

恶意干扰下的无人机辅助边缘计算加权能耗与时延智能优化

杨和林^{*①} 郑梦婷^① 刘帅^① 肖亮^① 谢显中^② 熊泽辉^③

^①(厦门大学信息学院 厦门 361005)

^②(重庆邮电大学重庆市计算机网络与通信技术重点实验室 重庆 400065)

^③(新加坡科技设计大学信息系统技术与设计学院 新加坡 487372)

摘要: 近年来, 将移动边缘计算(MEC)服务器搭载在无人机(UAV)上为地面移动用户提供服务备受学术界和工业界广泛的研究。但在恶意干扰环境下, 如何有效调度资源降低系统时延和能耗成为关键问题。为此, 针对干扰机影响下无人机辅助边缘计算的问题, 该文建立一个以最小化加权能耗与时延为目标的模型, 联合优化无人机飞行轨迹、资源调度和任务分配来提升无人机辅助移动边缘计算系统性能。鉴于优化问题难求解以及恶意干扰行为动态多变, 该文提出了一种基于双延迟深度确定性策略梯度(TD3)的资源调度算法, 同时结合优先经验回放(PER)机制提高算法收敛速度和稳定性, 高效对抗恶意干扰攻击。仿真结果表明所提算法较其他算法, 能够有效降低系统的时延和能耗, 并具有很好的收敛性与稳定性。

关键词: 无人机通信; 移动边缘计算; 资源调度; 抗干扰; 深度强化学习

中图分类号: TN92

文献标识码: A

文章编号: 1009-5896(2024)00-0001-09

DOI: 10.11999/JEIT230986

Intelligent Weighted Energy Consumption and Delay Optimization for UAV-Assisted MEC Under Malicious Jamming

YANG Helin^① ZHENG Mengting^① LIU Shuai^① XIAO Liang^①
XIE Xianzhong^② XIONG Zehui^③

^①(School of Informatics, Xiamen University, Xiamen 361005, China)

^②(Chongqing Key Laboratory of Computer Network and Communication Technology, Chongqing University of Posts and Telecommunications, Chongqing 400065, China)

^③(Pillar of Information Systems Technology and Design, Singapore University of Technology and Design, Singapore 487372)

Abstract: In recent years, mounting Mobile Edge Computing (MEC) servers on Unmanned Aerial Vehicle (UAV) to provide services for mobile ground users has been widely researched in academia and industry. However, in malicious jamming environments, how to effectively schedule resources to reduce system delay and energy consumption becomes a key challenge. Therefore, this paper considers a UAV-assisted MEC system under a malicious jammer, where an optimization model is established to minimize the weighted energy consumption and delay by jointly optimizing UAV flight trajectories, resource scheduling, and task allocation. As the optimization problem is difficult to be solved and the malicious jamming behavior is dynamic, a Twin Delayed Deep Deterministic policy gradient algorithm (TD3) is proposed to search for the optimal policy. At the same time, the Prioritized Experience Replay (PER) technique is added to improve the convergence speed and stability of the algorithm, which is highly effective against malicious interference attacks. The simulation results show that the proposed algorithm can effectively reduce the delay and energy consumption, and achieve

收稿日期: 2024-xx-xx; 改回日期: 2024-04-08; 网络出版: 2024-05-07

*通信作者: 杨和林 helinyang066@xmu.edu.cn

基金项目: 国家自然科学基金(62371408, 62301467, U21A20444, 61971366), 中央高校基本科研业务费专项资金(20720220080), 厦门市自然科学基金项目(3502Z202371010), “小米青年学者”项目, 科技部重点研发计划项目(2023YFB3107603)

Foundation Items: The National Natural Science Foundation of China (62371408, 62301467, U21A20444, 61971366), The Fundamental Research Funds for the Central Universities (20720220080), Natural Science Foundation of Xiamen(3502Z202371010), Xiaomi Young Talents Program. National Key Research and Development Program of China (2023YFB3107603)

good convergence and stability compared with other algorithms.

Key words: Unmanned Aerial Vehicle (UAV) communication; Mobile Edge Computing (MEC); Resource scheduling; Anti-jamming; Deep Reinforcement Learning (DRL)

1 引言

得益于第五代移动通信技术的发展,汽车自动驾驶、高清直播等新型技术的应用越来越普遍。这些新型应用产生的数据量呈爆炸式增长,对终端用户(User Equipment, UE)上的计算能力和存储能力提出很高的要求。为了解决用户在计算密集型任务时面临的高延迟和低可靠的问题,移动边缘计算(Mobile Edge Computing, MEC^[1])技术应运而生。MEC将计算资源部署在网络边缘,用户可以把计算密集型和时延敏感型任务卸载至边缘节点高效地处理,有效减少用户的能量损耗。然而,传统MEC网络依靠地面固定的基站进行任务卸载,造成系统通信灵活性差和应用范围受限等问题。无人机(Unmanned Aerial Vehicle, UAV)具有灵活、易部署、体积小等优点^[2],使得数据通信链路传输过程近似为视距传输^[3]。因此,无人机辅助的MEC系统可以实现更高质量的通信信道,从而减少传输时延和能量损耗。

目前在无人机部署MEC服务器作为移动基站为地面用户提供服务方面展开了大量的研究。按照不同的优化目标可以分为最小化能耗、最小化时延和最小化加权能耗与时延这3个部分。在以最小化能耗为目标的研究中,文献[4]针对MEC和无线信息与电力传输结合的多用户全双工通信系统,联合优化CPU频率、传输功率、传输速率和卸载计算量实现系统总能耗最小。文献[5]将非正交多址与MEC结合,以无人机和地面用户的加权能耗最小为目标展开研究。文献[6,7]均以最小化系统时延为优化目标,但仅研究系统延迟问题,并未讨论能耗对系统的影响。在以最小化能耗和时延的研究中,文献[8,9]虽取得不错的效果,但没有考虑到无人机能耗和本地能耗之间数量级不平衡的问题,也没有讨论实际过程中可能存在的干扰机干扰影响。无人机辅助的MEC系统的网络拓扑结构随着无人机的变化而变化,因此,目前研究高动态的通信链路也是一个巨大的挑战。文献[10,11]在优化系统能耗方面取得了不错的效果,但采用的是传统的优化算法。传统算法面对系统拓扑网络高度变化下难以解决多维资源混合决策的问题,并且还需要经过复杂的迭代后仅才得到次优解。随着机器学习的发展,许多研究人员应用深度强化学习(Deep Reinforcement Learning, DRL)来解决无人机辅助的MEC系

