轮机、储能等约束条件。IDC双层优化调度模型的基本架构如图3所示。

### 2.1 目标函数

#### 2.1.1 上层目标函数

考虑到可平移负荷的高负荷水平和功耗时间的连续性, 可以在较长的调度周期内进行调度, 它被用作IDC能源管理系统中的调度资源, 优化调度结果被分发到每个子IDC。

上层IDC系统优化调度模型以多IDC与配电网的交互成本最小为目标, 仅确定可平移负荷的调度, 为下层优化提供本地IDC每个节点净注入可平移负荷最优功率值。考虑时间维度的上层优化调度目标函数为:

$$F_{1t} = -C_{inc,t} + C_{grid,t} \quad (29)$$

$$C_{inc,t} = \sum_{i=1}^{24} \sum_{t=1}^N [c_{batch,i,t} (L_{batch,i,t} - L_{batch,i,t-1})] \Delta t \quad (30)$$

$$C_{grid,t} = \sum_{i=1}^{24} \sum_{t=1}^N (\beta \alpha_{buy,t} P_{grid,i,t} + (\beta - 1) \alpha_{sell,t} P_{grid,i,t}) \Delta t \quad (31)$$

式中:  $F_{1t}$  为多IDC与配电网的交互成本;  $\Delta t$  为调度周期时间; 根据文献[23-24],  $C_{inc,t}$  为配网  $t$  时段对可平移负荷的调度激励, 以起到削峰填谷的作用, 根据负荷高峰时段负荷平移给予相应的经济激励;  $C_{grid,t}$  为电网  $t$  时段购售电成本;  $C_{batch,i,t}$  为  $t$  时段单位可平移负荷调度成本,  $i$  为数据中心的编号;  $L_{batch,i,t}$  为  $t$  时段可平移负荷进行时间调度的数量;  $P_{grid,i,t}$  为第  $i$  个数据中心  $t$  时段与上级电网交互的功率; 当  $\beta = 1$ ,  $P_{grid,i,t} > 0$  时, 表示IDC从电网购电, 当  $\beta = 0$ ,  $P_{grid,i,t} < 0$  时, 表示  $t$  时段IDC向电网售电;  $\alpha_{buy,t}$ 、 $\alpha_{sell,t}$  分别为  $t$  时段购、售电价;  $N$  为数据中心

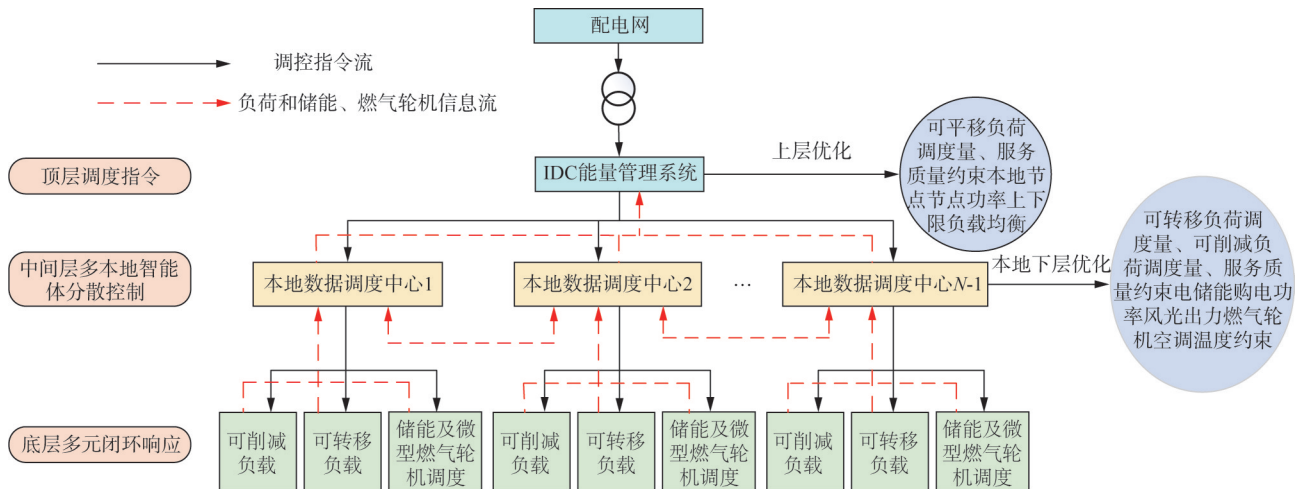


图3 IDC双层优化调度模型的基本架构

Fig. 3 Basic architecture of IDC dual-layer optimal scheduling model

的数量。

### 2.1.2 下层目标函数

下层以各 IDC 总运营成本最小为目标, 微型燃气轮机和储能具有高度的灵活性, 可转移负荷、可削减负荷可以快速响应, 具有较低的负荷水平, 因此可由本地 IDC 直接调度, 相应的结果可以返回到 IDC 能源管理系统。通过 IDC 之间的互补性可以实现协同运行, 降低 IDC 总运营成本, 间接减少上层系统与配电网之间的电力交互, 减少因访问 IDC 而导致的配电网负荷波动等不良影响。

上层优化得到可平移负荷结果后, 具有微型燃气轮机、储能设备、可削减、可转移负荷资源的本地 IDC 根据本地的风光和可转移、可削减负荷预测值在本地调度周期内进行优化调度。考虑空间维度的下层优化调度目标函数为:

$$F_{2t,i} = \sum_{i=1}^N (C_{res,i,t} + C_{sto,i,t} + C_{grid,i,t} + C_{gas,i,t} - C_{tran,i,t} - C_{cut,i,t}) \quad (32)$$

$$C_{tran,i,t} = \sum_{t=1}^{24} (L_{tran,i,t} c_{tran,i,t}) \Delta t \quad (33)$$

$$C_{cut,i,t} = \sum_{t=1}^{24} c_{cut,i,t} (L_{cut,i,t} - L_{cut0,i,t}) \quad (34)$$

$$C_{res,i,t} = \sum_{t=1}^{24} c_{pv,i,t} P_{pv,i,t} + c_{wt,i,t} P_{wt,i,t} \quad (35)$$

$$C_{sto,i} = \sum_{t=1}^{24} (C_{cha,i,t} P_{cha,i,t} + C_{dis,i,t} P_{dis,i,t}) \Delta t \quad (36)$$

$$C_{gas,i,t} = \sum_{t=1}^{24} \xi_{gas,i,t} M_{mt,i,t} \Delta t \quad (37)$$

式中:  $F_{2t,i}$  为各 IDC 总运营成本,  $i$  为数据中心的编号;  $\Delta t$  为调度周期时间;  $C_{res,i,t}$  为可再生能源运维成本;  $C_{sto,i,t}$  为储能设备充放电成本;  $C_{gas,i,t}$  为购气成本;  $c_{pv,i,t}$ 、 $c_{wt,i,t}$  分别为 IDC 光伏、风机运维成本系数;  $P_{pv,i,t}$ 、 $P_{wt,i,t}$  分别为  $t$  时段光伏、风机提供的功率;  $C_{cha,i,t}$ 、 $C_{dis,i,t}$  分别为电储能充/放电成本;  $P_{cha,i,t}$ 、 $P_{dis,i,t}$  分别为  $t$  时段储能系统充放电功率;  $L_{tran,i,t}$  为  $t$  时段可转移负荷进行空间调度的数量;  $L_{cut,i,t}$  为  $t$  时段可削减负荷数量;  $C_{tran,i,t}$ 、 $C_{cut,i,t}$  分别为给予 IDC 可转移/削减用户数据请求的激励成本;  $c_{tran,i,t}$ 、 $c_{cut,i,t}$  分别为给予 IDC 可转移/削减用户数据请求的激励成本系数;  $\xi_{gas,i,t}$  为单位购气价;  $M_{mt,i,t}$  为微型燃气轮机消耗的天然气量。

