SAG-MEC 网络下支持 WPT 的无人机动态任务卸载与资源分配

王怡，覃团发，韦睿，黄金宝

（广西大学计算机与电子信息学院广西多媒体通信与网络技术重点实验室，广西 南宁 530004）

摘　要： 针对偏远地区蜂窝网络覆盖不足且物联网（IoT）设备能量和计算能力低而无法满足大量延迟敏感型任务卸载和计算需求的问题，考虑将空天地一体化网络（SAGIN）和移动边缘计算（MEC）相结合，提出一种支持无线电力传输（WPT）技术的无人机辅助 IoT 设备的动态任务卸载和资源分配方案，其中无人机负责收集 IoT 设备产生的计算密集型任务，采用部分卸载模式将这些任务根据当前状态进行本地计算或动态卸载给基站和低地球轨道（LEO）卫星进一步处理。由于动态的异构网络和长期排队延迟与短期决策的耦合性，因此在排队延迟的约束下提出一种基于 Lyapunov 优化的双延迟深度确定性策略梯度（TD3PG）算法，该算法通过优化无人机动态关联、任务分配、计算资源分配和带宽分配来协调无人机学习最优卸载策略和资源分配。仿真结果表明，所提出的动态方案与其他对比方案相比能够有效降低无人机网络的能耗、网络积压总和及平均排队延迟，在 2 种学习率组合下 TD3PG 算法相对于深度确定性策略梯度（DDPG）算法和双重深度 Q 网络（DDQN）算法的奖励分别提高了 13.6%、24.0%和 20.4%、17.9%。

关键词： 空天地一体化网络；移动边缘计算；无人机；任务卸载；资源分配

中图分类号： TN929.5　　**文献标志码：** A　　**DOI：** 10.19678/j.issn.1000-3428.0070030

Dynamic Task Offloading and Resource Allocation of UAVs Supported by WPT in SAG-MEC Network

WANG Yi, QIN Tuanfa, WEI Rui, HUANG Jinbao

(Guangxi Key Laboratory of Multimedia Communications and Network Technology, School of Computer and Electronic Information, Guangxi University, Nanning 530004, Guangxi, China)

【Abstract】 Remote areas face problems such as insufficient cellular network coverage as well as low energy and computing power of Internet of Things (IoT) devices. Hence, the requirements for delay-sensitive task offloading and computing for a large number of tasks cannot be met. Considering the combination of the Space-Air-Ground Integrated Network (SAGIN) and Mobile Edge Computing (MEC), this paper proposes a strategy for dynamic task offloading and resource allocation for Unmanned Aerial Vehicle (UAV)-assisted IoT devices that support Wireless Power Transmission (WPT) technology, in which UAVs are responsible for collecting compute-intensive tasks generated by IoT devices. These tasks are locally calculated or dynamically unloaded to the base station and a Low Earth Orbit (LEO) satellite for further processing using a partial unloading mode, according to the current state. Given the dynamic heterogeneous network environment, as well as the tight coupling between long-term queuing delays and short-term decision-making, this paper proposes a Twin Delayed Deep Deterministic Policy Gradient (TD3PG) algorithm based on Lyapunov optimization under queuing delay constraints. The algorithm coordinates UAVs to learn the optimal offloading strategy and resource allocation by optimizing UAV dynamic association, task allocation, computing resource allocation, and bandwidth allocation. Simulation results show that, compared with other schemes, the proposed dynamic scheme can effectively reduce the energy consumption, network backlog sum, and average queue delay in the UAV network. Under different learning rate combinations, the reward of the TD3PG algorithm increases by 13.6% and 24.0% compared with that of the Deep Deterministic Policy Gradient (DDPG) algorithm, and by 20.4% and 17.9% compared with that of the Double Deep Q-Network (DDQN) algorithm.

