



专题：低空物联网与低空安全

面向协同感知的无人蜂群智能资源调度方案

赵鹏程, 李天扬, 冷甦鹏, 熊凯

(电子科技大学信息与通信工程学院, 四川 成都 611731)

摘 要: 随着低空经济的蓬勃发展, 无人机在监测和感知领域得到广泛应用。然而, 无人机有限的机载计算资源制约了感知数据的高效处理。此外, 协同感知中产生的观测区域重叠进一步增加了冗余的计算负担。同时, 无人机网络的高动态拓扑和节点资源的波动性大幅加剧了资源协同的难度。针对上述挑战, 提出了一种面向协同感知的无人蜂群智能资源调度方案, 通过自适应感知模式、分步卸载计算任务和竞价带宽策略, 实现了通信、感知、计算(通感算)异构资源协同互补, 提升协同感知效率。采用基于注意力机制的多智能体强化学习算法求解优化问题, 以增强智能体提取环境关键特征的能力。仿真结果表明, 与基准方案相比, 该方案不仅有效降低了感知任务的执行时间, 还提高了计算资源的利用率。

关键词: 协同感知; 无人蜂群; 资源调度; 多智能体强化学习; 注意力机制

中图分类号: TP393

文献标志码: A

doi: 10.11959/j.issn.1000-0801.2025050

Intelligent resource scheduling scheme for UAV swarm collaborative sensing

ZHAO Pengcheng, LI Tianyang, LENG Supeng, XIONG Kai

School of Information and Communication Engineering, University of Electronic Science and Technology of China, Chengdu 611731, China

Abstract: With the rapid development of the low-altitude economy, unmanned aerial vehicles (UAV) have been widely applied in monitoring and sensing tasks. However, the limited onboard computing resources of UAV constrain the efficient processing of sensing data. Moreover, overlapping observation areas in collaborative sensing introduce additional computational redundancy. Meanwhile, the highly dynamic network topology and fluctuating node resources significantly increase the complexity of resource coordination. To address these challenges, an intelligent resource scheduling scheme for UAV swarm collaborative sensing was proposed. Adaptive sensing mode selection, step-wise computation offloading, and competitive bandwidth allocation were integrated to achieve heterogeneous resource coordination across communication, sensing, and computation (CSC), thereby enhancing collaborative sensing

收稿日期: 2024-12-01; 修回日期: 2025-02-19

通信作者: 冷甦鹏, spleng@uestc.edu.cn

基金项目: 国家自然科学基金青年项目 (No.62201122); 四川省科技计划项目 (No.2024YFHZ0321)

Foundation Items: The National Natural Science Foundation of China Youth Project (No.62201122), Sichuan Science and Technology Program (No.2024YFHZ0321)



efficiency. Furthermore, a multi-agent reinforcement learning (MARL) algorithm with an attention mechanism was employed to solve the optimization problem, enabling agents to extract critical environmental features more effectively. Simulation results demonstrate that, compared with benchmark schemes, the proposed scheme significantly reduces the execution time of sensing tasks while improving computational resource utilization.

Key words: cooperative sensing, UAV swarm, resource scheduling, multi-agent reinforcement learning, attention mechanism

0 引言

近年来,随着低空经济的快速发展,无人机在城市交通监测、农业环境监测和电力巡检等感知任务中得到了广泛应用^[1-3]。借助机载传感器,无人机能够采集地面目标的图像并执行目标识别。然而,单个无人机的感知精度不足,容易产生错误识别结果。同时,由于无人机本身的机载计算芯片的功耗限制,难以在严格的时延约束下完成复杂的感知数据处理^[4-5]。传统方法中,无人机将感知数据传输至云端或边缘服务器处理来生成识别结果^[6-7]。虽然云端或边缘服务器充足的计算资源降低了计算时延,但是远距离通信会带来高传输时延和通信开销,尤其是高分辨率图像的回传。在野外或灾后地区,缺乏服务器为无人蜂群提供计算支持。因此,通过无人蜂群协同感知,提高覆盖范围和感知精度,并集合相邻无人机的计算能力进行协作计算,从而实现系统快速响应^[8-9]。

在基于无人机网络的协作计算中,计算任务可由中心节点集中式调度,其统计全网节点状态并分配计算子任务^[10-12]。然而,随着无人蜂群规模的增大,中心节点在求解任务分配问题时将面临维度爆炸和难以收敛的挑战。此外,集中式调度还可能带来通信瓶颈和中心无人机的单点失效问题。为克服这些限制,许多研究提出了分布式调度策略,每个无人机独立选择卸载目标,将计算子任务一次性分配^[13-14]。然而,这种方法忽略了节点资源的波动和无人机网络高度动态的拓扑结构,导致了资源浪费和额外时延。