### 2.2 约束条件

除了满足上述工作负荷服务质量约束外, 还应

满足如下约束。

#### 1) 服务器芯片温度上限约束

$$K_{IT,t} \leq K_{IT,max} \quad (38)$$

式中:  $K_{IT,t}$  为服务器在  $t$  时段的芯片温度;  $K_{IT,max}$  为服务器芯片温度的上限, 本文取 80 °C。

#### 2) 可再生能源机组约束

$$P_{pv,min} \leq P_{pv,t} \leq P_{pv,max} \quad (39)$$

$$P_{wt,min} \leq P_{wt,t} \leq P_{wt,max} \quad (40)$$

式中:  $P_{pv,min}$  与  $P_{wt,min}$  分别为光伏、风力机组出力的下限;  $P_{pv,max}$  与  $P_{wt,max}$  分别为光伏、风力机组出力的上限。

#### 3) 微型燃气轮机约束

$$P_{mt,t} = \eta_{mt} M_{mt,t} g_{gas} \quad (41)$$

$$0 \leq P_{mt,t} \leq P_{mt,max} \quad (42)$$

式中:  $P_{mt,t}$  为  $t$  时段微型燃气轮机的实际产电功率;  $M_{mt,t}$  为  $t$  时段的进气量;  $\eta_{mt}$  为产电效率;  $g_{gas}$  为天然气低位热值;  $P_{mt,max}$  为产电功率上限。

#### 4) 电储能约束

电储能约束主要包括交换功率上下限约束。

$$\begin{cases} s_{cha} + s_{dis} \leq 1 \\ 0 \leq S_{sto,t} \leq S_{sto,max} \\ 0 \leq P_{cha,t} \leq P_{cha,max} \\ 0 \leq P_{dis,t} \leq P_{dis,max} \end{cases} \quad (43)$$

式中:  $s_{cha}$ 、 $s_{dis}$  为 0-1 控制变量, 分别为电储能的充、放电标志, 电储能不能同时充、放电;  $S_{sto,t}$ 、 $S_{sto,max}$  分别为  $t$  时段系统储能容量和最大储能容量;  $P_{cha,t}$ 、 $P_{dis,t}$  分别为  $t$  时段电储能充、放电功率;  $P_{cha,max}$ 、 $P_{dis,max}$  分别为电储能最大充、放电功率。

#### 5) IT 设备入口气流温度范围约束

$$K_{out,min} \leq K_{out,t} \leq K_{out,max} \quad (44)$$

$$K_{in,min} \leq K_{in,t} \leq K_{in,max} \quad (45)$$

式中:  $K_{out,t}$  为机房在  $t$  时段的送风温度;  $K_{out,min}$ 、 $K_{out,max}$  分别为空调设备最小、最大送风温度;  $K_{in,max}$ 、 $K_{in,min}$  为 IT 设备入口气流温度上下限, 根据美国采暖、制冷和空调工程师协会 (ASHRAE) 规定<sup>[25]</sup>, 一般在 18~27 °C。

#### 6) 能量平衡约束

$$\begin{cases} P_{wt,i,t} + P_{pv,i,t} + P_{grid,i,t} + P_{dis,i,t} - P_{cha,i,t} = P_{tol,i,t} \\ P_{grid,min} \leq P_{grid,i,t} \leq P_{grid,max} \end{cases} \quad (46)$$

式中:  $P_{grid,i,t}$  为  $t$  时段从电网流入的电功率;  $P_{grid,max}$ 、 $P_{grid,min}$  分别为电网最大和最小交互功率。



### 3 基于MAPPO的IDC双层优化调度模型

多IDC系统的双层优化调度模型求解流程如图4所示。本文构建的双层模型之间存在耦合关系,这使得直接求解变得困难。首先,构造了下层模型,利用KKT条件将其转化为上层模型的约束条件,再利用大M法将互补松弛条件线性化,形成易于求解的单层混合整数线性规划问题<sup>[26]</sup>,由于预测的数据负载量庞大,利用MAPPO算法对其进行求解更为方便。

#### 3.1 MAPPO算法原理

IDC源荷储的最佳协同优化调度可以表示为时间序列马尔可夫决策过程(Markov decision process, MDP)。MAPPO是应用于多智能体环境(proximal policy optimization, PPO)算法的变体,MAPPO同样采用了Actor-Critic架构,不同的是Critic网络学习了一个集中的价值函数<sup>[27]</sup>,可以使收敛性能和样本复杂性更佳。

一个环境和环境中的多个智能体的集合共同组成了多智能体系统,目的是将一个庞大的、复杂的系统转变为一个小型的、方便协调管理的、相互通信的系统<sup>[28]</sup>。MAPPO算法中,每个智能体的损失函数和权重更新方式不同,它们的Actor网络和Critic网络相互独立,离线训练过程如图5所示。

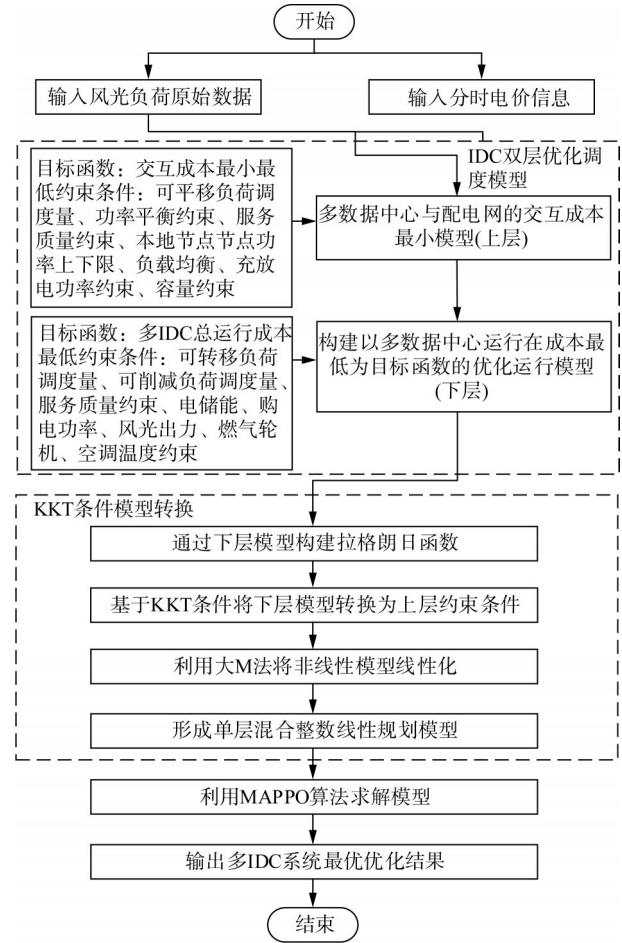


图4 多IDC系统的双层优化调度模型

Fig. 4 Dual-layer optimization model of multi-IDC system

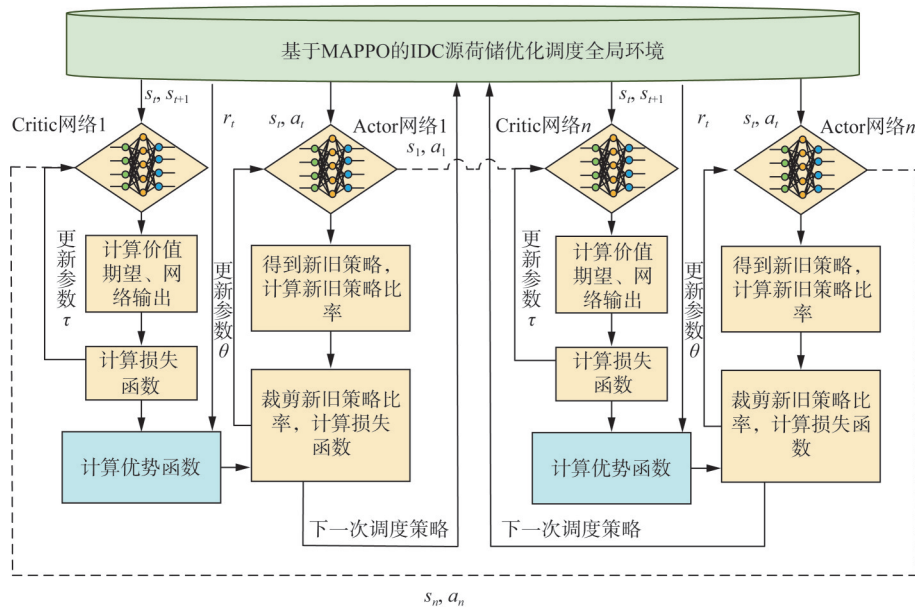


图5 基于MAPPO算法的IDC源荷储协同优化调度离线训练流程

Fig. 5 IDC source load storage collaborative optimization scheduling offline training process based on MAPPO algorithm

