

5-26-2022

## Low-carbon operation strategy of regional integrated energy system based on the Q learning algorithm

Jieyun ZHENG

*Economic Technology Research Institute, State Grid Fujian Electric Power Co., Ltd., Fuzhou 350012, China*

Qianyun SONG

*Economic Technology Research Institute, State Grid Fujian Electric Power Co., Ltd., Fuzhou 350012, China*

Guilian WU

*Economic Technology Research Institute, State Grid Fujian Electric Power Co., Ltd., Fuzhou 350012, China*

Hao CHEN

*Economic Technology Research Institute, State Grid Fujian Electric Power Co., Ltd., Fuzhou 350012, China*

Zhijian HU

*School of Electrical Engineering and Automation, Wuhan University, Wuhan 430074, China*

*See next page for additional authors*

Follow this and additional works at: <https://jepst.researchcommons.org/journal>

---

### Recommended Citation

ZHENG, Jieyun; SONG, Qianyun; WU, Guilian; CHEN, Hao; HU, Zhijian; CHEN, Zhi; WENG, Changhong; and CHEN, Jinpeng (2022) "Low-carbon operation strategy of regional integrated energy system based on the Q learning algorithm," *Journal of Electric Power Science and Technology*. Vol. 37: Iss. 2, Article 13.

DOI: 10.19781/j.issn.1673-9140.2022.02.014

Available at: <https://jepst.researchcommons.org/journal/vol37/iss2/13>

This Article is brought to you for free and open access by Journal of Electric Power Science and Technology. It has been accepted for inclusion in Journal of Electric Power Science and Technology by an authorized editor of Journal of Electric Power Science and Technology.

---

# Low-carbon operation strategy of regional integrated energy system based on the Q learning algorithm

## Authors

Jieyun ZHENG, Qianyun SONG, Guilian WU, Hao CHEN, Zhijian HU, Zhi CHEN, Changhong WENG, and Jinpeng CHEN

# 基于 Q 学习的区域综合能源系统 低碳运行策略

郑洁云<sup>1</sup>, 宋倩芸<sup>1</sup>, 吴桂联<sup>1</sup>, 陈浩<sup>1</sup>, 胡志坚<sup>2</sup>,  
陈志<sup>2</sup>, 翁菘宏<sup>2</sup>, 陈锦鹏<sup>2</sup>

(1. 国网福建省电力有限公司经济技术研究院, 福建 福州 350012; 2. 武汉大学电气与自动化学院, 湖北 武汉 430072)

**摘要:**考虑区域综合能源系统运行时的经济性与低碳性, 提出一种基于 Q 学习的区域综合能源系统低碳运行策略研究方法。首先, 基于能量枢纽的概念构建区域综合能源系统的基本结构模型; 接着, 以计及二氧化碳治理费用的日运行费用最低为目标函数, 提出区域综合能源系统的低碳经济运行策略; 然后, 针对低碳经济运行策略建立其马尔可夫决策问题, 并采用改进的 Q 学习进行求解。通过仿真验证 Q 学习算法求解区域综合能源系统运行策略的有效性, 结果表明, 所提运行策略能充分发挥区域综合能源系统的多能互补优势, 实现系统低碳经济运行, 为区域综合能源系统的运行优化问题提供思路和策略。

**关键词:**区域综合能源系统; 低碳运行; 经济运行; 马尔可夫决策; Q 学习

DOI:10.19781/j.issn.1673-9140.2022.02.014 中图分类号: TM73; TK018 文章编号: 1673-9140(2022)02-0106-10

## Low-carbon operation strategy of regional integrated energy system based on the Q learning algorithm

ZHENG Jieyun<sup>1</sup>, SONG Qianyun<sup>1</sup>, WU Guilian<sup>1</sup>, CHEN Hao<sup>1</sup>, HU Zhijian<sup>2</sup>,  
CHEN Zhi<sup>2</sup>, WENG Changhong<sup>2</sup>, CHEN Jinpeng<sup>2</sup>

(1. Economic Technology Research Institute, State Grid Fujian Electric Power Co., Ltd., Fuzhou 350012, China;

2. School of Electrical Engineering and Automation, Wuhan University, Wuhan 430074, China)

**Abstract:** Considering the economic and low-carbon performance of regional integrated energy systems, a method of low-carbon operation strategy of regional integrated energy systems based on the Q learning is proposed. Firstly, the basic operation model of such regional integrated energy system is constructed on the basis of the energy hub. Then, taking the minimum daily operating cost as the objective function, including the carbon dioxide treatment cost, a low-carbon economic operation strategy of regional integrated energy system is proposed. Then, the low-carbon economy operation strategy is modeled through the Markov decision problems, and the improved Q learning is utilized to solve those problems. The simulation results verify the effectiveness of Q learning algorithm for solving operation strategies

收稿日期: 2020-11-05; 修回日期: 2021-04-30

基金项目: 国家自然科学基金(51977156); 国网福建省电力有限公司科技项目(52130N19000P)

通信作者: 胡志坚(1969-), 男, 博士, 教授, 主要从事电力系统稳定分析与控制、主动配电网规划等研究; E-mail: zhujian\_hu@163.com

in the regional integrated energy system. It is shown that the proposed operation strategy can give full play to the multi-energy complementary advantage, and realize the economic and low-carbon objectives during operation of regional integrated energy system.

**Key words:** regional integrated energy system; low carbon operation; economic operation; Markov decision process; Q learning

随着能源短缺和环境污染问题日益凸显,如何通过节能减排、提高能源利用率、增加新能源占比等方式来实现可持续发展是当前研究中较为关注的问题<sup>[1-2]</sup>。区域综合能源系统(regional integrated energy system, RIES)在此背景下应运而生,RIES 直接面向终端能源用户,利用光伏、风力等发电方式可以提高清洁能源在终端能源消费中的占比,同时,通过热电联产、燃气锅炉等能源转化装置可以满足用户的多种用能需求<sup>[3-5]</sup>。在恰当的能源运行策略下,RIES 可以实现多能互补和能量的梯级利用<sup>[6]</sup>,最终实现能源的高效灵活利用。但 RIES 中负荷的波动性、风力和光伏的不确定性以及能流的多样性的存在,使得制定经济合理的运行策略具有较大的挑战<sup>[7-8]</sup>。

