

# 基于多无人机的空中计算网络资源分配算法

谈 玲<sup>1</sup>, 许 海<sup>1</sup>, 刘玉风<sup>2</sup>, 夏景明<sup>3</sup>

(1. 南京信息工程大学计算机学院, 江苏南京 210044; 2. 南京信息工程大学软件学院, 江苏南京 210044;  
3. 南京信息工程大学人工智能学院, 江苏南京 210044)

**摘 要:** 空中计算(over-the-Air Computation, AirComp)是一种有效提升分布式数据聚合效率的方法. 现有研究大多采用单无人机(Unmanned Aerial Vehicle, UAV)方案, 未考虑数据聚合质量和系统稳定性. 为此, 本文提出一种基于多 UAV 辅助的 AirComp 网络, 旨在实现多个地面移动传感器(Ground Mobile Sensor, GMS)的高效聚合. 为了改进数据采集质量并全面反映系统性能, 本文设计了一个多约束优化问题, 通过联合优化 UAV-GMS 关联、UAV 三维(Three Dimensional, 3D)部署、UAV 去噪因子以及传输功率分配, 以最大化系统的最小可达速率. 针对多约束优化问题的非线性特征, 本文提出一种 AirComp 网络下多 UAV 辅助的深度确定性策略梯度优化算法(Deep Deterministic Policy Gradient-based optimization algorithm for multi-UAV cooperation in AirComp network, AirDDPG-UAV), 用以协助多 UAV 在复杂环境下快速响应聚合任务. 该算法利用深度强化学习的确定性策略对网络中的状态、行为和奖励进行优化, 以最大化系统最小可达速率. 数值结果显示, AirDDPG-UAV 算法在保证较低的系统能耗和计算复杂度前提下, 能够使系统最小可达速率提高 15%, 表明本文所提方案适用于分布式数据聚合, 可以有效提高数据聚合效率.

**关键词:** 无人机; 空中计算; 3D 部署; 深度确定性策略梯度算法; 地面移动传感器; 数据聚合

**基金项目:** 国家自然科学基金(No.62271264)

**中图分类号:** TP391

**文献标识码:** A

**文章编号:** 0372-2112(2023)11-3070-09

**电子学报 URL:** <http://www.ejournal.org.cn>

**DOI:** 10.12263/DZXB.20230513

## Resource Allocation Algorithm of AirComp Network Based on Multiple UAVs

TAN Ling<sup>1</sup>, XU Hai<sup>1</sup>, LIU Yu-feng<sup>2</sup>, XIA Jing-ming<sup>3</sup>

(1. School of Computer, Nanjing University of Information Science and Technology, Nanjing, Jiangsu 210044, China;

2. School of Software, Nanjing University of Information Science and Technology, Nanjing, Jiangsu 210044, China;

3. School of Artificial Intelligence, Nanjing University of Information Science and Technology, Nanjing, Jiangsu 210044, China)

**Abstract:** Over-the-air computation (AirComp) is an effective method to improve the efficiency of distributed data aggregation, which can complete some task calculations while transmitting in the air. Most existing researches focus on the single unmanned aerial vehicle (UAV) scheme, without considering the quality of data aggregation and the stability of the system, making it unsuitable for practical AirComp environments. Therefore, this paper proposes an AirComp network based on multiple UAVs collaboration, which aims to achieve the efficient data aggregation for multiple ground mobile sensors (GMSs). In order to refine data acquisition and fully reflect system status, a multi-constraint non-convex optimization problem is constructed to jointly optimize UAV-GMS association, the three dimensional (3D) deployment of UAVs, UAV denoising factors, and transmission power allocation, aiming for maximizing the system's minimum achievable rate. Giving the nonlinear characteristics of multiple constraints optimization problems, a deep deterministic policy gradient-based optimization algorithm for multiple UAVs cooperation in AirComp network (AirDDPG-UAV) is proposed to assist UAVs rapidly responding to aggregation missions in complex environments. A deterministic policy in deep reinforcement is adopted to optimize the states, behaviors, and rewards of the AirComp network, aiming to maximize the minimal achievable rate. The numerical results show that the AirDDPG-UAV algorithm can significantly improve the system's minimum achievable rate by more than 15% compared to the benchmark meth-

ods, while ensuring suitable system energy consumption and computational complexity. The AirDDPG-UAV algorithm also obtains satisfactory results in optimizing the mean MSE, which illustrates our method has excellent performance in scaling signals and thus is helpful for fast data aggregation. The experiments indicate the proposed scheme is appropriate for the distributed data aggregation with low cost and can obviously improve the efficiency and stability of data aggregation.

**Key words:** unmanned aerial vehicle; over-the-air computation; 3D deployment; deep deterministic policy gradient algorithm; ground mobile sensors; data aggregation

**Foundation Item(s):** National Natural Science Foundation of China (No.62271264)

## 1 引言

随着物联网(Internet of Things, IoT)技术的广泛应用,大量传感器数据的收集和聚合正成为挑战. 受限于带宽频谱和计算能力等因素,传统的数据聚合方式难以满足当下各类应用需求. 为解决这一问题,空中计算(over-the-Air Computation, AirComp)技术应运而生<sup>[1,2]</sup>,它利用多址接入信道的信号叠加特性,在无线信道上进行计算,实现了通信和计算的一体化设计<sup>[2-5]</sup>.

AirComp的一个关键问题是高精度收发器的设计. 文献[6-8]最早研究了信号对齐误差和噪声抑制之间的关系,通过联合优化能量和波束赋形实现了AirComp误差最小化. 在此基础上文献[9]提出一种具有多天线集群传感器的AirComp结构,实现了多输入多输出的IoT传输. 为了提高AirComp的效率,数据融合的中心往往采用静态基站(Base Station, BS)<sup>[10-16]</sup>,但考虑到实际场景中的传感器通常被部署于可移动设备上,可能会随着移动而超出BS的有效覆盖范围<sup>[17,18]</sup>,因此,采用更灵活的BS部署是保证AirComp优势的关键.

