

8-25-2023

Energy Management Strategy of Multi-agent Microgrid Based on Q-learning Algorithm

Miaomiao Ma

School of Control and Computer Engineering, North China Electric Power University, Beijing 102206, China; State Key Laboratory of Alternate Electric Power System with Renewable Energy Sources, North China Electric Power University, Beijing 102206, China, mamm@ncepu.edu.cn

Lipeng Dong

School of Control and Computer Engineering, North China Electric Power University, Beijing 102206, China

Xiangjie Liu

School of Control and Computer Engineering, North China Electric Power University, Beijing 102206, China; State Key Laboratory of Alternate Electric Power System with Renewable Energy Sources, North China Electric Power University, Beijing 102206, China

Follow this and additional works at: <https://dc-china-simulation.researchcommons.org/journal>



Part of the [Artificial Intelligence and Robotics Commons](#), [Computer Engineering Commons](#), [Numerical Analysis and Scientific Computing Commons](#), [Operations Research, Systems Engineering and Industrial Engineering Commons](#), and the [Systems Science Commons](#)

This Paper is brought to you for free and open access by Journal of System Simulation. It has been accepted for inclusion in Journal of System Simulation by an authorized editor of Journal of System Simulation.

Energy Management Strategy of Multi-agent Microgrid Based on Q-learning Algorithm

Abstract

Abstract: This paper proposes a multi-agent microgrid energy management method for the energy trading and benefit distribution in the microgrid power market based on the Q-learning algorithm. Based on the electricity market, microgrid system and transaction process are constructed to clarify the responsibilities of each unit. The mathematical models of distributed power generations are established by considering the changes in wind speed, light intensity and ambient temperature, as well as the upper and lower limit constraints of the output power of each power generation unit. On this basis, *the distributed power generations and user loads are regarded as agents, and the Markov decision-making process is designed based on the Q-learning algorithm. Aiming at maximizing the benefits of distributed power generations and minimizing the costs of user loads, a microgrid energy management scheme based on Q-learning algorithm is proposed.* The results show that the proposed method can not only increase the benefits of distributed power generations but also reduce the costs of user loads in different scenarios.

Keywords

energy management, microgrid; multi-agents, benefit and cost, distributed power generation

Recommended Citation

Ma Miaomiao, Dong Lipeng, Liu Xiangjie. Energy Management Strategy of Multi-agent Microgrid Based on Q-learning Algorithm[J]. Journal of System Simulation, 2023, 35(7): 1487-1496.

基于 Q-learning 算法的多智能体微电网能量管理策略

马苗苗^{1,2}, 董利鹏¹, 刘向杰^{1,2}

(1. 华北电力大学 控制与计算机工程学院, 北京 102206; 2. 华北电力大学 新能源电力系统国家重点实验室, 北京 102206)

摘要: 针对微电网电力市场的能量交易和利益分配问题, 提出了一种基于 Q-learning 算法的多智能体微电网能量管理方法。构建基于电力市场的微电网系统及交易过程, 明确各单元的职责; 考虑风速、光照强度和环境温度的变化情况, 以及各发电单元的输出功率上下限约束, 建立分布式电源数学模型。在此基础上, 将分布式电源和用户负载视为智能体, 基于 Q-learning 算法设计它们的马尔可夫决策过程, 以最大化分布式电源收益和最小化用户负载成本为目标, 提出了基于 Q-learning 算法的微电网能量管理方案。研究表明, 在不同场景下所提方法既可以增加分布式电源的收益, 还可以降低用户负载的成本。

关键词: 能量管理; 微电网; 多智能体; 收益与成本; 分布式电源

中图分类号: TP391.9; TM732 文献标志码: A 文章编号: 1004-731X(2023)07-1487-10

DOI: 10.16182/j.issn1004731x.joss.22-0378

引用格式: 马苗苗, 董利鹏, 刘向杰. 基于 Q-learning 算法的多智能体微电网能量管理策略[J]. 系统仿真学报, 2023, 35(7): 1487-1496.

Reference format: Ma Miaomiao, Dong Lipeng, Liu Xiangjie. Energy Management Strategy of Multi-agent Microgrid Based on Q-learning Algorithm[J]. Journal of System Simulation, 2023, 35(7): 1487-1496.

Energy Management Strategy of Multi-agent Microgrid Based on Q-learning Algorithm

Ma Miaomiao^{1,2}, Dong Lipeng¹, Liu Xiangjie^{1,2}

(1. School of Control and Computer Engineering, North China Electric Power University, Beijing 102206, China; 2. State Key Laboratory of Alternate Electric Power System with Renewable Energy Sources, North China Electric Power University, Beijing 102206, China)

Abstract: This paper proposes a multi-agent microgrid energy management method for the energy trading and benefit distribution in the microgrid power market based on the Q-learning algorithm. Based on the electricity market, microgrid system and transaction process are constructed to clarify the responsibilities of each unit. The mathematical models of distributed power generations are established by considering the changes in wind speed, light intensity and ambient temperature, as well as the upper and lower limit constraints of the output power of each power generation unit. On this basis, *the distributed power generations and user loads are regarded as agents, and the Markov decision-making process is designed based on the Q-learning algorithm. Aiming at maximizing the benefits of distributed power generations and minimizing the costs of user loads, a microgrid energy management scheme based on Q-learning algorithm is proposed.* The results show that the proposed method can not only increase the benefits of distributed power generations but also reduce the costs of user loads in different scenarios.