目前,国内外关于 RIES 的运行策略优化已有一定的研究成果。文献[9-12]为了寻求综合能源系统的有效运行策略,或建立了混合整数线性模型,采用数学优化算法进行求解;或建立混合整数非线性模型,采用粒子群、遗传等智能算法进行求解。除了数学优化、智能算法外,也有研究将马尔可夫决策过程(Markov decision process, MDP)应用到综合能源系统的能量管理和运行策略优化中,并且达到了较好的效果<sup>[13-15]</sup>。文献[16]基于深度 Q 网络(deep Q network, DQN)对微能源网进行能量管理,验证了强化学习算法在能量管理中相较于启发式算法的优势;文献[17-18]在 Q 学习算法中引入双层模糊推理来对 Q 值表进行处理;文献[19-20]引入强化学习算法考虑多个智能体之间的博弈,制定了 RIES 的能源管理策略和负荷的调度策略,最终取得了较好的效果。

减少碳排放水平、实现低碳可持续发展是世界各国未来的重要发展战略目标,中国也根据国情提出了“双碳”目标<sup>[21]</sup>。而能源行业在碳排放行业中占比大,因此,在制定 RIES 的运行策略时,兼顾经济性和低碳性对节能减排尤为重要。目前,已有研

究在综合能源系统的运行优化中引入碳交易、碳税等机制来考虑系统的低碳性<sup>[22-24]</sup>,但鲜有研究将强化学习算法应用于兼顾经济性与低碳性的运行优化问题中。

基于上述研究,本文综合考虑区域综合能源系统的经济性和低碳性,将二氧化碳排放的治理费用转化为经济指标、优化 RIES 的运行策略。再者,传统的 Q 学习一般将 Q 值表的所有元素初始化为 0,收敛速度较慢。针对此不足,本文对 Q 学习算法的初始化方法进行改进,以提高算法的收敛速度。首先,本文对区域综合能源系统及其主要装置进行建模;接着以经济性和低碳性为目标建立 RIES 低碳运行策略的数学模型,决策 RIES 向上级电网、气网的购买量以及能源转化装置出力等变量;然后,阐述本文基于改进 Q 学习算法寻求 RIES 运行策略的过程;最后,在 Matlab 环境下,通过仿真验证本文所提模型和算法的有效性以及可行性。

## 1 RIES 结构及能源转化装置建模

RIES 通过能源转化装置及储能装置可以充分实现多种能源之间的互补及协同效益,满足用户的多种用能需求,本文构建的 RIES 结构如图 1 所示。

在能源转化装置中,热电联产(combined heat and power, CHP)的输入是天然气,输出是电能和热能;燃气锅炉(gas boiler, GB)的输入是天然气,其

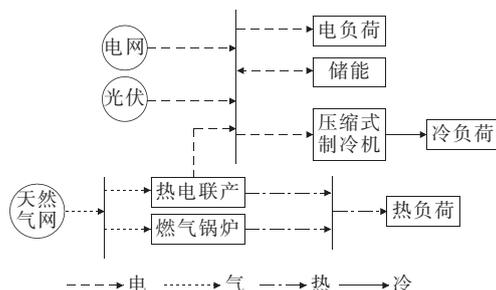


图 1 RIES 结构

Figure 1 Structure diagram of RIES

热输出与 CHP 的热输出共同为用户提供热负荷；压缩式制冷机 (compression refrigerator, CR) 的输入为电能, 其冷输出直接为用户提供冷负荷。若将能源转化装置的输入功率表示  $I_i$ , 输出功率表示为  $O_i$ , 则其模型即能源转化关系、额定功率限制可以表示为

$$\begin{cases} O_i = \eta_i I_i \\ \{I_i \leq O_i \leq \bar{I}_i \\ i \in \{\text{CHP, GB, CR}\} \end{cases} \quad (1)$$

式中  $\eta_i$  为能源转化装置的能源转化效率。

光伏 (PV) 发电具有较强的不确定性和波动性, 本文采用时序法<sup>[25]</sup> 对其进行建模, 要求 PV 实际输出的电功率小于其预测出力, 该约束可以表示为

$$0 \leq P_{\text{PV}}(t) \leq P_{\text{PV}}^{\text{max}}(t) \quad (2)$$

式中  $P_{\text{PV}}(t)$  为  $t$  时刻 PV 输出的电功率;  $P_{\text{PV}}^{\text{max}}(t)$  为  $t$  时刻 PV 的预测出力。

电储 (electricity storage, ES) 有充、放电 2 种状态, 合理配置 ES 的充、放电功率, 可有效缓解用电高峰期的用电压力。本文采用储能的通用模型对 ES 进行建模, 其模型及运行约束条件可以表示为

$$\begin{cases} 0 \leq P_{\text{ES}}^{\text{cha}}(t) \leq \bar{P}_{\text{ES}} \\ 0 \leq P_{\text{ES}}^{\text{dis}}(t) \leq \bar{P}_{\text{ES}} \\ P_{\text{ES}}^{\text{cha}}(t) \cdot P_{\text{ES}}^{\text{dis}}(t) = 0 \\ \left\{ \begin{aligned} W(t) &= W(t-1) + \\ &\frac{P_{\text{ES}}^{\text{cha}}(t)\eta_{\text{ES}}\Delta t - \frac{P_{\text{ES}}^{\text{dis}}(t)\Delta t}{\eta_{\text{ES}}}}{S_{\text{ES}}} \end{aligned} \right. \\ W_{\min} \leq W(t) \leq W_{\max} \end{cases} \quad (3)$$

式中  $P_{\text{ES}}^{\text{cha}}(t)$ 、 $P_{\text{ES}}^{\text{dis}}(t)$  分别为  $t$  时刻 ES 的充、放电功率;  $\bar{P}_{\text{ES}}$  为 ES 的充放电功率上限, 本文取值为 ES 额定容量的 25%; 第 4 式表示 ES 的充、放电行为不能同时进行;  $W(t)$  为  $t$  时刻 ES 的荷电状态;  $\eta_{\text{ES}}$  为 ES 的充放电效率, 本文取值为 1;  $S_{\text{ES}}$  为 ES 的额定容量;  $\Delta t$  取为 1 h;  $W_{\min}$ 、 $W_{\max}$  分别为 ES 荷电的下限、上限。