无人机(Unmanned Aerial Vehicle, UAV)可作为移动BS进行地面设备和UAV之间的通信<sup>[19]</sup>. 现有的研究主要集中于单UAV或单优化目标方案,忽略了单UAV辅助的AirComp网络在区域覆盖和资源方面的限制. 文献[20]提出了利用单UAV作为移动BS的方案,通过最小化AirComp的时均MSE(Mean-Squared Error)来提高系统性能. 文献[20~22]验证了单UAV辅助AirComp网络的价值,但都未考虑多UAV的效果,这限制了AirComp网络的性能和应用扩展. 而单优化目标MSE追求全局聚合质量最优,会使得一些设备的数据质量较差. 为解决以上问题,一些研究者提出多UAV<sup>[23,24]</sup>辅助多地面移动传感器(Ground Mobile Sensor, GMS)进行数据聚合的方案. 目前多UAV辅助AirComp网络的研究已有一些进展,但仍存在较多的局限性,例如UAV轨迹未考虑高度变化.

基于上述讨论,本文将重点关注符合实际通信需求的AirComp网络,利用多GMS和多UAV的协作来提升系统性能,并将UAV的二维部署拓展到三维(Three

Dimensional, 3D)部署,研究多UAV在不同高度飞行时的链路状态,充分考虑地面遮挡情况. 针对数据聚合质量问题,在单优化目标MSE基础上研究最大化系统最小可达速率的多约束优化问题,并借助深度确定性策略梯度优化算法(Deep Deterministic Policy Gradient-based optimization algorithm for multi-UAV cooperation in AirComp network, AirDDPG-UAV)应对复杂通信环境和高维非线性问题,以提高多GMS的聚合效率. 本文提出的多UAV辅助AirComp网络具有较强的可扩展性和适应性,可以为更广阔的研究提供有价值的启示.

## 2 系统模型与问题公式化

为了解决实际应用场景中多GMS数据聚合缓慢的问题,本文提出一种包含多UAV和多GMS的AirComp网络,如图1所示.

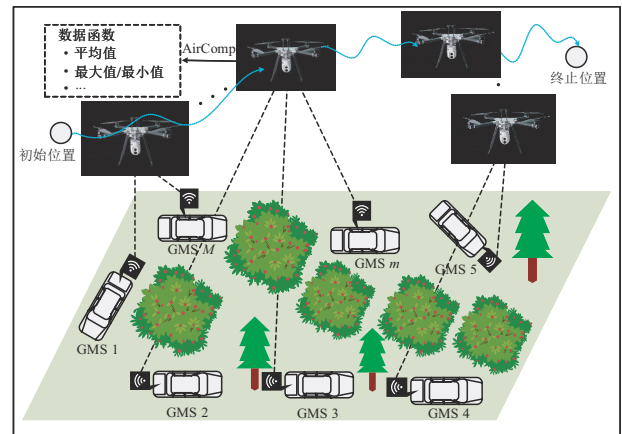


图1 基于多UAV和多GMS的AirComp网络

在该网络中,每架UAV充当部分GMS数据融合的中心. 其中每架UAV $p$ 均有特定编号 $p \in H = \{1, 2, \dots, P\}$ . 类似的,每个GMS $m$ 符合 $m \in J = \{1, 2, \dots, M\}$ . 考虑到设备尺寸和功率的限制,每架UAV和每个GMS均配备单个天线. 此外,为了加快数据聚合速度,每个GMS以预设的移动速度收集数据. 基于上述描述,本文提出的AirComp网络可以在BS缺失的情况下,实现快速而可靠的数据收集.

## 2.1 UAV 移动模型

本文在模型中设计了多 UAV 轨迹,以便在多 GMS 连续移动的情况下根据它们的位置进行优化. 为确保数据在传输过程中的连续性和稳定性,本文采用了时间离散技术对多 UAV 轨迹进行连续设计<sup>[25]</sup>. 将时间周期  $T$  均分为  $N$  份,每个时隙  $n, n \in I = \{1, 2, \dots, N\}$  的时长为  $\delta_t$ , 满足  $T = N\delta_t$ . 基于三维笛卡尔坐标系, UAV  $p$  在时隙  $n$  的位置坐标表示为

$$L_p^{\text{UAV}}(n) = \left\{ (L_{p_x}^{\text{UAV}}(n), L_{p_z}^{\text{UAV}}(n)) \right\}, \forall p \in H, \forall n \in N \quad (1)$$

其中,  $L_{p_x}^{\text{UAV}}(n)$  表示 UAV  $p$  在时隙  $n$  的水平位置,  $L_{p_z}^{\text{UAV}}(n)$  表示 UAV  $p$  在时隙  $n$  的飞行高度.

## 2.2 信道模型

在现实场景中, GMS  $m$  和 UAV  $p$  之间可能是视距连接信道也可能是非视距连接信道. 在第  $n$  个时隙中, GMS  $m$  和 UAV  $p$  之间的视距和非视距连接信道的概率为

$$p_{\text{LoS}}^{m,p}(n) = \frac{1}{1 + a \exp(-b(\theta_{m,p}(n) - a))} \quad (2)$$

$$p_{\text{NLoS}}^{m,p}(n) = 1 - p_{\text{LoS}}^{m,p}(n) \quad (3)$$

其中,  $\theta_{m,p}(n)$  为 GMS  $m$  和 UAV  $p$  之间的仰角<sup>[26]</sup>,  $a$  和  $b$  是取决于不同环境的参数<sup>[27]</sup>. 在时隙  $n$ , GMS  $m$  至 UAV  $p$  的信道变化可表示为

$$\bar{g}_{m,p}(n) = p_{\text{LoS}}^{m,p}(n) g_{\text{LoS}}^{m,p}(n) + p_{\text{NLoS}}^{m,p}(n) g_{\text{NLoS}}^{m,p}(n) \quad (4)$$

其中,  $g_{\text{LoS}}^{m,p}(n)$  和  $g_{\text{NLoS}}^{m,p}(n)$  分别是 GMS  $m$  和 UAV  $p$  之间的视距和非视距路径损耗.