在以往与无人蜂群协作计算或计算卸载相关文献中,通信时延通常在总任务时延中占显著地位^[10-11]。合理分配无人蜂群的带宽资源可以缓解瓶颈节点的压力,提高协同效率。此外,无人机需要通过自适应优化感知的时机和方法,在保证对目标高精度感知的同时减少感知冗余^[15-16]。在无人蜂群协同感知任务中,计算、通信与感知资源互相耦合,通过资源调度实现异构资源的协同互补,进一步提高系统性能^[17]。随着网络规模扩大,为避免调度策略求解的维度爆炸问题,分布式调度策略成为无人蜂群资源管理的重要发展方向^[10-12]。此外,多智能体强化学习(multi-agent reinforcement learning, MARL)算法已被广泛应用于无人蜂群协同任务优化,并且注意力机制能够帮助智能体关注关键的环境特征,降低无关信息干扰,从而提高策略收敛效率。

为了应对上述挑战,本文提出了一种面向协同感知的无人蜂群智能资源调度方案。该方案通过智能调度感知任务中无人蜂群的通信、感知与计算资源,实现异构资源的协同互补,完成对目标的高精度感知与系统快速响应。感知数据处理过程中,无人机采用分步卸载的协作计算策略,能够适应节点资源波动及网络拓扑的高度动态变化。同时,采用了嵌入注意力单元的MARL算法求解异构资源联合优化问题。本文主要贡献如下。

(1) 提出了一种基于无人蜂群的动态资源调度框架,适用于高动态环境下的多目标感知任务,为异构资源联合调度的优化奠定了基础。该框架针对高动态网络拓扑和节点资源波动性,各

无人机根据自身状态、感知目标及邻居节点的动态信息、智能决策感知模式选择、带宽竞价和计算任务卸载，实现异构资源的协同互补，提高了资源利用率和感知任务效率。

(2) 面向无人机网络高动态拓扑与资源波动性，设计了一种基于分步卸载的协作计算算法。该算法中，推理模型被分割成多个串行的计算子任务，并从初始节点开始逐跳卸载至各自的邻居节点，直至所有子任务完成。对于多目标感知任务，该算法构建多条并行计算流水线，提高了感知数据推理效率。同时，该算法充分利用了当前邻居节点剩余资源，并避免了拓扑变化引发的额外时延。

(3) 设计了基于行动者-评论家 (actor-critic) 框架的多智能体注意力 (multi-agent critic with attention, MACA) 算法，用于求解异构资源调度的联合优化问题。智能体通过评论家 (critic) 网络评估策略价值，并通过行动者 (actor) 网络输出联合动作。该算法通过引入注意力单元和分裂式输出层，增强了智能体的环境感知能力和策略收敛性能。

1 系统模型

本文研究的面向协同感知的无人蜂群智能资源调度场景如图 1 所示，无人蜂群对下方多个目标进行协同感知，并协作推理感知图像。无人机集合表示为 $\mathcal{M}=\{1,2,\dots,M\}$ 。每架无人机具有感

知、计算与通信能力，并处于同一高度飞行。定义无人机可以使用两种图像推理方法，分别为高精度识别的卷积神经网络 (convolutional neural network, CNN) 和轻量推理支持向量机 (support vector machine, SVM)。

1.1 感知模型

任务区域由边长为 d 的单元格组成，单元格的集合表示为 $\mathcal{G}=\{1,2,\dots,G\}$ 。假设目标随机分布在不同单元格内，并且位于单元格中心。定义无人机 i 对单元格 g 中目标的感知精度^[18-20]为：

$$S_{i,g}(t) = \begin{cases} e^{-\lambda d_{i,g}(t)}, & d_{i,g}(t) \leq R_S \\ 0, & d_{i,g}(t) > R_S \end{cases} \quad (1)$$

其中， $d_{i,g}(t)$ 为在时隙 t 时无人机 i 到单元格 g 中心的距离， R_S 是最远感知距离， λ 是感知性能参数。

无人机 i 的图像推理方法由变量 β_i 表示，当 $\beta_i=0$ 时表示选择 SVM 推理， $\beta_i=1$ 时表示选择 CNN 推理。感知效用取决于无人机的感知精度和推理方法的准确率，定义无人机 i 对单元格 g 中目标的感知效用为：

$$U_{i,g}(t) = [(1-\beta_i)P_S + \beta_i P_C] S_{i,g}(t) \quad (2)$$

其中， P_S 和 P_C 分别为 SVM 与 CNN 两种图像推理方法的准确率。当多个无人机同时感知单元格 g 中的目标时，该目标的总感知效用定义为：

$$U_g(t) = 1 - \exp \left[-\alpha_S \sum_{i=1}^M \varepsilon_i U_{i,g}(t) \right] \quad (3)$$