统计算卸载问题,提供一种新的辅助和管理方案^[12]。文献[13]提出了针对离散空间的强化学习算法,但该算法难以应用在高维连续变量的环境中。文献[14]应用深度确定性策略梯度(Deep Deterministic Policy Gradient, DDPG)制定计算卸载方法,所用的DDPG算法能够解决连续动作空间问题,仿真结果表明该算法优于其他算法。但是DDPG算法在实际过程中还存在高估Q值的情况,并且可能出现错误Q估值引导下的错误策略。

上述文献尚未考虑恶意干扰机影响下的无人机辅助边缘计算加权能耗和时延优化问题。此外,无人机辅助边缘计算环境和恶意干扰行为分别具有高动态性和不确定性的特点,进一步增大系统方案的设计难度。因此,本文提出面向恶意干扰机影响下系统加权能耗和时延最小化的资源智能调度方案,设计了改进的TD3深度强化学习算法,有效地捕捉学习环境动态特性,联合优化多域资源来提升系统性能。本文主要的研究贡献如下:

(1)本文研究了在恶意干扰攻击下无人机辅助边缘计算系统的资源优化问题,构建了加权能耗和传输时延最小化为目标的模型,并通过联合优化无人机飞行轨迹、资源调度和任务分配来降低系统传输延迟和能耗,有效抵抗恶意干扰。

(2)系统中恶意干扰行为动态多变,所提的问题是混合整数规划问题,难以用传统的优化算法进行解决。针对该问题,本文首先将问题转化为马尔科夫决策过程,然后提出双延迟深度确定性策略梯度(Twin Delayed Deep Deterministic Policy Gradient, TD3)的算法解决该优化问题,同时引入优先经验回放(Prioritized Experience Replay, PER)机制充分挖掘环境动态变化因素,加快网络收敛速度与稳定性,与传统优化算法相比可以较好地抵抗恶意干扰的攻击,所提PER-TD3算法可高效求得无人机辅助边缘计算下智能资源分配策略。

(3)通过仿真实验表明,所提算法在恶意干扰下能高效训练,拥有更好的收敛性能。并且在不同情况下所提算法均能够有效降低系统的传输时延和能耗,使系统总开销最低,具有较好的稳定性。

2 系统模型及问题描述

2.1 系统模型

本文考虑一个在恶意干扰影响下无人机辅助移动边缘计算系统。如图1所示,该系统由1个搭载了

MEC服务器的无人机、 I 个用户以及1个干扰机组成。整个系统周期 T 被等长划分为 N 个时隙，无人机在每一个时隙中仅为一个用户提供服务。同时，无人机与用户在通信过程存在建筑物等障碍物遮挡，以及干扰机影响。其中 I 个用户的位置是在限制的区域范围内随机生成的，本文采用3维欧几里得坐标系，则第 $n(n \in \{1, 2, \dots, N\})$ 个时隙第 $i(i \in \{1, 2, \dots, I\})$ 个用户的坐标可以用 $\mathbf{m}_i(n) = [x_i(n), y_i(n), 0]$ 表示。无人机在固定的高度 H 上飞行，其在第 n 个时隙开始的坐标为 $\mathbf{u}(n) = [x_{\text{uav}}(n), y_{\text{uav}}(n), H]$ ，停止的坐标为 $\mathbf{u}(n+1) = [x_{\text{uav}}(n+1), y_{\text{uav}}(n+1), H]$ 。设置无人机的飞行速度为 $v(n) \in [0, v_{\text{max}}]$ ，飞行角度为 $\theta(n) \in [0, 2\pi]$ ，从起始位置到停止位置飞行时间为 t_{fly} ，则其停止位置又可以表示为 $\mathbf{u}(n+1) = [x_{\text{uav}}(n) + v(n)t_{\text{fly}} \cos \theta(n), y_{\text{uav}}(n) + v(n)t_{\text{fly}} \sin \theta(n), H]$ 其中无人机和用户的水平坐标均在长度为 L ，宽为 W 的区域范围内动态变化。本文设置用 L_J 来表示干扰机的坐标。

本文设置无人机与用户之间以及无人机与干扰机之间的通信为视距传输(Line of Sight, LoS)链路^[3]，则在第 n 个时隙中无人机和第 i 个用户以及无人机与干扰机之间的信道增益可以分别用 g_i^U 和 g^J 来表示

$$g_i^U(n) = \beta_0 d_i^{U-2}(n) = \frac{\beta_0}{\|x_{\text{uav}}(n+1) - x_i(n)\|^2 + \|y_{\text{uav}}(n+1) - y_i(n)\|^2 + H^2} \quad (1)$$

$$g^J(n) = \beta_0 d^J(n) = \frac{\beta_0}{\|u(n+1) - L_J\|^2 + H^2} \quad (2)$$

其中， β_0 是信道增益， d_i^U 和 d^J 分别为无人机与第 i 个用户的距离以及无人机与干扰机之间的距离。

那么，第 n 个时隙在干扰机影响下无人机与第 i 个用户之间传输速率为

$$R_i(n) = W \log_2 \left(1 + \frac{g_i^U(n) P_U}{\sigma^2 + g^J(n) P_J + h_i(n) P_N} \right) \quad (3)$$

其中， W 表示信道带宽， σ^2 表示无人机处的噪声功率， P_N 表示有障碍物遮挡下的功率损耗， P_U 和 P_J 分别表示用户和干扰机当前信号的发射功率。 $h_i(n)$ 表示无人机与第 i 个用户之间障碍物遮挡的有无， $h_i(n)$ 是一个只能为1或0的均匀分布变量，本文设置有障碍物遮挡和无障碍物遮挡概率均为0.5。