## 2.3 基于 AirComp 的数据聚合

每架 UAV 的目标是计算来自多 GMS 聚合数据的目标函数, 其中, GMS  $m$  在时隙  $n$  收集到的数据用  $z_m(n)$  表示. 因此, 在时隙  $n$ , UAV  $p$  计算的目标函数可表示为

$$\begin{aligned} & f_p(n)(z_1(n), \dots, z_M(n)) \\ &= \phi_p \left( \sum_{m=1}^M a_m^p(n) \psi_m(z_m(n)) \right) \quad (5) \\ & , \forall m \in J, p \in H, n \in I \end{aligned}$$

其中,  $\phi_p$  表示 UAV  $p$  的后处理函数.  $a_m^p(n)$  表示 GMS  $m$  在时隙  $n$  的决策变量, 满足以下约束:

$$a_m^p(n) \in \{0, 1\}, \forall m \in J, p \in H, n \in I \quad (6)$$

$$\sum_{p=1}^P a_m^p(n) = 1, \forall m \in J, n \in I \quad (7)$$

$$\sum_{m=1}^M a_m^p(n) \geq 1, \forall p \in H, n \in I \quad (8)$$

其中, 式(6)表示 GMS  $m$  在时隙  $n$  的决策变量; 式(7)表示 GMS  $m$  的计算任务只能于一架 UAV 处进行处理; 式(8)表示每架 UAV 可以处理来自多个 GMS 的计算任

务.  $\psi_m$  为 GMS  $m$  的预处理函数, 本文定义  $s_m(n) = \psi_m(z_m(n))$  表示 GMS  $m$  在时隙  $n$  传输的信号.

与文献[12, 20]类似, 考虑 UAV  $p$  对多个 GMS 收集的数据取平均. 则 UAV  $p$  获取的函数可表示为

$$f_p(n)(z_1(n), \dots, z_M(n)) = \frac{1}{\text{Num}_p(n)} \sum_{m=1}^M a_m^p(n) s_m(n) \quad (9)$$

其中,  $\frac{1}{\text{Num}_p(n)}$  表示 UAV  $p$  的后处理函数.

本文进一步考虑各个 GMS 的平均发射功率, 满足以下约束:

$$\sum_{N=1}^N p_{m,p}(n) \leq \bar{P}_m, \forall p \in H, m \in J \quad (10)$$

其中,  $p_{m,p}(n)$  表示 GMS  $m$  在时隙  $n$  的传输功率,  $\bar{P}_m$  表示平均传输功率. 为了保证式(10)的有效性, 本文设置  $\bar{P}_m \leq P_m$ ,  $P_m$  为 GMS  $m$  的最大传输功率. 在此基础上, 本文引入去噪因子  $\eta_p(n)$  用于缩放 UAV  $p$  在时隙  $n$  接收到的信号  $y_p(n)$ <sup>[13, 20]</sup>, 缩放后的信号表达式为

$$\hat{f}_p(n) = \frac{y_p(n)}{\text{Num}_p(n) \sqrt{\eta_p(n)}}, \forall p \in H, n \in I \quad (11)$$

## 2.4 性能指标

在时隙  $n$ , UAV  $p$  对应的 MSE 计算为

$$\text{MSE}_p(n) = E \left[ \left| \hat{f}_p(n) - f_p(n) \right|^2 \right] \quad (12)$$

因此, UAV  $p$  在噪声干扰下可达到的数据聚合速率可定义为

$$R_p(n) = B \log_2 \left( 1 + \frac{E[|\hat{f}_p(n)|^2] - E[|f_p(n)|^2]}{E[|\hat{f}_p(n) - f_p(n)|^2]} \right) \quad (13)$$

其中,  $E[|\hat{f}_p(n)|^2]$ <sup>[13, 20]</sup> 表示信号重构后 UAV  $p$  接收到的总功率.

根据式(13)的数据聚合速率表达式,  $N$  个时隙 UAV  $p$  的平均速率可定义为

$$R_p = \frac{1}{N} \sum_{n=1}^N R_p(n), \forall p \in H \quad (14)$$

## 2.5 问题构建

本文的研究目标是在考虑实际通信需求的 AirComp 网络环境下最大化系统最小可达速率, 重点研究 UAV 和 GMS 之间的关联  $A = \{a_m^p(n), \forall m \in J, p \in H, n \in I\}$ , 3D 部署 UAV 轨迹  $L = \{L_p^{\text{UAV}}(n), \mathbf{v}_{p_x}(n), \mathbf{a}_{p_x}(n), \mathbf{v}_{p_z}(n), \mathbf{a}_{p_z}(n), \forall p \in H, n \in I\}$ , UAV 去噪因子  $\eta = \{\eta_p(n), \forall p \in H, n \in I\}$  和传输功率  $P = \{p_{m,p}(n), \forall m \in J, p \in H, n \in I\}$ . 基于上述设置, 本文的研究问题描述为 Problem 1:

Problem 1:  $\max_{A,L,\eta,P} \phi$

s.t.  $C_1: 0 \leq p_{m,p}(n) \leq P_m, \forall m \in J, p \in H$

$C_2: 0 \leq \frac{1}{N} \sum_{n=1}^N p_{m,p}(n) \leq \bar{P}_m, \forall m \in J, p \in H$

$C_3: \bar{P}_m \leq P_m, \forall m \in J$

$C_4: \eta_p(n) \geq 0, \forall p \in H, \forall n \in I$

$C_5: L_{p_{xy}}^{\text{UAV}}(n+1) = L_{p_{xy}}^{\text{UAV}}(n) + v_{p_{xy}}(n)\delta_t + \frac{1}{2} a_{p_{xy}}(n)\delta_t^2,$   
 $\forall p \in H, n \in I$

$C_6: |v_{p_{xy}}(n)| \leq v_{p_{xy}}^{\max}, \forall p \in H, n \in I$

$C_7: |a_{p_{xy}}(n)| \leq a_{p_{xy}}^{\max}, \forall p \in H, n \in I$

$C_8: L_p^{\text{UAV}}(0) = L_{p,1}^{\text{UAV}}, L_p^{\text{UAV}}(N) = L_{p,F}^{\text{UAV}}, \forall p \in H$

$C_9: \|L_p^{\text{UAV}}(n) - L_q^{\text{UAV}}(n)\|^2 \geq d_{\min}^2, \forall p, q \in H, p \neq q, \forall n \in I$

$C_{10}: L_{p_z}^{\text{UAV}}(n+1) = L_{p_z}^{\text{UAV}}(n) + v_{p_z}(n)\delta_t + \frac{1}{2} a_{p_z}(n)\delta_t^2,$   
 $\forall p \in H, n \in I$