## 2 RIES 低碳运行策略数学建模

本文综合考虑 RIES 的经济性和低碳性, 以考

虑碳治理费用的日运行成本最小为目标函数, 建立 RIES 低碳运行的数学模型, 决策 RIES 向上级电网、天然气网的购买量以及能源转化装置出力等变量, 最终实现 RIES 的多能互补以及协同效益。

### 2.1 目标函数

本文基于不轻易改变 RIES 用户侧消费意愿的原则, 即不轻易切掉负荷, 以 RIES 的日运行费用最小为目标函数, 该目标函数由 RIES 向上级电网购电、向上级天然气网购气、二氧化碳的治理费用组成。目标函数可以表示为

$$\begin{cases} F = F_{\text{P}} + F_{\text{G}} + F_{\text{C}} \\ F_{\text{P}} = \sum_{t=1}^{24} P(t)\Delta t \cdot f_{\text{P}}(t) \\ \left\{ \begin{aligned} F_{\text{G}} &= \sum_{t=1}^{24} G(t)\Delta t \cdot f_{\text{G}}(t) \\ F_{\text{C}} &= f_{\text{C}} \sum_{t=1}^{24} C(t) \end{aligned} \right. \end{cases} \quad (4)$$

式中  $F$  为目标函数;  $F_{\text{P}}$ 、 $F_{\text{G}}$ 、 $F_{\text{C}}$  分别为 RIES 向上级电网购电、向上级天然气网购气、二氧化碳的治理费用;  $P(t)$ 、 $G(t)$  为  $t$  时刻 RIES 向上级电网的购电、向上级气网的购气功率;  $f_{\text{P}}(t)$ 、 $f_{\text{G}}(t)$  分别为  $t$  时刻的电价、气价;  $C(t)$  为  $t$  时刻 RIES 的  $\text{CO}_2$  排放量;  $f_{\text{C}}$  为单位  $\text{CO}_2$  排放量的治理价格。

RIES 向上级电网所购买的电量大多通过燃烧煤炭而产生, 此产电过程会排放  $\text{CO}_2$ 。因此, 本文考虑 RIES 的  $\text{CO}_2$  治理成本由向上级电网购电、CHP、GB 构成。 $C(t)$  的具体描述可以表示为

$$C(t) = C_{\text{P}}(t) + C_{\text{CHP}}(t) + C_{\text{GB}}(t) \quad (5)$$

$$C_{\text{P}}(t) = P(t)\Delta t\beta_{\text{e}} \quad (6)$$

$$C_{\text{CHP}}(t) = \{H_{\text{CHP}}(t)\Delta t + \varphi_{\text{e,h}}P_{\text{CHP}}(t)\Delta t\}\beta_{\text{h}} \quad (7)$$

$$C_{\text{GB}}(t) = H_{\text{GB}}(t)\Delta t\beta_{\text{h}} \quad (8)$$

式(5)~(8)中  $C_{\text{P}}(t)$ 、 $C_{\text{CHP}}(t)$ 、 $C_{\text{GB}}(t)$  分别为  $t$  时刻购电、CHP、GB 的  $\text{CO}_2$  排放量;  $\beta_{\text{e}}$ 、 $\beta_{\text{h}}$  为 RIES 单位产电量、产热量所产生的  $\text{CO}_2$  排放量;  $\varphi_{\text{e,h}}$  为单位产电量向单位产热量的折算系数;  $H_{\text{CHP}}(t)$  为 CHP 的输出热功率;  $P_{\text{CHP}}(t)$  为 CHP 的输出电功率;  $H_{\text{GB}}(t)$  为 GB 的输出热功率。

### 2.2 约束条件

基于所构建的目标函数, 本文考虑 2 类约束: 各类装置的运行约束与电、热、气、冷功率平衡的约束,

以达到供需平衡的目的。

### 2.2.1 装置运行约束

RIES 涉及的 CHP、GB、CR、PV、ES 等装置在运行时需要满足的约束条件与式(1)~(3)一致。

### 2.2.2 功率平衡约束

RIES 运行时不仅需要约束设备的状态,也必须实现功率的平衡,根据图 1,本文考虑以下约束。

#### 1) 电功率平衡。

$$P(t) + P_{PV}(t) + P_{CHP}(t) =$$

$$P_{load}(t) + P_{CR}(t) + P_{ES}^{cha}(t) - P_{ES}^{dis}(t) \quad (9)$$

式中  $P_{load}(t)$  为  $t$  时刻的电负荷;  $P_{CR}(t)$  为  $t$  时刻 CR 电能输入功率。

#### 2) 热功率平衡。

$$H_{CHP}(t) + H_{GB}(t) = H_{load}(t) \quad (10)$$

式中  $H_{load}(t)$  为  $t$  时刻的热负荷。

#### 3) 天然气平衡。

$$G(t) = G_{CHP}(t) + G_{GB}(t) \quad (11)$$

式中  $G_{CHP}(t)$  为  $t$  时刻 CHP 的天然气输入;  $G_{GB}(t)$  为  $t$  时刻 GB 的天然气输入。

#### 4) 冷功率平衡。

$$F_{CR}(t) = F_{load}(t) \quad (12)$$

式中  $F_{CR}(t)$  为  $t$  时刻 CR 的输出;  $F_{load}(t)$  为  $t$  时刻的冷负荷。

## 3 基于 Q 学习的 RIES 运行策略求解

### 3.1 Q 学习简介

Q 学习算法与模型无关,具有良好的后效性,以 MDP 为理论基础。在 Q 学习算法中,智能体通过不断与外部环境进行交互学习获得经验,如果智能体在当前状态下采取的某个动作导致环境给予积极的奖赏,则智能体后续再次处于这个状态时选中该行为的趋势增强。总的来说,Q 学习算法是一种基于值函数迭代的在线学习和动态最优技术<sup>[14,16,20]</sup>。Q 值的迭代过程可以表示为

$$Q^*(s_t, a_t) = (1 - \alpha)Q(s_t, a_t) + \alpha[r(s_t, a_t) + \gamma \max_{a'} Q(s_{t+1}, a')] \quad (13)$$

式中  $(s_t, a_t)$  为  $t$  时刻的状态—动作对;  $Q(s_t, a_t)$  为采取动作  $a_t$  之前的 Q 值;  $Q^*(s_t, a_t)$  为执行动作  $a_t$  之后更新的 Q 值;  $\max_{a'} Q(s_{t+1}, a')$  为执行

