

# 基于改进 Dueling DQN 的无人机网络 多域资源分配策略

黄嘉伟, 黎海涛, 吕鑫

(北京工业大学 信息学部, 北京 100124)

**摘要:** 针对无人机(Unmanned Aerial Vehicle, UAV)网络节点间干扰导致通信质量降低问题,文中从空域与功率域出发,构建以最大化 UAV 系统频谱效率和能量效率为目标的波束成形与发射功率联合优化模型。针对该多域资源分配模型,提出参数化动作空间的加权 Dueling DQN(P-wDDQN)学习算法,该算法适用于包含连续功率分配和离散波束成形的混合动作空间,且解决了 Dueling DQN 的目标  $Q$  值过低估计问题。利用所提 P-wDDQN 算法设计了联合波束与功率资源分配策略。仿真结果表明,所提多域资源分配策略提升了无人机网络的频谱效率和能量效率,且具有快速收敛的优势。

**关键词:** UAV 网络;深度强化学习;多域资源分配;参数化动作空间

**中图分类号:** TN92      **文献标志码:** A      **文章编号:** 1673-5692(2023)07-646-06

## A Multidomain Resource Allocation Strategy for UAV Network Based on Improved Dueling DQN

HUANG Jia-wei, LI Hai-tao, LV Xin

(Faculty of Information Technology, Beijing University of Technology, Beijing 100124, China)

**Abstract:** To address the problem of communication quality degradation due to inter-link interference in unmanned aerial vehicle (UAV) network, this paper constructs a joint beamforming and transmit power optimization model with the objective of maximizing the spectral efficiency and energy efficiency of UAV systems from the space and power domains. A weighted Dueling DQN learning algorithm based on parameterized action space (P-wDDQN) is proposed for the multidomain resource allocation model. This algorithm is suitable for mixed action spaces containing continuous power allocation and discrete beamforming, and solves the problem of overestimating the target  $Q$  value of Dueling DQN. A beamforming and power resource allocation strategy was designed using the proposed P-wDDQN algorithm. Simulation results show that the proposed multidomain resource allocation strategy improves the spectral efficiency and energy efficiency of UAV network and has the advantage of rapid convergence.

**Key words:** UAV network; deep reinforcement learning; multidomain resource allocation; parameterized action space

## 0 引言

随着无线通信技术的快速发展,由无人机组成的

空中网络能够建立视距通信链路而实现信号的定向可靠传输。然而,由于无线环境的开放特性,UAV 通信链路易受到强的射频干扰,导致系统频谱效率降低。另一方面,无人机飞行过程也受限于其能量

收稿日期:2023-06-26      修订日期:2023-08-01

基金项目:航空科学基金资助项目(2018ZC15003)

供给能力。因此,亟需研究有效的 UAV 抗干扰通信技术,在保障通信质量的前提下提高频谱效率与能源利用率<sup>[1-2]</sup>。

基于波束成形和功率控制的多域资源分配技术能对 UAV 网络进行有效的干扰管理。目前,一些学者对 UAV 联合波束成形与功率控制进行了深入研究。文献[3]针对无人机辅助蜂窝网络的场景,提出基于 Dinkelbach 算法和逐次凸近似法的 UAV 发射功率和波束成形优化策略来抑制干扰。文献[4]针对无人机中继网络,提出基于交替干扰抑制算法的联合波束成形与功率控制的设计方法,可最大化目标信号的波束成形增益并降低干扰。文献[5]针对无人机全双工中继网络,提出了基于块坐标下降法的联合波束成形与功率控制优化策略。

这些 UAV 多域资源分配技术均采用基于规则的优化算法,需要大量迭代和矩阵反演计算,故计算复杂度较高。随着人工智能技术的发展,基于深度强化学习(Deep Reinforcement Learning, DRL)的无线网络资源分配受到广泛关注。数据驱动的深度强化学习算法通过智能体与环境交互来选择最优控制策略的方式更为高效。譬如,文献[6]提出基于分层 DDPG 的联合波束成形与功率控制算法,以最大化时变下行链路的和速率。文献[7]提出基于 DQN 算法的下行链路波束成形和功率控制优化方案,可有效平衡联合优化问题中的系统性能与时间复杂度。