$C_{11}: |v_{p_z}(n)| \leq v_{p_z}^{\max}, \forall p \in H, n \in I$

$C_{12}: |a_{p_z}(n)| \leq a_{p_z}^{\max}, \forall p \in H, n \in I$

$C_{13}$ : 式(6)~式(8)

(15)

其中,  $\phi$  的描述如下:

$$\phi(A, L, \eta, P) = \min_{p \in H} R_p \quad (16)$$

约束  $C_1$  表示 GMSm 与 UAVp 之间传输功率的最大值和最小值, 约束  $C_2$  表示各个 GMS 平均发射功率的范围, 约束  $C_3$  保证式(10)的有效性, 约束  $C_4$  表示 UAV 去噪因子为正值, 约束  $C_5$  表示 UAVp 横向位置、速度和加速度间的关系, 约束  $C_6$  和  $C_7$  表示 UAVp 横向飞行速度和加速度的最大值, 约束  $C_8$  表示 UAVp 的初始位置和终止位置, 约束  $C_9$  表示不同 UAV 之间避免碰撞的最小距离, 约束  $C_{10}$  表示 UAVp 纵向位置、速度和加速度间的关系, 约束  $C_{11}$  和  $C_{12}$  表示 UAVp 纵向飞行速度和加速度的最大值, 约束  $C_{13}$  表示 GMSm 在时隙  $n$  的决策变量限制。

### 3 基于 AirDDPG-UAV 算法的 UAV 轨迹优化和计算资源分配策略

优化问题 Problem 1 具有混合整数非线性特征, 为了获得 Problem 1 的最优解, 同时加快求解速度, 本文提出一种 AirComp 网络下多 UAV 辅助的深度确定性策略梯度优化算法 AirDDPG-UAV. 该算法能够处理连续状态和动作空间, 适用于动态变化的场景. 并且该算法计算复杂度低、训练时间短, 是解决优化问题 Problem 1 的理想选择。

#### 3.1 AirDDPG-UAV 算法要素定义

AirDDPG-UAV 算法基于深度强化学习, 利用智能体和环境之间的相互作用, 通过不断学习得到最优策略. 本文将采集数据的多 GMS 和处理数据的多 UAV 视为智能体, 通过最大化未来潜在奖励来优化系统性能. AirDDPG-UAV 算法的三个关键要素为状态、动作和奖励, 具体定义如下:

状态空间定义:

$$S = \{L_m^{\text{GMS}}(n), L_p^{\text{UAV}}(n), z_m(n), \forall m \in J, n \in I, p \in H\} \quad (17)$$

其中,  $L_m^{\text{GMS}}(n)$  和  $L_p^{\text{UAV}}(n)$  分别表示 GMSm 和 UAVp 在时隙  $n$  的实时位置,  $z_m(n)$  表示 GMSm 在时隙  $n$  采集到的数据。

动作空间定义:

$$A = \{a_m^p(n), v_{p_{xy}}(n), v_{p_z}(n), a_{p_{xy}}(n), a_{p_z}(n), \eta_p(n), p_{m,p}(n), m \in J, n \in I, p \in H\} \quad (18)$$

其中, 变量  $a_m^p(n)$  表示 GMSm 和 UAVp 在时隙  $n$  的关联, 将其转换为连续空间, 即  $0 \leq a_m^p(n) \leq 1$ ,  $v_{p_{xy}}(n)$  和  $v_{p_z}(n)$  分别表示 UAVp 在时隙  $n$  的横、纵飞行速度,  $a_{p_{xy}}(n)$  和  $a_{p_z}(n)$  分别表示 UAVp 在时隙  $n$  的横、纵加速度,  $\eta_p(n)$  表示 UAVp 在时隙  $n$  的去噪因子,  $p_{m,p}(n)$  表示 GMSm 和 UAVp 之间在时隙  $n$  的传输功率。

奖励函数定义: 奖励函数与目标函数相关联, 本文的即时奖励为

$$r(n+1) = \phi - \zeta \quad (19)$$

其中,  $\zeta$  表示惩罚项。

#### 3.2 AirDDPG-UAV 算法

由于问题存在非凸性, 为降低求解复杂度, 优化问题 Problem 1 可重构为 Problem 2:

$$\text{Problem 2: } \arg \max_{\pi^*} \sum_{n=0}^{N-1} r(n+1) \quad (20)$$

s.t.  $C_{14}: 0 \leq a_m^p(n) \leq 1, \forall m \in J, p \in H, n \in I$

$C_{15}: C_1 \sim C_{12}$ , 式(7), 式(8)

其中,  $\pi^*$  表示策略集  $\{A, L, \eta, P\}$  的最优策略。

图 2 展示了 AirDDPG-UAV 算法框架, 通过智能体与环境间的不断交互实现策略优化. 该算法包括两组双深度 Q 网络 (Double Deep Q-Network, DDQN), 即参数分别设置为  $\theta^a$ 、 $\theta^q$  的 Actor 网络和 Critic 网络, 以及参数分别设置为  $\theta^{a'}$ 、 $\theta^{q'}$  的 Actor 目标网络和 Critic 目标网络。

Actor 网络能够根据环境状态产生一个动作, 从而探索状态空间, 以此学习一个有效的策略函数. AirDDPG-UAV 算法用确定性策略  $\mu(s(n)|\theta^a)$  生成当前状态  $s(n)$  下的动作  $a(n)$ , 具体表示为

$$a(n)=\mu(s(n)|\theta^\mu) \quad (21)$$

Actor网络使整体折现收益的期望值最大化,并应用 $Q$ 函数 $Q(s(n),a(n)|\theta^Q)$ 近似该总折现收益.因此,在策略 $\mu$ 的指导下,性能目标函数可表示为

$$J(\theta^\mu)=E_{\rho^\mu}[Q(s(n),\mu(s(n)|\theta^\mu)|\theta^Q)] \quad (22)$$

为实现逼近最优解的目标,Actor网络采用梯度上升法更新策略,可表示为

$$\nabla_{\theta^\mu} J = \frac{1}{W} \sum_{w=1}^W \nabla_a Q(s(w),a(w)|\theta^Q) \nabla_{\theta^\mu} \mu(s(w)|\theta^\mu) \quad (23)$$

其中, $W$ 表示抽取样本的数量.