Keywords: energy management; microgrid; multi-agents; benefit and cost; distributed power generation

收稿日期: 2022-04-18 修回日期: 2022-06-06

基金项目: 中央高校基本科研业务费专项资金(2023JC002); 国家自然科学基金(61873091)

第一作者: 马苗苗(1982-), 女, 教授, 博士, 研究方向为模型预测控制、新能源电力系统优化与控制、车辆控制与智能化。

E-mail: mamm@ncepu.edu.cn

0 引言

微电网包含多种分布式能源和储能系统，可以通过恰当的系统配置，并采用合适的运行策略，使其在满足负荷需求的前提下，实现能量的梯级利用，最终实现能源的高效灵活利用^[1]。能源类型的多样性、可再生能源的间歇性和波动性、分布式发电单元的功率约束以及能量的双向流动等问题加剧了微电网能量管理的复杂性^[2]，因此，寻求一种有效的能量优化管理策略对提高微电网系统性能至关重要^[3]。

近几年微电网能量优化管理问题受到了许多研究者的广泛关注^[4-9]。文献[4]建立了基于风光储互补微电网系统的优化调度模型，针对孤岛和并网两种运行模式提出了以最小化运行成本为目标的能量管理策略。文献[5]构建了能源侧与负荷侧的互动关系，考虑经济性和用户满意度，从而引导用户改变用电策略，通过制定负荷计划来优化微网出力。上述研究在优化电源侧出力时大多忽略了可再生能源随机性和负荷可削减性的影响。文献[6]考虑可再生能源的不确定性，基于实时市场构建微网环境并采用强化学习算法在连续的状态空间和动作空间中进行学习提升。文献[7]考虑可再生能源和分时电价的不确定性，提出了多能源微电网优化运行模型。文献[8]利用协调控制策略减小了可再生能源随机性和波动性给微电网稳定经济运行带来的影响。文献[9]以经济性为目标，利用深度Q网络对预测负荷、发电功率等信息进行学习，但由于没有考虑环境的可变性，故不能充分体现学习策略的有效性。

针对含有多种分布式电源的微电网电力系统，考虑可再生能源的间歇性和波动性，以及各发电单元输出功率的上下限约束，本文提出了基于Q-learning算法的多智能体微电网能量管理策略。该策略以最大化分布式电源收益和最小化用户负载成本为目标，可以在复杂环境中自主决策并学习最优策略。

1 基于电力市场的微电网系统

1.1 微电网系统结构

根据微电网是否与主电网相连接可将其分为并网型和孤立型两种，本文研究并网型微电网的能量管理问题，微电网系统结构如图1所示。分布式发电单元包括风电机组、光伏机组、微型燃气轮机、柴油发电机和燃料电池，储能系统作为备用电源，需求侧包括3个用户负载。

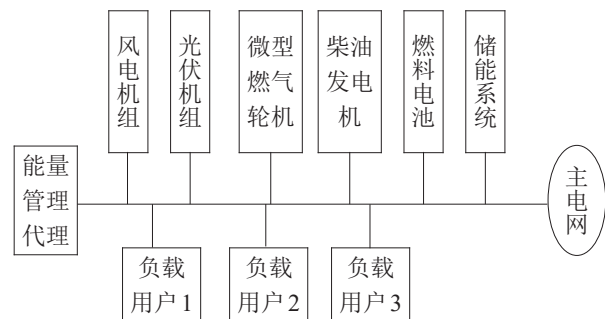


图1 微电网系统结构

Fig. 1 Microgrid system structure

在微电网电力市场中引入能量管理代理，负责监督管理电力市场，但不参与电力交易行为。分布式电源以盈利为目的来寻求最佳策略，而用户负载则以用电成本最小化为目的，根据市场电价来决定需求量。微电网系统的能量供需平衡关系为

$$P_{load}^t = P_{WT,m}^t + P_{PV,m}^t + P_{MT,m}^t + P_{DE,m}^t + P_{FC,m}^t + P_{ESS}^t + P_{maingrid}^t \quad (1)$$

式中： P_{load}^t 为 t 时刻所有用户负载的总需求量； $P_{WT,m}^t$ 、 $P_{PV,m}^t$ 、 $P_{MT,m}^t$ 、 $P_{DE,m}^t$ 和 $P_{FC,m}^t$ 分别为 t 时刻风电机组、光伏机组、微型燃气轮机、柴油发电机和燃料电池出售到微电网的电量； P_{ESS}^t 值为正表示储能系统处于放电模式； P_{ESS}^t 值为负表示储能系统处于充电模式； $P_{maingrid}^t$ 为 t 时刻微电网与主网的能量交互。

1.2 电力市场交易过程

电力市场能量交易包括2个阶段，二者存在

先后次序并相互影响: ①能量管理代理根据能量供需关系和市场信息, 确定市场电价; ②分布式电源和用户负载根据市场电价, 确定发电量和需求量(考虑可再生能源的随机性, 风电机组和光伏机组的发电量不可控)。

2 分布式电源模型

风电机组输出功率为^[10-11]

$$P_{WT} = \begin{cases} 0, & 0 \leq v \leq v_{ci} \\ P_{WT}^r \frac{v - v_{ci}}{v_r - v_{ci}}, & v_{ci} \leq v \leq v_r \\ P_{WT}^r, & v_r \leq v \leq v_{co} \\ 0, & v_{co} \leq v \end{cases} \quad (2)$$

式中: P_{WT}^r 为额定功率; v 为环境风速; v_{ci} , v_r 和 v_{co} 分别为风机运行的切入风速、额定风速和切出风速。

光伏机组输出功率计算公式可表示为^[12]

$$P_{PV} = FF \cdot V \cdot I \quad (3)$$

式中: FF 为填充因子; V 为光伏电池输出电压; I 为光伏电池输出电流。上述各变量表示如下:

$$\begin{cases} FF = \frac{V_{MPP} \cdot I_{MPP}}{V_{OC} \cdot I_{SC}} \\ V = V_{OC} - K_V \cdot T_c \\ I = s [I_{SC} + K_I (T_c - 25)] \end{cases} \quad (4)$$