其中， α_S 为聚合参数， ε_i 表示无人机 i 是否参与感知，其取值为 0 或 1。

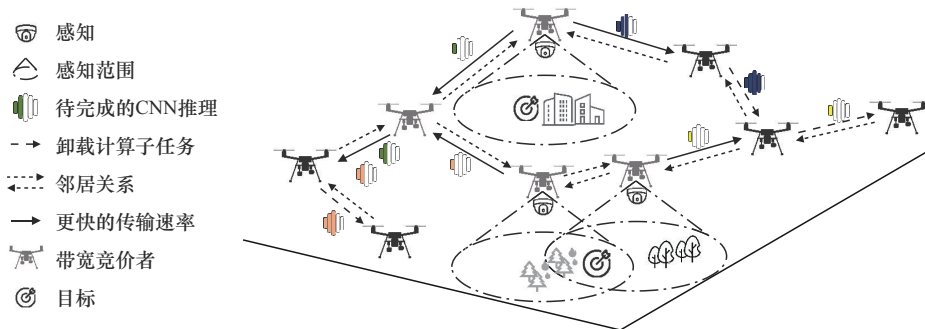


图1 面向协同感知的无人蜂群智能资源调度场景



1.2 计算模型

获得感知图像后，无人机可以选择基于SVM的轻量推理或者基于CNN的深度推理来识别图像。假设无人机使用的是基于径向基核函数的SVM模型^[21]。令 N_S 和 d_S 分别表示支持向量数量和输入特征维度， c_{SVM} 表示核计算常数^[22]，则无人机 i 的SVM推理时延计算式为：

$$\sigma_S^C = \frac{c_{SVM} N_S (2d_S + 2)}{c_i} \quad (4)$$

其中， c_i 为无人机 i 的浮点运算速度。

CNN模型由一系列卷积层、池化层和全连接层构成，其中卷积操作占据主要的计算负担^[23]。因此，将CNN模型以卷积层为单位串行分割成多个计算子任务，即将一个卷积层及其后的池化层或全连接层作为一个计算子任务。定义CNN模型中卷积层的索引集为 $\mathcal{L}=\{1, 2, \dots, L\}$ 。对于卷积层 l ，其计算复杂度由输出通道数 k_{l-1} 、卷积核宽度 w_l 与输出特征图宽度 o_l 共同决定。无人机 i 对卷积层 l 的推理时延为：

$$\sigma_{l,i}^C = \frac{c_{CNN} k_{l-1} w_l^2 o_l^2 k_l}{c_i} \quad (5)$$

其中， c_{CNN} 为卷积计算常数。

1.3 优化模型

本文提出的面向协同感知的智能资源调度方案旨在通过无人蜂群中感知、计算和通信资源的协同互补，在保证对目标高感知效用的同时，最小化感知任务时延。定义多无人机对同一目标感知所形成的图像集合为 $\mathcal{P}=\{1, 2, \dots, P\}$ 。无人机默认在本地执行SVM推理，而CNN推理的每一卷积层可作为计算子任务进行卸载。每架无人机维护计算队列和传输队列，分别用于存储待处理的计算子任务和待传输的计算信息。令符号 μ_l^p 表示负责计算图像 p 的第 l 卷积层的无人机编号。当 $\mu_l^p \neq \mu_{l+1}^p$ 时，无人机 μ_l^p 需将中间计算结果传递至无人机 μ_{l+1}^p ，从而产生时延 $t_{\mu_l^p, \mu_{l+1}^p}^l$ ；当 $\mu_l^p = \mu_{l+1}^p$

时，时延 $t_{\mu_l^p, \mu_{l+1}^p}^l$ 为0。因识别结果的数据量极小，这里忽略识别结果回传的时延。

图像 p 的总处理时延包括计算时延、传输时延和等待时延3部分，表示为 $T_p(t) = C_p(t) + D_p(t) + W_p(t)$ ，其中计算时延为 $C_p(t) = (1 - \beta_i) \sigma_S^C + \beta_i \sum_{l=1}^L \sigma_{l, \mu_l^p}^C$ ，传输时延为 $D_p(t) = \sum_{l=1}^{L-1} t_{\mu_l^p, \mu_{l+1}^p}^l$ ，等待时延 $W_p(t)$ 包含图像 p 所有计算子任务的计算排队时延和传输排队时延。

定义对某一目标的感知任务时延为 $\max_{p \in \mathcal{P}} T_p(t)$ ，即图像集 $\mathcal{P}$ 中图像最大的处理时延。本文以最小化目标感知任务时延为目标，协同调度无人蜂群感知、计算与通信资源，构建了以下优化模型：

$$\begin{aligned} \min_{\rho_i, \phi_i, \varphi_i} \quad & \max_{p \in \mathcal{P}} C_p(t) + D_p(t) + W_p(t) \\ \text{s.t.} \quad & \begin{cases} \text{C1: } \rho_i \in \mathcal{M}_i, i \in M \\ \text{C2: } \phi_i \in \{0, 1, 2\}, i \in M \\ \text{C3: } \varphi_i \in \{0, 1\}, i \in M \\ \text{C4: } T_p(t) < T_0, p \in P \\ \text{C5: } U_g(t) \geq U_0, g \in G \end{cases} \end{aligned} \quad (6)$$

其中， ρ_i 为无人机 i 对当前计算子任务的卸载决策， $\mathcal{M}_i$ 为包含无人机 i 和其邻居的集合， ϕ_i 是无人机 i 的感知决策， φ_i 是无人机 i 的带宽竞价决策， T_0 和 U_0 分别是感知任务时延和感知效用的阈值。约束C1~C3定义了变量的取值范围，C4保证每张图像在约束延迟内完成处理，C5要求对目标的感知效用大于规定阈值。由于C1~C3为整数约束，该优化问题属于NP-hard问题。