此外,Actor目标网络根据下一个状态 $s(n+1)$ 输出下一个动作 $a(n+1)$ ,并定期将参数 $\theta^\mu$ 应用软更新的方式进行更新,如下所示:

$$\theta^\mu \leftarrow \tau\theta^\mu + (1-\tau)\theta^{\mu'} \quad (24)$$

其中, $\tau$ 表示软更新因子.

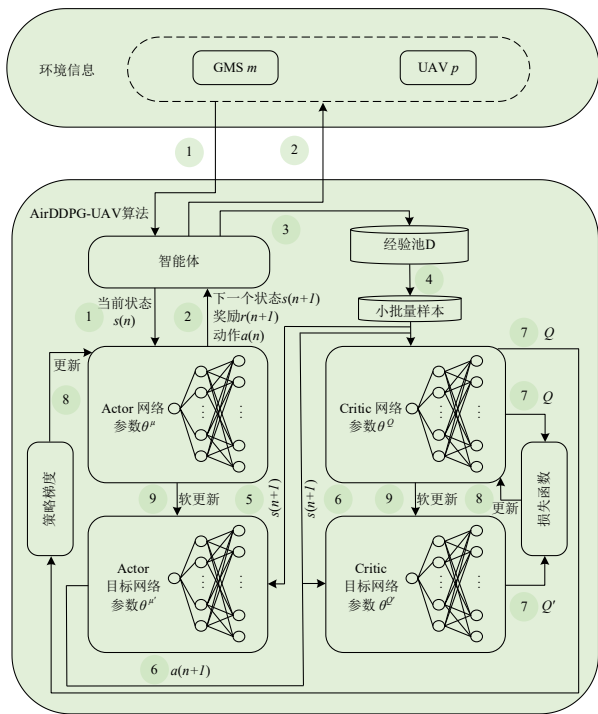


图2 基于 AirDDPG-UAV 算法的分配策略

Critic网络能够学习一个值函数,用以预测在某个状态下采取某个动作的预期累积奖励值. AirDDPG-UAV算法针对当前状态对动作进行评估,使用参数为 $\theta^Q$ 的神经网络来模拟目标 $Q$ 函数的值.该目标 $Q$ 函数定义为

$$Q'(n)=r(n+1)+\gamma Q'(n+1, \mu'(s(n+1)|\theta^{\mu'})|\theta^Q) \quad (25)$$

其中, $\gamma \in [0,1]$ 表示折扣因子.同时,Critic网络的目标是使损失最小化,该损失可定义为

$$L = \frac{1}{W} \sum_{w=1}^W (Q'(w) - Q(s(w), a(w)|\theta^Q))^2 \quad (26)$$

此外,Critic目标网络根据下一个状态 $s(n+1)$ 和下一个动作 $a(n+1)$ 计算目标 $Q$ 函数中的部分数值 $Q'(s(n+1),a(n+1))$ ,并定期将参数 $\theta^Q$ 以软更新的方式进行更新,可表示为

$$\theta^Q \leftarrow \tau\theta^Q + (1-\tau)\theta^{Q'} \quad (27)$$

### 3.3 收敛性和计算复杂度分析

本节主要探讨 AirDDPG-UAV 算法的收敛性和计算复杂度.

我们首先对 AirDDPG-UAV 算法在不同学习率下的收敛性进行了比较,结果如图3所示.该算法的平均奖励会随着迭代次数的增加而不断增加,最终趋于稳定.

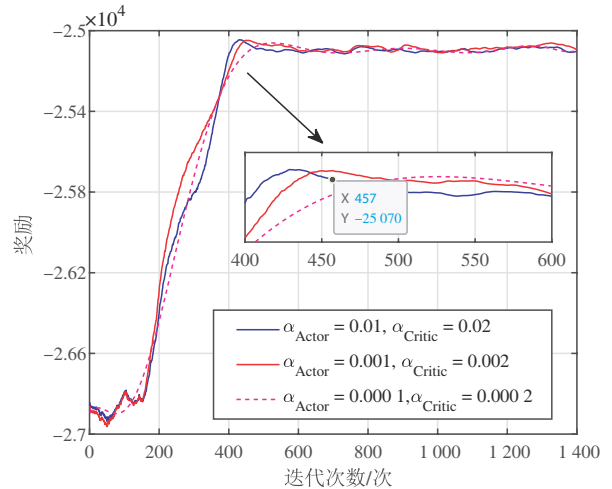


图3 不同学习率下 AirDDPG-UAV 算法的收敛性

神经网络的计算复杂度受多种因素影响,本文重点关注生成最优动作所需的计算复杂度.在每次迭代中,AirDDPG-UAV算法中的智能体会遍历所有动作,以找到具有最大 $Q$ 函数值的最优动作.假设 $l$ 表示网络层数, $n_i$ 和 $n_{i+1}$ 分别表示第 $i$ 层和 $i+1$ 层的神经元数,Number $_c$ 表示迭代次数,则生成动作的计算复杂度可表示为

$$O\left(\text{Number}_c N \sum_{i=0}^l n_i n_{i+1}\right).$$

## 4 实验结果与分析

### 4.1 实验环境

为了验证 AirDDPG-UAV 算法的有效性,本节进行数值仿真.使用 Python 3.7 和 TensorFlow 框架,同时设置 GMS 数量  $M=10$ , UAV 数量  $P=2$ .表1列举了其他重要参数.此外,本文应用以下三个基准方案来检验 AirDDPG-UAV 算法的优势.