动作  $a_t$  进入状态  $s_{t+1}$  后,在  $s_{t+1}$  状态下智能体能执行的所有动作中的最大 Q 值;  $\alpha$  为学习率,其值越大,表示智能体学习新经验的概率越大,保留以往经验的概率越小;  $\gamma$  为折现因子;  $r(s_t, a_t)$  为动作对  $(s_t, a_t)$  所对应的奖励函数值。

### 3.2 RIES 运行策略求解的 MDP 建模

在满足供需平衡的前提下,本文建立的 RIES 低碳运行策略通过决策优化各类装置的出力、购电量、购气量,来达到 RIES 低碳经济运行的目的。而在 MDP 问题中,RIES 智能体在当前状态下所采取的下一步动作仅仅与其当前状态有关,即当前时刻下的负荷功率大小、ES 的荷电状态、PV 出力等,当智能体执行某一决策动作后,会获得与该动作相对应的奖赏,智能体的动作、状态、奖赏即对应 MDP 问题的 3 个要素:动作空间、状态空间和奖励函数。

#### 3.2.1 动作空间

RIES 的运行策略目的在于优化装置的出力、电和气的购买量。因此,基于图 1,本文考虑 MDP 的动作空间由 CHP/GB/CR 这 3 类能源转化装置的输入、ES 充放电功率、向上级电网购电量、向上级气网购气量组成,动作空间可以表示为

$$A = [G_{CHP}(t), G_{GB}(t), P_{CR}(t), P_{ES}(t), P(t), G(t)] \quad (14)$$

式中  $P_{ES}(t)$  为 ES 的充放电功率,其取值为正时表示 ES 充电,取值为负时表示 ES 放电。

由于电、热、冷负荷以及 PV 出力可以通过预测而在决策前得知,因此,结合热功率平衡约束式(10)和天然气平衡约束式(11),若 CHP 的天然气输入已知,则 GB 的天然气输入、天然气购买量可以通过平衡式(10)、(11)推算出来。同理,冷负荷仅仅由 CR 提供,CR 的电输入功率可由冷平衡式(12)推算而得;若 ES 的充放电功率已知,则通过电功率平衡式(9)同样可以推算出 RIES 的购电量。综上所述,若动作空间  $A$  仅由 CHP 的天然气输入、ES 的充放电功率组成,式(14)中的其余变量均可由 CHP 天然气输入、电储充放电功率根据相应约束条件推算而得。因此,动作空间可以简化为

$$A = [G_{CHP}(t), P_{ES}(t)] \quad (15)$$

#### 3.2.2 状态空间

对于智能体 RIES 来说,外界环境能提供的状

态信息包括电价以及 ES 的荷电状态。因此,本文将分时电价信息以及 ES 的荷电状态作为 MDP 问题的状态空间,可以表示为

$$S = [f_p(t), W(t)] \quad (16)$$

### 3.2.3 奖励函数

奖励函数和模型的目标函数有关,同时,动作空间的确定考虑到了 CHP、GB、CR 的设备约束条件以及电、热、气、冷的功率平衡,却没有涉及 ES 的荷电状态约束。本文考虑对 ES 荷电状态越限进行惩罚,MDP 问题的奖励函数除了和目标函数有关,也和该惩罚函数有关。Q 学习算法考虑的是最终的奖赏最大化,因此,本文构建的奖励函数  $R$  的具体描述可以表示为

$$R = -(F + F_2) \quad (17)$$

$$F_2 =$$

$$\begin{cases} 0, & \text{if } x = 0 \\ \{100f_p(t)S_{ES}\{W_{\min} - W(t+1)\}, & \text{if } x = -1 \\ \{100f_p(t)S_{ES}\{W(t+1) - W_{\max}\}, & \text{if } x = 1 \end{cases} \quad (18)$$

$$\begin{cases} 0, & \text{if } W_{\min} \leq W(t+1) \leq W_{\max} \\ x = -1, & \text{if } W(t+1) < W_{\min} \end{cases} \quad (19)$$

$$\begin{cases} 1, & \text{if } W(t+1) > W_{\max} \end{cases}$$

其中  $x$  用于判断 ES 的荷电状态是否越限,  $t$  时刻采取动作后 ES 荷电状态变为  $W(t+1)$ ,若 ES 荷电状态越限,将按照越限量乘以当前时刻电价的 100 倍来进行惩罚。

## 3.3 RIES 运行策略求解的 MDP 建模

### 3.3.1 动作空间及状态空间离散化

Q 学习算法只能用于解决离散问题,文 3.2 中构建的 MDP 问题的动作和状态变量均为连续变量,因此,本文需要首先对其进行离散化。动作空间涉及 CHP 的天然气输入和 ES 充放电功率 2 个变量,若 CHP 的输入离散为  $a+1$  个值,ES 的充放电功率离散为  $2b+1$  个值,则动作空间共有  $(a+1) \times (2b+1)$  维。CHP 的天然气输入应当在其所能承受的最小值(0)和最大值  $\bar{G}_{\text{CHP}}$  之间取值,按公式进行离散,即

$$0, \frac{1}{a}\bar{G}_{\text{CHP}}, \frac{2}{a}\bar{G}_{\text{CHP}}, \dots, \frac{a-1}{a}\bar{G}_{\text{CHP}}, \bar{G}_{\text{CHP}} \quad (20)$$

同样的,ES 有充、放电 2 种行为,允许的最大充电功

率和最大放电功率通常取值一致,都为  $\bar{P}_{\text{ES}}$ ,因此,ES 的充放电功率可以按照公式进行离散,即

$$\begin{aligned} & -\bar{P}_{\text{ES}}, -\frac{b-1}{b}\bar{P}_{\text{ES}}, \dots, -\frac{1}{b}\bar{P}_{\text{ES}}, 0, \\ & \frac{1}{b}\bar{P}_{\text{ES}}, \dots, \frac{b-1}{b}\bar{P}_{\text{ES}}, \bar{P}_{\text{ES}} \end{aligned} \quad (21)$$

由于电价通常在决策之前给定,因此,状态空间仅有  $W(t)$  一个变量需要进行离散。 $W(t)$  在  $W_{\min}$  和  $W_{\max}$  之间取值,可以按照公式离散为  $c+1$  个值,即

$$\begin{aligned} & W_{\min}, \frac{1}{c}(W_{\max} - W_{\min}), \dots, \frac{c-1}{c} \cdot \\ & (W_{\max} - W_{\min}), W_{\max} \end{aligned} \quad (22)$$