2 基于MARL的智能资源协同调度

为求解上述优化问题，本文提出了MACA算法，其基于actor-critic网络模型，并嵌入了注意力单元^[24]。该算法中，每架无人机作为独立智能体，根据自身观察做出感知模式、计算卸载与带宽竞价的联合决策，实现无人蜂群异构资源的协

同互补, 提高协同感知效率。在 MACA 算法中, 智能体利用 critic 网络对目标和其他智能体进行建模, 降低状态交换产生的通信开销, 并通过分裂输出层的 actor 网络生成联合动作。在大规模无人蜂群中, 学习模型难以处理高维输入数据。本文引入注意力机制, 使智能体能够关注对其决策影响更大的同伴, 增强模型收敛性能。

2.1 马尔可夫决策过程

在协同感知任务中, 将无人机蜂群的异构智能资源调度建模为部分可观测马尔可夫决策过程 (partially observable Markov decision process, POMDP), 定义为元组 $\mathcal{D}=\{\mathcal{S}, \mathcal{A}, \mathcal{T}, \mathcal{R}, \Omega, \mathcal{O}, \gamma\}$, 其中 $\mathcal{S}$ 、 $\mathcal{A}$ 和 $\mathcal{R}$ 分别是系统状态空间、动作空间和奖励函数, $\mathcal{T}$ 和 Ω 分别是状态转移函数和观测函数, $\mathcal{O}$ 是系统观察空间, γ 是未来奖励的折扣因子。每个时隙中, 智能体 i 从当前系统状态 $s_t \in \mathcal{S}$ 更新自身观察 $o_t^i \in \mathcal{O}$, 观察包含感知范围内目标状态与其他智能体状态。智能体依据自身策略选择感知、卸载与带宽竞价的联合动作 a_t^i ($a_t^i \in \mathcal{A}$)。

状态空间: 系统状态由无人机状态与目标状态两个子集组成, 表示为 $\mathcal{S}=\{\mathcal{S}^u, \mathcal{S}^{\text{tar}}\}$ 。 $\mathcal{S}_t^u=\{s_{t,i}^u\}_{i \in \mathcal{M}}$ 是时隙 t 中所有无人机状态信息的集合。在时隙 t 时, 无人机 i 的状态向量表示为 $s_t^i=(p_t^i, l_t^D, l_t^C)$, 其中 p_t^i 为无人机 i 的位置, l_t^D 和 l_t^C 分别是无人机 i 的传输和计算队列长度。 $\mathcal{S}_t^{\text{tar}}=\{s_{t,g}^g\}_{g \in \mathcal{G}}$ 是时隙 t 中所有目标状态信息的集合。对于目标 g , 其在时隙 t 中的状态向量表示为 $s_t^g=(p_t^g, U_g)$, 包含了当前的位置与感知效用。

观察空间: 相较于完整的状态空间, 智能体的观察范围是受限的。定义时隙 t 下智能体 i 的观察空间为 $\mathcal{O}_t^i=\{\mathcal{T}_t^i, \mathcal{U}_t^i\}$ ($\mathcal{O}_t^i \in \mathcal{O}$), 其中 $\mathcal{T}_t^i \subseteq \mathcal{S}_t^{\text{tar}}$ 是智能体 i 感知范围内的目标状态信息集合, $\mathcal{U}_t^i \subseteq \mathcal{S}_t^u$ 是自身状态与广播范围内其他智能体状态信息的集合。

动作空间: 定义时隙 t 下智能体 i 的动作空间为 $a_t^i=\{\rho_i, \phi_i, \varphi_i\}$, 其中感知决策 $\phi_i \in \{0, 1, 2\}$ 包含不感知、感知后使用 SVM 推理或 CNN 推理 3 种动作, 卸载决策 $\rho_i \in \mathcal{M}_i$ 表示将当前计算子任务进行本地推理或卸载至邻居之一, 带宽竞价决策 $\varphi_i \in \{0, 1\}$ 表示智能体 i 是否参与当轮的带宽竞价。

奖励函数: 总奖励 r_t^i 为 3 个动作对应奖励函数之和。卸载奖励为 $r_{t,i}^C=w_1(T_0-T_p)$ 。带宽竞价奖励为 $r_{t,i}^D=-w_2 D_t^L$ 。若约束 C5 满足时, 感知奖励为 $r_{t,i}^S=w_3 U_{i,g}(t)$; 若不满足, 则 $r_{t,i}^S=-R_p$ 作为惩罚。上述奖励函数中, w_1 、 w_2 和 w_3 是系数, R_p 是常数, D_t^L 是此竞价轮次中无人机 i 的平均传输时延。

定义每个竞价轮次包含 κ 个时隙, 在每个轮次的第一个时隙内进行带宽竞价并分配带宽, 直至下一个轮次带宽被重新分配。参与带宽竞价时, 无人机 i 的竞价值为 $P_i=\varphi_i(\phi_i+w_4 l_i^C)+w_5 l_i^D$, 分配的带宽为 $B_i=B \frac{P_i}{\sum_{j=1}^M P_j}$, 其中 w_4 和 w_5 是系数,