【Key words】 Space-Air-Ground Integrated Network (SAGIN); Mobile Edge Computing (MEC); Unmanned Aerial Vehicle (UAV); task offloading; resource allocation

基金项目： 国家自然科学基金（61563004）。

作者简介： 王怡（CCF 学生会员），女，硕士研究生，主研方向为空天地一体化、无人机通信；覃团发（通信作者），教授、博士生导师；韦睿、黄金宝，硕士研究生。

收稿日期： 2024-06-24　　**修回日期：** 2024-09-07　**E-mail：** tfqin@gxu.edu.cn

计算机工程 www.ecice06.com

0 引言

随着 5G 和 6G 网络的发展,物联网(IoT)设备受到越来越多的关注[1],但这些 IoT 设备功耗和计算能力低,难以执行计算密集型任务。无人机具有部署灵活性高、传输速率快等优点,可作为特殊的空中平台应用在各种场景中进行任务采集[2-3],例如农业监测、森林火灾预防等。但是,无人机能量和计算能力有限,面对突发计算密集型任务时可能导致计算资源短缺并恶化网络延迟性能[4]。虽然蜂窝网络可解决此类问题,但蜂窝网络并不能提供无处不在的覆盖[5-6],因此将计算任务直接卸载到蜂窝网络和云服务器变得不切实际,这对无人机来说是一个巨大的挑战。

卫星因覆盖范围广、可靠性高和成本低等特点,能够为地面网络提供一定的计算资源来减轻网络负担。因此,将卫星、无人机和地面蜂窝等多个网络系统集成形成一个覆盖范围广、速度快的空天地一体化网络(SAGIN)以实现更高效的通信和数据交换,达到优化决策、增强任务效率的目的[7-11]。通常,云计算能够灵活地按需分配资源、提供多样化服务,是 SAGIN 处理复杂计算任务的首选。然而,在 SAGIN 环境下,云计算由于传输延迟大、通信带宽有限,难以保证延迟敏感型任务的服务质量。移动边缘计算(MEC)作为 5G 和 6G 的新兴技术,将计算任务从云端分配到边缘服务器,不仅能提供高效、灵活的计算服务,还能减少传输延迟和缩短响应时间,故本文考虑将 MEC 与 SAGIN 相结合,设计一种为计算密集型任务提供计算服务的空天地一体化的边缘计算网络架构 SAG-MEC[12-14]。虽然 SAG-MEC 网络具有优势,但无人机在不确定的异构网络环境、有限的计算资源以及严格的延迟要求下对计算密集型任务进行卸载分配和计算是另一大挑战。

近年来,国内外学者对优化卸载和计算性能进行了广泛的研究。文献[15]考虑了结合云服务器的天地一体化网络,通过卫星中继任务实现云计算,但忽视了卫星部署 MEC 服务器的可行性,还增加了网络的传输延迟。文献[16]基于 SAGIN 场景,在最大任务数和延迟约束下以最小化能耗为目标优化无人机的卸载策略,但忽略了无人机的飞行能耗。文献[17-18]考虑在 SAGIN 场景中通过约束系统延迟及风险来优化无人机的任务卸载比例和网络性能,但仅考虑了单架无人机,当计算资源不足时无法保证延迟敏感型任务的服务质量。文献[19]在考虑随机信道条件、无人机覆盖和任务队列的情况下研究了 SAGIN 场景下无人机辅助物联网任务卸载的能耗问题,但研究仅涉及将计算任务作为一个整体进行本地计算或卸载到服务器执行的二元卸载,没有考虑将其分成两个或多个独立的部分,进而通过局部卸载和计算并行的部分卸载模式。

由于异构网络的复杂性,因此设计一种高效、稳定的卸载计算方案至关重要。文献[20]提出了一种交替迭代优化算法,该算法将问题分解为 3 个子问题,利用松弛变量和连续凸逼近方法将问题转化为凸问题进行求解,但容易陷入局部最优,而且需要较长的计算时间,从而增大网络延迟。文献[21]在 SAGIN 中提出了一种深度确定性策略梯度(DDPG)算法为用户设备进行卸载和计算,虽然该算法可以有效学习最佳卸载策略,但依然会出现过估计的情况,容易陷入局部最优。

此外,任务卸载和计算会使无人机消耗过多的能量,通过对无人机进行无线电力传输(WPT)以保持网络的可持续性是极其关键的[22-23]。而且大多数的研究忽略了任务卸载和计算的动态任务积压,排队延迟会对端到端延迟性能产生重大影响,因此在具有严格延迟要求的无人机任务卸载中不容忽视[24]。