式中: V_{MPP} 为光伏电池最大功率点电压; I_{MPP} 为光伏电池最大功率点电流; V_{OC} 为光伏电池开路电压; I_{SC} 为光伏电池短路电流; s 为光照强度; K_I 为电流系数; T_c 为光伏电池的温度。

$$T_c = T_a + \frac{s(N_{OT} - 20)}{0.8}$$

式中: K_V 为电压系数; T_a 为环境温度; N_{OT} 为光伏电池标准工作温度。

微型燃气轮机输出功率应满足如下约束^[13]:

$$0 \leq P_{MT} \leq P_{MT}^{\max} \quad (5)$$

式中: P_{MT}^{\max} 为微型燃气轮机输出功率上限。根据微型燃气轮机的变工况特性, 其发电成本可表示为

$$Cost_{MT} = a_1 P_{MT}^2 + b_1 P_{MT} + c_1 \quad (6)$$

式中: a_1 , b_1 , c_1 为微型燃气轮机成本系数^[14]。

柴油发电机输出功率应满足如下约束^[15]:

$$0 \leq P_{DE} \leq P_{DE}^{\max} \quad (7)$$

式中: P_{DE}^{\max} 为柴油发电机输出功率上限。柴油发电机的发电成本以二次函数表示^[16]:

$$Cost_{DE} = a_2 P_{DE}^2 + b_2 P_{DE} + c_2 \quad (8)$$

式中: a_2 , b_2 , c_2 为柴油发电机成本系数。

燃料电池输出功率应满足^[17]:

$$0 \leq P_{FC} \leq P_{FC}^{\max} \quad (9)$$

式中: P_{FC}^{\max} 为燃料电池输出功率上限。

燃料电池的发电成本以二次函数表示^[18]:

$$Cost_{FC} = a_3 P_{FC}^2 + b_3 P_{FC} + c_3 \quad (10)$$

式中: a_3 , b_3 , c_3 为燃料电池成本系数。

储能系统的荷电状态与其工作模式有关, 其表达式为

$$SOC_{ESS}^{t+1} = SOC_{ESS}^t + P_{ESS}^t \quad (11)$$

3 基于 Q-learning 算法的能量管理策略

3.1 Q-learning 算法

Q-learning 算法是一种基于价值的强化学习算法, 智能体通过从环境状态到动作策略集的学习, 获得最大累积奖励, 其学习过程如图 2 所示^[19]。

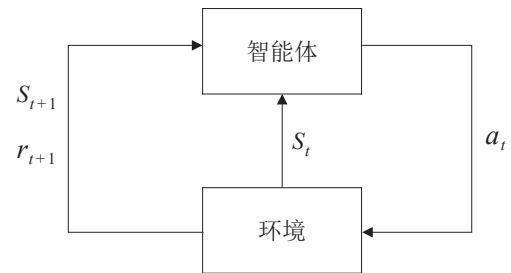


图 2 马尔可夫决策过程
Fig. 2 Markov decision making process

智能体与环境的互动是一种马尔可夫决策过程, 该过程可由元组 $(\mathbf{S}, \mathbf{A}, R, \pi)$ 定义, 其中, \mathbf{S} 为智能体的状态空间; $s_t \in \mathbf{S}$ 是智能体在 t 时刻感知到的状态; \mathbf{A} 为智能体的动作空间, $a_t \in \mathbf{A}$ 为智能体在 t 时刻执行的动作; R 为奖励函数; r_t 为智能

体在状态 s_t 执行动作 a_t 获得的实时奖励； π 为智能体的策略集，表示状态空间 \mathcal{S} 到动作空间 \mathcal{A} 的映射^[20-21]。

智能体利用 Q-learning 寻找最佳策略的思路为：定义智能体的状态动作价值函数 $Q(s, a)$ ，即 Q 函数，通过将观测到的数据代入式(12)中对自身 Q 函数进行迭代学习^[22]。

$$Q(s_t, a_t) \leftarrow Q(s_t, a_t) + \alpha [r_{t+1} + \gamma \max_{a'} Q(s_{t+1}, a') - Q(s_t, a_t)] \quad (12)$$

式中： $\alpha \in (0, 1)$ 为学习率， α 值越大，智能体表现得越贪婪，获得更多的实时奖励； $\gamma \in (0, 1)$ 为未来的 Q 值在当前的衰减率； a' 为智能体在状态 s_{t+1} 可以执行的动作。

3.2 状态空间描述

微电网环境提供给智能体的信息主要是实时的市场电价，因此智能体的状态空间为

$$\mathcal{S} = (MCP_1, MCP_2, \dots, MCP_t) \quad (13)$$

式中： MCP_t 为 t 时刻的市场电价。

3.3 动作空间描述

观测到环境的状态信息以后，智能体根据 Q 值表使用 ϵ -greedy 策略在动作空间中选择一个动作， ϵ -greedy 策略选择动作步骤如下：

- (1) 选择 $\epsilon \in (0, 1)$ 的初始值；
- (2) 选择一个随机数 $rand \in (0, 1)$ ；
- (3) 若 $rand < \epsilon$ ，则从该状态下的动作空间中随机选择一个动作，否则，将在动作空间中选择具有最高 Q 值的动作。

用户负载的动作空间为

$$\mathcal{A}_{L,i} = (\delta^1 P_{L,i}^1, \delta^2 P_{L,i}^2, \dots, \delta^t P_{L,i}^t) \quad (14)$$

式中： $\mathcal{A}_{L,i}$ 为用户负载 i 的动作空间； $P_{L,i}^t$ 为用户负载 i 在 t 时刻的总需求量； δ 为削减系数，且 $\delta \in \{1, 0.95, 0.9, \dots, 0.4\}$ 。

分布式电源感知到市场电价后，在满足相应的约束条件下选择发电量，并从价格集合中选择报价，提交能量管理代理，价格集合为

$$q \in \{5, 10, \dots, 50\} \quad (15)$$

考虑到可再生能源发电的随机性，本文假设风电机组和光伏机组的发电量可以全部售出，风电机组动作空间为

$$\mathcal{A}_{WT} = (q_{WT}^1, q_{WT}^2, \dots, q_{WT}^t) \quad (16)$$

式中： q_{WT}^t 为风电机组在 t 时刻的报价。

光伏机组动作空间为

$$\mathcal{A}_{PV} = (q_{PV}^1, q_{PV}^2, \dots, q_{PV}^t) \quad (17)$$

式中： q_{PV}^t 为光伏机组在 t 时刻的报价。

微型燃气轮机的动作空间为

$$\mathcal{A}_{MT} = [(P_{MT}^1, q_{MT}^1), (P_{MT}^2, q_{MT}^2), \dots, (P_{MT}^t, q_{MT}^t)] \quad (18)$$