上述研究中采用深度强化学习算法仅优化 UAV 网络频谱效率,且没有考虑到学习算法的动作空间同时存在离散、连续动作的情形,以及 DQN 算法的  $Q$  值过估计问题。为进一步提高无人机网络通信性能,本文综合考虑 UAV 系统频谱效率与能量效率,提出一种参数化动作空间的加权 Dueling DQN 算法,并据此研究 UAV 网络的多域资源分配策略。该策略针对多域资源分配中由连续功率和离散波束组成的混合动作空间,设计了动作—参数  $Q$  网络以实现连续—离散动作决策,避免了连续功率动作离散化导致的量化误差。同时考虑 Dueling DQN 的  $Q$  值低估计问题,设计加权双估计器避免单估计量的高估和双估计量的低估问题,以更为准确地估计目标  $Q$  值,从而提供更优的资源分配策略。

## 1 系统模型

本文采用长机—僚机(Leader-Follower)的分布

式编队控制方法,构建如图 1 的无人机网络,其根据任务需求划分为若干联盟,相互协同执行任务。每个 UAV 联盟内部包括配置有均匀线性阵列天线阵元数为  $M$  的  $I$  个长机和单天线的  $K$  个从属僚机,部署于长机的指挥平台,能够主动收集各 UAV 状态信息,并利用所提算法进行功率和波束资源的分配。



图 1 UAV 系统模型

UAV 网络中,长机发送信号经过信道传输后到达僚机  $k$  的接收信号为

$$y_k = \mathbf{h}_{k,i}^H \mathbf{w}_i x_i + \sum_{j \neq i} \mathbf{h}_{k,j}^H \mathbf{w}_j x_j + n_k \quad (1)$$

式中:  $x_i, x_j \in \mathbb{C}^{M \times 1}$  为联盟内长机  $i$  和联盟外长机  $j$  的发射信号,其功率满足约束  $E[|x_i|^2] = P_i$ ;  $\mathbf{w}_i, \mathbf{w}_j \in \mathbb{C}^{M \times k}$  表示长机  $i$  和  $j$  的波束成形向量;  $\mathbf{h}_{i,k}^H, \mathbf{h}_{j,k}^H$  分别为长机  $i$  和  $j$  到僚机  $k$  之间的信道增益;  $n_k \sim CN(0, \sigma^2)$  表示均值为零、方差为  $\sigma^2$  的加性复高斯噪声。

式(1)的首项表示僚机  $k$  接收的期望信号,第二项为其他 UAV 发射信号对僚机  $k$  的干扰。

波束成形向量的波束成形权重由恒模相位移位器  $[\mathbf{w}_n(t)]_m = e^{j\theta_m}$  所确定,其第  $n$  个元素为

$$\mathbf{w}_n = \frac{1}{\sqrt{M}} [1, e^{jnd \cos(\theta_n)}, \dots, e^{jnd(M-1) \cos(\theta_n)}]^T \quad (2)$$

式中:  $n = 2\pi/\lambda$ ;  $\lambda$  为载波波长;  $d$  为天线阵列元件间距;方位角  $\theta_n$  由  $0 \sim \pi$  的天线角度空间除以天线阵元数  $M$  得到。

UAV 间通信链路性能由接收信号与干扰噪声功率比(Signal to Interference Plus Noise Ratio, SINR)决定,则第  $k$  个 UAV 接收的 SINR 为

$$\gamma_{k,i} = \frac{P_i |\mathbf{h}_{k,i}^H \mathbf{w}_i(t)|^2}{\sigma^2 + \sum_{j \neq i} P_j |\mathbf{h}_{k,j}^H \mathbf{w}_j(t)|^2} \quad (3)$$

式中:  $P_i$  为 UAV 长机  $i$  的电台发射功率,利用得到的 SINR 值,可计算该链路的频谱效率(容量),为

$$\eta_{SE}^{k,i} = \log_2 \left( 1 + \frac{P_i |\mathbf{h}_{k,i}^H \mathbf{w}_i(t)|^2}{\sigma^2 + \sum_{j \neq i} P_j |\mathbf{h}_{k,j}^H \mathbf{w}_j(t)|^2} \right) \quad (4)$$