(1)系统时延：由于移动边缘计算服务器处理后的数据量比较小，因此在通信损耗问题上不考虑返回结果过程下行链路传输时延和能耗^[8]。在无人机辅助边缘计算系统时延包括本地计算时延、用户传输时延和无人机处理时延这3个部分。

本文设置无人机辅助的边缘计算系统计算采用部分卸载模式。设置在第 n 个时隙第 i 个用户任务卸载率为 $\lambda_i(n)$ ，生成总量为 $l_i(n)$ 需要处理任务，单位比特计算任务所需的CPU循环次数为 C ，用户的计算能力为 f_{UE} 。因此，在第 n 个时隙第 i 个用户的本地计算时延 $T_i^{\text{loc}}(n)$ 为

$$T_i^{\text{loc}}(n) = \frac{(1 - \lambda_i(n)) l_i(n) C}{f_{\text{UE}}} \quad (4)$$

配置了服务器的无人机上的时延包括传输和计算卸载任务时延，用 f_{UAV} 表示边缘计算服务器的计算频率，则传输时延 $T_i^{\text{tr}}(n)$ 和计算卸载任务时延 $T_i^{\text{UAV}}(n)$ 可以分别表示为

$$T_i^{\text{tr}}(n) = \frac{\lambda_i(n) l_i(n)}{R_i(n)} \quad (5)$$

$$T_i^{\text{UAV}}(n) = \frac{\lambda_i(n) l_i(n) C}{f_{\text{UAV}}} \quad (6)$$

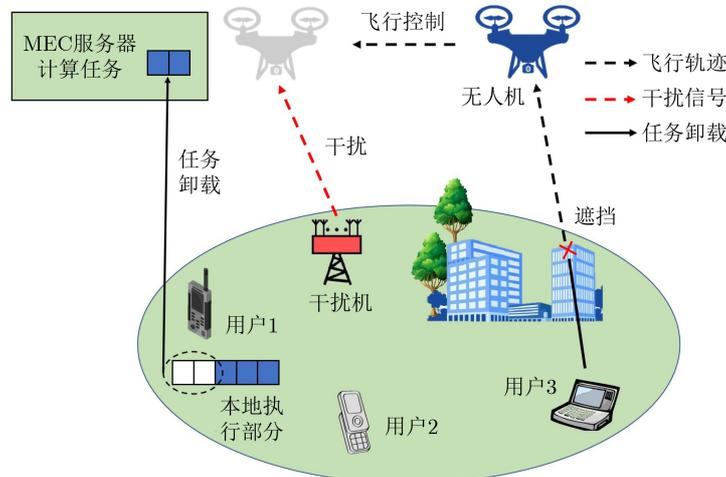


图1 干扰环境下无人机辅助边缘计算系统模型

假设 $\rho_i(n)$ 表示在第 n 个时隙中无人机与第 i 个用户的关联系数, 1 代表进行服务, 0 代表不服务。考虑无人机和用户可以同时处理任务, 则第 n 个时隙总延迟 $T_{\text{total}}(n)$ 可以表示为

$$T_{\text{total}}(n) = \sum_{i=1}^I \rho_i(n) \max \{ T_i^{\text{loc}}(n), T_i^{\text{tr}}(n) + T_i^{\text{UAV}}(n) \} \quad (7)$$

(2) 系统能耗: 无人机能量损耗包括两个部分, 一部分是无人机计算卸载任务的能耗, 另一部分是无人机的飞行能耗。而用户的能耗也包括两个部分, 一部分是将任务卸载边缘计算服务器系统传输能耗, 另一部分是本地计算能耗。

用 ξ 表示与实际芯片结构有关的功耗系数, 边缘计算服务器的计算功率为 $P_{\text{UAV}}^{\text{com}} = \xi f_{\text{UAV}}^3$, 本地计算功率为 $P_{\text{UE}}^{\text{com}} = \xi f_{\text{UE}}^3$ 。 M_{UAV} 为无人机及其载重的质量, 则第 n 时刻服务器计算能耗 $E_i^{\text{UAV}}(n)$ 、本地计算能耗 $E_i^{\text{loc}}(n)$ 、无人机飞行能耗 $E_i^{\text{fly}}(n)$ 、本地传输能耗 $E_i^{\text{tr}}(n)$ 可以分别表示为

$$E_i^{\text{UAV}}(n) = P_{\text{UAV}}^{\text{com}} T_i^{\text{UAV}}(n) \quad (8)$$

$$E_i^{\text{loc}}(n) = P_{\text{UE}}^{\text{com}} T_i^{\text{loc}}(n) \quad (9)$$

$$E_i^{\text{fly}}(n) = 0.5 M_{\text{UAV}} t_{\text{fly}} \|v(t)\|^2 \quad (10)$$

$$E_i^{\text{tr}}(n) = P_{\text{U}} T_i^{\text{tr}}(n) \quad (11)$$

用 ν 代表无人机能耗的权重因子, 则第 n 时刻总能耗 $E_{\text{total}}(n)$ 可以表示为

$$E_{\text{total}}(n) = \sum_{i=1}^I \rho_i(n) (E_i^{\text{tr}}(n) + E_i^{\text{loc}}(n) + \nu (E_i^{\text{fly}}(n) + E_i^{\text{UAV}}(n))) \quad (12)$$