B 表示无人蜂群总带宽。

2.2 MACA 模型

MACA 模型由深度神经网络 (deep neural network, DNN) 构成, 分别为 DNN actor 网络与 DNN critic 网络。actor 网络包括在线 actor 网络与目标 actor 网络, 分别表示为 DNN $\pi(o_t|\theta_\pi)$ 和 DNN $\pi'(o_t|\theta_\pi^T)$, 类似地, 在线与目标 critic 网络分别表示为 DNN $Q(O_t, A_t|\theta_Q)$ 和 DNN $Q'(O_t, A_t|\theta_Q^T)$ 。actor 网络采用分裂式输出层, 即多头结构, 分别输出感知、卸载与动作。模型前面的共享网络部分接收输入的观察, 初步提取特征后再输入各头部。分裂输出的结构减少了动作维度, 提高了网络收敛效率。

在时隙 t 中, 智能体 i 根据自身观察 $\mathcal{O}_t^i$ 生成动作 a_t^i 并执行。在获得延时奖励后, 无人机 i 生成经验元组 $(O_t, A_t, r_t^i, O_{t+1})$, 并存储到经验池。训



练时, 智能体 i 从经验池中随机选择一批经验样本, 更新 actor 网络与 critic 网络的参数。在线 actor 网络利用 critic 网络提供的价值评估进行策略改进, 其损失函数计算式如下:

$$L(\theta_\pi) = - \sum_{h=1}^3 E \left[\ln \pi_{\theta_\pi}^h(a_{t,h}|o_t) \right] Q(O_t, A_t | \theta_Q) \quad (7)$$

其中, h 是动作头部索引, $a_{t,h}$ 是头部 h 对应的单一动作, $\pi_{\theta_\pi}^h(a_{t,h}|o_t)$ 为头部 h 输出的动作概率分布。在线 critic 网络的损失函数计算式如下:

$$L(\theta_Q) = E \left[(y_t - Q(O_t, A_t | \theta_Q))^2 \right] \quad (8)$$

其中, y_t 为目标 Q 值, 其计算式如下:

$$y_t = r_t^i + \gamma Q'(O_{t+1}, A'_{t+1} | \theta_Q^T)_{A'_{t+1} \in \{\pi(O_{t+1} | \theta_\pi^T)\}} \quad (9)$$

其中, γ 为折扣因子。训练后, 将在线网络的参数软同步至目标网络, 计算式如下:

$$\theta_\pi^T = \tau \theta_\pi + (1 - \tau) \theta_\pi^T \quad (10)$$

$$\theta_Q^T = \tau \theta_Q + (1 - \tau) \theta_Q^T \quad (11)$$

其中, τ 为同步参数。目标网络保证了学习模型的平稳收敛。

2.3 注意力机制

在 MACA 算法中, 智能体将观测输入 actor 网络以获得动作, critic 网络负责对环境建模并评估 actor 网络的输出。随着系统中智能体数量的增加, critic 网络的输入维度也随之提高, 导致其收敛缓慢, 进而影响 actor 网络的训练。在协同感知场景中, 无人机的感知决策与卸载决策主要取决于下方目标和邻居节点的状态。因此, 通过在 critic 网络中嵌入注意力单元, 使其成为注意力-评论家 (attention-critic) 网络, 让智能体能够优先关注这些关键因素, 从而提升策略的准确性。

attention-critic 网络由输入层、隐藏层和输出层组成。为提取环境的多维特征, 在输入层与隐藏层之间嵌入了多个注意力单元。隐藏层同时接

收智能体自身特征与其他智能体的注意力值。对于智能体 i , 为获得其 attention-critic 网络的输出 Q 值, 首先将所有智能体的观测-动作对 (o^j, a^j) 输入至输入层提取特征, 表示为:

$$g_j = g_i(o^j, a^j) \quad (12)$$

其中, g_i 为输入层函数。然后, 其他智能体特征值被输入注意力单元以计算智能体 i 的注意力值。首先, 注意力单元使用双线性映射将智能体 i 与 j 的特征进行比较, 得出相似值 sim_j , 其计算式如下:

$$\text{sim}_j = g_j^T W_k^T W_q g_i \quad (13)$$

其中, W_q 和 W_k 分别为“查询”与“键”的转换矩阵。将相似值传递至 softmax 层中作归一化处理, 得出各智能体 j 对智能体 i 的注意力权重 α_j 。最终, 利用注意力权重求得其他智能体特征的加权和, 即注意力值:

$$e_i = \sum_{j \neq i} \alpha_j W_v g_j \quad (14)$$

其中, W_v 为“值”的转换矩阵。当使用多个注意力单元时, 需要将并行产生的多个注意力值拼接并输入多层感知器 (multi-layer perceptron, MLP)。因此, attention-critic 网络输出 Q 值的计算式如下:

$$Q_i^v(O, A) = f_i(g_i, \text{concat}(e_i^1, e_i^2, \dots, e_i^K)) \quad (15)$$

其中, f_i 为多层感知器函数, K 为注意力单元个数。

综上, 在所提算法中, 每个时隙内, 智能体根据观察到的无人机与目标状态, 通过自身 actor 网络获取感知模式、计算卸载与带宽竞价的联合动作。获得奖励后, 生成训练样本存储至经验池。每隔 η 时隙后, 智能体选取一部分训练样本用于更新网络参数。MACA 算法的具体步骤如算法 1 所示。