表 1 仿真参数表

参数	取值
时隙数量 $N$	1 000
GMS 数量 $M$	10
系统带宽 $B$	100 MHz
软更新因子 $\tau$	0.000 1
迭代次数 $\text{Number}_e$	1 400
小批量样本大小 $W$	500
加性高斯白噪声 $\sigma^2$	-80 dBm
奖励折扣因子 $\gamma$	0.95
随机噪声 $\zeta$	0.001
Actor 网络学习率 $\alpha_{\text{Actor}}$	0.000 1
Critic 网络学习率 $\alpha_{\text{Critic}}$	0.000 2
信道功率增益 $\beta_0$	-40 dB
UAV 纵向最大飞行速度 $v_{p_z}^{\max}$	30 m/s
UAV 横向最大飞行速度 $v_{p_y}^{\max}$	30 m/s
UAV 纵向最大飞行加速度 $a_{p_z}^{\max}$	4 m/s <sup>2</sup>
UAV 横向最大飞行加速度 $a_{p_y}^{\max}$	4 m/s <sup>2</sup>

方案 1: UAV-GMS 关联、UAV 3D 部署、UAV 去噪因子以及传输功率分配变量均由演员-评论家算法优化。

方案 2: UAV-GMS 关联、UAV 去噪因子以及传输功率分配变量均由深度确定性策略梯度算法优化,而所有 UAV 均悬停在预设位置。

方案 3: UAV-GMS 关联、UAV 3D 部署以及 UAV 去噪因子变量均由深度确定性策略梯度算法优化,而所有 GMS 的传输功率统一设置为  $p_{m,p}(n) = \bar{P}_m, \forall p \in H, n \in I$ 。

#### 4.2 对比实验

由图 4 可见,本文所提方案最佳。方案 1 收敛困难,影响了系统的性能,而 AirDDPG-UAV 算法利用经验回放技术,能够提高强化学习训练的效率和稳定性。方案 2 在系统最小传输速率方面明显不如本文算法,这是因为静止状态的 UAV 无法适应动态环境,使得传输速率和数据聚合效率受到影响。而 AirDDPG-UAV 算法利用 UAV 3D 部署能够动态调整 UAV 的位置和速度,可以克服遮挡或复杂地形的影响,最终提高了系统的最小传输速率。方案 3 在提高系统最小传输速率方面表现最差,因其没有进行功率优化。方案 3 会使得多 UAV 的电池更快耗尽而将无法完成任务,从而影响系统稳定性。

值得注意的是,当数据量由 10 Mbps 增至 20 Mbps 时,方案 2 和方案 3 下的系统最小传输速率出现急剧下降的现象。这是由于在初始数据量增大时,这些方案下多 UAV 的位置和传输功率并未做出相应调整,导致信号质量下降,从而使得系统的最小传输速率骤降。相比之下, AirDDPG-UAV 算法采用了更智能的策略,能够根据数据量的增大自适应地学习和调整,从而始终保持

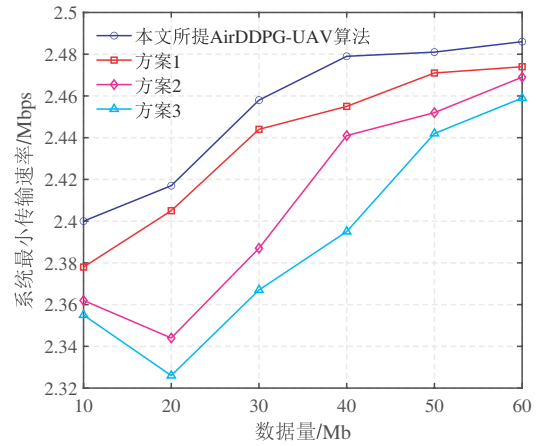
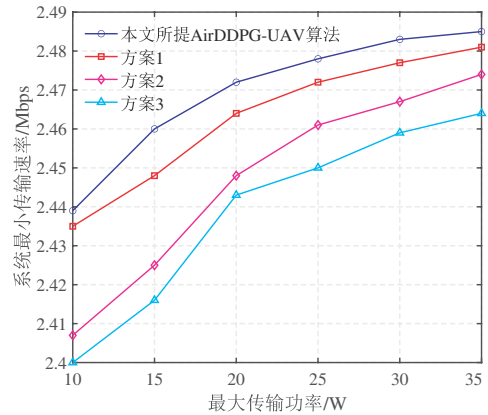


图 4 不同数据量下的系统最小传输速率对比

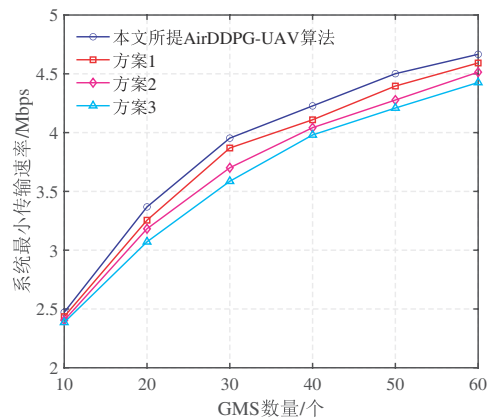
系统最小传输速率增长的趋势。

在图 5(a)中,我们展示了系统最小传输速率随着最大传输功率变化而变化的情况。图 5(b)给出了数据量为 30 Mbps 时,各方案在不同数量 GMS 下的系统最小传输速率情况。

从图 5(a)中可以观察到,与其他三种方案相比,本



(a) 不同最大传输功率下的系统最小传输速率对比



(b) 不同 GMS 数量下的系统最小传输速率对比

图 5 最大传输功率和 GMS 数量对系统最小传输速率的影响

文所提 AirDDPG-UAV 算法在提高系统最小传输速率方面表现最佳. 这是因为 AirDDPG-UAV 算法采用了确定性策略方案, 同时优化了 UAV 轨迹和计算资源, 能够确保复杂环境中的传输可靠性, 更好地控制当前状态下的传输功率, 进而提高系统性能. 在图 5(a) 中还可以看到, 随着最大传输功率从 10 W 增加到 15 W, 本文提出的 AirDDPG-UAV 算法在系统最小传输速率方面显著提升. 这是由于传输功率的增加改变了信道状态, 提高了信噪比和信道容量, 从而使得传输速率迅速增长.

在图 5(b) 中, 方案 3 具有最低的系统最小传输速率, 这是由于在全功率传输的情况下, 信道容易发生过载和干扰. 而方案 2 的最小传输速率相对较低, 原因在于方案

2 中 UAV 的位置一直保持静止, 无法获取更多的传感数据. 方案 1 采用与本文算法类似的联合优化措施, 故其最小传输速率能够达到较高的水平, 但由于方案 1 存在收敛困难的情况, 系统的性能受到一定的影响. 随着 GMS 数量的增加, 本文所提 AirDDPG-UAV 算法与其他方案之间的性能差距逐步增大, 进一步证明了本文算法的优越性, 其能够通过优化 UAV 的轨迹和计算资源, 在复杂环境下确保信号传输的可靠性, 从而提高系统的性能.