2.2 问题描述

本文基于无人机辅助的移动边缘计算系统, 由于无人机能耗和用户能耗不在一个数量级, 无人机的飞行功率可以达到千瓦, 而用户的传输功率仅毫瓦级别, 同时无人机的计算功率也远大于用户的计算功率。因此, 本文在计算能耗设置一个权重因子来平衡无人机能耗与用户能耗。本文联合优化无人机飞行轨迹 u 、资源调度 λ 和任务分配 ρ 实现加权能耗和时延最小。该优化问题可以表示为

$$\min_{\{u, \lambda_i, \rho_i\}} \sum_{n=1}^N T_{\text{total}}(n) + E_{\text{total}}(n) \quad (13)$$

$$\text{s.t. C1: } \rho_i(n) \in \{0, 1\}, \forall i, \forall n; \sum_{i=1}^I \rho_i(n) = 1, \forall n$$

$$\text{C2: } 0 \leq \lambda_i(n) \leq 1, \forall i, n$$

$$\text{C3: } h_i(n) \in \{0, 1\} \forall i, n$$

$$\text{C4: } 0 \leq x_{\text{uav}}(n) \leq L, 0 \leq y_{\text{uav}}(n) \leq W, \forall n$$

$$\text{C5: } \sum_{n=1}^N (E_i^{\text{fly}}(n) + E_i^{\text{UAV}}(n)) \leq E_{\text{UAV}}$$

$$\text{C6: } \sum_{n=1}^N \sum_{i=1}^I \rho_i(n) l_i(n) = L$$

$$\text{C7: } v(n) t_{\text{fly}} = \|u(n+1) - u(n)\| \leq d_{\text{max}}, \forall n$$

$$\text{C8: } 0 \leq v(n) \leq v_{\text{max}}, \forall n; 0 \leq \theta(n) \leq 2\pi, \forall n$$

其中, C1 表示在一个时隙中无人机仅对地面其中一个用户提供服务, C2 表示卸载率的约束范围, C3 表示无人机和每个用户通信过程障碍物遮挡的情况用 0 或者 1 来表示, C4 是无人机的坐标的约束范围, C5 表示无人机的电池容量是有限的, C6 表示在规定的时间内需要完成总的任务量, C7 限制了无人机在每个时隙中最大的飞行距离, C8 分别是无人机的最大飞行速度和角度的约束。

由于该优化问题是一个非凸优化问题, 采用传统的优化算法需要经过复杂的迭代后才仅能得到次优解, 并且在一些需要观察环境状态的边缘计算场景中传统优化算法存在很多的局限性。强化学习是机器学习的重要分支, 它具有强大的数据计算能力, 可以解决传统优化算法无法解决的在未知环境统计信息下的动态决策问题[15]。

3 基于双延迟深度确定性策略梯度的优化方案

双延迟深度确定性策略梯度(TD3)算法是一种基于AC框架的面向连续动作空间的深度强化学习算法, 其可以在复杂的恶意干扰动态环境中探索到最优策略[16]。本文提出基于PER-TD3算法在TD3的基础上加入PER机制, 增加学习效率。本文用该算法设计无人机飞行轨迹、用户资源调度以及任务分配实现系统时延和加权能耗最小化, 抵抗恶意干扰攻击。

3.1 MDP基本要素

深度强化学习用深度学习来感知环境特征, 利用强化学习根据当前状态做出动作决策, 然后转移到下一个新的状态, 不断循环此过程最后得到一个最优策略。而马尔可夫决策过程(Markov Decision Process, MDP)是一种智能体和环境交互过程建立数学模型的一种方法。MDP可以表示为 $\langle S, A, P, r \rangle$, S 表示具有MDP属性的状态集合, A 表示智能体的动作集合, P 表示状态转移函数, r 表示智能体得到的奖励集合。针对MDP问题, 本文根据上述的

资源调度优化问题定义了状态空间、动作空间和奖励函数。

(1) 状态空间 $s(n)$ ：在无人机辅助边缘计算系统中，状态空间由一个搭载边缘计算服务器的无人机， I 个用户和一个干扰机共同确定，由于干扰机的位置等参数不发生变化，因此在状态空间中仅考虑用户和无人机即可。无人机的状态包括第 n 个时隙无人机的剩余电池电量 $E_{\text{uav}}^b(n)$ 和坐标 $\mathbf{u}(n)$ ，用户的状态包括第 n 个时隙 I 个用户的坐标 $\mathbf{m}_i(n)$ 和任务量 $l_i(n)$ ，以及剩余的任务量 $l^b(n)$ 、无人机和 I 个用户的遮挡情况 $h_i(n)$ 。可以将的状态空间表示为

$$s(n) = (E_{\text{uav}}^b(n), \mathbf{u}(n), \mathbf{m}_i(n), l_i(n), l^b(n), h_i(n)) \quad (14)$$

(2) 动作空间 $a(n)$ ：本文根据联合优化对象来定义动作空间。无人机飞行轨迹可以分解为飞行角度 $\theta(n)$ 和飞行速度 $v(n)$ ，资源分配指任务卸载比率 $\lambda_i(n)$ ，用户调度表示当前无人机服务用户情况 $p(n)$ 。

$$a(n) = (\theta(n), v(n), \lambda_i(n), p(n)) \quad (15)$$

(3) 奖励函数 $r(n)$ ：由于本文以最小化系统和能耗加权最小为目标，而TD3算法以最大化累计奖励，奖励函数和总处理延迟呈负相关关系，则奖励函数可以表示为

$$r(n) = -(T_{\text{total}}(n) + E_{\text{total}}(n)) \quad (16)$$

3.2 基于PER-TD3实现时延和加权能耗最小化

TD3算法主要由智能体、环境、经验池以及6个网络组成的。其中，TD3算法采用的6个网络分

别为：Actor在线网络 $u'(s|\theta')$ ，Actor目标网络 $Q_1(s, a|w_1)$ ，两个Critic在线网络 $Q_2(s, a|w_2)$ 和 $Q'_1(s, a|w'_1)$ ，两个Critic目标网络 $Q'_2(s, a|w'_2)$ 和 s 。