综上所述,本文在排队延迟的约束下,以最大限度地降低无人机网络的能耗为目标,研究了 SAG-MEC 网络下支持 WPT 的无人机辅助 IoT 设备动态任务卸载和资源分配方案,具体工作如下:

1)在 SAG-MEC 网络下,无人机采用部分卸载模式将收集的任务进行动态分配卸载,同时为了保持网络的可持续性,基站采用 WPT 技术为无人机供电。

2)由于问题的复杂性,首先采用虚拟队列将长期约束转换为队列稳定性约束,然后提出基于 Lyapunov 优化的深度强化学习(DRL)算法联合优化动态关联、任务分配、计算资源分配和带宽分配以最小化无人机网络的能耗。

3)仿真结果表明,相较于其他对比方案,本文方案能够有效地降低无人机网络能耗、队列延迟,并提高系统性能。

1 系统模型与框架

1.1 SAG-MEC 网络框架

SAG-MEC 网络框架如图 1 所示(彩色效果见《计算机工程》官网 HTML 版,下同),低地球轨道

(LEO)卫星、K 个基站和 M 架无人机均配备 MEC 服务器,为 IoT 设备提供计算服务,其中 LEO 卫星提供无缝覆盖。在任务期间,多架无人机沿着同一中心以 r_u 为半径的圆形轨道飞行并收集 IoT 设备的计算密集型任务,然后将任务本地处理或卸载到附近基站和 LEO 卫星,以降低无人机网络的能耗。无人机和基站的集合分别为 $m\in\{1,\cdots,M\}$,$k\in\{1,\cdots,K\}$,其中,$M\geqslant 1$,$K\geqslant 1$。设置任务时间为 $\mathcal{T}$,其中 $\mathcal{T}$ 小于或等于卫星的覆盖时间。为了便于分析,将 $\mathcal{T}$ 划分为 T 个间隙为 δ_t 的时隙,用集合表示为 $t\in\{1,2,\cdots,T\}$,其中 δ_t 足够小,使得无人机坐标可以被认为在 1 个时隙内近似不变。本文采用部分卸载模式,定义 $\chi_m(t)\in[0,1]$和 $\gamma_{m,k}(t)\in[0,1]$ 分别表示无人机 m 分配到本地和卸载到基站 k 的任务系数,$1-\chi_m(t)-\gamma_{m,k}(t)$表示卸载到卫星的任务系数。无人机的位置在 t 时隙表示为 $q_m(t)=[x_m(t),y_m(t),H_m(t)]$,基站的位置表示为 $g_k=[x_k,y_k,0]$。系统参数定义如表 1 所示。

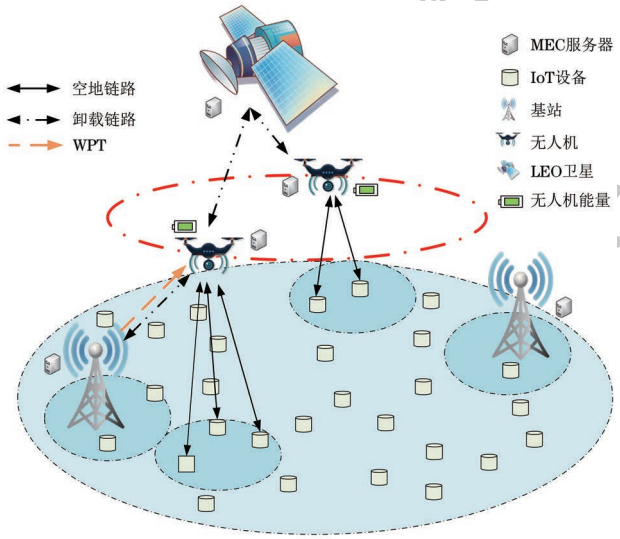


图 1 面向 SAG-MEC 网络的任务卸载与计算模型

Fig. 1 Task offloading and computing model for SAG-MEC network

1.2 信道模型

无人机在飞行中除了视距链路以外,通常还会受到非视距链路的影响[25]。假定无人机飞行高度 H 大于区域中最高障碍物的高度,在 t 时隙无人机 m 与基站 k 的路径损耗可表示如下:

$$G_{m,k}(t)=20\lg\left(\frac{4\pi f_c d_{m,k}(t)}{\lambda}\right)+P_{m,k}^{\mathrm{LOS}}(t)\times g_{m,k}^{\mathrm{LOS}}(t)+(1-P_{m,k}^{\mathrm{LOS}}(t))\times g_{m,k}^{\mathrm{NLOS}}(t) \tag{1}$$