### 3.3.2 Q 值表初始化

传统的 Q 学习算法将 Q 值表的全部元素初始化为零,本文的目标函数为 RIES 的日运行费用最小,将奖励函数  $R$  定义为负值,则需把 Q 值表的全部元素初始化为一个较小的负值。无论是初始化为零还是一个较小的负值,收敛的速度都会较慢,为了提高算法的收敛速度,本文将 Q 值表初始化为每个动作对  $(s_t, a_t)$  相应的奖励函数值,即

$$Q_{\text{in}}(s_t, a_t) = r(s_t, a_t) \quad (23)$$

式中  $Q_{\text{in}}(s_t, a_t)$  为动作对  $(s_t, a_t)$  对应的初始 Q 值。

### 3.3.3 求解流程

本文提出的基于 Q 学习的 RIES 低碳运行策略的求解流程如图 2 所示,其求解流程如下。

1) 按照式(23)初始化 Q 值表,迭代次数  $n=1$ 。

2) 决策周期  $t$  初始化为 0,并初始化 RIES 在开始时刻的取值,本文认为 ES 在 0 点时储有 ES 额定容量的 10% 电量。

3) 根据 RIES 当前的状态  $s_t$ ,基于  $\epsilon$ -greedy 贪心算法选择合适的动作  $a_t$ , $\epsilon$ -greedy 贪心算法的要点是探索和利用:初始化  $\epsilon$  的取值,智能体在学习过程以  $\epsilon$  的概率进行探索,以  $1-\epsilon$  的概率利用,探索即将所有的尝试机会均匀地分到每个可以执行的动作中去,利用即选择奖赏最大的动作。本文为了在探索和利用中实现较好的折中,首先将  $\epsilon$  设置为一个较大的值以获得多个动作的奖赏经验值,然后随着迭代次数的增加逐渐减小  $\epsilon$  的取值,使得选取奖赏最大的动作的概率增大。执行动作  $a_t$  后便可以

获得下一决策周期的状态  $s_{t+1}$ 。

4) 根据步骤 3 所选择的动作  $a_t$ , 按照式 (17) 计算相应的奖赏  $r(s_t, a_t)$ , 并按照式 (13) 更新动作对  $t=t+1$  的 Q 值。

5) 令  $t=t+1$ ,  $T=24$  表示总的决策周期数, 若  $t \leq T$ , 则返回步骤 4, 否则进入步骤 6。

6) 令  $n=n+1$ ,  $N$  表示迭代次数最大值, 若  $n > N$  或者学习过程已收敛, 则结束迭代, 得到最终的 Q 值表和运行策略, 否则返回步骤 2。

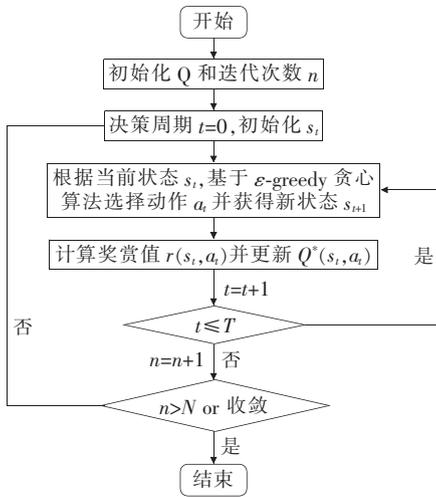


图 2 基于 Q 学习算法的求解流程

Figure 2 The solution flow based on Q learning algorithm

## 4 仿真分析

### 4.1 仿真数据说明

过度季典型日 RIES 的日负荷、PV 出力曲线如图 3 所示, 电、热、冷负荷的最大值分别为 100、175、75 kW。

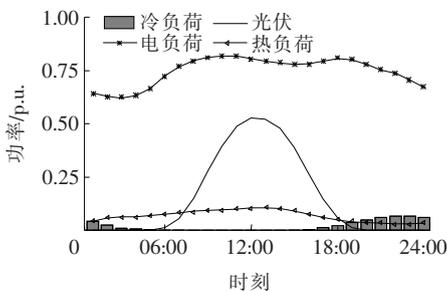


图 3 RIES 在过度季典型日的负荷、PV 出力曲线

Figure 3 Load and PV output curves of RIES on a typical day during the transition season

本文采用文献[25]中的分时电价模式, 峰、平、谷时段的电价分别为 1.35、0.90、0.47 元/(kW·h), 峰时段为 11:00—16:00、19:00—21:00, 平时段为 08:00—11:00、16:00—19:00、21:00—00:00, 谷时段为 00:00—08:00。天然气价为 0.35 元/(kW·h), 二氧化碳的治理费用为 100 元/t。能源转化装置、设备的碳排放以及 Q 学习算法的相关参数分别如表 1~3 所示。

表 1 设备参数

Table 1 The equipment parameters

设备类型	额定容量/kW	产能效率
CHP	200	产电 0.3, 产热 0.45
GB	200	0.95
CR	300	5

表 2 碳排放相关参数

Table 2 Related parameters of carbon emission

单位发电量的 CO <sub>2</sub> 排放量 $\beta_e /$ (t/(MW·h))	单位产热量的 CO <sub>2</sub> 排放量 $\beta_h /$ (t/GJ)	产电向产热的折算系数 $\varphi_{e,h} /$ (MJ/(kW·h))
1.08	0.065	6

表 3 Q 学习算法相关参数

Table 3 Related parameters of Q learning algorithm

学习率 $\alpha$	折现因子 $\gamma$	贪心概率 $\epsilon$	最大迭代次数 $N$	决策周期数 $T$
0.1	0.9	0.99	$10^5$	24

储能 ES 的额定容量为 200 kW·h, 最大的充放电功率为容量的 25%, 荷电状态的取值为 0.1~0.9, 因此, 本文以 0.0125 为区间长度, 按照式 (22) 将荷电状态离散为 65 个值, 作为 MDP 问题的状态空间, 即状态空间共有  $65 \times 24 = 1560$  维。ES 的最大充放电功率为 50 kW, 本文以 2.5 kW 为区间长度, 按照式 (21) 将动作空间的充放电功率离散为 41 个取值。同理, 在 0~200 kW 之间以 2.5 kW 为区间长度将 CHP 的输入离散为 81 个取值, 共有  $41 \times 81 = 3321$  种动作组合。