TD3算法是DDPG算法基础上改进得到的，相比于DDPG算法，TD3算法不仅减轻过估计情况，还具有较稳定的学习过程。如图2所示为TD3算法的框架，TD3的算法流程可以分成采样、训练、参数更新这三个部分。算法在采样阶段随机抽取状态 s 输入到Actor在线网络后得到四元组 (s_n, a_n, r_n, s_{n+1}) 放入经验复用池。当经验复用池满则开始训练，否则继续采样。在训练阶段，从经验池中抽取一定批次的元组 (s_t, a_t, r_t, s_{t+1}) 作为样本数据输入到网络中进行训练。然后可以得到输出的动作 a_{t+1} ，进而计算得到目标Q值 y_t ，最后得到计算TD误差。在参数更新阶段，基于延迟更新机制首先进行更新两个Critic在线网络，等Critic1和Critic2网络更新 d 步再更新Actor在线网络，并且用软更新的方式更新两个Critic目标网络和Actor目标网络。其中两个Critic在线网络通过最小化损失函数来更新

$$L(w_k) = \frac{1}{N} \sum (y_t - Q_k(s_t, a_t|w_k))^2 \quad (17)$$

Actor在线网络通过计算策略梯度来更新：

$$\nabla_{\theta} J(\theta) = \frac{1}{N} \sum \nabla_{a_t} Q_k(s_t, a_t)|_{a_t=u(s_t)} \nabla_{\theta} u(s_t) \quad (18)$$

为了增强学习效率，本文在TD3算法的基础上增加了PER机制。PER通过优先抽样高优先级的经验进行训练，使得重要的经验能够频繁地被学

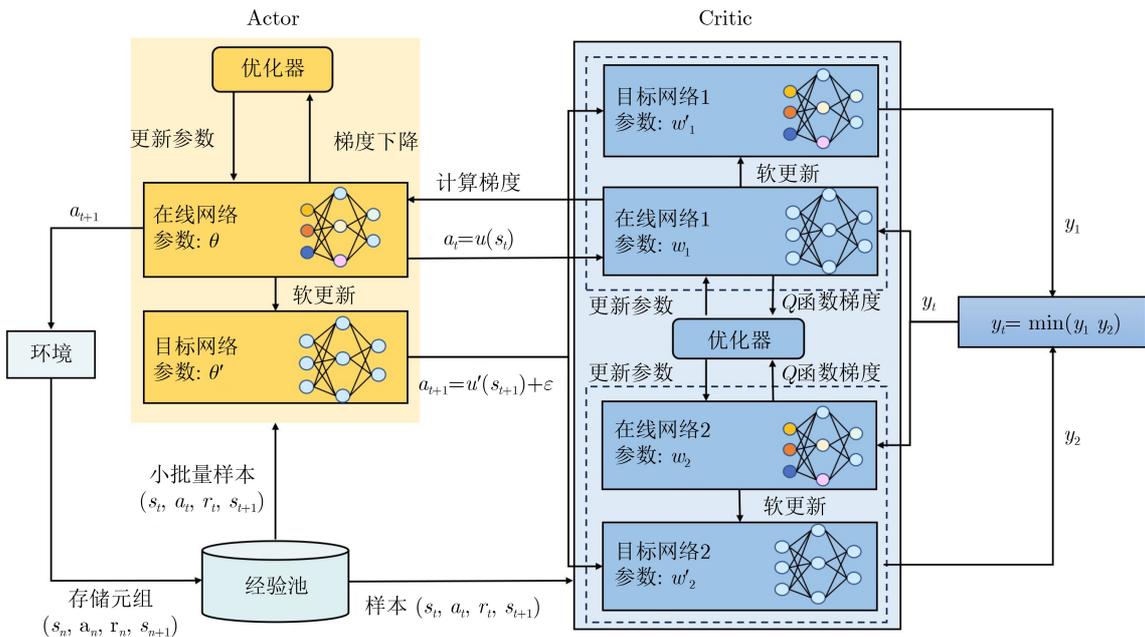


图2 无人机辅助边缘计算系统中基于TD3的资源调度算法框

习, 实现训练过程更快收敛。本文提出无人机辅助边缘计算系统中基于PER-TD3加权能耗和时延最小化算法流程如**算法1**所示。

4 仿真结果

本节给出无人机辅助移动边缘计算系统仿真及分析结果。实验的仿真硬件平台为CPU型号Intel Core i9-12900H、内存16GB、显卡RTX-3060的笔记本电脑。仿真实验的集成开发环境为Python3.7, 在该环境下调用Tensorflow深度学习库作为本文深度学习的框架。本文设置训练轮数Episode=600, 策略和价值网络学习率分别为0.001和0.002, 折扣因子 γ 为0.99。本文系统环境的参数设置基于文献[17–19], 其中用户数量 $I = 20$, 周期 $T = 80$ s, 时隙数 $N = 40$, 带宽 $W = 1$ MHz, 信道增益 $\beta_0 = -50$ dB, 处理单位比特所需的CPU周期 $C = 1000$, 用户计算频率 $f_{UE} = 0.8$ GHz, 无人机计算频率 $f_{UAV} = 1.4$ GHz, 权重因子 $\nu = 0.002$, 无人机电池容量 $E_{UAV} = 80$ kJ, 任务总量 $L = 10$ Mbit, 无人机最大飞行距离 $d_{\max} = 30$ m, 干扰机坐标 $L_J = [300, 300]$, 用户设备发射功率 $P_U = 100$ mW, 干扰机发射功率 $P_J = 100$ mW, 无人机飞行高度 $H = 100$ m。为了说明PER-TD3算法在无人机辅助的边缘计算系统的性能, 设置了以下方案用于比较: (1)基于TD3算法: 基于TD3的资源调度、任务分配和轨迹规划联合优化算法; (2)基于DDPG算法: 通过DDPG算法进行资源调度、任务分配和轨迹规划, 可以用于连续的动作空间; (3)迭代算法: 该算法通过在每一步选择当前最优的选择, 即问题的最优解通过

子问题的最优解来构建; (4)遗传算法: 该算法模拟自然进化的自然选择、交叉和变异的过程, 逐步演化出优秀的个体, 实现逼近问题的最优解。

4.1 算法的收敛性能对比

图3评估了所提算法的收敛性能。**图3**中纵坐标cost为系统加权能耗和时延之和, 用于衡量整个系统的总开销。算法在200轮后网络参数开始更新, 最终均收敛到最优值。DDPG算法在训练过程中存在异常值, 并且在训练过程抖动明显, 且最终收敛总开销还略大于TD3和PER-TD3算法。这是因为DDPG算法在学习Q价值函数时存在高估真实的Q值导致学习错误策略。TD3算法在400轮左右收敛到最优值, PER-TD3算法则是在300轮左右收敛, 因此PER-TD3算法收敛速度较快。PER-TD3算法是在TD3算法的基础上加上经验优先回放机制, 能够高效利用优先级较高的数据, 提高算法的收敛速度和稳定性。综上, 本文所用PER-TD3算法在干扰影响下拥有更好的收敛性能, 能够快速学习到最优的策略。