UAV 长机作为信号发射端与所属联盟内僚机

进行通信的同时,需降低联盟外其他长机-僚机间的数据传输对其造成的干扰。为此,UAV 发射端可以调整发射功率和波束成形向量,以满足 UAV 通信服务质量要求。

另一方面,随着 UAV 传输业务的增长,其能量消耗急剧增加。因此,UAV 网络在追求频谱效率与服务质量的同时,需要考虑能量供给的可持续性。定义能量效率(Energy Efficiency, EE)为无人机通信时供能可持续性的评价指标,其为僚机  $k$  传输数据时所消耗的能量与传输的数据量之比,即

$$\eta_{EE}^{k,i} = \frac{C}{p_i + P_c + P_{RF}} \quad (5)$$

式中: $P_c$ 和 $P_{RF}$ 分别为基带电路和射频天线的功耗。

注意到增大 UAV 发射功率能在提高频谱效率的同时降低能量效率,故需在 UAV 频谱效率和能量效率之间进行合理折衷。为在满足数据传输服务质量的(Quality of Service, QoS)前提下最大化 UAV 系统的谱效率和能效,建立 UAV 网络多域资源分配的联合优化模型为

$$\begin{cases} \max \lambda \sum_{k=1}^K \sum_{i=1}^I \eta_{SE}^{k,i} + (1-\lambda) \sum_{k=1}^K \sum_{i=1}^I \eta_{EE}^{k,i} \\ \text{Subject to:} \\ 0 \leq P_i \leq P_{\max}, \forall i \\ \gamma_{k,i} > \gamma_{\text{target}}, \forall i, k \end{cases} \quad (6)$$

式中: $P_{\max}$ 为无人机发射功率阈值; $\gamma_{\text{target}}$ 为链路传输的目标 SINR; $\lambda$ 为调节系数。

## 2 基于 P-wDDQN 的多域资源分配

由于优化模型式(6)中目标函数中存在干扰项,故该问题为一个非凸的优化问题。采用传统优化算法求解的搜索空间过大,易陷入局部最优而难以获得全局最优解。考虑到无人机通信链路状态动态变化的特性,本文采用解决序列决策问题的强化学习算法,通过与 UAV 网络进行环境交互,并在训练过程中根据环境反馈动态调整资源分配策略,直至算法收敛得到最佳策略。

### 2.1 RL 元组定义

所构建的优化问题模型中,UAV 运动轨迹、发射功率和波束状态等都具有马尔可夫性,即下一个时刻的状态仅与当前状态有关,与之前的状态均无关,可将其建立为离散时间马尔可夫决策过程(Markov Decision Process, MDP)。一个完整的马尔

科夫决策过程由四元组  $\langle s, a, p_{s,a}, r \rangle$  表示,其中  $s$  为状态空间、 $a$  为动作空间、 $p_{s,a}$  为状态转移概率、 $r$  为奖励值。考虑到状态转移概率在本问题中是未知的,故将 MDP 重新定义为三元组  $\langle s, a, r \rangle$ 。利用深度强化学习求解该 MDP 问题时,元组中状态、动作和奖励函数的定义如下。

#### 1) 状态空间

状态空间用于表征当前无人机网络的环境状态,由 UAV 空间位置坐标 $k_{t,i}$ 、波束成形向量 $s_t^w$ 和发射功率 $s_t^p$ 组成,智能体在与环境交互过程中根据当前状态实时做出决策。若无人机飞行高度提前预设,则空间位置坐标只需记录僚机的水平相对位置。在  $t$  时刻,UAV 状态空间定义为

$$s_t = [s_{t,1}^w, s_{t,1}^p, k_{t,1}, s_{t,2}^w, s_{t,2}^p, k_{t,2}, \dots, s_{t,i}^w, s_{t,i}^p, k_{t,i}], \forall i \in I \quad (7)$$

#### 2) 动作空间

智能体在时刻  $t$  的动作空间为

$$a_t = [w^1(t), p^1(t), w^2(t), p^2(t), \dots, w^i(t), p^i(t)], \quad \forall i \in I \quad (8)$$

式中: $p(t)$ 为 UAV 发射功率; $w(t)$ 为波束成形向量对应的码本索引值。

#### 3) 奖励函数

奖励函数用于评估当前状态-动作对的学习策略性能,本文以最大化 UAV 谱效和能效为目标,定义奖励函数为

$$r(t) = \lambda \sum_{k=1}^K \sum_{i=1}^I \eta_{SE}^{k,i} + (1-\lambda) \sum_{k=1}^K \sum_{i=1}^I \eta_{EE}^{k,i} \quad (9)$$