式中： P_{MT}^t 为微型燃气轮机在 t 时刻的发电量； q_{MT}^t 为微型燃气轮机在 t 时刻的报价。

柴油发电机的动作空间为

$$\mathcal{A}_{DE} = [(P_{DE}^1, q_{DE}^1), (P_{DE}^2, q_{DE}^2), \dots, (P_{DE}^t, q_{DE}^t)] \quad (19)$$

式中： P_{DE}^t 为柴油发电机在 t 时刻的发电量； q_{DE}^t 为柴油发电机在 t 时刻的报价。

燃料电池的动作空间为

$$\mathcal{A}_{FC} = [(P_{FC}^1, q_{FC}^1), (P_{FC}^2, q_{FC}^2), \dots, (P_{FC}^t, q_{FC}^t)] \quad (20)$$

式中： P_{FC}^t 为燃料电池在 t 时刻的发电量； q_{FC}^t 为燃料电池在 t 时刻的报价。

储能系统的动作空间为

$$\mathcal{A}_{ESS} = [(P_{ESS}^1, q_{ESS}^1), (P_{ESS}^2, q_{ESS}^2), \dots, (P_{ESS}^t, q_{ESS}^t)] \quad (21)$$

式中： P_{ESS}^t 为储能系统 t 时刻的充放电功率； q_{ESS}^t 为储能系统 t 时刻的报价。

3.4 奖励函数

用户负载使用 Q-learning 寻求最佳策略，在市场电价较高时会削减负荷需求来减少成本，以防止用户负载无限制地削减需求，导致微电网供需不平衡，用户负载的成本为

$$Cost_{L,i}^t = \delta^t \cdot P_{L,i}^t \cdot MCP_t + PF(1 - \delta^t) P_{L,i}^t \quad (22)$$

式中： PF 为惩罚系数。

分布式电源的奖励为售电利润减去发电成本。为了鼓励可再生能源发电, 本文假设风电机组和光伏机组无发电成本。

风电机组奖励函数为

$$R_{WT}^t = P_{WT,m}^t \cdot MCP_t + P_{WT,M}^t \cdot P_{maingrid}^t \quad (23)$$

式中: $P_{maingrid}^t$ 为 t 时刻主电网上网电价; $P_{WT,m}^t$ 为 t 时刻风电机组出售到微电网的电量; $P_{WT,M}^t$ 为 t 时刻风电机组出售到主电网的电量。

光伏机组的奖励函数为

$$R_{PV}^t = P_{PV,m}^t \cdot MCP_t + P_{PV,M}^t \cdot P_{maingrid}^t \quad (24)$$

微型燃气轮机的奖励函数为

$$R_{MT}^t = P_{MT,m}^t \cdot MCP_t + P_{MT,M}^t \cdot P_{maingrid}^t - Cost_{MT}^t \quad (25)$$

式中: $Cost_{MT}^t$ 为微型燃气轮机 t 时刻发电成本。

柴油发电机的奖励函数为

$$R_{DE}^t = P_{DE,m}^t \cdot MCP_t + P_{DE,M}^t \cdot P_{maingrid}^t - Cost_{DE}^t \quad (26)$$

式中: $Cost_{DE}^t$ 为柴油发电机 t 时刻发电成本。

燃料电池的奖励函数为

$$R_{FC}^t = P_{FC,m}^t \cdot MCP_t + P_{FC,M}^t \cdot P_{maingrid}^t - Cost_{FC}^t \quad (27)$$

式中: $Cost_{FC}^t$ 为燃料电池 t 时刻的发电成本。

储能系统的奖励函数为

$$R_{ESS}^t = P_{ESS}^t \cdot MCP_t \quad (28)$$

3.5 能量管理策略

市场电价的确定过程如下: 首先能量管理代理收到用户负载提交的总负荷需求 L , 以及分布式电源提交的可出售电量 x_i 和出价 q_i 形成的报价块 (x_i, q_i) , 之后能量管理代理将这些报价块按照价格 q_i 从低到高进行排序, 然后从低价到高价累积这些报价块的发电量, 直至满足负荷需求, 则最后一个报价块对应的价格即为市场电价。

基于上述内容, 提出基于 Q-learning 算法的微电网能量管理策略, 具体步骤如下:

- (1) 设置各智能体学习率 α 和衰减率 γ ;
- (2) 初始化各智能体的 Q 函数;
- (3) 根据环境风速、光照强度和温度预测风光输出功率;

(4) 智能体获取当前环境状态信息, 利用 ϵ -greedy 策略选择一个动作并提交给能量管理代理;

(5) 能量管理代理根据供需关系确定市场电价, 根据式(22)~(28)计算智能体的奖励, 并根据式(12)更新智能体 Q 函数;

(6) 若相邻两次迭代的智能体策略相同, 则认为算法达到纳什均衡, 否则返回步骤(4)。

上述智能体的博弈过程如下: 对于分布式电源, 开始其收益较低, 然后分布式电源在满足约束条件的前提下增加发电量并提高报价, 市场电价随之提高, 接着分布式及电源的收益也随之增加, 直至达到均衡状态, 任何一方改变发电量或报价, 都不会影响交易量和市场电价, 因此其收益不会增加。对于用户负载, 开始其成本较高, 然后用户负载在满足约束条件的前提下逐渐减少负荷需求, 其成本随之降低, 直至达到均衡状态, 用户负载无论怎么减少负荷需求, 其成本都不会降低, 因为用户负载在减少负荷需求的同时也需要支付惩罚费用。

4 算例仿真

4.1 仿真参数设置

风电机组和光伏机组的参数如表1, 2所示, 微电网所在地区 24 h 风速变化曲线如图3所示, 光照强度和温度变化曲线如图4所示, 各分布式电源的额定功率如表3所示, 用户负载负荷需求曲线如图5所示, 选取用户负载的惩罚系数为10。