4.2 算法的稳定性

图4(a)和**4(b)**显示了在不同用户的CPU计算能力下, 不同方案完成任务所需的时间和能耗对比。当用户计算能力增大, 系统的卸载率随其减小, 用户本地计算更多的任务, 使得系统时延减少, 能耗增加。迭代算法是通过子问题的最优解来逼近问题最优解, 可以看到其所需时延是最小的, 但是能耗开销确是最大的, 陷入了局部最优解。遗传算法其能耗和时延均大于PER-TD3算法。DDPG和TD3算法收敛速度均小于PER-TD3算法, 故这两

算法 1 基于PER-TD3的时延和能耗加权最小化的算法流程

- (1) 设置无人机辅助边缘计算系统环境, 初始化在线网络参数 w_1, w_2, θ 和目标网络参数: $w'_1 \leftarrow w_1, w'_2 \leftarrow w_2, \theta' \leftarrow \theta$ 。
- (2) 初始化经验回放池。
- (3) 循环训练轮数Episode = 1, 2, ..., E:
- (4) 重置参数并得到初始状态 s_1 ;
- (5) 循环训练步数Step = 1, 2, ..., N:
- (6) 通过Actor选择加入噪声的动作: a_n ;
- (7) 无人机执行动作 a_n , 进入下一个状态 s_{n+1} , 并从环境中获得奖励 r_n , 并计算当前经验的优先级 p_n ;
- (8) 如果经验池未滿, 将四元组 (s_n, a_n, r_n, s_{n+1}) 及其优先级存储至经验池中;
- (9) 如果经验池满, 在经验池按照优先级选取小批量样本 (s_t, a_t, r_t, s_{t+1}) 输入网络中;
- (10) 基于策略平滑由Actor目标网络输出动作: $a_{t+1} = u'(s_{t+1} | \theta') + \varepsilon$;
- (11) 计算目标值: $y_t = r_t + \gamma \min(Q'_1(s_{t+1}, a_{t+1} | w'_1), Q'_2(s_{t+1}, a_{t+1} | w'_2))$;
- (12) 根据TD值更新优先级, 计算重要性采样权重, 更新损失函数, 更新网络参数;
- (13) 直到Step = N;
- (14) 直到Episode = E;
- (15) 计算获得无人机飞行轨迹、资源调度和任务分配策略, 输出系统的能耗和时延。

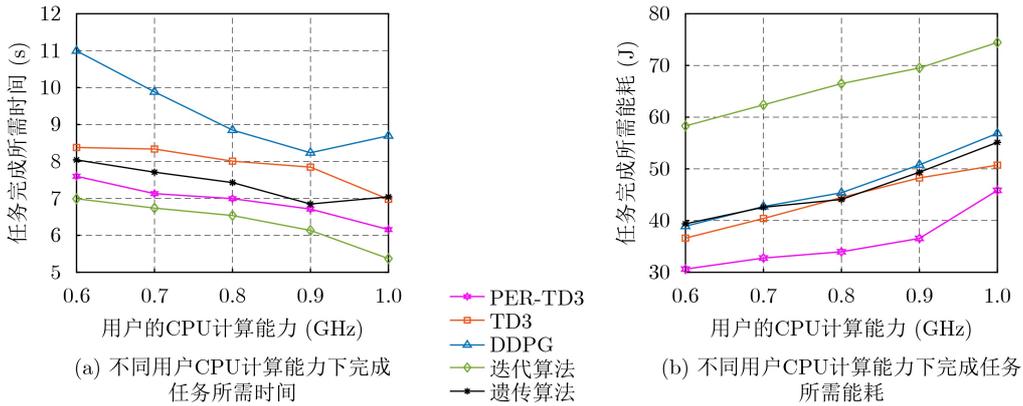


图 4 不同用户CPU计算能力下5个方案开销对比

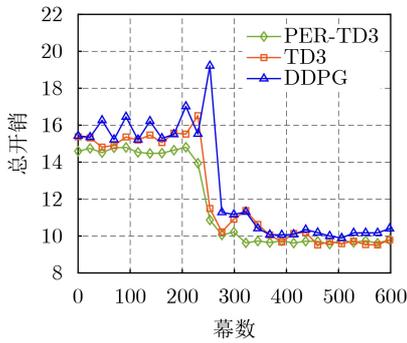


图 3 PER-TD3算法与其他算法的收敛性能比较

个算法在时延和能耗开销均大于所提算法。因此，在干扰环境中，所提的PER-TD3算法相比于其他算法能够获得最小的总开销，同时收敛也更快更加稳定。

图5(a)和5(b)展示了在不同计算任务下，不同算法完成任务所需时间和能耗的对比。贪心算法仅实现了时延最小，却在能耗上有较大的开销。遗传算法性能依赖于种群内多样性程度，算法不够稳定。DDPG算法和TD3算法均比PER-TD3算法收敛速度慢，因此在能耗和时延上均大于所提算法。以上结果说明了PER-TD3算法不仅所需开销最小，而且也更稳定，进一步体现了PER-TD3算法的优越性。