算法 1 MACA 算法

初始化所有无人机 actor 主网络 DNN $\pi(o_i | \theta_\pi)$ 与 attention-critic 主网络 DNN $Q(O_t, A_t | \theta_Q)$;

初始化所有无人机目标网络 DNN $\pi'(o_i|\theta_\pi^T)$ 与 DNN $Q'(O_t, A_t|\theta_Q^T)$, 同步参数 $\theta_\pi^T \leftarrow \theta_\pi$ 和 $\theta_Q^T \leftarrow \theta_Q$:

for each episode do

for each time slot t do

for each UAV do

无人机 i 通过自身 actor 网络获得联合动作 $a_t^i = \{\rho_i, \phi_i, \varphi_i\}$;

end for

执行所有动作 $A_t = \{a_t^1, a_t^2, \dots, a_t^M\}$;

系统进入下一状态, $S_t = S_{t+1}$, $O_t = O_{t+1}$;

获得奖励后, 无人机 i 将样本 $(O_t, A_t, r_t^i, O_{t+1})$ 存入经验池;

if $t \bmod \eta = 0$ then

for each UAV do

从经验池选取一批训练样本;

通过式 (8) 更新 attention-critic 主网络参数;

通过式 (7) 更新 actor 主网络参数;

通过式 (10) 和式 (11) 软更新目标网络参数;

end for

end if

end for

end for

3 仿真结果与分析

3.1 仿真环境

本节使用 Python 3.11 设计仿真实验平台, 并利用 Pytorch 包构建了 MACA 模型。在不同无人蜂群规模和目标数量下, 利用仿真平台评估了 MACA 算法在动态异构资源协同调度上的表现。仿真参数见表 1。

表 1 仿真参数

仿真参数	取值
任务区域	5 000 m × 5 000 m
单元格边长	$d = 500$ m
无人机速度	10 ~ 20 m/s
总带宽	$B = 2$ MHz
无人机运算速度	$c_i = 2$ GFLOPS
支持向量数量	$N_S = 10$
数据向量维度	$d_S = 4\,500$
CNN 卷积层数量	5
CNN 识别准确度	$P_C = 90\%$
SVM 识别准确度	$P_S = 70\%$
actor 网络学习率	0.000 5
attention-critic 网络学习率	0.000 1
折扣因子	$\gamma = 0.9$

3.2 实验结果分析

为了验证所提 MACA 算法的优越性, 将其与以下 4 个基准方案相比较。

(1) 独立推理方案: 无人机产生一批图像后, 独自进行 CNN 或 SVM 推理, 且带宽均匀分配。

(2) 均匀调度方案: 无人机将计算子任务优先卸载至全网计算队列最短的无人机, 带宽同样均匀分配。

(3) 无竞价多智能体注意力 (multi-agent critic with attention no bidding, MACA-NoB) 算法: 去除 MACA 算法中的带宽竞价动作, 仅联合优化感知和计算资源调度, 带宽均匀分配。

(4) 多智能体深度确定性策略梯度 (multi-agent deep deterministic policy gradient, MADDPG) 算法: 每个智能体拥有独立的 critic 网络, actor 网络依然采用分裂输出结构。

MACA 算法的收敛性如图 2 所示, 展示了 MACA 算法与 MADDPG 算法在训练迭代过程中平均感知任务时延的收敛情况。可以看到, MACA 算法在初期更快探索出更优策略并降低时延, 而 MADDPG 算法收敛速度稍慢。训练收敛后, MACA 算法与 MADDPG 算法的时延分别稳定在 5.50 s 和 6.81 s, 且 MADDPG 算法的时延波动更大。在 MACA 算法中, 注意力机制帮助智能体聚焦目标



和邻居节点状态,减少了无关信息的干扰,从而相较MADDPG算法找到更高效的资源调度方案。

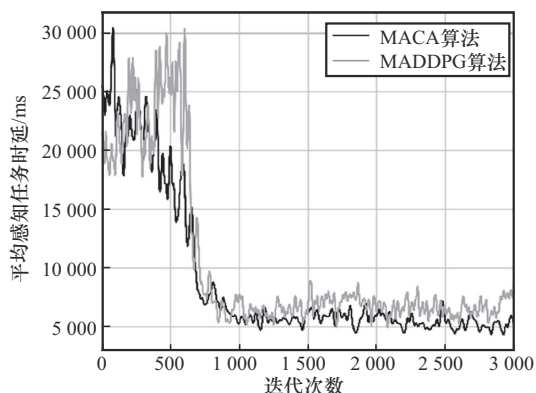


图2 MACA算法的收敛性

不同无人机数量下各方案的平均感知任务时延如图3所示,展示了在不同规模的网络中,MACA算法的平均感知任务时延明显优于其他方案。均匀调度方案只侧重计算任务均匀分配,造成了过高的传输时延。随着无人机数量增加,MADDPG算法因高维输入挑战而学习效率下降,无法做出更优的决策;MACA-NoB由于缺乏带宽分配策略,无法有效缓解传输瓶颈。相比之下,MACA算法联合优化感知、计算与通信资源调度,在不同规模的网络中展现出更强的适应能力。