算法训练主要包括 3 个部分：首先，Critic 网络计算状态价值和折扣奖励，然后计算优势函数并更新 Critic 网络的权重；其次，Actor 网络与环境交互，将获得的有用信息存储在经验池中；最后在采样网络中计算重要性采样率，更新 Actor 网络的权重。通过同环境的不断交互学习，并采集环境状态组成的集合  $\Omega = \{\rho_1, \rho_2, \dots, \rho_n\}$ ，计算优势函数，训练出最优联合策略  $\rho'$ 。

### 1) Critic 网络

在 Critic 网络训练中，首先在经验样本集合中抽取一个序列  $(s_t, s_{t+1}, a_t, r_t)$ ，使用时间差分算法<sup>[29]</sup>计算目标价值函数  $y_{\text{target}}$ ，使用前向传播机制<sup>[30]</sup>计算当前状态  $s_t$  下的价值函数  $V(s_t; \tau)$ 。最后，更新 Critic 网络给出的策略参数  $\tau$  梯度如式 (47) — (49) 所示。

$$y_{\text{target}} = r_t + \sigma V(s_{t+1}; \tau) \quad (47)$$

$$L(\tau) = E[y_{\text{target}} - V(s_t; \tau)]^2 \quad (48)$$

$$\tau \leftarrow \tau - \theta \nabla L(\tau) \quad (49)$$

式中： $\tau$  为 Critic 网络的权重； $\theta$  为学习率； $L(\tau)$  为损失函数； $\nabla$  为梯度计算函数； $V(s_t; \tau)$  为  $t$  时刻 Critic 网络的输出； $y_{\text{target}}$  为  $t+1$  时刻价值函数的估计值； $r_t$  为  $t$  时刻奖励； $E(\cdot)$  为期望回报函数。

### 2) Actor 网络

在 Actor 网络中，在 TRPO<sup>[28]</sup> 的基础上进一步简化，通过优化损失函数  $J(\theta)$  来更新网络权重。通过引入优势函数  $A(s_t, a_t)$ ，可以有效地指导策略网络，引导其通过强化优势动作进一步提升网络性能<sup>[31]</sup>。策略网络的参数更新如式 (50) — (54) 所示。

$$Q(s_t, a_t) = E(r_t | s_t = s, a_t = a; \rho) \quad (50)$$

$$V(s) = E(r_t | s_t = s; \rho) \quad (51)$$

$$A(s_t, a_t) = Q(s_t, a_t) - V(s) \quad (52)$$

$$\begin{cases} J(\theta) = \max_{\theta} E_t \{ \min [ \frac{\rho_{\theta_{\text{new}}}(a_t | s_t)}{\rho_{\theta_{\text{old}}}(a_t | s_t)} A(s_t, a_t), \\ \text{clip}(\frac{\rho_{\theta_{\text{new}}}(a_t | s_t)}{\rho_{\theta_{\text{old}}}(a_t | s_t)}, 1 - \varepsilon, 1 + \varepsilon) A(s_t, a_t) ] \}, \end{cases} \quad (53)$$

$$\theta \leftarrow \theta + s \nabla_{\theta} J(\theta) \quad (54)$$

式中： $Q(s_t, a_t)$  为采取策略时状态  $s_t$  下采取动作  $a_t$  的价值； $V(s)$  为当前状态下按照策略  $\rho$  执行的价值期望； $\text{clip}(\cdot)$  函数为截断函数，引导新旧策略变化在  $[1 - \varepsilon, 1 + \varepsilon]$  之间，保证了 TRPO 算法的稳定性和可靠性； $s$  为学习率； $\theta$  为 Actor 网络的权重；

$\rho_{\theta_{\text{new}}}(a_t | s_t)$ 、 $\rho_{\theta_{\text{old}}}(a_t | s_t)$  分别为新、老策略对  $(s, a)$  被访问的概率； $E(r_t | s_t = s, a_t = a; \rho)$  和  $E(r_t | s_t = s; \rho)$  分别为新、老策略对  $(s, a)$  被访问的期望； $\nabla_{\theta}$  为折扣因子； $E_t(\cdot)$  为期望回报函数。

得到训练好的智能体后，再进行在线调度决策，输入系统状态得到最佳调度策略。MAPPO 算法的伪代码如表 1 所示。

表 1 MAPPO 伪代码

Tab. 1 MAPPO Pseudo-code

| MAPPO 伪代码                                                                     |
|-------------------------------------------------------------------------------|
| 1. 初始化 Critic 和 Actor 网络相关参数；                                                 |
| 2. 初始化经验样本池 $E$ ；                                                             |
| 3. for $n$ in total numbers                                                   |
| (1) for $t=1, 2, \dots, T$                                                    |
| I. 智能体 $i$ 执行策略，得到动作 $a_{i,t}$ ；                                              |
| II. 执行动作 $a_{i,t}$ ，计算奖励 $r_{i,t}$ ，得到新状态 $s_{i,t+1}$ ；                       |
| III. 把 $s_{i,t}, a_{i,t}, s_{i,t+1}, r_{i,t}$ 放回经验样本池 $E$ ；                   |
| End for                                                                       |
| (2) 通过式 (50) 和式 (52) 分别计算 $Q_i(s_{i,t}, a_{i,t})$ 和 $A_i(s_{i,t}, a_{i,t})$ ； |
| (3) for $i=1, 2, \dots, H$ ( $H$ 为智能体的数量)                                     |
| I. 随机从经验样本池 $E$ 中抽取经验组 $E_h$ ；                                                |
| II. 通过式 (49) 和式 (54) 分别更新 $\tau_i$ 和 $\theta_i$ ；                             |
| End for                                                                       |
| End for                                                                       |

## 3.2 基于 MAPPO 算法的多 IDC 源荷储优化调度

首先将多 IDC 源荷储协同优化调度优化问题转化成 MDP，设计基于 MAPPO 的状态空间和动作空间，并根据目标函数设计奖励函数。

### 3.2.1 状态空间

选取 IDC 工作负荷量、冷负荷、可再生能源出力、分时购售电价作为状态变量，由此可建立 IDC 调度模型的状态空间  $S_{i,t}$  如式 (55) 所示。

$$S_{i,t} = \left\{ L_{\text{ba}, i, t}, L_{\text{tran}, i, t}, L_{\text{cut}, i, t}, P_{\text{tol}, i, t}, Q_{\text{air}, i, t}, T_{\text{in}, i, t}, T_{\text{out}, i, t}, P_{\text{wt}, i, t}, P_{\text{pv}, i, t}, \alpha_{\text{buy}, t}, \alpha_{\text{sell}, t}, S_{\text{sto}, i, t} \right\} \quad (55)$$

式中： $L_{\text{ba}, i, t}, L_{\text{tran}, i, t}, L_{\text{cut}, i, t}$  分别为  $t$  时刻到达 IDC 的可平移、可转移、可削减数量； $T_{\text{in}, i, t}$  为  $t$  时刻 IT 设备入口气流温度，其中下标  $i$  为数据中心的编号； $T_{\text{out}, i, t}$  为  $t$  时刻空调送风温度； $P_{\text{tol}, i, t}$  为  $t$  时刻 IDC 实时总运行电功率； $Q_{\text{air}, i, t}$  为  $t$  时刻 IDC 实时总运行冷负荷； $P_{\text{wt}, i, t}, P_{\text{pv}, i, t}$  分别为  $t$  时刻风机和光伏出力； $\alpha_{\text{buy}, t}, \alpha_{\text{sell}, t}$  分别为分时买卖电价； $S_{\text{sto}, i, t}$  为  $t$  时刻系统储能容量。



### 3.2.2 动作空间

在接收到环境给出的状态后,智能体将根据当前时刻的策略从动作空间中采取动作。动作空间包括燃气轮机的发电功率、电网流入的电功率、电储能充/放电功率。因此,动作空间可定义为:

$$A_{i,t} = \{P_{mt,i,t}, P_{grid,i,t}, P_{cha,i,t}, P_{dis,i,t}\} \quad (56)$$

式中:  $P_{mt,i,t}$  为  $t$  时刻燃气轮机输出的电功率,下标  $i$  为数据中心的编号;  $P_{grid,i,t}$  为  $t$  时刻从电网流入的电功率;  $P_{cha,i,t}$ 、 $P_{dis,i,t}$  分别为  $t$  时刻电储能充、放电功率。