### 2.2 加权 Dueling DQN 算法

针对定义的多域资源分配 MDP 问题,可用经典的 DQN 等深度强化学习算法来求解。当采用 DQN 算法,对于许多状态而言,无需估计每个动作的  $Q$  值,即值函数的大小与动作无关。因此,基于 DQN 进行改进得到的 Dueling DQN 算法,能够快速识别最佳动作而提高学习效率,其将 DQN 算法中  $Q$  网络的输出解耦为基于状态的价值函数和基于动作的优势函数,即

$$Q(s, a; \theta, \alpha, \beta) = V(s; \theta, \beta) + A(s, a; \theta, \alpha) \quad (10)$$

式中:状态价值函数  $V(s; \theta, \beta)$  仅与状态有关,不考虑动作对状态的影响;动作价值函数  $A(s, a; \theta, \alpha)$  与动作和状态都有关; $\theta, \alpha, \beta$  分别表示公共网络、状态价值网络和优势网络的网络参数。若仅用该公式进行更新,会存在不可辨识性问题,即若  $V(s; \theta, \beta)$  和  $A(s, a; \theta, \alpha)$  分别加上或减去一个值,能够得到相同



的  $Q$  值,但反之无法由相同的  $Q$  得到唯一的  $V(s; \theta, \beta)$  和  $A(s; a, \theta, \alpha)$ 。为解决该问题, Dueling DQN 对同一状态采用固定的状态值函数来更新动作优势函数,即

$$Q(s, a; \theta, \alpha, \beta) = V(s; \theta, \beta) + [A(s, a; \theta, \alpha) - \frac{1}{|A|} \sum_{a'} A(s, a'; \theta, \alpha)] \quad (11)$$

式中:  $\frac{1}{|A|} \sum_{a'} A(s, a'; \theta, \alpha)$  表示当前状态下所有可能动作优势函数的平均值。

在训练过程中,每执行  $T$  时间步,当前  $Q$  网络的相关参数复制到目标  $Q$  网络进行更新。目标  $Q$  值更新如下:

$$\tilde{Q} = r_t + \gamma \max_{a'} Q(s', a'; \theta^-, \alpha^-, \beta^-) \quad (12)$$

式中:  $s', a'$  分别为状态和动作;  $\theta^-, \alpha^-, \beta^-$  分别表示目标网络、状态价值函数和动作价值函数的网络参数; 奖励折扣因子  $\gamma \in [0, 1]$  表示未来奖励值所占比重。

Dueling DQN 算法的双重估计器会导致低估计,增加目标值估计的偏差。为解决该问题,本文设计了加权 Dueling DQN 算法 wDDQN,其目标网络输出为

$$\bar{Q} = r_t + \gamma [\mu \cdot Q(s', a'; \theta^-, \alpha^-, \beta^-) + (1 - \mu) \cdot Q(s', a''; \theta^-, \alpha^-, \beta^-)] \quad (13)$$

其中,

$$a' = \arg\max Q(s', a; \theta, \alpha, \beta)$$

$$a'' = \arg\min Q(s', a; \theta, \alpha, \beta)$$

$$\rho = |Q(s', a'; \theta^-, \alpha^-, \beta^-) - Q(s', a''; \theta^-, \alpha^-, \beta^-)|$$

式中:  $a', a''$  分别为下一状态时最大和最小化当前  $Q$  值对应的动作;  $\mu$  为平衡  $Q(s', a'; \theta^-, \alpha^-, \beta^-)$  与  $Q(s', a''; \theta^-, \alpha^-, \beta^-)$  的权重系数,  $c$  为可调超参数。

### 2.3 参数化 Dueling DQN 算法

所构建的多域资源分配 MDP 决策模型中,动作空间内同时存在连续发射功率和离散波束码本索引两类动作,而一般 Dueling DQN 算法仅适用于离散动作空间。为此,本文提出一种基于参数化动作空间的方法,以联合优化波束成形与发射功率。参数化动作空间可表示为  $A^{\text{PF}} = \{(u, v_u) | u \in A_d, v_u \in A_c\}$ , 其中,  $A_c$  和  $A_d$  分别代表连续动作空间(即发射功率)