式中:λ 为波长;f_c 为载波频率;$g_{m,k}^{\mathrm{LOS}}(t)$和 $g_{m,k}^{\mathrm{NLOS}}(t)$分别为视距和非视距链路的路径损耗;$d_{m,k}(t)$和 $P_{m,k}^{\mathrm{LOS}}(t)$分别表示无人机 m 与基站 k 的水平距离和视距链路概率。$d_{m,k}(t)$和 $P_{m,k}^{\mathrm{LOS}}(t)$可表示如下:

$$d_{m,k}(t)=\|(x_m(t)-x_k),(y_m(t)-y_k)\| \tag{2}$$

表 1 系统参数

Table 1 System parameters

参数	含义
K	地面基站数量
M	无人机数量
$\chi_m(t)$	无人机 m 分配到本地的任务系数
$\gamma_{m,k}(t)$	无人机 m 卸载到基站 k 的任务系数
$h_{m,k}(t)$	无人机 m 与基站 k 的信道系数
$R_{m,s}(t)$	无人机 m 与卫星的传输速率
$R_{m,k}(t)$	无人机 m 与基站 k 的传输速率
$\alpha_{m,k}(t)$	无人机 m 与基站 k 的带宽分配系数
$D_m(t)$	无人机 m 收集的计算任务量
$Q_m^{\mathrm{local}}(t)$	无人机 m 的本地任务缓冲队列
$Q_m^{\mathrm{BS}}(t)$	无人机 m 卸载到基站 k 的任务缓冲队列
$Q_m^{\mathrm{LEO}}(t)$	无人机 m 卸载到卫星的任务缓冲队列
$Y_m^{\mathrm{BS}}(t)$	无人机 m 卸载到基站 k 的任务量
$Y_m^{\mathrm{LEO}}(t)$	无人机 m 卸载到卫星的任务量
$f_m(t)$	无人机 m 的计算频率资源
$A_{m,k}(t)$	基站 k 为无人机 m 提供的计算关联因子
$A_{m,s}(t)$	卫星为无人机 m 提供的计算关联因子
$\overline{A}_{m,k}(t)$	基站 k 为无人机 m 充电的关联因子
$S_{m,k}^{\mathrm{BS}}(t)$	基站 k 为无人机 m 提供的计算队列
$X_{m,k}^{\mathrm{BS}}(t)$	基站 k 为无人机 m 计算的任务量
$X_m^{\mathrm{LEO}}(t)$	卫星为无人机 m 计算的任务量
$f_{m,k}(t)$	基站 k 的计算频率资源
$f_{m,s}(t)$	卫星的计算频率资源
$\tau_{m,k}^{\mathrm{BS}}$	基站 k 任务卸载的排队延迟
$\tau_{m,s}^{\mathrm{LEO}}$	卫星任务卸载的排队延迟
τ_m^{local}	无人机 m 本地任务计算排队延迟
$\overline{\tau}_{m,k}^{\mathrm{BS}}$	基站 k 任务计算排队延迟
$e_m^{\mathrm{local}}(t)$	无人机 m 本地计算能耗
$e_m^{\mathrm{BS}}(t)$	无人机 m 与基站 k 的卸载能耗
$e_m^{\mathrm{LEO}}(t)$	无人机 m 与卫星的卸载能耗
$e_m^{\mathrm{fly}}(t)$	无人机 m 的飞行能耗
$E_m^{\mathrm{In}}(t)$	无人机 m 的充电能量
$E_m(t)$	无人机 m 的能量队列

$$P_m^{\mathrm{LOS}}(t)=\frac{1}{1+\bar{a}\exp[-\bar{b}(\bar{\theta}_{m,k}(t)-\bar{a})]} \tag{3}$$