### 3.2.3 奖励函数

奖励函数用来引导智能体学习,各IDC共享一个奖励函数,并将系统内违反系统正常运行的约束条件以惩罚函数的形式添加到奖励函数中。

1) 根据上、下层成本函数  $F_{1t}$ 、 $F_{2t}$ , 经过 KKT 变换后形成单层混合整数规划模型, 得到奖励函数  $r_1(t)$ ;

2) 以光伏为例,  $t$  时刻光伏违反设备运行约束的惩罚函数  $\delta_{pv,i,t}$  为:

$$\delta_{pv,i,t} = -( |P_{pv,i,t} - P_{pv,min}| + |P_{pv,max} - P_{pv,i,t}| + |P_{pv,max} - P_{pv,min}| ) \quad (57)$$

$t$  时刻风机、燃气轮机、送风温度、IT 设备入口流温度、电储能、电网交互约束的惩罚函数与光伏具有相同的形式,在此不再赘述。

参考文献[32],  $t$  时刻运行设备违反运行约束的罚函数为:

$$r_2(t) = -\xi_1 \times (\delta_{pv,i,t} + \delta_{wt,i,t} + \delta_{mt,i,t} + \delta_{T_{in,i,t}} + \delta_{T_{out,i,t}} + \delta_{cha,i,t} + \delta_{dis,i,t} + \delta_{sto,i,t} + \delta_{grid,i,t}) \quad (58)$$

式中:  $\xi_1$  为部分运行设备的惩罚系数,本文取 0.001;  $\delta_{wt,i,t}$ 、 $\delta_{mt,i,t}$ 、 $\delta_{T_{in,i,t}}$ 、 $\delta_{T_{out,i,t}}$ 、 $\delta_{cha,i,t}$ 、 $\delta_{dis,i,t}$ 、 $\delta_{sto,i,t}$ 、 $\delta_{grid,i,t}$  分别为第  $i$  个数据中心  $t$  时刻风机、燃气轮机、IT 设备入口流温度、送风温度、电储能充电功率、电储能放电功率、电储能容量、电网交互功率约束的惩罚函数。

3) 在保证服务质量的情况下,  $t$  时刻违反工作负荷迁移、削减约束的惩罚函数为:

$$05_3(t) = -\xi_2 \times ( |L_{ba,i,t} - 0| + |L_{ba,max} - L_{ba,i,t}| + |L_{ba,max} - 0| + |L_{tran,i,t} - 0| + |L_{tran,max} - L_{tran,i,t}| + |L_{tran,max} - 0| + |L_{cut,i,t} - 0| + |L_{cut0,i,t} - L_{cut,i,t}| + |L_{cut0,i,t} - 0| ) \quad (59)$$

式中  $\xi_2$  为惩罚系数,本文取 0.001。

综上,总奖励函数  $R(t)$  可表示为:

$$R(t) = -\sum_{t=1}^{24} \sum_{i=1}^N \xi_3 (r_1(t) + r_2(t) + r_3(t)) \quad (60)$$

式中:  $r_1(t)$  为系统运行成本;  $r_2(t)$ 、 $r_3(t)$  分别为运行设备、工作负荷调度的惩罚;  $R(t)$  为总奖励函数;  $\xi_3$  为奖励的缩放系数,本文取 0.000 01。

## 4 算例分析

### 4.1 算例数据设置

算例采用由 IDC 运营商管理的 3 个不同地理位置、不同电力市场环境的 IDC 进行仿真验证。调度周期内总时段数为 24 h,相邻调度时段间隔为 1 h。每个 IDC 内有 2 000 台同构服务器、10 台空调;假设 IDC 中所有服务器型号相同,且任意  $t$  时段负荷均匀分配给所有开机状态下的服务器。各 IDC 日前电价曲线如附图 A1 所示,售电价格为购电价格的一半,固定气价为 2.5 元/m<sup>3</sup>; IDC 机房送风温度参数如附图 A2 所示;各 IDC 可再生能源机组日前预测曲线如附图 A3 所示;根据负荷均衡原则,将可转移负荷平均分配给各子 IDC,各 IDC 空间调度前的可转移负荷需求曲线如附图 A4 所示;服务器集群特性见附表 A1;系统各设备参数见附表 A2;不同地域单位负荷转移成本见附表 A3。根据文献[18],单个负荷任务大小为 0.2 GB,可转移负荷最大可容忍延迟时间  $T_{max}$  为 0.3 s,可平移负荷最大延迟时间为 8h。仿真计算设备基于 MATLAB R2022b 实现,在 64 位 Windows 系统、Intel Core i7-10750H CPU@2.6 GHz 的环境下运行。

### 4.2 在线调度及仿真结果分析

为探究 IDC 双层优化调度策略对系统效益和成本的影响,依据数据负荷时空分配特性设立 4 种场景对比分析。

场景 1: 既不进行负荷空间转移和削减调度,也不进行负荷时间平移。

场景 2: 只进行负荷空间转移和削减,但不进行负荷平移。

场景 3: 只进行负荷平移,但不进行负荷空间转移和削减。

场景 4: 本文所提优化策略,进行负荷空间转移和削减,并进行负荷平移。

采用 MAPPO 算法,假设风电、光伏、工作负荷、机房温度的预测误差满足正态分布,标准差为 10%,均值为各时段出力预测值。根据预测误差,

通过蒙特卡洛抽样随机生成 4 000 个训练曲线, 作为 MAPPO 算法的训练集。Clip 函数中的参数  $\varepsilon$  为 0.2, 折扣因子  $\gamma$  为 0.96, Critic 网络和 Actor 网络的学习率均为 0.000 01, 计划迭代 5 000 次。得到多智能体累积奖励的曲线如图 6 所示。

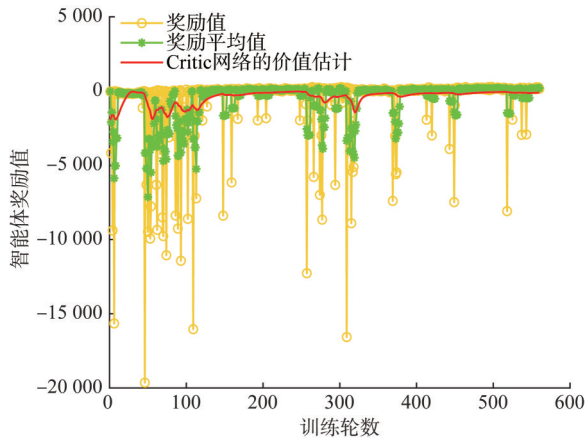


图 6 多智能体奖励值曲线

Fig. 6 Curves of multi-agent reward value

由图 6 可以看出, 深度强化学习智能体在训练 560 轮后实现了收敛。在训练开始时, 由于智能体不了解环境因而作出的动作违反约束导致奖励值大幅波动。随着不断训练智能体不断与环境学习, 逐渐减少动作越限的情况。随着每回合训练得到的奖励值 (episodes reward)、回合训练奖励的平均值 (average reward) 和 Critic 网络对状态行为的价值估计 (episode  $Q_0$ , 即  $y_{\text{target}}$ ) 几乎接近, 表示智能体已经学习到最佳的调度策略。

优化前的 IDC 可平移/可削减负荷和空间调度后的可转移负荷情况如图 7—8 所示, 使用训练完成的智能体对日前预测的数据负荷优化, 得到各 IDC 最佳调度方案如图 9—12 所示。

为验证算法有效性, 对 MAPPO 算法求解结果具体分析。以 IDC1 为例, 由图 9 可以看出, IDC 上层传至下层的可平移负荷不需要立即处理, 对时间敏感性弱, 在确保用户服务质量的同时, 将电价高、数据量大的时段 (10—12、17—18) 尽量分配给其他电价低、可再生能源产量高的时段; 由于夜间风电强劲, 可再生能源生产充足, 应用负荷平移特性不仅可以降低电力消耗, 还可以降低系统中的购电成本, 实现可再生能源消耗; IDC 下层的可转移负荷则需要立即处理, 对时间敏感性强, 可削减负