### 4.2 仿真结果及分析

#### 4.2.1 算法收敛分析

算法的收敛过程如图 4 所示, 给出了学习过程中的最佳动作和平均最佳序列奖励的迭代过程。本文设定的最大迭代次数为  $10^5$  次, 在最终的最终动

作序列下,RIES 的日运行费用为 2 586.91 元。由图 4 可以发现,基于奖励函数的初始化方法,相较于采用将 Q 值表所有元素置为同一个取值的传统初始化方法,Q 学习算法更快收敛,经过 7 000 次迭代后,采用本文改进初始化方法的 Q 学习算法的最佳动作和平均最佳序列奖励已收敛,说明 Q 学习算法已收敛,同时,也证明 Q 学习算法可以有效解决本文所提出能量运行策略问题。

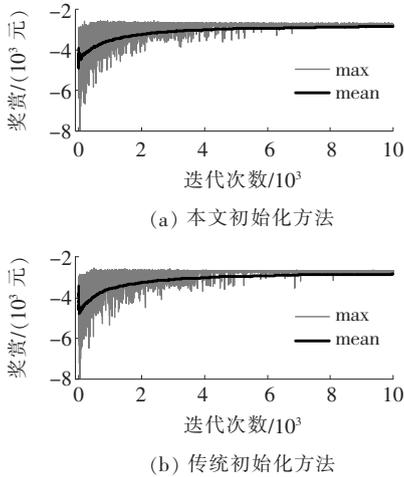


图 4 算法收敛过程

Figure 4 Algorithm convergence process

采用改进 Q 值初始化方法的 Q 学习、遗传、粒子群算法的优化结果对比如表 4 所示,可以发现,不论是目标函数日运行费用还是碳排放量,Q 学习算法具有较好的优化结果。这是因为 Q 学习算法与模型无关,具有良好的后效性,算法会考虑当前操作对未来的影响,会通过牺牲目前的部分收益以换取整体的更大收益。而遗传、粒子群等启发式算法后效性较差,以至于优化结果相对较差。

表 4 不同算法的优化结果

Table 4 Optimization results of different algorithms

算法	日运行费用/元	碳排放量/t
Q 学习	2 586.91	2.240
遗传	2 709.44	2.368
粒子群	2 744.45	2.370

在文 4.1 中,动作空间的 2 个变量  $G_{CHP}(t)$ 、 $P_{ES}(t)$  分别被离散为 81、41 个取值,即在式(20)、(21)中, $a = 80, b = 20$ ,同样可以由式(22)推算出  $c = 64$ (定义为离散情形 2)。为了研究不同离散化程度,即不同  $a, b, c$  取值对算法收敛的影响,本文

进一步对  $a = 40, b = 10, c = 32$ (定义为离散情形 1)和  $a = 160, b = 40, c = 128$ (定义为离散情形 3)2 种情形进行仿真,各情形下算法收敛过程如图 5 所示,各情形下目标函数即 RIES 的日运行费用如表 5 所示。

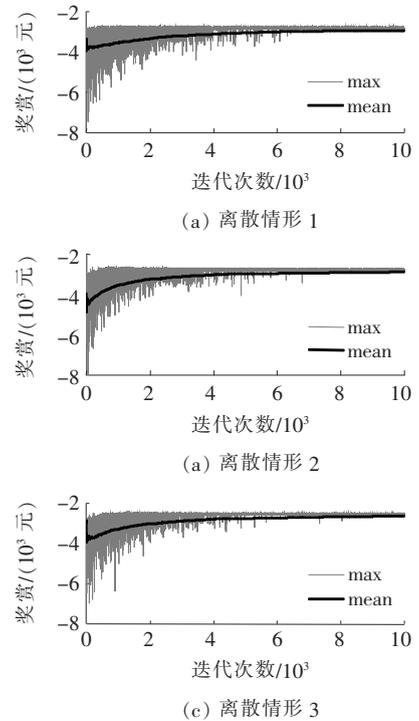


图 5 不同离散化程度下的算法收敛过程

Figure 5 Algorithm convergence process of different degrees of dispersion

表 5 不同离散化程度下的目标函数值

Table 5 Values of objective function within different degrees of dispersion

$a, b, c$ 取值	日运行费用/元
$a = 40, b = 10, c = 32$	2 603.05
$a = 80, b = 20, c = 64$	2 586.91
$a = 160, b = 40, c = 128$	2 543.32

由图 5、表 5 可以看出,随着  $a, b, c$  值的增大,收敛速度变慢,目标函数值增大。而  $a, b, c$  的值越大,离散度也越高。由此可知,随着离散化程度的提高,RIES 的日运行费用也会减小,这意味着 RIES 有更好的运行策略;但同时,随着离散化程度的提高,收敛速度变慢,导致 RIES 运算策略求解变慢。这是因为  $a, b, c$  的取值增大了动作和状态空间的维度,导致算法不易收敛,而  $a, b, c$  取值的增大使得被离散的连续变量更能接近原本的连续取值范围,因此能为系统求解得到更优的运行策略。综上,