式中：$\|\cdot\|$表示欧氏范式运算；$\bar{a}$ 和 $\bar{b}$ 为传播环境的常数因子；$\bar{\theta}_{m,k}(t)$为基站与无人机之间的仰角。无人机 m 与基站 k 的信道系数 $h_{m,k}(t)$可表示如下：

$$h_{m,k}(t)=\delta_m\cdot 10^{-G_m(t)/10} \tag{4}$$

式中：δ_m 为小尺度信道衰落系数且服从瑞利分布。

为了减少干扰，本文采用频分多址技术进行通信，故无人机 m 到基站 k 的速率可表示如下：

$$R_{m,k}(t)=\alpha_{m,k}(t)B\mathrm{lb}\left(1+\frac{P_{m,k}h_{m,k}^2(t)}{\hat{N}+\sigma^2}\right) \tag{5}$$

式中：$\alpha_{m,k}(t)$为带宽分配系数；B 为带宽；$\hat{N}$ 为其他无人机的干扰；σ^2 为接收机收到的加性高斯白噪声。

对于无人机和 LEO 卫星而言，在任务期间，LEO 卫星相对于无人机的位置几乎没有变化，因此无人机和 LEO 卫星的传输速率采用固定值 $R_{m,s}(t)$表示。此外，LEO 卫星云计算相比于无人机和基站边缘计算具有更高的计算能力，因而每个任务处理结果的数据量相对较小，故忽略了计算结果从 LEO 卫星传输到无人机所涉及的传输延迟[23]。

1.3 飞行模型

根据无人机的移动模型，无人机在下一时刻的水平坐标可表示如下：

$$\begin{cases}x_m(t+1)=x_m(t)+r_{\mathrm{u}}(1-\cos\beta)\\ y_m(t+1)=y_m(t)+r_{\mathrm{u}}\sin\beta\end{cases} \tag{6}$$

式中：r_{u} 为无人机的飞行半径；$\beta=180^\circ\times v_m(t)\delta_t/\pi r_{\mathrm{u}}$ 为单位时间内的飞行弧长对应的圆心角大小；$v_m(t)$为无人机的飞行速度。

为了保证无人机的飞行以及避免无人机之间发生碰撞，有以下约束：

$$0\leqslant v_m(t)\leqslant V_{\max},\ \forall m \tag{7}$$

$$\|q_m(t)-q_{\hat{m}}(t)\|^2\geqslant\left(2r_{\mathrm{u}}\sin\frac{\bar{\beta}}{2}\right)^2,\ \forall\hat{m}\neq m \tag{8}$$

式中：$\bar{\beta}=180^\circ\times l_{\min}/\pi r_{\mathrm{u}}$ 为无人机之间的最小安全弧长 $l_{\min}$ 对应的圆心角大小。

1.4 卸载模型

假设任意无人机 m 在 t 时隙收集的计算任务量为 $D_m(t)$，并且无人机 m 有 3 个缓冲队列用于存储分配到本地、卸载到基站和卸载到 LEO 卫星的任务。假设当前时隙的任务是从下一个时隙开始执行的，在部分卸载模式下每个缓冲队列的动态积压模型可表示如下：

$$Q_m^{\mathrm{local}}(t+1)=\{Q_m^{\mathrm{local}}(t)-Y_m^{\mathrm{local}}(t)\}^+ +\chi_m(t)D_m(t) \tag{9}$$

$$Q_m^{\mathrm{BS}}(t+1)=\{Q_m^{\mathrm{BS}}(t)-Y_m^{\mathrm{BS}}(t)\}^+ +\gamma_{m,k}(t)D_m(t) \tag{10}$$

$$Q_m^{\mathrm{LEO}}(t+1)=\{Q_m^{\mathrm{LEO}}(t)-Y_m^{\mathrm{LEO}}(t)\}^+ +D_{m,s}(t) \tag{11}$$

式中：$D_{m,s}(t)=(1-\chi_m(t)-\gamma_{m,k}(t))D_m(t)$；$\{x\}^+=\max(x,0)$；$Y_m^{\mathrm{local}}(t)$表示无人机 m 本地计算的任务量；$Y_m^{\mathrm{BS}}(t)$和 $Y_m^{\mathrm{LEO}}(t)$为无人机 m 卸载到基站 k 和 LEO 卫星的任务量。