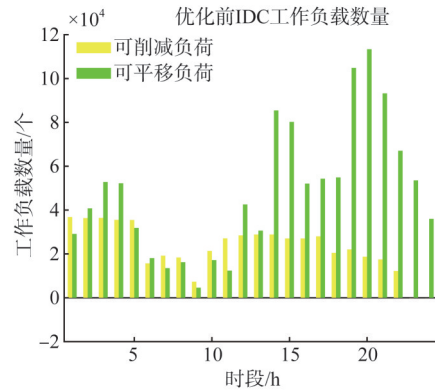


图 7 优化前的 IDC 可平移/可削减负荷调度情况

Fig. 7 Scheduling situation of IDC shift/reduction reduction before optimization

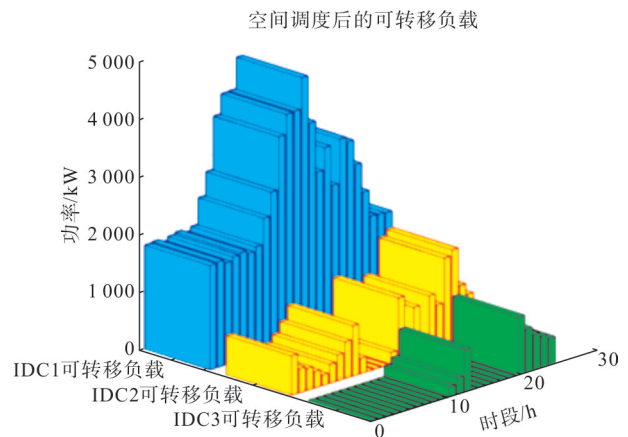


图 8 IDC 空间调度后的可转移负荷调度情况

Fig. 8 Scheduling situation of transferable loads after IDC space scheduling

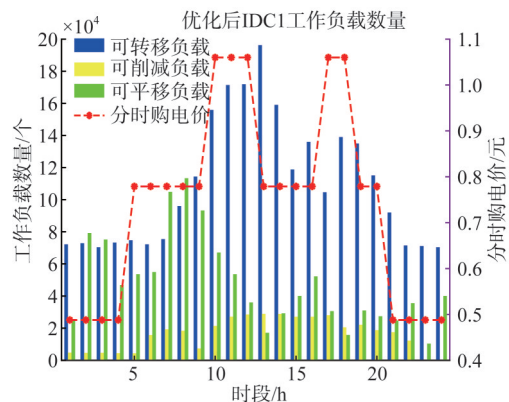


图 9 时空调度后的 IDC1 工作负荷调度情况

Fig. 9 Scheduling situation of IDC1 workloads after spatiotemporal scheduling

荷在需求高峰期 (6—22) 减少数据任务, 减少了服务器等 IT 设备配置容量, 降低了系统运营成本。由图 12 可以看出, 仅考虑时间调度时, 对于时段 8—

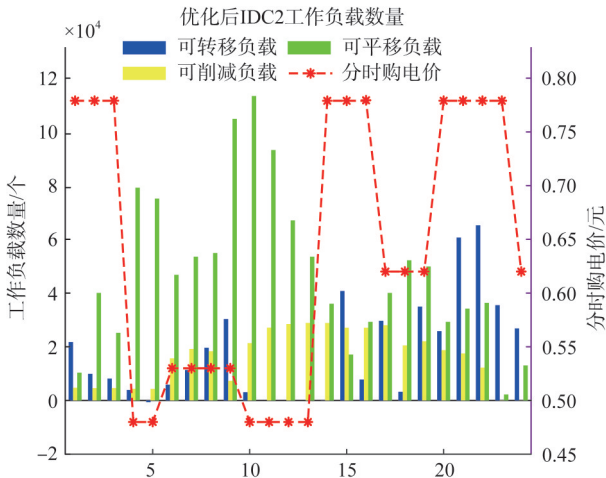


图 10 时空调度后的 IDC2 工作负荷调度情况  
Fig. 10 Scheduling situation of IDC2 workloads after spatiotemporal scheduling

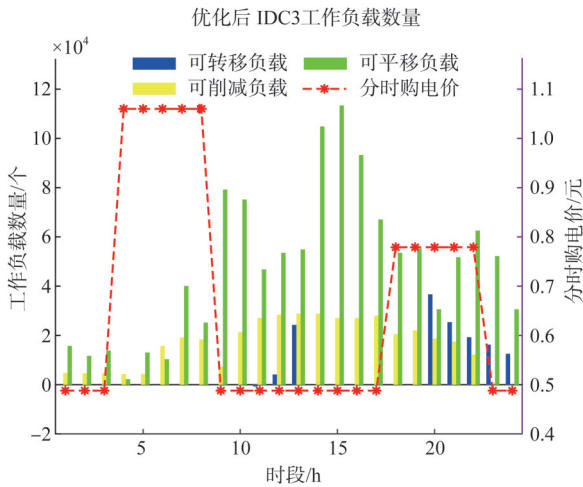


图 11 时空调度后的 IDC3 工作负荷调度情况  
Fig. 11 Scheduling situation of IDC3 workloads after spatiotemporal scheduling

10, 由于此时段 3 个 IDC 电价相对较低, 且风光出力充足, 优化后三个数据中心的电功率显著增加; 对于时段 18—22, 由于此时段可再生能源出力不足, 电网需大量购电, 此时段的三个数据中心优化后的电功率显著减小。

由于 IDC 下层的可转移负荷则需要立即处理, 由附图 A4 和图 8 对比看出, 各 IDC 可转移负荷最优调度计划的制定主要取决于可再生能源出力及单位负荷转移成本。在 10—20 时段, IDC1 向 IDC2、3 转移了部分可转移数据负荷, 这是因为 IDC1 的单位负荷转移成本较小, 对用户的转移成本越小, IDC 运营商的收益越大; 在 15—20 时段, 此时

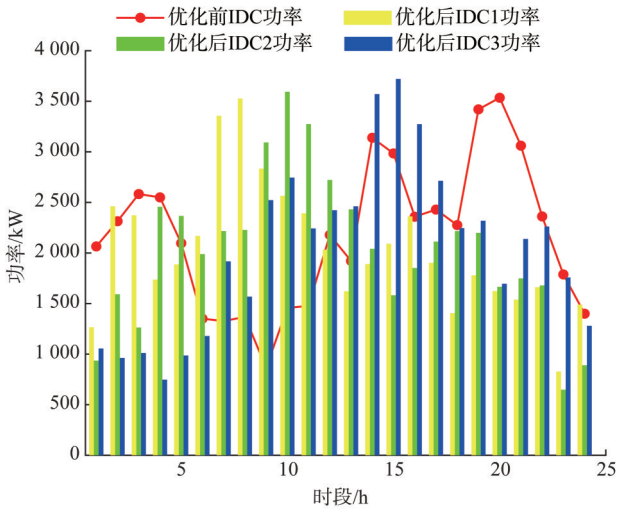


图 12 时间优化前后的 IDC 电功率变化情况  
Fig. 12 Scheduling situation of IDC power change before and after time optimization

IDC2 的可再生能源出力较大, 故该时段分配的可转移负荷较多, 有利于促进可再生能源的消纳, 降低弃风弃光成本; 由于 IDC3 可再生能源机组出力相对较少, 且其单位负荷转移成本相对较大, 故分配给该 IDC 的可转移负荷相对较少; 可以看出利用 IDC 用能空间可调特性对整个系统中分布式光伏消纳的促进作用更为明显。

4.3 多 IDC 运行成本分析

表 2 为不同场景的运行成本, 由场景 2 可以看出, 经过空间调度和削减, 与配电网交互成本明显下降, 即 IDC 的收益明显增大。由于上下层是相互耦合, 相互依赖的, 因此考虑负荷时间平移、空间转移和削减在绝大多数时刻对多 IDC 与配电网之间的交互功率有降低作用。这是因为相对于独立 IDC, 多 IDC 系统间进行数据传输, 互为备用, 减小了各子 IDC 处理数据负载的压力。