与离散动作空间(即波束码本索引)。每个离散动作都有对应的连续动作参数  $v_u \in V_u$ , 采用参数化的动作空间,智能体不仅可以从离散动作集合中选取离散动作,而且可以根据该离散动作对应的连续动作参数,选择与其共同作用的连续动作参数,并执行相应的连续动作。为实现参数化动作空间,需在当前  $Q$  网络的基础上构造一个动作-参数网络  $v_u(s; \omega)$ , 如图 2 所示。

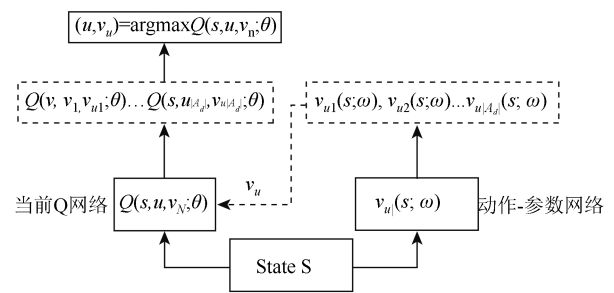


图 2 参数化动作空间  $Q$  网络设计

具体地,参数化动作空间  $Q$  网络实现流程为:先将当前状态  $s$  输入到动作-参数网络  $v_u(s; \omega)$ , 使离散动作空间与连续动作参数  $\chi_c$  相互映射,输出离散动作空间对应的连续动作参数集合  $v_{u1}(s; \omega), v_{u2}(s; \omega), \dots, v_{u_{|A_d|}}(s; \omega)$ ; 然后,将该参数集作为当前  $Q$  网络  $Q(s, u, v_u; \theta)$  的输入,经处理后输出每个  $(u, v_u)$  对应的  $Q$  值,利用  $\arg\max Q(s_i, u, v_u; \theta)$  选取最优动作  $(u, v_u)$ 。此外,参数化 Dueling DQN 对经验回放池采样得到随机样本,用于计算动作网络和动作-参数网络的损失函数为

$$L^v(\omega) = -\frac{1}{N} \sum_{i=1}^N Q(s_i, u_i, v_{u_i}(s_i; \omega); \theta) \quad (14)$$

$$L^Q(\theta) = \frac{1}{N} \sum_{i=1}^N (\bar{Q} - Q(s_i, u_i, v_{u_i}(s_i; \omega); \theta))^2 \quad (15)$$

式中:  $N$  表示训练样本大小。迭代计算过程中,通过损失函数更新神经网络权重,损失函数更新如下:

$$\omega = \theta - \alpha_{a,p} \nabla_{\theta} L^v(\omega) \quad (16)$$

$$\theta = \omega - \alpha_a \nabla_{\theta} L^Q(\theta) \quad (17)$$

式中:  $\alpha_{a,p}$  和  $\alpha_a$  分别代表动作网络和动作-参数网络的学习率。

当前状态下,  $v_u(\omega)$  将输出连续参数以最大化每个离散动作的  $Q$  值。在训练过程中,智能体基于  $\varepsilon$ -greedy 策略选择动作并进行探索。

基于上述加权 Dueling DQN 和参数化 Dueling DQN 两方面的改进,所提基于 P-wDDQN 的多域资

源分配策略实现流程如算法 1。

算法 1 基于 P-wDDQN 的多域资源分配策略	
输入: UAV 网络初始状态	
输出: 波束成形码本序列与发射功率值	
初始化: 模型参数、状态空间、经验池 D	
1: for episode = 0 to maximum episode do	
2:   for t = 1 to T do	
3:   观测当前状态 $s$ , 初始化策略 $e$	
4:   根据贪婪算法选择动作 $a_t = (u_t, v_{u_t})$	
5:   计算 $\gamma_{i,j}$ 并判断是否其值是否大于 $\gamma_{\min}$ , 若小于最小信噪比, 则直接赋最小奖励值 $r_{\min}$	
6:   基于当前状态与环境交互后计算与并对奖励 $r$ 进行更新	
7:   观测下一时刻状态 $s'$ , 将四元组 $(s, a, r(s, a), s')$ 存储到经验池 D 中	
8:   随机批量采样小样本 $d$ , 对模型进行训练	
9:   计算目标 Q 值:	
$\bar{Q} = r_t + \gamma[\mu \cdot Q(s', a'; \theta^-, \alpha^-, \beta^-) + (1 - \mu)Q(s', a''; \theta^-, \alpha^-, \beta^-)]$	
10:   end for	
11:   基于式(16), (17)更新 $Q(s, u, v_u; \theta)$ 和 $v_u(s; \omega)$ 的网络参数	
12:   每隔 T 时间步复制当前 Q 网络参数到目标 Q 网络进行更新	
13:   end for	