本文采用二进制因子 $A_{m,k}(t)\in\{0,1\}$、$A_{m,s}(t)\in\{0,1\}$分别表示无人机 m 是否将任务卸载到基站 k 和 LEO 卫星，当 $A_{m,k}(t)=1$ 时无人机将任务卸载到基站 k，否则 $A_{m,k}(t)=0$；同理当 $A_{m,s}(t)=1$ 时无人机 m 将任务卸载到 LEO 卫星，否则 $A_{m,s}(t)=0$。

在 t 时隙，任意无人机最多有一个基站为其服务，约束如下：

$$\sum_{k=1}^{K}A_{m,k}(t)\leqslant 1,\ \forall m,t \tag{12}$$

因此，$Y_m^{\mathrm{local}}(t)$、$Y_m^{\mathrm{BS}}(t)$和 $Y_m^{\mathrm{LEO}}(t)$分别表示如下：

$$Y_m^{\mathrm{local}}(t)=\min\left\{Q_m^{\mathrm{local}}(t),\frac{\delta_t f_m(t)}{\rho_m}\right\} \tag{13}$$

$$Y_m^{\mathrm{BS}}(t)=\sum_{k=1}^{K}A_{m,k}(t)\min\{Q_m^{\mathrm{BS}}(t),\delta_t R_{m,k}(t)\} \tag{14}$$

$$Y_m^{\mathrm{LEO}}(t)=A_{m,s}(t)\min\{Q_m^{\mathrm{local}}(t),\delta_t R_{m,s}(t)\} \tag{15}$$

式中：$f_m(t)$为分配给无人机 m 的计算频率资源；ρ_m 为 CPU 计算密度。

设置基站 k 有缓冲队列 $S_{m,k}^{\mathrm{BS}}(t)$用来存储无人机卸载的任务，故基站 k 的动态缓冲队列表示如下：

$$S_{m,k}^{\mathrm{BS}}(t+1)=\{S_{m,k}^{\mathrm{BS}}(t)-X_{m,k}^{\mathrm{BS}}(t)\}^+ +Y_m^{\mathrm{BS}}(t) \tag{16}$$

式中：$X_{m,k}^{\mathrm{BS}}(t)$表示基站 k 计算的任务量。

设置基站 k 提供的计算频率资源为 $f_{m,k}(t)$，故 $X_{m,k}^{\mathrm{BS}}(t)$可表示如下：

$$X_{m,k}^{\mathrm{BS}}(t)=\min\left\{Q_m^{\mathrm{BS}}(t),\frac{\delta_t f_{m,k}(t)}{\rho_m}\right\} \tag{17}$$

由于 LEO 卫星具有丰富的计算资源，因此到达 LEO 卫星的任务将立即执行，LEO 卫星计算的

计算机工程 www.ecice06.com

任务量可表示如下：

$$X_m^{\mathrm{LEO}}(t)=\min\left\{Q_m^{\mathrm{LEO}}(t),\frac{\delta_t f_{m,s}(t)}{\rho_m}\right\} \tag{18}$$

1.5 排队延迟

长期排队延迟包括任务卸载排队延迟和计算排队延迟。根据利特尔法则[25]，排队延迟与队列积压成正比，与平均任务到达量成反比。本文设置基站 k 和 LEO 卫星任务卸载的排队延迟 $\tau_{m,k}^{\mathrm{BS}}$、$\tau_{m,s}^{\mathrm{LEO}}$ 有以下约束：

$$\tau_{m,k}^{\mathrm{BS}}=\lim_{T\to\infty}\frac{1}{T}\sum_{t=1}^{T}\frac{Q_m^{\mathrm{BS}}(t)}{\overline{C}_k(t)}\leqslant\tau_{k,\max}^{\mathrm{BS}} \tag{19}$$

$$\tau_{m,s}^{\mathrm{LEO}}=\lim_{T\to\infty}\frac{1}{T}\sum_{t=1}^{T}\frac{Q_m^{\mathrm{LEO}}(t)}{\overline{C}(t)}\leqslant\tau_{\max}^{\mathrm{LEO}} \tag{20}$$