通过对比场景 2 与场景 4 看出, 在考虑 IDC 间空间转移的基础上, 进一步考虑负荷平移, 可以

表 2 不同场景结果对比

| Tab. 2 Comparison of results in different scenarios 万元 |          |            |            |            |
|--------------------------------------------------------|----------|------------|------------|------------|
| 场景                                                     | 与配电网交互成本 | IDC1 运行总成本 | IDC2 运行总成本 | IDC3 运行总成本 |
| 1                                                      | 135.63   | 128.95     | 154.63     | 86.53      |
| 2                                                      | -25.34   | 117.38     | 149.69     | 83.65      |
| 3                                                      | 58.53    | 46.76      | 57.42      | 53.23      |
| 4                                                      | -38.67   | 35.19      | 54.48      | 50.35      |



降低 IDC 与配电网的交互功率。因为用户侧需求响应可以达到数据负荷的削峰填谷,降低峰值负载,降低功耗。并且由表 2 对比可以看出,负荷如果进行空间转移对系统成本的影响大于时间维度平移的影响;若系统不进行负荷的空间转移,系统总运行总成本将大幅增加;负荷无法进行区域互补,无法分担其他 IDC 区域内数据负荷的压力, IDC 无法起到削峰填谷的作用,只能在电价高峰期购电以满足需求,使购电、购气成本大大增加。对比可知,本文所提出的优化策略(即场景 4),在经济成本方面具有显著优势,在满足用户服务质量的条件下实现了削峰填谷,减少了与配电网交互成本,提高了运行经济性,促进可再生能源消纳,满足了 IDC 源荷储调度模型经济、环保的运行要求。图 13 为 IDC1 源荷储双层调度模型电负荷平衡曲线。

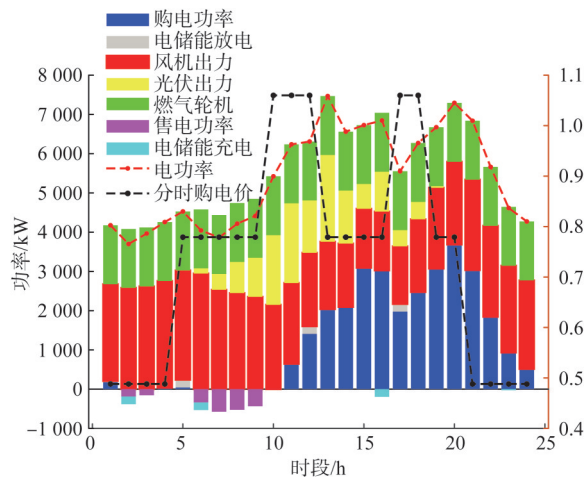


图 13 IDC1 源荷储双层调度模型电负荷平衡曲线

Fig. 13 Electric load balance curves of IDC1 dual-layer scheduling model

以 IDC1 为例进行分析,由图 13 可以看出,在电价高峰期(10—12、17—18),电能主要由微型燃气轮机和风能和太阳能发电机组生产,为了促进可再生能源消纳, IDC 向电网购电较少;在电价低谷期(13—16),由于可平移负荷主要被分配到这些时段,通过平移部分数据负荷,减少了其他时段的服务器开放数量,合理地分配了能量,当机组出力不足时从主电网购买的电能比例显著增加;若风光出力大于所需负荷时服务器部分关闭,多余的风光出力被储存起来,并在负荷高峰期、风光发电不足时排放,从而降低了 IDC 购电

成本。

IDC 机房 IT 设备运行过程中会产生大量废热,以 IDC1 为例,系统在运行周期内的空调设备耗电调度前后对比如图 14 所示。由于空调设备耗电量与 IT 设备能耗有关,空调耗电和数据调度具有相似性,在不考虑可转移负荷的情况下,即在电价低、数据量大且可再生能源出力充足的时段耗能多,在电价高、数据量低且可再生能源出力充足的时段耗能低,实现对服务器周围环境的散热功能。

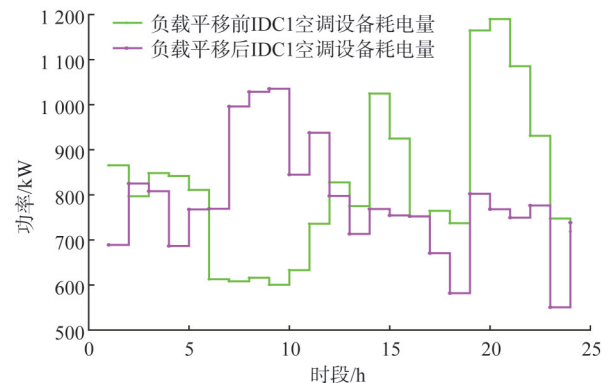


图 14 IDC1 调度前后空调耗电量对比

Fig. 14 Comparison of power consumptions of air conditioner before and after IDC1 dispatching

#### 4.4 本文算法与其他算法对比分析

##### 4.4.1 算法有效性分析

为验证本文所提方法的有效性,将本文方法与深度强化学习方法 MADDPG (muti-agent deep deterministic policy gradient, MADDPG) 和 ADMM 传统优化算法进行对比,对比结果如表 3 所示。

表 3 不同算法结果对比

Tab. 3 Comparison of different algorithms results

| 方法     | 求解时间/s | IDC1 运行成本/万元 |
|--------|--------|--------------|
| MAPPO  | 1.2    | 35.19        |
| MADDPG | 1.3    | 36.25        |
| ADMM   | 1 562  | 41.36        |

由表 3 可以看出,ADMM 算法求解时间最长,这是因为本文求解模型较为复杂,数据量庞大,实验过程中波动较大,调度结果部分存在设备越限的情况,而深度强化学习算法求解相对迅速,调度结果满足约束限制,在处理大数据问题中占据很大优势;深度强化学习算法相对其他方法求解的运行成

表4 鲁棒性结果对比

Tab. 4 Comparison of robustness results

万元

| 预测误差         | 方法     | 配网交互成本 | 时间调度激励收益 | 空间调度成本 | 运维/储能充放电成本 | 购能成本  | IDC1运行成本 |
|--------------|--------|--------|----------|--------|------------|-------|----------|
| $\sigma=0.1$ | MAPPO  | -38.67 | 53.46    | 45.19  | 24.20      | 17.68 | 35.19    |
|              | MADDPG | -36.36 | 51.35    | 36.25  | 22.91      | 23.02 | 36.25    |
|              | ADMM   | -33.31 | 49.38    | 41.36  | 36.20      | 28.15 | 41.36    |
| $\sigma=0.2$ | MAPPO  | -36.48 | 49.58    | 36.56  | 18.97      | 18.89 | 36.56    |
|              | MADDPG | -32.57 | 45.37    | 37.69  | 29.24      | 24.12 | 37.69    |
|              | ADMM   | -31.65 | 42.69    | 40.75  | 36.83      | 27.73 | 40.75    |

本较低,本文MAPPO算法相对MADDPG运行成本降低了2.92%;实验中发现MADDPG算法对超参数十分敏感,若超参数选择不恰当则不容易收敛,且离线训练收敛时间比本文算法长10.7 min;虽然本文算法离线训练时长为5.5 min,但在线决策只需1.2 s,可以进行长时间在线应用。

#### 4.4.2 算法鲁棒性分析

为了验证所提出的优化运算方法在示例IDC1上的稳定性,假设3种数据负荷的预测误差满足相同的正态分布 $N(0, \sigma^2)$ ,优化结果如表4所示。

随着不断增大数据负荷不确定性,IDC运行的经济性也在下降。然而,在相同的不确定性水平下本文提出的算法具有最佳的经济性,这表明MAPPO算法对计算任务的不确定性带来的影响更小,鲁棒性更好。

## 5 结论

本文提出了一种基于MAPPO算法的IDC双层优化调度模型。将IDC动态调度问题转化为具有连续状态和动作的深度强化学习问题来处理,与传统的优化算法对比分析发现,MAPPO算法能够快速在线决策,具有良好的收敛性和经济性。

对IDC的工作负载、服务器、空调进行机理建模,同时受服务质量约束,IDC的可平移负载被分配到电价较低、可再生能源功率较高的时段进行处理。将IDC的可转移负荷分配到可转移成本较低的区域IDC进行处理,减少了与配电网交互成本,提高能源利用率,促进可再生能源消纳。

通过算例仿真分析,综合考虑服务质量和供能成本,对比4种不同场景,分析得出采用本文所提出规划方案可以充分发挥IDC工作负荷灵活性潜力,实现调峰填谷,降低系统运行成本,提升系统能效降低能耗,减少环境污染,且具有良好