### 3 仿真结果

假设 UAV 联盟内僚机随机分布位于长机周围半径为  $R$  的覆盖区域, UAV 悬停高度 100 m。仿真参数如表 1 和表 2 所示。利用这些参数对基于 P-wDDQN 的多域资源分配策略进行性能仿真。

表 1 仿真参数

仿真参数	数值
折扣因子 $\gamma$	0.995
初始探索率 $\varepsilon$	1
最小探索率	0.10
学习率	0.01
激活函数	ReLu
Minibatch 大小	32
经验池容量 $N$	2 000

图 3 给出了 UAV 链路平均频谱效率与天线阵元数量的关系, 可以看到随着天线数的增加, UAV 链路的平均频谱效率增大。且所提 P-wDDQN 多域资源分配策略由于避免了连续动作离散化而引入的量化误差, 故其频谱效率优于 DQN 和 Dueling DQN 方法。

表 2 环境参数

仿真参数	数值
最大传输功率/dBm	46
下行频带/GHz	28
覆盖半径/m	150
UAV 移动速度/(m · s <sup>-1</sup> )	2
接收天线增益/dBi	0
发射天线增益/dBi	3
天线数量	[4, 8, 16, 32, 64]
无线帧长/ms	10

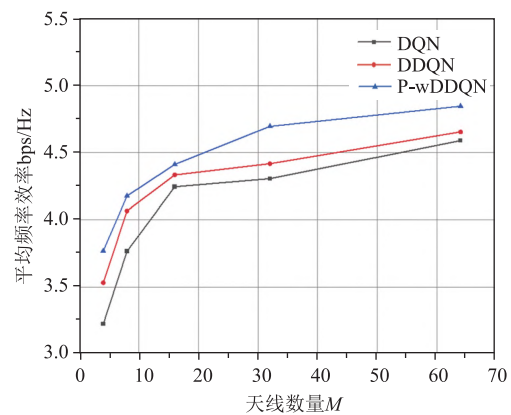


图 3 频谱效率对比

图 4 表明不同天线数量时无人机的平均能量效率。由于天线功耗与其数量成正比, 能量效率将随着天线数量的增加而降低。同时, 基于 P-wDDQN 资源分配策略的能量效率优于 DQN 和 Dueling DQN。

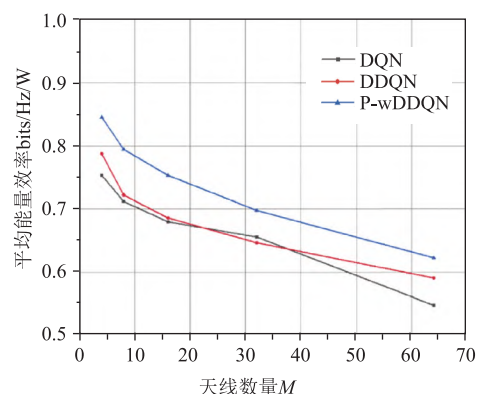


图 4 能量效率对比

图 5 对比了不同算法的收敛性, 本文将收敛回合定义为 UAV 接收 SINR 达到目标信噪比所需要的迭代次数。可以看到, 随着天线的数量增加, 波束成形向量对应的码本维度变高, 学习算法动作空间增