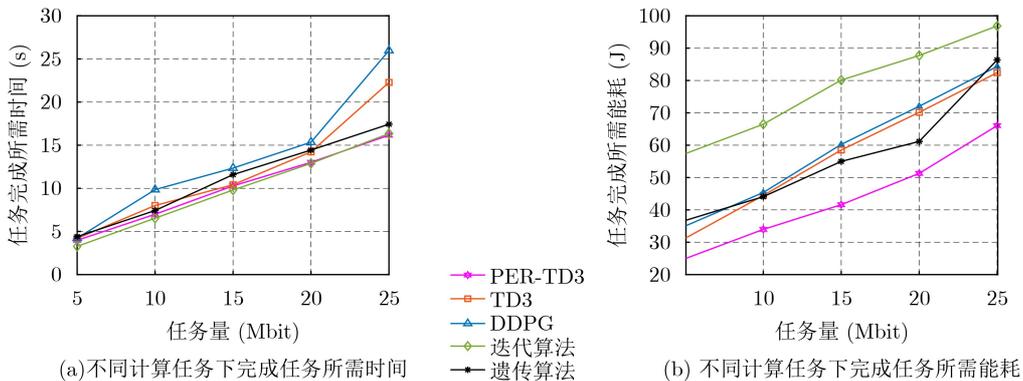


图 5 不同计算任务下5个方案开销对比

如图6(a)和6(b)比较了不同用户数量下不同算法完成任务的时间和能耗。由图可知迭代算法和遗传算法陷入局部最优解。而DDPG, TD3, PER-TD3算法通过训练合理分配资源可以很好地平衡系统时延和能耗问题。PER-TD3能够更快收敛到最优策略，因此在不同的用户数目下能够以最小的开销完成任务。由以上实验可以证明所提算法在不同的用户数目下更加稳定，所需的总开销也更低。

图7为不同用户计算能力以及无人机边缘计算能力下无人机的2维飞行轨迹对比图。当无人机计算能力大小从 $f_{UAV} = 1.4$ GHz 增大到 $f_{UAV} = 1.6$ GHz 时，系统的卸载率增大，有更多的任务卸载到无人机上，因此无人机飞行轨迹更加靠近用户，降低系统传输时延。而当用户的计算大小从 $f_{UE} = 0.8$ GHz 增大到 $f_{UE} = 1$ GHz 时，系统的卸载率变低，无人机飞行轨迹距离变短，从而减少飞行能耗。

5 结束语

本文考虑了在干扰环境下无人机辅助移动边缘计算系统，通过联合优化无人机飞行轨迹、任务分配和资源调度实现系统加权能耗和时延之和最小化。由于所提的问题是混合整数规划问题，本文采用基于深度强化学习的PER-TD3算法得到最优的

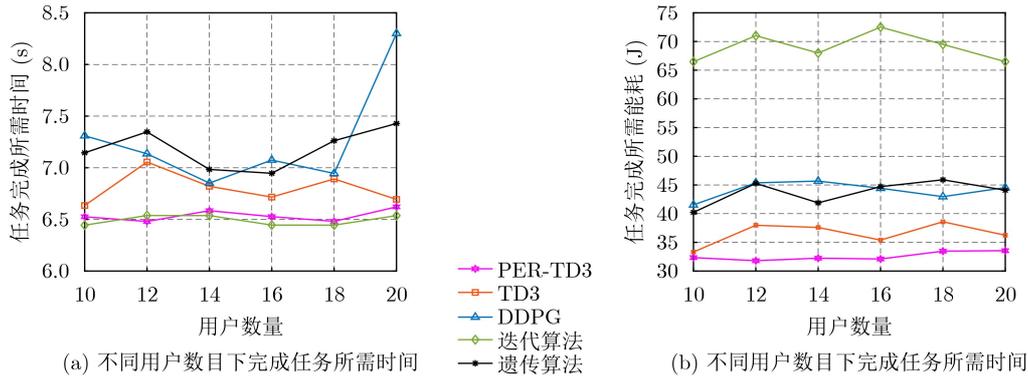


图 6 不同用户数目下5个方案开销对比

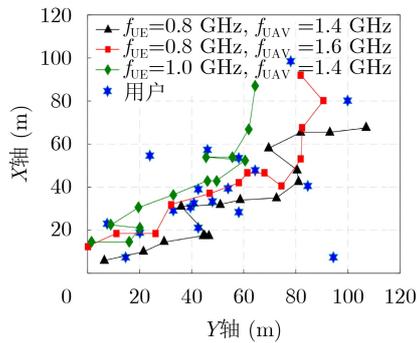


图 7 无人机飞行轨迹

卸载方案。仿真结果表明，本文所用的PER-TD3算法较对比算法能够快速地收敛到最优值，且在不同的用户数目、不同用户的计算能力以及不同总任务量下，PER-TD3算法能够有效降低系统时延和能耗使系统的总开销最小，证明了PER-TD3算法的优越性。