表1 风电机组特征参数
Table 1 Characteristic parameters of wind turbine

参数名称	参数值
切入风速 $v_{ci}/(m \cdot s^{-1})$	4
额定风速 $v_r/(m \cdot s^{-1})$	24
切出风速 $v_{co}/(m \cdot s^{-1})$	14
额定功率 P_{WT}^r/kW	8

表2 光伏机组特征参数

Table 2 Characteristic parameters of photovoltaic unit

参数名称	参数值
最大功率点电压 V_{MPP}/V	17.32
最大功率点电流 I_{MPP}/A	4.76
开路电压 V_{OC}/V	21.98
短路电流 I_{SC}/A	5.32
电池标称工作温度 $N_{OT}/(^{\circ}C)$	43
电压系数 K_V	14.4
电流系数 K_I	1.22

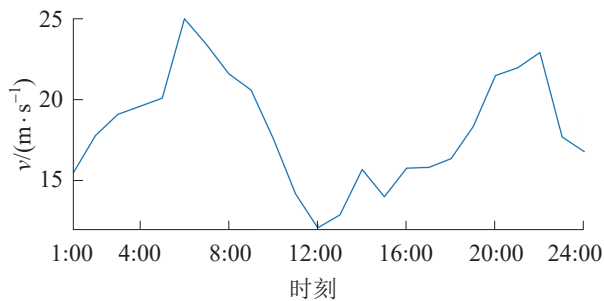


图3 某地区24 h风速变化

Fig. 3 24 hours wind speed change in an area

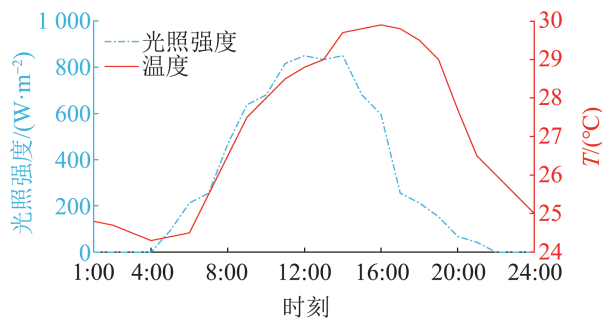


图4 某地区24 h光照强度和温度变化

Fig. 4 24 hours illumination intensity and temperature change in an area

表3 各分布式电源的额定功率

Table 3 Rated power of each distributed power source

电源	额定功率 kW
风电机组	8
光伏机组	6
微型燃气轮机	10
柴油发电机	3
燃料电池	8
储能系统	3

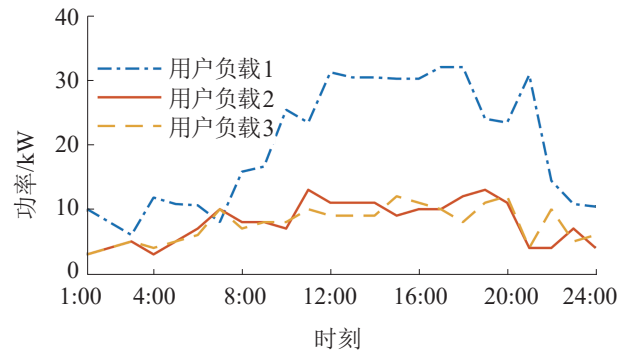


图5 用户负载24 h负荷需求

Fig. 5 24 hours load demand of user loads

利用仿真软件分别在4种场景下对微电网能量管理策略进行仿真,并记录智能体1天24 h的数据,设置仿真周期为30天,步长为1天。这4种场景分别为:

- (1) 所有智能体都没有学习能力(nobody learn, NL);
- (2) 除了用户负载,其他智能体都没有学习能力(customers learn, CL);
- (3) 除了分布式电源,其他智能体都没有学习能力(DER learn, DL);
- (4) 所有智能体都有学习能力(all learn, AL)。

4.2 关于学习参数的灵敏度分析

为了研究学习率 α 和衰减率 γ 对智能体学习性能的影响,本文对智能体进行了灵敏度分析。首先,将其他智能体的参数设置为 $\alpha=0.5$ 和 $\gamma=0.5$,不同学习率 α 和衰减率 γ 的储能系统平均利润如图6所示,在高衰减率和低学习率情况下的储能系统获得的利润最高,这也验证了学习率和衰减率分别反映当前和未来状态对智能体学习性能的影响。由于储能系统目标是在电力交易中获益而不是产能,因此,储能系统应该在电价较低时充电,电价较高时放电。

对其他智能体进行同样的灵敏度分析,各类智能体获得最高利润的学习参数如表4所示,其中,可再生能源发电单元包括风电机组和光伏机组,化石燃料发电单元包括微型燃气轮机、柴油

发电机和燃料电池。由表 4 可知, 化石燃料发电单元类似于储能系统, 在较高 γ 值和较低 α 值的情况下平均利润最高。而可再生能源发电单元更看重实时奖励, 这是因为可再生能源具有随机性, 只能出售当前所有电能来获取利润。

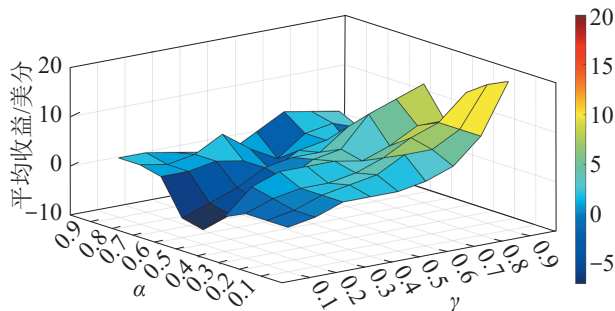


图 6 不同参数下储能系统的平均利润
Fig. 6 Average profit of energy storage system under different parameters