离散化程度过高或过低都有一定的缺陷,需要恰当选择  $a$ 、 $b$ 、 $c$  的取值以实现 RIES 的日运行费用与收敛速度之间的折衷。

#### 4.2.2 功率平衡分析

RIES 的运行策略如图 6~8 所示。

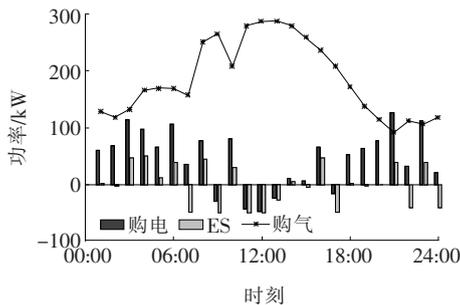


图 6 RIES 能源购买量

Figure 6 Energy purchases of RIES

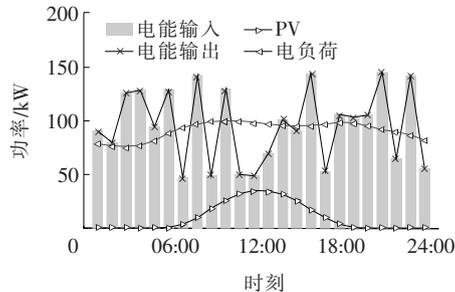


图 7 电能优化结果

Figure 7 Optimization results of power

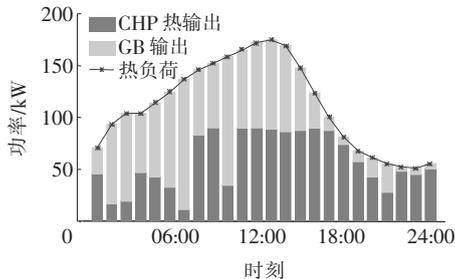


图 8 天然气优化结果

Figure 8 Optimization results of natural gas

图 6 展示了 RIES 向上级电网、上级气网的购电量和购气量以及电储充放电情况。由图 6 可以发现,相比于电价的平价、谷时段,RIES 在电价的高峰时段(11:00—16:00)购电量较少、购气量较多,并且电储装置基本处于“低储高发”的状态。图 7 展示了电能优化结果,其中电能输入表示购电量、PV 出力、CHP 电输出之和,电能输出表示电负荷、CR 输入、ES 充放电功率之和。

对比图 6、7 可知,在满足供需平衡的前提下,电价高峰时段 RIES 不仅减少了向上级电网的购电量,还可以通过增加天然气的购买量进而增大 CHP 的电功率输出;同时,电价高峰时段 ES 处于放电状态,进一步减少了 RIES 的购电量,甚至可以向电网售电,提高 RIES 运行时的经济性。说明通过合理配置 CHP、ES 的运行策略,可以充分发挥 RIES 的多能互补优势,进而实现 RIES 的经济运行。

图 8 给出了最佳动作序列下热功率的平衡情况,CHP 机组的输出热功率与 GB 的热输出功率一起满足用户的热负荷需求。

结合图 6~8 分析可知,合理配置 CHP、ES、GB 等装置的运行策略是实现 RIES 供需平衡和多能互补的基本前提。总的来说,本文采用 Q 学习算法学习得到的运行策略具有以下特点:

1) 电储能在低负荷的谷时段或者平时段利用光伏和谷价电能充电,然后在电价高峰时段集中放电,可以提升 RIES 运行时的经济性能,也能在一定程度上起到缓解主网供电压力的作用;

2) 在电价高峰时段,除了电储放电,RIES 的购气量也相对其他时段有一定的增长,从而 CHP 的电能输出增加,可以进一步减少 RIES 的购电量,甚至可以和电储配合向主网售电;

3) 在满足供需平衡的基础上,求得的运行策略可以充分发挥 RIES 的多能互补优势。

#### 4.2.3 碳排放分析

为了分析 RIES 运行策略的低碳性,本文分析比较目标函数不考虑二氧化碳排放治理成本(情形 1)、二氧化碳的治理费用为 100 元/t(情形 2)这 2 种情况下的 RIES 总运行成本、RIES 与上级电/气网的能源交易成本、碳治理费用以及二氧化碳排放量,优化结果如表 6 所示,可以发现,如果以 100 元/t 的治理费用来处理情形 1 下的二氧化碳排放量,则情形 1 的总成本为  $2\ 350.01 + 2.330 \times 100 = 2\ 580.01$  元,以 0.27% 的降幅略低于情形 2 下的总成本,而情形 1 的碳排放量却相较情形 2 增加了 4.02%。因此,当碳治理成本被纳入目标函数时,尽管目标函数值会小幅增加,碳排放却会相对大幅减少。由此可以得出结论,本文利用 Q 学习算法求解

得到的 RIES 运行策略可以在一定程度上同时兼顾系统的经济性和低碳性。

表 6 碳排放优化结果

Table 6 Optimization results of carbon emissions

情形 1 优化结果			
总运行成本/元	能源交易成本/元	碳治理费用/元	二氧化碳排放量/t
2 350.01	2 350.01	0	2.330
情形 2 优化结果			
总运行成本/元	能源交易成本/元	碳治理费用/元	二氧化碳排放量/t
2 586.91	2 362.91	224.00	2.240

为了进一步分析碳治理价格对 RIES 运行策略的影响,本文将碳治理价格分别重置为 50、100、200 元/t 进行仿真。不同价格下的仿真结果如图 9 所示,其展示了不同价格下系统的碳排放量和日运行费用。从图 9 可以看出,随着碳治理价格的增加,碳排放量随之减少,RIES 的日运行费用随之增加。当碳治理价格定义为 100 元/t 时,碳排放和日常运营成本之间存在折衷。

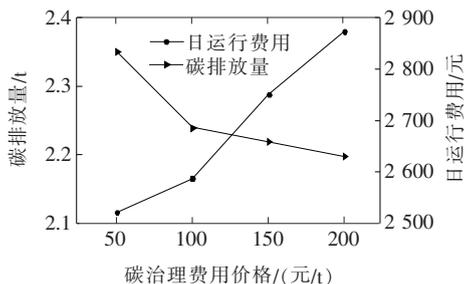


图 9 不同碳治理价格下的仿真结果

Figure 9 Simulation results under different carbon management prices

## 5 结语

基于 Q 学习算法,本文对区域综合能源系统的低碳运行策略展开了研究,得出以下结论:

1) 利用本文给出的改进 Q 学习算法求解区域综合能源系统的运行策略,算法能较快地收敛;

2) 针对区域综合能源系统所建立的低碳运行策略,本文通过合理配置多种能源转化和储能装置的

输出功率,达到了充分发挥区域综合能源系统多能互补优势的目的,进而实现了能源高效利用的目的;

3) 计及二氧化碳排放,本文把二氧化碳排放的治理费用考虑进区域综合能源系统的日运行总费用中,虽然总费用因此而增加,但是区域综合能源系统可以因此以相对低碳的模式运行,即可以在一定程度上兼顾经济性与低碳性。