在图 6(a) 中, 我们分析了时间周期对系统最小传输速率的影响. 为了定量地分析 AirDDPG-UAV 算法在数据聚合方面的优势, 我们开展了不同时间周期下缩放信号和原信号的时均 MSE 统计实验, 结果见图 6(b).

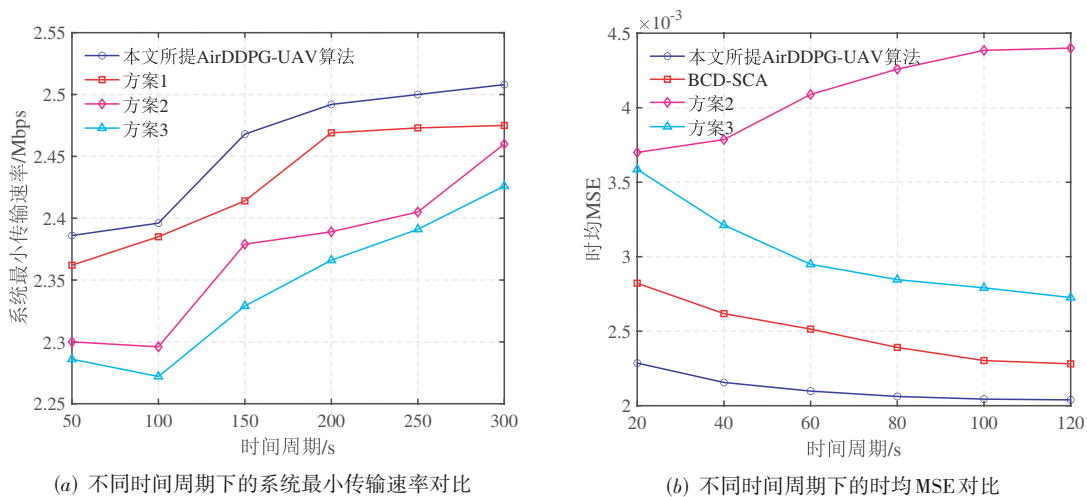


图6 几种方案在不同时间周期下的系统最小传输速率和时均MSE对比

从图 6(a) 中可以看出, 方案 3 的系统最小传输速率最低. 这主要是由于全功率传输会导致信号失真的问题, 从而降低系统的整体性能. 方案 2 采用固定 UAV 位置的做法限制了多 UAV 的移动性. 方案 1 的系统最小传输速率略低于本文所提 AirDDPG-UAV 算法, 方案 1 虽将 UAV 的移动性和信道干扰等因素考虑在内, 但在探索策略时可能会陷入局部最优解. 此外, 另外, 我们发现在时间周期从 50 s 增加到 100 s 的过程中, 方案 2 和方案 3 下的系统最小传输速率出现了突然下降的趋势. 相比之下, AirDDPG-UAV 算法和方案 1 能够不断调整策略, 从而始终保持系统最小传输速率增长的趋势.

从图 6(b) 的观察结果可以得知时间周期对时均 MSE 产生的影响. 方案 2 的时均 MSE 随着时间周期增加呈现上升趋势, 这是由于随着时间增长, GMS 会移动得更远, 另外遮挡也会导致信道状态变差. 与之相对, 其余方案的时均 MSE 随着时间周期的增加而不断降低. 其中方案 3 的时均 MSE 虽然呈下降趋势, 但依然高于 AirDDPG-UAV 算法和 BCD-SCA 算法<sup>[20]</sup>, 其原因在上

文已经解释, 在此不做赘述. 方案 3 的表现也说明了联合优化措施在降低时均 MSE 方面的重要意义. AirDDPG-UAV 算法的时均 MSE 性能在所有算法中最优, 与 BCD-SCA 算法相比, AirDDPG-UAV 算法能获得每个子问题的最优解, 而 BCD-SCA 算法只能获得多 UAV 轨迹问题的近似最优解.

综上所述, 本文所提 AirDDPG-UAV 算法不仅能够优化系统最小传输速率, 在优化时均 MSE 方面也展现出优越的性能, 其通过多 UAV 协作和数据聚合, 有效提高了系统的吞吐量, 从而确保了网络整体性能稳定.

## 5 结论

本文提出了一种多 UAV 辅助的 AirComp 网络, 旨在解决分布式数据聚合速度缓慢的问题. 该网络以系统最小可达速率最大化作为优化目标, 并对 UAV-GMS 关联、UAV 3D 部署、UAV 去噪因子和传输功率分配等因素进行了联合优化, 实现了多 GMS 采集数据的高效聚合. 为应对联合优化问题中的非凸性挑战, 本文提出

了一种基于深度强化学习的 AirDDPG-UAV 算法. 数值结果显示,相较于对比方案, AirDDPG-UAV 算法的系统最小可达速率显著提高,同时其具有较低的能耗和计算复杂度,适合于快速训练和部署. 接下来将在本文研究基础上考虑更加复杂的信道环境,进一步探究多 UAV 辅助 AirComp 网络的策略优化和性能提升,如应用智能反射面进行信号反射,以提升信号质量并避免遮挡问题.