式中：$\tau_{k,\max}^{\mathrm{BS}}$、$\tau_{\max}^{\mathrm{LEO}}$ 为基站和 LEO 卫星的卸载延迟上限；$\overline{C}_k(t)$、$\overline{C}(t)$为序列 $Q_m^{\mathrm{BS}}(t)$和 $Q_m^{\mathrm{LEO}}(t)$的平均任务到达率。$\overline{C}_k(t)$、$\overline{C}(t)$分别表示如下：

$$\overline{C}_k(t)=\frac{1}{T}\sum_{t=0}^{T-1}\gamma_{m,k}(t)D_m(t) \tag{21}$$

$$\overline{C}(t)=\frac{1}{T}\sum_{t=0}^{T-1}D_{m,s}(t) \tag{22}$$

此外，无人机 m 和基站 k 的任务计算排队延迟 τ_m^{local}、$\overline{\tau}_{m,k}^{\mathrm{BS}}$ 的约束如下：

$$\tau_m^{\mathrm{local}}=\lim_{T\to\infty}\frac{1}{T}\sum_{t=1}^{T}\frac{Q_m^{\mathrm{local}}(t)}{\overline{C}_m(t)}\leqslant\tau_{m,\max}^{\mathrm{local}} \tag{23}$$

$$\overline{\tau}_{m,k}^{\mathrm{BS}}=\lim_{T\to\infty}\frac{1}{T}\sum_{t=1}^{T}\frac{S_{m,k}^{\mathrm{BS}}(t)}{\overline{C}_k^*(t)}\leqslant\overline{\tau}_{k,\max}^{\mathrm{BS}} \tag{24}$$

式中：$\tau_{m,\max}^{\mathrm{local}}$、$\overline{\tau}_{k,\max}^{\mathrm{BS}}$ 为无人机 m 和基站 k 的计算排队延迟上限；$\overline{C}_m(t)$、$\overline{C}_k^*(t)$为无人机 m 和基站 k 的任务计算序列 $Q_m^{\mathrm{local}}(t)$和 $S_{m,k}^{\mathrm{BS}}(t)$的平均任务到达率。$\overline{C}_m(t)$、$\overline{C}_k^*(t)$可分别表示如下：

$$\overline{C}_m(t)=\frac{1}{T}\sum_{t=0}^{T-1}\chi_m(t)D_m(t) \tag{25}$$

$$\overline{C}_k^*(t)=\frac{1}{T}\sum_{t=0}^{T-1}A_{m,k}(t)Y_m^{\mathrm{UAV}}(t) \tag{26}$$

1.6 能耗

本文的目标是最小化无人机网络的总能耗，针对无人机网络的能耗具体为本地计算、卸载到基站、卸载到 LEO 卫星以及飞行消耗。

1)本地计算。

此时，无人机自身对任务进行计算，无人机 m 的计算能耗 $e_m^{\mathrm{local}}(t)$可表示如下：

$$e_m^{\mathrm{local}}(t)=\kappa[f_m(t)]^3\min\left\{\frac{\rho_m Q_m^{\mathrm{local}}(t)}{f_m(t)},\delta_t\right\} \tag{27}$$

式中：κ 是与 CPU 有关的开关电容系数。

2)卸载到基站。

无人机 m 在基站 k 的覆盖范围内，无人机 m 将任务卸载到基站 k 的能耗可表示如下：

$$e_m^{\mathrm{BS}}(t)=A_{m,k}(t)P_{m,k}(t)\min\left\{\frac{Q_{m,k}^{\mathrm{BS}}(t)}{R_{m,k}(t)},\delta_t\right\} \tag{28}$$

式中：$P_{m,k}(t)$为无人机卸载到基站的功率。

3)卸载到 LEO 卫星。

在任务时间内，无人机 m 和 LEO 卫星之间的卸载能耗可表示如下：

$$e_m^{\mathrm{LEO}}(t)=A_{m,s}(t)P_{m,s}(t)\min\left\{\frac{Q_m^{\mathrm{LEO}}(t)}{R_{m,s}(t)},\delta_t\right\} \tag{29}$$

式中：$P_{m,s}(t)$为无人机卸载到卫星的功率。

4)无人机飞行。

根据文献[26]，无人机的飞行能耗可表示如下：

$$e_m^{\mathrm{fly}}(t)=c_1\left(1+\frac{3v_m(t)^2}{U_{\mathrm{tip}}^2}\right)+\frac{\rho_0}{2}\psi s\overline{A}v_m(t)^3+c_2\left(\sqrt{v_m(t)^{-4}+\frac{1}{4v_0^4}}-\frac{1}{2v_0^2}\right)^{\frac{1}{2}} \tag{30}$$