Q 学习算法只能处理离散问题,后续的研究可以进一步针对处理连续变量而展开。

## 参考文献:

- [1] CHEN J W, YU T, XU Y, et al. Fast analytical method for reliability evaluation of electricity-gas integrated energy system considering dispatch strategies[J]. Applied Energy, 2019, 242: 260-272.
- [2] 龚晓琴,王进,王珑,等. 含电转气的电—气互联综合能源系统低碳经济运行[J]. 电力科学与技术学报, 2020, 35(2): 76-83.  
GONG Xiaoqin, WANG Jin, WANG Long, et al. Low-carbon economic operation for integrated electricity and natural-gas energy system with power-to-gas[J]. Journal of Electric Power Science and Technology, 2020, 35(2): 76-83.
- [3] 高强,刘畅,金道杰,等. 考虑综合需求响应的园区综合能源系统优化配置[J]. 高压电器, 2021, 57(8): 159-168.  
GAO Qiang, LIU Chang, JIN Daojie, et al. Optimal configuration of Park-level integrated energy system considering integrated demand response[J]. High Voltage Apparatus, 2021, 57(8): 159-168.
- [4] 张利军,王一铮,陈飞,等. 计及能源网络特性的综合能源系统最优协调规划[J]. 电力科学与技术学报, 2020, 35(1): 3-13.  
ZHANG Lijun, WANG Yizheng, CHEN Fei, et al. Optimal coordinated planning of an integrated energy system considering characteristics of energy network[J]. Journal of Electric Power Science and Technology, 2020, 35(1): 3-13.
- [5] 白雪婷,杨瑞琪,陈泽春,等. 计及综合需求响应的园区综合能源系统优化运行[J]. 电力科学与技术学报, 2021, 36(3): 27-35.  
BAI Xueting, YANG Ruiqi, CHEN Zechun, et al. Optimal operation of the park integrated energy system con-

- sidering integrated demand response[J]. Journal of Electric Power Science and Technology, 2021, 36(3): 27-35.
- [6] HOU L, SU K, LIU Y, et al. Modeling and simulation of regional integrated energy system operation based on refined thermal equipment model[C]//IEEE Sustainable Power and Energy Conference (ISPEC), Beijing, China, 2019.
- [7] 王海洋, 李珂, 张承慧, 等. 基于主从博弈的社区综合能源系统分布式协同优化运行策略[J]. 中国电机工程学报, 2020, 40(17): 5435-5445.
- WANG Haiyang, LI Ke, ZHANG Chenghui, et al. Distributed coordinative optimal operation of community integrated energy system based on Stackelberg game [J]. Proceedings of the CSEE, 2020, 40(17): 5435-5445.
- [8] 徐文涛, 李明轩. 园区综合能源服务利益方运营关系和经营战略[J]. 中国电力, 2021, 54(4): 151-157.
- XU Wentao, LI Mingxuan. Operation relationship and business strategy of stakeholders of park-level integrated energy services[J]. Electric Power, 2021, 54(4): 151-157.
- [9] 徐晨博, 薛友, 林紫菡, 等. 计及可转移负荷的电一气综合能源系统多目标优化[J]. 电力科学与技术学报, 2019, 34(3): 13-23.
- XU Chenbo, XUE You, LIN Zihan, et al. Multi-objective optimization of an integrated electricity-gas energy system considering transferable loads[J]. Journal of Electric Power Science and Technology, 2019, 34(3): 13-23.
- [10] ZHU X, YANG J, LIU Y, et al. Optimal scheduling method for a regional integrated energy system considering joint virtual energy storage[J]. IEEE Access, 2019, 7: 138260-138272.
- [11] YANG H Z, LI M L, JIANG Z Y, et al. Multi-time scale optimal scheduling of regional integrated energy systems considering integrated demand response[J]. IEEE Access, 2020, 8: 5080-5090.
- [12] 白宏坤, 尹硕, 李虎军, 等. 计及碳交易成本的多能源站综合能源系统规划[J]. 电力科学与技术学报, 2019, 34(1): 11-19.
- BAI Hongkun, YIN Shuo, LI Hujun, et al. Optimal planning of multi-energy stations considering carbon-trading cost[J]. Journal of Electric Power Science and Technology, 2019, 34(1): 11-19.
- [13] 杨挺, 赵黎媛, 王成山. 人工智能在电力系统及综合能源系统中的应用综述[J]. 电力系统自动化, 2019, 43(1): 2-14.
- YANG Ting, ZHAO Liyuan, WANG Chengshan. Review on application of artificial intelligence in power system and integrated energy system[J]. Automation of Electric Power Systems, 2019, 43(1): 2-14.
- [14] LU R Z, HONG S H, ZHANG X F. A dynamic pricing demand response algorithm for smart grid: reinforcement learning approach[J]. Applied Energy, 2018, 220: 220-230.
- [15] 瞿凯平, 张孝顺, 余涛, 等. 基于知识迁移 Q 学习算法的多能源系统联合优化调度[J]. 电力系统自动化, 2017, 41(15): 18-25.
- QU Kaiping, ZHANG Xiaoshun, YU Tao, et al. Knowledge transfer based Q-learning algorithm for optimal dispatch of multi-energy system[J]. Automation of Electric Power Systems, 2017, 41(15): 18-25.
- [16] 金璐, 何伟, 闫华光, 等. 考虑多评价指标的可再生微能源网双层优化配置[J]. 电测与仪表, 2022, 59(2): 112-119.
- JIN Lu, HE Wei, YAN Huaguang, et al. Double layer optimal allocation of renewable micro energy network considering multiple evaluation indexes[J]. Electrical Measurement & Instrumentation, 2022, 59(2): 112-119.
- [17] 张宁, 孙秋野, 马大中. 基于模糊 Q 学习的自能源综合能量管理策略[J]. 全球能源互联网, 2019, 2(6): 530-537.
- ZHANG Ning, SUN Qiuye, MA Dazhong. Comprehensive energy management strategy of we-energy based on fuzzy Q-learning[J]. Journal of Global Energy Interconnection, 2019, 2(6): 530-537.
- [18] ALFAVERH F, DENAI M, SUN Y. Demand response strategy based on reinforcement learning and fuzzy reasoning for home energy management[J]. IEEE Access, 2020, 8: 39310-39321.
- [19] 刘洪, 李吉峰, 葛少云, 等. 基于多主体博弈与强化学习的并网型综合能源微网协调调度[J]. 电力系统自动化, 2019, 43(1): 40-50.
- LIU Hong, LI Jifeng, GE Shaoyun, et al. Coordinated scheduling of grid-connected integrated energy micro-grid based on multi-agent game and reinforcement learning[J]. Automation of Electric Power Systems, 2019, 43(1): 40-50.