鲁棒性。为日后IDC源荷储能量管理提供了新思路。

## 参考文献

- [1] WAN Tong, TAO Yuechuan, QIU Jing, et al. Internet data centers participating in electricity network transition considering carbon-oriented demand response [J]. Applied Energy, 2023(329): 120305.1 – 120305.10.
- [2] 可思为,董萍,马铭宇,等.考虑风光荷时空互补的多能源绿色数据中心多目标配置方法[J].电力系统保护与控制,2024,52(22):22 – 33.  
KE Siwei, DONG Ping, MA Mingyu, et al. A multi-objective allocation method for multi-energy green data centers considering wind, solar and load spatial-temporal complementarity [J]. Power System Protection and Control, 2024, 52(22): 22 – 33.
- [3] LINDBERG J, LESIEUTRE B C, ROALD L A. Using geographic load shifting to reduce carbon emissions [J]. Electric Power Systems Research, 2022 (212): 108586.1 – 108586.11.
- [4] 王新令.积极应对气候变化,推动清洁低碳转型:《2030年前碳达峰行动方案》解读[J].中国电业,2021(12):22 – 25.  
WANG Xinling. Actively respond to climate change and promote clean and low-carbon transformation: interpretation of peak carbon dioxide emissions action plan to 2030 [J]. China Electric Power, 2021(12): 22 – 25.
- [5] 闫龙川,白东霞,刘万涛,等.人工智能技术在云计算数据中心能量管理中的应用与展望[J].中国电机工程学报,2019,39(1):31 – 42, 318.  
YAN Longchuan, BAI Dongxia, LIU Wantao, et al. Application and prospect of artificial intelligence technology in energy management and optimization for cloud computing data center [J]. Proceedings of the CSEE, 2019, 39(1): 31 – 42, 318.
- [6] 丁肇豪,曹雨洁,张素芳,等.能源互联网背景下数据中心与电力系统协同优化(一):数据中心能耗模型[J].中国电机工程学报,2022,42(9):3161 – 3177.  
DING Zhaohao, CAO Yujie, ZHANG Sufang, et al. Coordinated operation for data center and power system in the context of energy Internet ( I ): energy demand management model of data center [J]. Proceedings of the CSEE, 2022, 42 (9): 3161 – 3177.
- [7] 李彬,杜亚彬,曹望璋,等.考虑风光储互补与工作负载分配的数据中心优化调度[J].现代电力,2022,39(3):356 – 363.  
LI Bin, DU Yabin, CAO Wangzhang, et al. Optimal scheduling of data center considering wind-solar-storage complementary and workload distribution [J]. Modern Electric

- Power, 2022, 39(3): 356 – 363.
- [8] 朱一昕, 王靖芸, 毕恺韬, 等. 基于随机模型预测控制的数据中心微电网能量优化调度研究[J]. 智慧电力, 2022, 50(8): 102 – 108.
- ZHU Yixin, WANG Jingyun, BI Kaitao, et al. Energy optimization dispatch of data center microgrid based on stochastic model predictive control[J]. Smart Power, 2022, 50(8): 102 – 108.
- [9] 温正楠, 刘继春. 风光水互补发电系统与需求侧数据中心联动的优化调度方法[J]. 电网技术, 2019, 43(7): 2449 – 2460.
- WEN Zhengnan, LIU Jichun. A optimal scheduling method for hybrid wind-solar-hydro power generation system with data center in demand side[J]. Power System Technology, 2019, 43(7): 2449 – 2460.
- [10] LIU H Z, LIU B S, YANG L T, et al. Thermal-aware and DVFS-enabled big data task scheduling for data centers [J]. IEEE Transactions on Big Data, 2018, 4(2): 177 – 190.
- [11] 郭嘉, 杨洪明, 张鹏兴, 等. 面向需求响应的数据中心供能系统优化调度[J]. 电力科学与技术学报, 2018, 33(2): 43 – 49.
- GUO Jia, YANG Hongming, ZHANG Pengxing, et al. Optimal dispatching of data center energy-supplying system with demand response [J]. Journal of Electric Power Science and Technology, 2018, 33(2): 43 – 49.
- [12] 严干贵, 阚天洋, 杨玉龙, 等. 基于深度强化学习的分布式电采暖参与需求响应优化调度[J]. 电网技术, 2020, 44(11): 4140 – 4149.
- YAN Gangui, KAN Tianyang, YANG Yulong, et al. Demand response optimal scheduling for distributed electric heating based on deep reinforcement learning [J]. Power System Technology, 2020, 44(11): 4140 – 4149.
- [13] 刘威, 张东霞, 王新迎, 等. 基于深度强化学习的电网紧急控制策略研究[J]. 中国电机工程学报, 2018, 38(1): 109 – 119, 347.
- LIU Wei, ZHANG Dongxia, WANG Xinying, et al. A decision making strategy for generating unit tripping under emergency circumstances based on deep reinforcement learning [J]. Proceedings of the CSEE, 2018, 38(1): 109 – 119, 347.
- [14] 余宏晖, 林声宏, 朱建全, 等. 基于深度强化学习的微电网在线优化[J]. 电测与仪表, 2023, 43(6): 1884 – 1892.
- YU Honghui, LIN Shenghong, ZHU Jianquan et al. Online optimization of microgrid based on deep reinforcement learning [J]. Electrical Measurement and Instrumentation, 2023, 43(6): 1884 – 1892.
- [15] 李怡瑾, 唐昊, 吕凯, 等. 源荷不确定冷热电联供微网能量调度的建模与学习优化[J]. 控制理论与应用, 2018, 35(1): 56 – 64.
- LI Yijin, TANG Hao, LÜ Kai, et al. Modeling and learning-based optimization of the energy dispatch for a combined cooling, heat and power microgrid system with uncertain sources and loads [J]. Control Theory & Applications, 2018, 35(1): 56 – 64.
- [16] 方和平, 刘曙光, 冉泳屹, 等. 基于深度强化学习的多数据中心一体化调度优化[J], 2023, 43(6): 1884 – 1892.
- FANG Heping, LIU Shuguang, RAN Yongyi, et al. Multi-data center integrated scheduling optimization based on deep reinforcement learning [J]. Computer Applications, 2023, 43(6): 1884 – 1892.
- [17] 王晴, 刘友波, 黄杨, 等. 考虑数据中心需求响应的城市电网阻塞管理[J]. 电网技术, 2020, 44(8): 3129 – 3138.
- WANG Qing, LIU Youbo, HUANG Yang, et al. Congestion management of urban power grid considering demand response of data center [J]. Power System Technology, 2020, 44(8): 3129 – 3138.
- [18] 曾博, 刘一贤, 张卫翔, 等. 综合考虑内外生不确定性的绿色数据中心多域资源协同规划方法[J]. 中国电机工程学报, 2024, 44(12): 4732 – 4748.
- ZENG Bo, LIU Yixian, ZHANG Weixiang, et al. Multi-domain resource collaborative planning method for green data center considering internal and external uncertainty [J]. Proceedings of the CSEE, 2024, 44(12): 4732 – 4748.
- [19] 詹勇. 数据中心高效与公平的能源性能管理[D]. 成都: 电子科技大学, 2017.
- [20] 王兴华, 可思为, 潘柏崇, 等. 考虑子模块特性的多站融合光储容量优化配置方法[J]. 南方电网技术, 2025, 19(1): 106 – 117.
- WANG Xinghua, KE Siwei, PAN Baichong, et al. Optimal configuration method of PV-storage capacity of multi-station integration considering submodule characteristics [J]. Southern Power System Technology, 2025, 19(1): 106 – 117.
- [21] 黄义成. 数据中心空调系统能耗分析[J]. 智能建筑, 2014(10): 52 – 55.
- HUANG Yicheng. Analysis of energy consumption of air conditioning system in data center [J]. Intelligent Building, 2014(10): 52 – 55.
- [22] 谷丽君, 白焰. 基于芯片温度感知负载调度的数据中心能耗优化[J]. 华北电力大学学报(自然科学版), 2018, 45(6): 100 – 106.
- GU Lijun, BAI Yan. Chip temperature-aware workload scheduling-based power consumption optimization of data center [J]. Journal of North China Electric Power University (Natural Science Edition), 2018, 45(6): 100 – 106.
- [23] 兰洲, 蒋晨威, 谷纪亭, 等. 促进可再生能源发电消纳和碳减排的数据中心优化调度与需求响应策略[J]. 电力建设, 2022, 43(4): 1 – 9.
- LAN Zhou, JIANG Chenwei, GU Jiting, et al. Optimal dispatch and demand response strategies of data centers for promoting accommodation of renewable energy generation and reducing carbon emission [J]. Electric Power Construction, 2022, 43(4): 1 – 9.
- [24] 张鹏. 考虑需求响应的虚拟数据中心集群优化运营方法研究[D]. 北京: 华北电力大学, 2021.
- [25] 吴云芸, 方家琨, 艾小猛, 等. 计及需求响应的数据中心联盟共享储能规划[J]. 电力系统自动化, 2023, 47(7): 42 – 50.
- WU Yunyun, FANG Jiakun, AI Xiaomeng, et al. Shared energy storage planning of data center alliance with demand response [J]. Automation of Electric Power Systems, 2023, 47(7): 42 – 50.
- [26] 张儒峰, 姜涛, 李国庆, 等. 考虑电转气消纳风电的电-气综合能源系统双层优化调度[J]. 中国电机工程学报, 2018, 38(19): 5668 – 5678, 5924.
- ZHANG Rufeng, JIANG Tao, LI Guoqing, et al. Bi-level optimization dispatch of integrated electricity-natural gas systems considering P2G for wind power accommodation [J]. Proceedings of the CSEE, 2018, 38(19): 5668 – 5678, 5924.
- [27] GUO Delin, TANG Lan, ZHANG Xinggan, et al. Joint optimization of handover control and power allocation based on multi-agent deep reinforcement learning [J]. IEEE Transactions on Vehicular Technology, 2020, 69(11): 13124 – 13138.