#### 参考文献

- [1] NAZER B, GASTPAR M. Computation over multiple-access channels[J]. *IEEE Transactions on Information Theory*, 2007, 53(10): 3498-3516.
- [2] ZHU G X, XU J E, HUANG K B, et al. Over-the-air computing for wireless data aggregation in massive IoT[J]. *IEEE Wireless Communications*, 2021, 28(4): 57-65.
- [3] LI Y Q, JIANG M A, ZHANG G C, et al. Joint optimization for multi-antenna AF-relay aided over-the-air computation[J]. *IEEE Transactions on Vehicular Technology*, 2022, 71(6): 6744-6749.
- [4] NI W L, LIU Y W, YANG Z H, et al. Over-the-air federated learning and non-orthogonal multiple access unified by reconfigurable intelligent surface[C]//*IEEE INFOCOM 2021 - IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS)*. Piscataway: IEEE, 2021: 1-6.
- [5] MOLINARI F, AGRAWAL N, STANCZAK S, et al. Max-consensus over fading wireless channels[J]. *IEEE Transactions on Control of Network Systems*, 2021, 8(2): 791-802.
- [6] LIU W C, ZANG X, LI Y H, et al. Over-the-air computation systems: Optimization, analysis and scaling laws[J]. *IEEE Transactions on Wireless Communications*, 2020, 19(8): 5488-5502.
- [7] CAO X W, ZHU G X, XU J E, et al. Optimized power control for over-the-air computation in fading channels[J]. *IEEE Transactions on Wireless Communications*, 2020, 19(11): 7498-7513.
- [8] LI X Y, ZHU G X, GONG Y, et al. Wirelessly powered data aggregation for IoT via over-the-air function computation: Beamforming and power control[J]. *IEEE Transactions on Wireless Communications*, 2019, 18(7): 3437-3452.
- [9] WEN D Z, ZHU G X, HUANG K B. Reduced-dimension design of MIMO AirComp for data aggregation in clustered IoT networks[C]//*2019 IEEE Global Communications Conference (GLOBECOM)*. Piscataway: IEEE, 2019: 1-6.
- [10] FU M, ZHOU Y, SHI Y M, et al. UAV aided over-the-air computation[J]. *IEEE Transactions on Wireless Communications*, 2022, 21(7): 4909-4924.
- [11] MA X A, SUN H J, WANG Q, et al. User scheduling for federated learning through over-the-air computation[C]//*2021 IEEE 94th Vehicular Technology Conference (VTC2021-Fall)*. Piscataway: IEEE, 2021: 1-5.
- [12] NI W L, LIU Y W, YANG Z H, et al. Federated learning in multi-RIS-aided systems[J]. *IEEE Internet of Things Journal*, 2022, 9(12): 9608-9624.
- [13] YANG K, JIANG T, SHI Y M, et al. Federated learning via over-the-air computation[J]. *IEEE Transactions on Wireless Communications*, 2020, 19(3): 2022-2035.
- [14] CHEN L, QIN X W, WEI G. A uniform-forcing transceiver design for over-the-air function computation[J]. *IEEE Wireless Communications Letters*, 2018, 7(6): 942-945.
- [15] HU Y T, CHEN M, CHEN M Z, et al. Energy minimization for federated learning with IRS-assisted over-the-air computation[C]//*ICASSP 2021 - 2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. Piscataway: IEEE, 2021: 3105-3109.
- [16] WANG F, LAU V K N. Multi-level over-the-air aggregation of mobile edge computing over D2D wireless networks[J]. *IEEE Transactions on Wireless Communications*, 2022, 21(10): 8337-8353.
- [17] ZHU G X, HUANG K B. MIMO over-the-air computation for high-mobility multimodal sensing[J]. *IEEE Internet of Things Journal*, 2019, 6(4): 6089-6103.
- [18] 崔玉亚, 张德干, 张婷, 等. 一种面向移动边缘计算的多用户细粒度任务卸载调度方法[J]. *电子学报*, 2021, 49(11): 2202-2207.  
CUI Y Y, ZHANG D G, ZHANG T, et al. A multi-user fine-grained task offloading scheduling approach of mobile edge computing[J]. *Acta Electronica Sinica*, 2021, 49(11): 2202-2207. (in Chinese)
- [19] LU W D, DING Y, GAO Y A, et al. Resource and trajectory optimization for secure communications in dual unmanned aerial vehicle mobile edge computing systems[J]. *IEEE Transactions on Industrial Informatics*, 2022, 18(4): 2704-2713.
- [20] FU M, ZHOU Y, SHI Y M, et al. UAV-assisted over-the-air computation[C]//*ICC 2021 - IEEE International Conference on Communications*. Piscataway: IEEE, 2021: 1-6.
- [21] CAO X W, ZHU G X, XU J E, et al. Cooperative interference management for over-the-air computation networks[J]. *IEEE Transactions on Wireless Communications*,

- 2021, 20(4): 2634-2651.
- [22] ZENG X A, ZHANG X A, WANG F. Optimized UAV trajectory and transceiver design for over-the-air computation systems[J]. IEEE Open Journal of the Computer Society, 2022, 3: 313-322.
- [23] FU M, ZHOU Y, SHI Y M, et al. UAV-assisted multi-cluster over-the-air computation[J]. IEEE Transactions on Wireless Communications, 2023, 22(7): 4668-4682.
- [24] JOUNG J, FAN J C. Over-the-air computation strategy using space-time line code for data collection by multiple unmanned aerial vehicles[J]. IEEE Access, 2021, 9: 105230-105241.
- [25] SUN L, WAN L T, WANG X P. Learning-based resource allocation strategy for industrial IoT in UAV-enabled MEC systems[J]. IEEE Transactions on Industrial Informatics, 2021, 17(7): 5031-5040.
- [26] ZHANG X C, ZHANG J A, XIONG J, et al. Energy-efficient multi-UAV-enabled multiaccess edge computing incorporating NOMA[J]. IEEE Internet of Things Journal, 2020, 7(6): 5613-5627.
- [27] AL-HOURANI A, KANDEEPAN S, LARDNER S. Optimal LAP altitude for maximum coverage[J]. IEEE Wireless Communications Letters, 2014, 3(6): 569-572.



夏景明(通讯作者) 男,1980年生. 现为南京信息工程大学人工智能学院教授、博士生导师. 主要研究方向为气象大数据、物联网应用.  
E-mail: xiajingming@nuist.edu.cn

#### 作者简介



谈玲 女,1979年生. 现为南京信息工程大学计算机学院教授、硕士生导师. 研究方向为数据处理、边缘计算.  
E-mail: cillatan0@nuist.edu.cn



许海 男,1998年生. 现为南京信息工程大学计算机学院硕士生. 主要研究方向为无人机和智能反射面辅助边缘计算. 中国电子学会会员编号:E190016078M.  
E-mail: 202212490302@nuist.edu.cn



刘玉凤 女,1996年生. 2023年毕业于南京信息工程大学软件学院. 主要研究方向为无人机辅助边缘计算.  
E-mail: yufengliu@nuist.edu.cn