参考文献

- [1] MAO Yuyi, YOU Changsheng, ZHANG Jun, *et al.* A survey on mobile edge computing: The communication perspective[J]. *IEEE Communications Surveys & Tutorials*, 2017, 19(4): 2322–2358. doi: [10.1109/COMST.2017.2745201](https://doi.org/10.1109/COMST.2017.2745201).
- [2] WANG Di, TIAN Jie, ZHANG Haixia, *et al.* Task offloading and trajectory scheduling for UAV-enabled MEC networks: An optimal transport theory perspective[J]. *IEEE Wireless Communications Letters*, 2022, 11(1): 150–154. doi: [10.1109/LWC.2021.3122957](https://doi.org/10.1109/LWC.2021.3122957).
- [3] LIN Xingqin, YAJNANARAYANA V, MURUGANATHAN S D, *et al.* The sky is not the limit: LTE for unmanned aerial vehicles[J]. *IEEE Communications Magazine*, 2018, 56(4): 204–210. doi: [10.1109/MCOM.2018.1700643](https://doi.org/10.1109/MCOM.2018.1700643).
- [4] CHEN Fangni, FU Jiafei, WANG Zhongpeng, *et al.* Joint communication and computation resource optimization in FD-MEC cellular networks[J]. *IEEE Access*, 2019, 7: 168444–168454. doi: [10.1109/ACCESS.2019.2954622](https://doi.org/10.1109/ACCESS.2019.2954622).
- [5] ZHANG Xiaochen, ZHANG Jiao, XIONG Jun, *et al.* Energy-efficient multi-UAV-enabled multiaccess edge computing incorporating NOMA[J]. *IEEE Internet of Things Journal*, 2020, 7(6): 5613–5627. doi: [10.1109/JIOT.2020.2980035](https://doi.org/10.1109/JIOT.2020.2980035).
- [6] ZHANG Liang and ANSARI N. Latency-aware IoT service provisioning in UAV-aided mobile-edge computing networks[J]. *IEEE Internet of Things Journal*, 2020, 7(10): 10573–10580. doi: [10.1109/JIOT.2020.3005117](https://doi.org/10.1109/JIOT.2020.3005117).
- [7] HU Qiyu, CAI Yunlong, YU Guanding, *et al.* Joint offloading and trajectory design for UAV-enabled mobile edge computing systems[J]. *IEEE Internet of Things Journal*, 2019, 6(2): 1879–1892. doi: [10.1109/JIOT.2018.2878876](https://doi.org/10.1109/JIOT.2018.2878876).
- [8] YU Zhe, GONG Yanmin, GONG Shimin, *et al.* Joint task offloading and resource allocation in UAV-enabled mobile edge computing[J]. *IEEE Internet of Things Journal*, 2020, 7(4): 3147–3159. doi: [10.1109/JIOT.2020.2965898](https://doi.org/10.1109/JIOT.2020.2965898).
- [9] ZHOU Wen, XING Ling, XIA Junjuan, *et al.* Dynamic computation offloading for MIMO mobile edge computing systems with energy harvesting[J]. *IEEE Transactions on Vehicular Technology*, 2021, 70(5): 5172–5177. doi: [10.1109/TVT.2021.3075018](https://doi.org/10.1109/TVT.2021.3075018).
- [10] 余雪勇, 朱焯, 邱礼翔, 等. 基于无人机辅助边缘计算系统的节能卸载策略[J]. *系统工程与电子技术*, 2022, 44(3): 1022–1029. doi: [10.12305/j.issn.1001-506X.2022.03.35](https://doi.org/10.12305/j.issn.1001-506X.2022.03.35).
YU Xueyong, ZHU Ye, QIU Lixiang, *et al.* Energy efficient offloading strategy for UAV aided edge computing systems[J]. *Systems Engineering and Electronics*, 2022, 44(3): 1022–1029. doi: [10.12305/j.issn.1001-506X.2022.03.35](https://doi.org/10.12305/j.issn.1001-506X.2022.03.35).
- [11] 李安, 戴龙斌, 余礼苏, 等. 加权能耗最小化的无人机辅助移动边缘计算资源分配策略[J]. *电子与信息学报*, 2022, 44(11): 3858–3865. doi: [10.11999/JEIT210832](https://doi.org/10.11999/JEIT210832).
- [12] LI An, DAI Longbin, YU Lisu, *et al.* Resource allocation for unmanned aerial vehicle-assisted mobile edge computing to minimize weighted energy consumption[J]. *Journal of Electronics & Information Technology*, 2022, 44(11): 3858–3865. doi: [10.11999/JEIT210832](https://doi.org/10.11999/JEIT210832).
- [12] LI Yuxi. Deep reinforcement learning: An

- overview[EB/OL].<https://arxiv.org/abs/1701.07274>, 2017.
- [13] CHENG Nan, LYU Feng, QUAN Wei, *et al.* Space/aerial-assisted computing offloading for IoT applications: A learning-based approach[J]. *IEEE Journal on Selected Areas in Communications*, 2019, 37(5): 1117–1129. doi: [10.1109/JSAC.2019.2906789](https://doi.org/10.1109/JSAC.2019.2906789).
- [14] SEID A M, BOATENG G O, ANOKYE S, *et al.* Collaborative computation offloading and resource allocation in multi-UAV-assisted IoT networks: A deep reinforcement learning approach[J]. *IEEE Internet of Things Journal*, 2021, 8(15): 12203–12218. doi: [10.1109/JIOT.2021.3063188](https://doi.org/10.1109/JIOT.2021.3063188).
- [15] XIAO Liang, DING Yuzhen, HUANG Jinhao, *et al.* UAV anti-jamming video transmissions with QoE guarantee: A reinforcement learning-based approach[J]. *IEEE Transactions on Communications*, 2021, 69(9): 5933–5947. doi: [10.1109/TCOMM.2021.3087787](https://doi.org/10.1109/TCOMM.2021.3087787).
- [16] FUJIMOTO S, HOOF H, and MEGER D. Addressing function approximation error in actor-critic methods[C]. Proceedings of the 35th International Conference on Machine Learning, Stockholm, Sweden, 2018: 1587–1596.
- [17] LIN Na, TANG Hailun, ZHAO Liang, *et al.* A PDDQNLP algorithm for energy efficient computation offloading in UAV-assisted MEC[J]. *IEEE Transactions on Wireless Communications*, 2023, 22(12): 8876–8890. doi: [10.1109/TWC.2023.3266497](https://doi.org/10.1109/TWC.2023.3266497).
- [18] LIU Boyang, WAN Yiyao, ZHOU Fuhui, *et al.* Resource allocation and trajectory design for MISO UAV-assisted MEC networks[J]. *IEEE Transactions on Vehicular Technology*, 2022, 71(5): 4933–4948. doi: [10.1109/TVT.2022.3140833](https://doi.org/10.1109/TVT.2022.3140833).
- [19] ZHOU Yi, PAN Cunhua, YEOH P L, *et al.* Secure communications for UAV-enabled mobile edge computing systems[J]. *IEEE Transactions on Communications*, 2020, 68(1): 376–388. doi: [10.1109/TCOMM.2019.2947921](https://doi.org/10.1109/TCOMM.2019.2947921).
- 杨和林：男，副教授，博士生导师，研究方向为无线通信、边缘计算、智能通信、资源分配。
- 郑梦婷：女，硕士生，研究方向为边缘计算、无人机通信、资源调度。
- 刘 帅：男，博士生，研究方向为边缘计算、抗干扰通信、无线资源调度、无人机通信。
- 肖 亮：女，教授，博士生导师，研究方向为无线通信、智能通信、通信安全、物理层安全。
- 谢显中：男，教授，博士生导师，研究方向为通信网理论与技术、物联网、可见光通信、无人机通信。
- 熊泽辉：男，助理教授，博士生导师，研究方向为边缘计算、无线网络技术、强化学习、无线通信。

责任编辑：陈 倩