大,导致收敛回合随之增加。同时,相较于 DQN、Dueling DQN 算法,所提 P-wDDQN 算法的收敛回合数更低,即收敛性能更优。

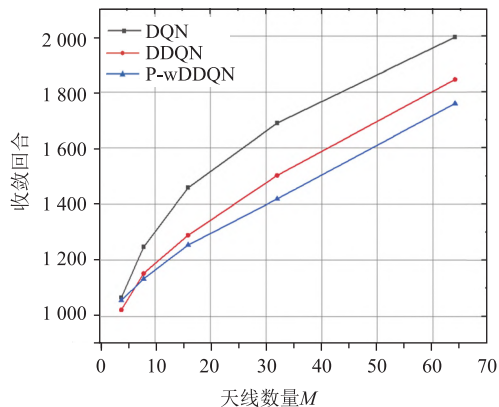


图 5 收敛回合对比

## 4 结 语

本文研究了无人机网络的多域资源分配问题,提出一种参数化动作空间的加权 Dueling DQN 的联合波束成形与功率控制算法。其通过加权 Dueling 的方法解决了 Dueling DQN 的低估计问题,且设计参数化动作空间,避免了将连续动作空间离散化导致的量化误差。仿真结果表明,与现有 DQN 与 DDQN 学习策略相比,所提 P-wDDQN 多域资源分配策略具有更优的能量效率和频谱效率,且收敛性能更优。

### 参考文献:

- [1] 陈新颖,盛敏,李博,等. 面向 6G 的无人机通信综述[J]. 电子与信息学报, 2022, 44(3): 781-789.
- [2] GUPTA L, JAIN R, VASZKUN G. Survey of important issues in UAV communication networks[J]. IEEE Communications Surveys & Tutorials, 2016, 18(2): 1123-1152.
- [3] RAHMATI A, HOSSEINALIPOUR S, YAPICI Y, et al. Energy-efficient beamforming and power control for uplink NOMA in mmWave UAV networks[C]// Global Communications Conference. Taipei: IEEE press, 2020: 1-6.
- [4] ZHU L, ZHANG J, XIAO Z, et al. Millimeter-wave full-duplex UAV relay: Joint positioning, beamforming, and power control[J]. IEEE Journal on Selected Areas in Communications, 2020, 38(9): 2057-2073.
- [5] SONG Q, ZHENG F C, ZENG Y, et al. Joint beamforming and power allocation for UAV-enabled full-duplex re-

lay[J]. IEEE Transactions on Vehicular Technology, 2018, 68(2): 1657-1671.

- [6] LIU J, LIN CR, HU YC, DONTA PK. Joint beamforming, power allocation, and splitting control for SWIPT-Enabled IoT networks with deep reinforcement learning and game theory. Sensors (Basel). 2022, 17; 22(6):2328.
- [7] MISMAR F B, EVANS B L, ALKHATEEB A. Deep reinforcement learning for 5G networks: Joint beamforming, power control, and interference coordination[J]. IEEE Transactions on Communications, 2020, 68(3): 1581-1592.
- [8] 姚昌华,高泽邵,韩贵真,等. 基于数据辅助的无人机集群协同空域抗干扰[J]. 电子测量技术, 2022, 45(16):164-170.
- [9] 王晶,顾维博,窦立亚. 基于 Leader-Follower 的多无人机编队轨迹跟踪设计[J]. 航空学报, 2020, 41(S1): 88-98.
- [10] LYU Z H, ZHU G X, XU J. Joint maneuver and beamforming design for UAV-enabled integrated sensing and communication[EB/OL]. [2023-03-24]. [http://arXiv:2110.02857\[cs.IT\]](http://arXiv:2110.02857[cs.IT]).
- [11] LU Y, WEI L, HE S. et al. Resource efficient beamforming design for heterogeneous multiuser systems[C]// International Conference on Wireless Communications & Signal Processing (WCSP). Nanjing: IEEE Press, 2015:1-5.
- [12] LEE C, KIM Y H. Receive beamforming and resource allocation for wireless powered non-orthogonal multiple access[J]. IEEE Transactions Veh. Technol., 2020, 69(4):4563-4568.
- [13] XIAOZ, ZHU L, CHOI J, et al. Joint power allocation and beamforming for non-orthogonal multiple access (noma) in 5G millimeter wave communications[J]. IEEE Transactions. Wireless Commun., 2018, 17(5): 2961-2974.
- [14] CHENG L, YUE G, XIONG X, et al. Tensor decomposition-aided time-varying channel estimation for millimeter wave MIMO systems[J]. IEEE Wireless Communications Letters, 2019, 8(4): 1216-1219.

### 作者简介

黄嘉伟(1998—),硕士研究生,主要研究方向为网络通信与计算;

黎海涛(1972—),通信作者,副教授,主要研究方向为网络通信与智能计算、信号处理等;

吕鑫(1997—),硕士研究生,主要研究方向为网络通信与智能计算。