[28] 叶佩文, 贾向东, 杨小蓉, 等. 面向车联网的多智能体强化学习边云协同卸载[J]. 计算机工程, 2021, 47(4): 13 – 20.  
YE Peiwen, JIA Xiangdong, YANG Xiaorong, et al. Collaborative edge and cloud offloading for Internet of vehicles using multi-agent reinforcement learning [J]. Computer Engineering, 2021, 47(4): 13 – 20.

[29] 陈新. 基于强化学习的调度与导航策略研究[D]. 杭州: 浙江大学, 2021.

[30] SCHULMAN J, LEVINE S, MORITZ P, et al. Trust region policy optimization[J]. Computer Science, 2015(10): 1889 – 1897.

[31] 李涛, 胡维昊, 李坚, 等. 基于深度强化学习算法的光伏-抽蓄互补系统智能调度[J]. 电工技术学报, 2020, 35(13): 2757 – 2768.  
LI Tao, HU Weihao, LI Jian, et al. Intelligent economic dispatch for PV-PHS integrated system: a deep reinforcement learning-based approach [J]. Transactions of China Electrotechnical Society, 2020, 35(13): 2757 – 2768.

[32] 蒯伟山, 王小君, 孙庆凯, 等. 计及安全约束的综合能源系统深度强化学习优化调度策略研究[J]. 电网技术, 2023, 47(5): 1970 – 1983.  
LIN Weishan, WANG Xiaojun, SUN Qingkai, et al. Optimal dispatch strategy of integrated energy system based on deep reinforcement learning considering security constraints [J]. Power System Technology, 2023, 47(5): 1970 – 1983.

收稿日期: 2023-12-04;网络首发日期:2024-06-28  
作者简介:  
杨秀(1972), 男, 教授, 博士, 从事分布式发电与微电网技术、主动配电网与能源互联网运行仿真研究, angxiu721102@126.com;  
张相寅(1999), 男, 通信作者, 硕士研究生, 研究方向为数据中心负荷响应特性及优化调度, 2271098983@qq.com;  
黄海涛(1978), 女, 教授, 博士, 研究方向为电力市场和能源互联, hbd1\_huanghaitao@sina.com。

附录

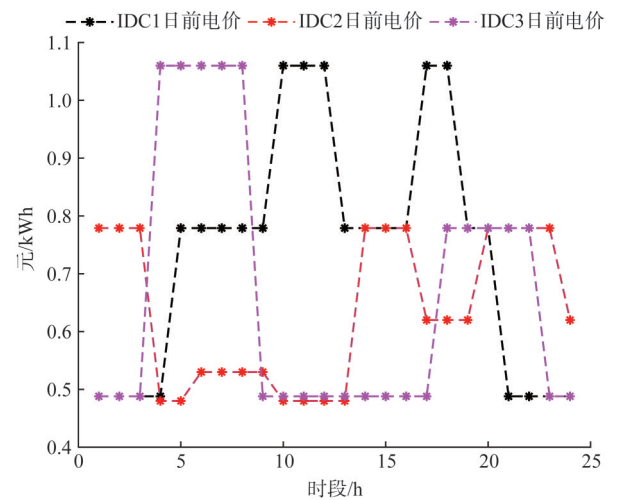


图 A1 各 IDC 日前电价曲线  
Fig. A1 Day-ahead electricity price curves of various IDCs

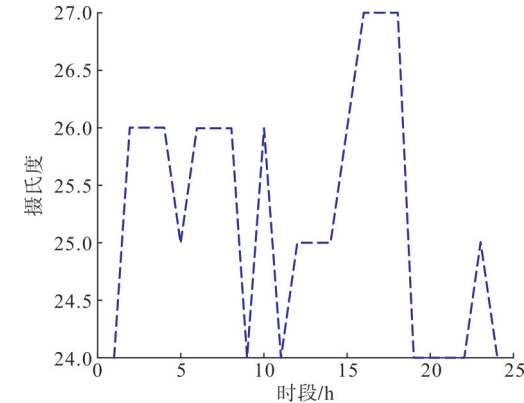


图 A2 机房送风温度参数曲线  
Fig. A2 Curve of temperature parameters in the equipment room

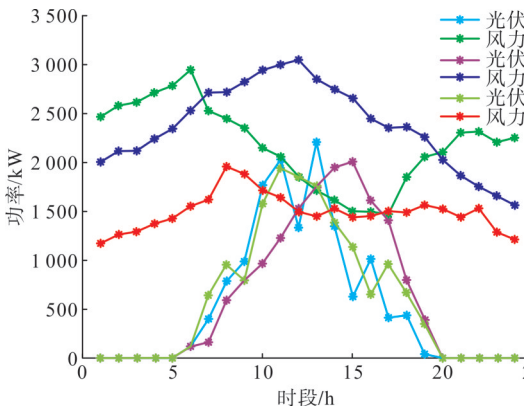


图 A3 风机、光伏出力曲线  
Fig. A3 Output curves of wind turbines and photovoltaics

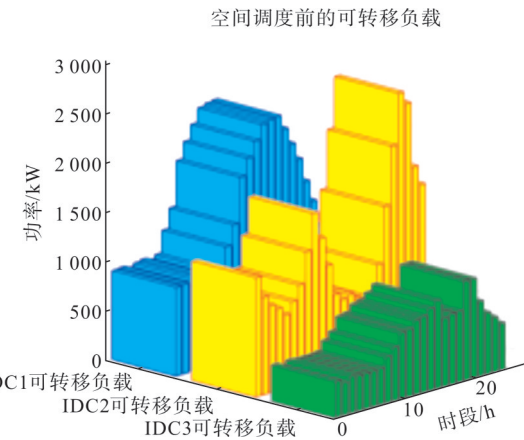


图 A4 IDC 空间调度前的工作负荷调度情况  
Fig. A4 Workload schedulings before IDC space scheduling

| 表 A1 服务器相关参数                             |             |             |             |                        |
|------------------------------------------|-------------|-------------|-------------|------------------------|
| Tab. A1 Related parameters of the server |             |             |             |                        |
| 参数                                       | 服务器数<br>量/台 | 满载能耗/<br>kW | 空载能耗/<br>kW | 平均服<br>务率              |
| 数量                                       | 2 000       | 500         | 100         | 13.5                   |
|                                          |             |             |             | 可平移工作负荷集<br>群的最大延迟时间/h |
|                                          |             |             |             | 8                      |

(下转第 131 页 Continued on Page 131)