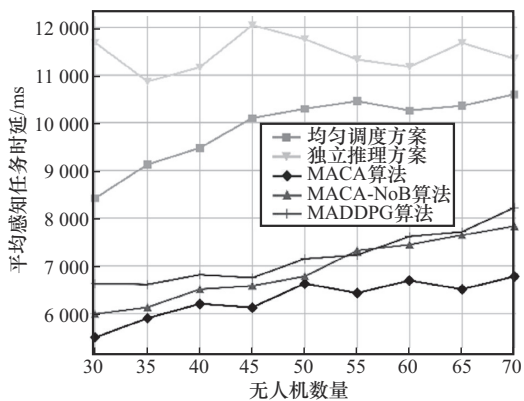


图3 不同无人机数量下各方案的平均感知任务时延

不同目标数量下各方案的平均感知任务时延如图4所示,对比了在目标数量(计算负载)增加时,5种方案的平均感知任务时延变化趋势。由于

缺乏有效协作计算策略,独立推理方案和均匀调度方案在面对高计算负载时平均感知任务时延增长更快。相比MADDPG算法,MACA算法通过注意力机制更好地理解环境并利用邻近资源,在高计算负载下进一步降低了时延。MACA-NoB算法缺乏带宽分配策略,瓶颈节点无法及时完成卸载,导致较高的传输时延,因而性能不及MACA算法。

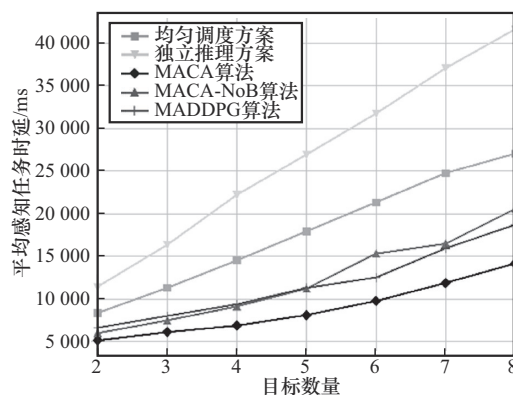


图4 不同目标数量下各方案的平均感知任务时延

不同目标数量下各方案的平均计算资源利用率如图5所示,平均计算资源利用率定义为感知任务期间执行推理任务的无人机数量占总无人机数量的平均比例。当目标数量较少时,MACA算法因综合考虑通信和计算资源,利用率略低于仅关注计算资源的均匀调度方案。随着目标数量增加,MACA算法的平均计算资源利用率超过其他基准方案。MACA-NoB算法的计算卸载效率受限于通信资源,而MADDPG算法无法充分利用邻近资源,整体表现不及MACA算法。

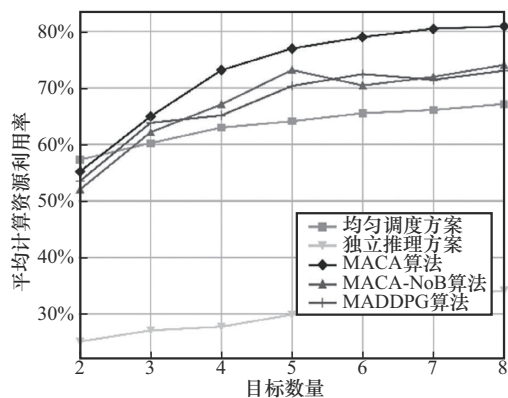


图5 不同目标数量下各方案的平均计算资源利用率

4 结束语

本文研究了协同感知场景中无人蜂群的异构资源调度问题。针对在高精度感知约束下实现系统快速响应的挑战,提出了一种面向协同感知的无人蜂群智能资源调度方案,实现计算、感知与通信资源的协同互补。本文无人蜂群采用了基于分步卸载的协作推理、多模式感知选择和基于竞价的带宽分配。在确保目标感知效用高于阈值的前提下,以最小化感知任务时延为目标,构建了异构资源联合优化问题,并利用嵌入注意力单元的MACA算法求解。数值分析表明,提出的智能资源调度方案在平均感知任务时延和平均计算资源利用率方面都优于基准方案。

参考文献:

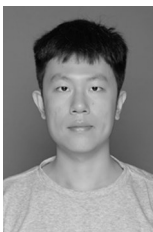
- [1] MENG K T, WU Q Q, XU J, et al. UAV-enabled integrated sensing and communication: opportunities and challenges[J]. IEEE Wireless Communications, 2024, 31(2): 97-104.
- [2] XIONG K, LENG S P, CHEN L Y, et al. RIS-aided trajectory optimization in layered urban air mobility[J]. IEEE Internet of Things Journal, 2025, 12(2): 2188-2202.
- [3] XIONG K, YU H Q, LENG S P, et al. Multi-hop RIS-aided learning model sharing for urban air mobility[J]. IEEE Transactions on Intelligent Transportation Systems, 2024(99): 1-13.
- [4] ITKIN M, KIM M, PARK Y. Development of cloud-based UAV monitoring and management system[J]. Sensors, 2016, 16(11): 1913.
- [5] CHEN X, FENG Z Y, WEI Z Q, et al. Performance of joint sensing-communication cooperative sensing UAV network[J]. IEEE Transactions on Vehicular Technology, 2020, 69(12): 15545-15556.
- [6] YIN N H, LIU R A, ZENG B B, et al. A review: UAV-based remote sensing[J]. IOP Conference Series: Materials Science and Engineering, 2019, 490: 062014.
- [7] LI B, LIU W S, XIE W C, et al. Adaptive digital twin for UAV-assisted integrated sensing, communication, and computation networks[J]. IEEE Transactions on Green Communications and Networking, 2023, 7(4): 1996-2009.
- [8] CAI T, YANG Z H, CHEN Y F, et al. Cooperative data sensing and computation offloading in UAV-assisted crowdsensing with multi-agent deep reinforcement learning[J]. IEEE Transactions on Network Science and Engineering, 2022, 9(5): 3197-3211.
- [9] ZHOU L Y, LENG S P, WANG Q, et al. Integrated sensing and communication in UAV swarms for cooperative multiple targets tracking[J]. IEEE Transactions on Mobile Computing, 2023, 22(11): 6526-6542.
- [10] DHUHEIR M A, BACCOUR E, ERBAD A, et al. Deep reinforcement learning for trajectory path planning and distributed inference in resource-constrained UAV swarms[J]. IEEE Internet of Things Journal, 2023, 10(9): 8185-8201.
- [11] JOUHARI M, AL-ALI A K, BACCOUR E, et al. Distributed CNN inference on resource-constrained UAVs for surveillance systems: design and optimization[J]. IEEE Internet of Things Journal, 2021, 9(2): 1227-1242.
- [12] LI T Y, CHENG Z K, LENG S P. A UAV swarm sensing oriented distributed computing cooperation scheme[C]//Proceedings of the 2021 IEEE 6th International Conference on Computer and Communication Systems (ICCCS). Piscataway: IEEE Press, 2021: 168-173.
- [13] LI T Y, LENG S P, WANG Z H, et al. Intelligent resource allocation schemes for UAV-swarm-based cooperative sensing[J]. IEEE Internet of Things Journal, 2022, 9(21): 21570-21582.
- [14] WEI D W, MA J F, LUO L B, et al. Computation offloading over multi-UAV MEC network: a distributed deep reinforcement learning approach[J]. Computer Networks, 2021(199): 108439.
- [15] SEIFERT E, SEIFERT S, VOGT H, et al. Influence of drone altitude, image overlap, and optical sensor resolution on multi-view reconstruction of forest images[J]. Remote Sensing, 2019, 11(10): 1252.
- [16] TMUŠIĆ G, MANFREDA S, AASEN H, et al. Current practices in UAS-based environmental monitoring[J]. Remote Sensing, 2020, 12(6): 1001.
- [17] MENG K T, HE X F, WU Q Q, et al. Multi-UAV collaborative sensing and communication: joint task allocation and power optimization[J]. IEEE Transactions on Wireless Communications, 2023, 22(6): 4232-4246.
- [18] SHAKHOV V V, KOO I. Experiment design for parameter estimation in probabilistic sensing models[J]. IEEE Sensors Journal, 2017, 17(24): 8431-8437.
- [19] CHAKRABORTY A, ROUT R R, CHAKRABARTI A, et al. On network lifetime expectancy with realistic sensing and traffic generation model in wireless sensor networks[J]. IEEE Sen-



sors Journal, 2013, 13(7): 2771-2779.

- [20] ZHANG S H, ZHANG H L, DI B Y, et al. Cellular cooperative unmanned aerial vehicle networks with sense-and-send protocol[J]. IEEE Internet of Things Journal, 2019, 6(2): 1754-1767.
- [21] ABE S. Support vector machines for pattern classification[M]. London: Springer, 2005.
- [22] SHILTON A, PALANISWAMI M, RALPH D, et al. Incremental training of support vector machines[J]. IEEE Transactions on Neural Networks, 2005, 16(1): 114-131.
- [23] SIMONYAN K, ZISSERMAN A. Very deep convolutional networks for large-scale image recognition[EB]. 2014.
- [24] IQBAL S, SHA F. Actor-attention-critic for multi-agent reinforcement learning[EB]. 2018.

[作者简介]



赵鹏程 (1999-), 男, 电子科技大学信息与通信工程学院博士生, 主要研究方向为无人蜂群协同感知、资源分配、机器学习等。



李天扬 (1997-), 女, 电子科技大学信息与通信工程学院博士生, 主要研究方向为无线网络、无人机自组网、多智能体强化学习、数字孪生、网络演算、空中计算等。



冷甦鹏 (1973-), 男, 电子科技大学信息与通信工程学院教授、博士生导师, 主要研究方向为物联网、车联网、新一代宽带无线网络、无线自组织网络、智能交通信息系统的资源管理、介质访问控制、路由、组网与互联、智能算法理论及技术应用等。



熊凯 (1991-), 男, 电子科技大学信息与通信工程学院副研究员, 主要研究方向为无人机编队资源分配、移动边缘计算和机器学习。