表 4 各智能体的最佳学习参数

Table 4 Optimal learning parameters for each agent

智能体	α	γ
可再生能源发电单元	0.7	0.4
用户负载	0.5	0.5
化石燃料发电单元	0.3	0.9

4.3 仿真结果

用户负载的学习过程如图 7 所示, 可知其学习过程在经过 200 次迭代之后收敛。以用户负载 3 为例分析用户负载的成本情况, 用户负载 3 在 4 种场景下的每日成本情况如图 8 所示, 可以看出用户负载 3 的平均成本分别为 8 087.3, 6 227.1, 8 850.1 和 6 575.2, 单位为美分。由图 8 可知, 在 NL 和 DL 场景下, 没有学习能力的用户负载的成本较高, 这是因为在这两种场景下他们满足了所有的负荷需求。然而, 由于在 DL 场景中分布式电源的策略提高了市场电价, 从而导致用户负载的平均成本比 NL 场景高出了 762.8。同理可知, 用户负载 3 在 AL 场景下的平均成本比 CL 高了 348.1。不过, 从图 8 中也可以看出, 用户负载 3 在 CL 和 AL 场景下利用 Q-learning 算法学会了削

减负荷需求来降低成本。

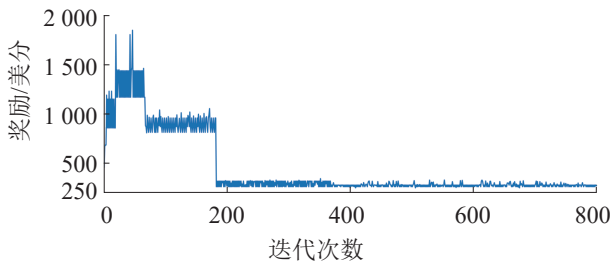


图 7 用户负载的学习过程
Fig.7 Learning process of user load

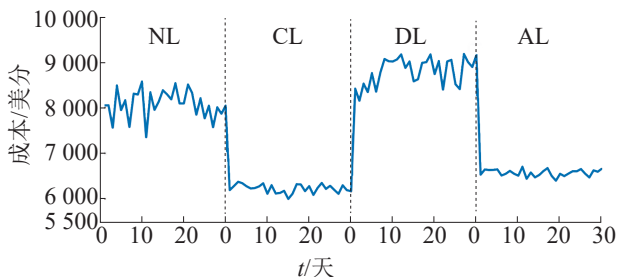


图 8 用户负载 3 的成本
Fig. 8 Cost of user load 3

分布式电源的学习过程如图 9 所示, 可知其学习过程在经过 400 次迭代之后收敛。以风电机组为例分析可再生发电单元的收益情况, 风电机组的利润曲线如图 10 所示, 可以看出风电机组的平均利润分别为 4 739.9, 3 919.5, 5 931.4 和 5 097.8 美分。由图 10 可知, 只要分布式电源具有学习能力, 利润便会明显增加。由于风电机组在 DL 场景下能够与环境交互并学习最佳策略, 因此获得的平均利润比 NL 和 CL 都高出了至少 20%, 同样地, AL 场景下的风电机组平均利润比 NL 和 CL 高了 357.9 和 1 178.3 美分。由于用户负载在市场电价较高时学会了减少消耗, 从而导致风电机组在 AL 场景下的平均利润比在 DL 场景下减少了 14%。此外, 风能的不确定性导致风电机组的平均利润每日波动幅度较为明显, 但由于在 DL 和 AL 场景下启用了学习能力来探索最佳策略, 因此风电机组的平均利润波动幅度明显比 NL 和 CL 时低。

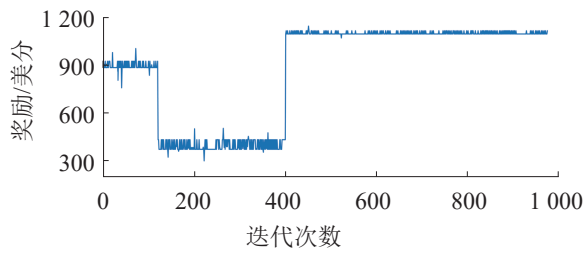


图9 分布式电源的学习过程
Fig. 9 Learning process of distributed power source

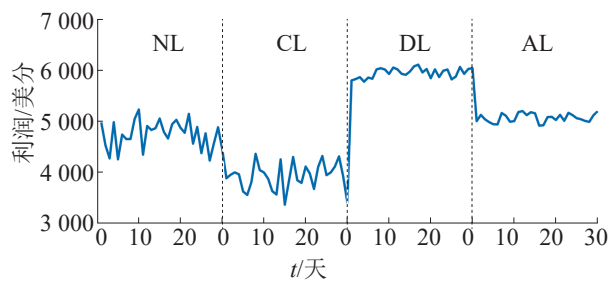


图10 风电机组的利润
Fig. 10 Profit of wind turbine

以燃料电池为例分析化石燃料发电单元的收益情况, 燃料电池的利润曲线如图11所示, 可以看出燃料电池的平均利润在NL和CL场景下分别为3941.0和2733.2美分, 使用学习算法之后达到了4661.3和4017.2美分, 可见燃料电池的平均利润有了明显的提升。然而, 在用户负载有学习能力的情况下, 燃料电池代理的利润有所减少。

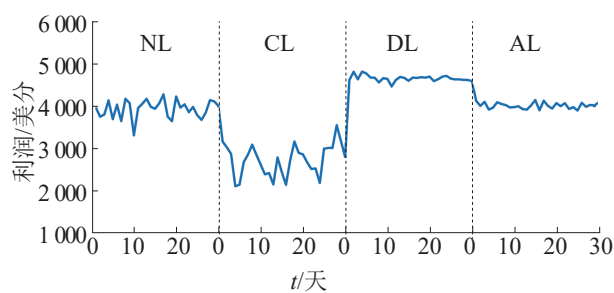


图11 燃料电池的利润
Fig. 11 Profit of fuel cell

储能系统的利润曲线如图12所示, 由图可知储能系统的平均利润在CL和NL场景下为负值, 这是因为储能系统没有最佳运营策略, 从而导致储能系统出现高买贱卖的情况。由于使用学习算

法的储能系统可以优化自身行为, 购买廉价电能, 而在市场电价较高时再将其售出, 因此储能系统可以获取可观的利润。

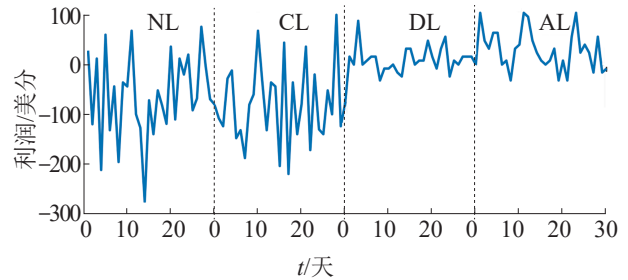


图12 储能系统的利润
Fig. 12 Profit of energy storage system

微电网与主电网的每日交换功率如图13所示, 可以看到, 在NL场景下, 由于用户负载的需求较高, 并且分布式电源没有最佳的运营策略, 因此微电网每日需要从主电网购入439.1 kW的电能。然而, 用户负载追求低成本的学习行为大幅降低了微电网对主电网的依赖, 同时分布式能源代理的策略也会减少微网与主电网的交换功率。在AL场景下各分布式电源的出力情况如图14所示, 可知分布式电源均满足各自的功率约束条件。

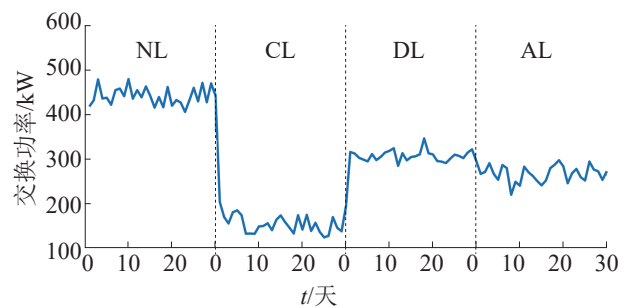
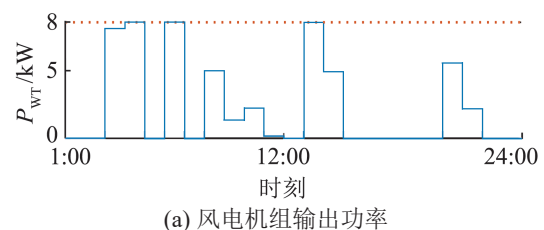


图13 微电网与主电网的每日交换功率
Fig. 13 Daily exchanged power between microgrid and main grid



(a) 风电机组输出功率

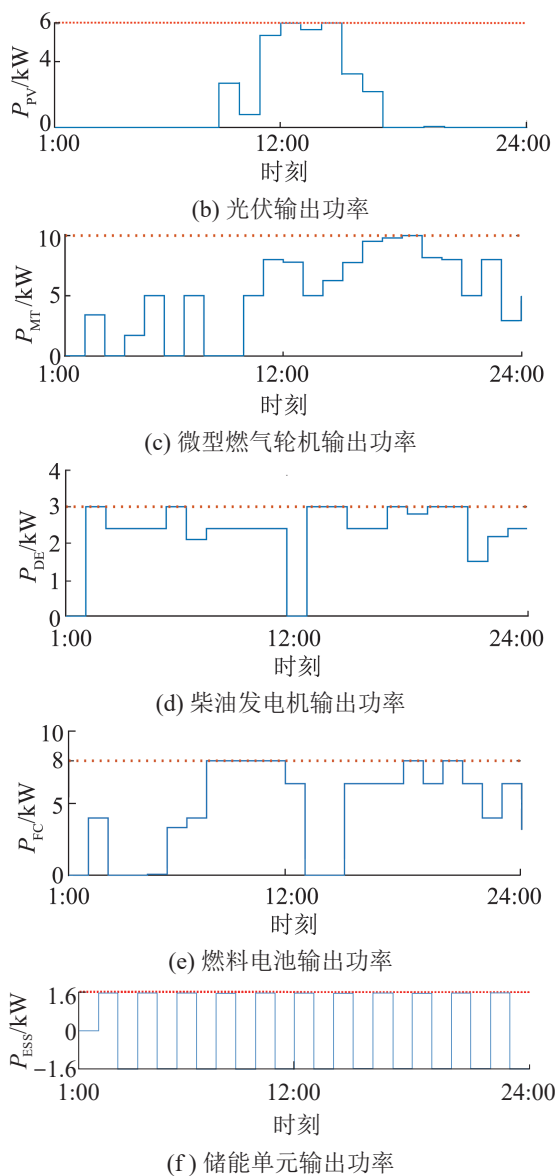


图 14 各分布式电源的出力情况
Fig. 14 Output of each distributed power source

5 结论

可再生能源的随机性和负荷需求的不确定性使微电网能量管理面临诸多挑战。本文构建了基于电力市场的微电网系统结构并建立了分布式电源模型, 基于 Q-learning 算法设计用户负载和分布式电源的状态空间、动作空间和奖励函数, 并将用户负载和分布式电源视为智能体, 以最大化收益和最小化成本为目标, 使其在微电网电力市场环境下学习最优策略, 在此基础上提出了基于

Q-learning 算法的多智能体微电网能量管理策略。仿真结果表明, 该策略可以有效增加分布式电源的收益, 降低用户负载的成本, 为研究微电网能量管理问题提供了新思路。然而在实际情况下, 单一微电网往往不是解决大规模电力系统供能问题的唯一办法, 因此多个微电网在电力市场的竞价是将来研究的一个课题。

参考文献:

- [1] Hao Ran, Jiang Ziqing, Ai Qian, et al. Hierarchical Optimisation Strategy in Microgrid Based on the Consensus of Multi-agent System[J]. IET Generation, Transmission & Distribution, 2018, 12(10): 2444-2451.
- [2] 李利明, 李征. 基于分层多目标优化的微电网能量管理算法[J]. 电力建设, 2018, 39(4): 75-82.
Li Liming, Li Zheng. Microgrid Energy Management Based on Hierarchical Multi-objective Optimization Algorithm[J]. Electric Power Construction, 2018, 39(4): 75-82.
- [3] Shah P, Mehta B. Microgrid Optimal Scheduling with Renewable Energy Sources Considering Islanding Constraints[J]. Iranian Journal of Science and Technology, Transactions of Electrical Engineering, 2020, 44(2): 805-819.
- [4] 郭骏. 风光储互补微电网的优化调度与能量管理策略研究[J]. 电气工程, 2020, 8(1): 36-47.
Guo Jun. Research on Optimal Scheduling and Energy Management Strategy of Wind-PV-ES Complementary Micro-grid[J]. Journal of Electrical Engineering, 2020, 8(1): 36-47.
- [5] 李华, 于潇. 基于DDPG的光储微电网系统能量决策优化[J]. 可再生能源, 2021, 39(5): 687-693.
Li Hua, Yu Xiao. Energy Decision Optimization of Photovoltaic Energy Storage System Based on Deep Deterministic Strategy Gradient Algorithm[J]. Renewable Energy Resources, 2021, 39(5): 687-693.
- [6] 郭国栋, 龚雁峰. 电力市场环境下基于深度强化学习的微网能量管理系统实时自动控制算法[J]. 电测与仪表, 2021, 58(9): 78-88.
Guo Guodong, Gong Yanfeng. Real-time Automatic Control Algorithm of Micro-grid Energy Management System Based on Deep Reinforcement Learning in Electricity Market Environment[J]. Electrical Measurement & Instrumentation, 2021, 58(9): 78-88.
- [7] Li Zhengmao, Xu Yan. Temporally-coordinated Optimal Operation of a Multi-energy Microgrid Under Diverse

- Uncertainties[J]. *Applied Energy*, 2019, 240: 719-729.
- [8] 刘俊峰, 陈剑龙, 王晓生, 等. 基于深度强化学习的微能源网能量管理与优化策略研究[J]. *电网技术*, 2020, 44(10): 3794-3803.
- Liu Junfeng, Chen Jianlong, Wang Xiaosheng, et al. Energy Management and Optimization of Multi-energy Grid Based on Deep Reinforcement Learning[J]. *Power System Technology*, 2020, 44(10): 3794-3803.
- [9] Lee S, Seon J, Kyeong C, et al. Novel Energy Trading System Based on Deep-reinforcement Learning in Microgrids[J]. *Energies*, 2021, 14(17): 5515.
- [10] Mojica-Nava E, Barreto C, Quijano N. Population Games Methods for Distributed Control of Microgrids[J]. *IEEE Transactions on Smart Grid*, 2015, 6(6): 2586-2595.
- [11] 刘帅. 基于智能算法的微电网能量协调管理研究[D]. 上海: 东华大学, 2019.
- Liu Shuai. Research on Energy Coordination Management of Microgrid Based on Intelligent Algorithms[D]. Shanghai: Donghua University, 2019.
- [12] 黄代政, 龚仁喜, 阎昌国, 等. 免疫约束多目标算法在微电网能量优化的应用[J]. *系统仿真学报*, 2014, 26(1): 225-229.
- Huang Daizheng, Gong Renxi, Yan Changguo, et al. Immune Constrained Multi-objective Optimization for Microgrids Energy Management[J]. *Journal of System Simulation*, 2014, 26(1): 225-229.
- [13] Foo Eddy Y S, Gooi H B, Chen S X. Multi-agent System for Distributed Management of Microgrids[J]. *IEEE Transactions on Power Systems*, 2015, 30(1): 24-34.
- [14] 高雪莹, 唐昊, 苗刚中, 等. 储能系统能量调度与需求响应联合优化控制[J]. *系统仿真学报*, 2016, 28(5): 1165-1172.
- Gao Xueying, Tang Hao, Miao Gangzhong, et al. Joint Optimization Control of Energy Storage System Management and Demand Response[J]. *Journal of System Simulation*, 2016, 28(5): 1165-1172.
- [15] Ju Liwei, Huang Liling, Tan Qinliang, et al. Three-level Energy Flexible Management Strategy for Micro Energy Grids Considering Multiple Uncertainties at Different Time Scales[J]. *International Journal of Energy Research*, 2020, 45(1): 316-341.
- [16] Cintuglu M H, Martin H, Mohammed O A. Real-time Implementation of Multiagent-based Game Theory Reverse Auction Model for Microgrid Market Operation[J]. *IEEE Transactions on Smart Grid*, 2015, 6(2): 1064-1072.
- [17] Mocanu E, Mocanu D C, Nguyen P H, et al. On-line Building Energy Optimization Using Deep Reinforcement Learning[J]. *IEEE Transactions on Smart Grid*, 2019, 10(4): 3698-3708.
- [18] Zhang Yuxiang, Ma Rui, Zhao Dongdong, et al. A Novel Energy Management Strategy Based on Dual Reward Function Q-learning for Fuel Cell Hybrid Electric Vehicle[J]. *IEEE Transactions on Industrial Electronics*, 2022, 69(2): 1537-1547.
- [19] 黄星源, 李岩屹. 基于双Q学习算法的干扰资源分配策略[J]. *系统仿真学报*, 2021, 33(8): 1801-1808.
- Huang Xingyuan, Li Yanyi. The Allocation of Jamming Resources Based on Double Q-learning Algorithm[J]. *Journal of System Simulation*, 2021, 33(8): 1801-1808.
- [20] 朱志斌, 王付永, 尹艳辉, 等. 基于Q-learning的离散时间多智能体系统一致性[J]. *控制理论与应用*, 2021, 38(7): 997-1005.
- Zhu Zhibin, Wang Fuyong, Yin Yanhui, et al. Consensus of Discrete-time Multi-agent System Based on Q-learning[J]. *Control Theory & Applications*, 2021, 38(7): 997-1005.
- [21] Zhang Qingchen, Lin Man, Yang L T, et al. Energy-efficient Scheduling for Real-time Systems Based on Deep Q-learning Model[J]. *IEEE Transactions on Sustainable Computing*, 2017, 4(1): 132-141.
- [22] Kofinas P, Dounis A I, Vouros G A. Fuzzy Q-learning for Multi-agent Decentralized Energy Management in Microgrids[J]. *Applied Energy*, 2018, 219: 53-67.