式中：c_1 和 c_2 为有关飞行的常数；U_{tip} 是转子叶片的叶尖速度；v_0 是平均悬停速度；ψ 是阻力比；s 是转子的坚固度；$\overline{A}$ 是转子盘面积；ρ_0 是空气密度。

1.7 网络可持续模型

当无人机电池电量降至阈值 Π_e 以下且在基站覆盖范围内时，基站可以通过 WPT 技术为无人机充电。采用二进制变量 $\overline{A}_{m,k}(t)\in\{0,1\}$表示基站在 t 时隙是否为无人机充电，$\overline{A}_{m,k}(t)=1$ 表示无人机成功连接到基站充电，反之 $\overline{A}_{m,k}(t)=0$。同时，无人机采用线性能量收集模型，无人机 m 在 t 时隙收获的能量 $E_m^{\mathrm{In}}(t)$可表示如下：

$$E_m^{\mathrm{In}}(t)=\sum_{m=1}^{M}\overline{A}_{m,k}(t)\eta_0 P_{\mathrm{BS}}G_{\mathrm{B}}^{\mathrm{tr}}G_{\mathrm{U}}^{\mathrm{re}}h_{m,k}(t)\delta_t \tag{31}$$

式中：$\eta_0\in(0,1]$为节能效率；P_{BS} 为传输功率；$G_{\mathrm{B}}^{\mathrm{tr}}$和 $G_{\mathrm{U}}^{\mathrm{re}}$ 分别是基站发射机和无人机接收机的天线功率增益。

设 $E_m(t)$表示无人机 m 在 t 时隙的能量队列积压，因此 $E_m(t)$的动态模型可表示如下：

$$E_m(t+1)=[E_m(t)-E_m^{\mathrm{cons}}(t)]^+ +E_m^{\mathrm{In}}(t) \tag{32}$$

式中：$E_m^{\mathrm{cons}}(t)$为无人机 m 在 t 时隙的能量消耗。$E_m^{\mathrm{cons}}(t)$可表示如下：

www.ecice00.com 计算机工程

$$E_m^{\text{cons}}(t)=e_m^{\text{local}}(t)+e_m^{\text{BS}}(t)+e_m^{\text{LEO}}(t)+e_m^{\text{fly}}(t) \tag{33}$$

$$e_m^{\text{local}}(t)+e_m^{\text{BS}}(t)+e_m^{\text{LEO}}(t)+e_m^{\text{fly}}(t)\leqslant E_m(t) \tag{34}$$

假设无人机的电池容量为 E，$E_m(0)=E$。由于无人机电池的能限，能量队列被限制如下：

$$0\leqslant E_m(t)\leqslant E \tag{35}$$

1.8 问题公式化

在卸载和计算排队延迟及充电约束的情况下，本文通过联合优化动态关联、任务卸载、带宽分配和计算资源分配来最小化无人机网络的能耗。问题 P1 可建模如下：

$$\min\sum_{t=1}^{T}\sum_{m=1}^{M}E_m^{\text{cons}}(t)$$

s. t. C1：$\chi_m(t)\in(0,1),\gamma_{m,k}(t)\in[0,1]$

$1-\chi_m(t)-\gamma_{m,k}(t)\in[0,1]$，

$\forall m,k,t$

C2：$A_{m,k}(t),A_{m,s}(t),\overline{A}_{m,k}(t)\in\{0,1\}$，

$\forall m,k,t$

C3：$\sum_{k=1}^{K}A_{m,k}(t)\leqslant 1,\forall m,t$

C4：$0\leqslant\alpha_{m,k}(t)\leqslant 1,\forall m,k$

C5：$0\leqslant f_m(t)\leqslant f_m^{\max}$，$0\leqslant f_{m,k}(t)\leqslant f_{m,k}^{\max}$，

$0\leqslant f_{m,s}(t)\leqslant f_{m,s}^{\max},\forall m,k,t$

C6：$\dfrac{\chi_m(t)D_m(t)}{f_m}+\tau_m^{\text{local}}\leqslant\mathcal{T},\ \forall m$

C7：$\dfrac{\gamma_{m,k}(t)D_m(t)}{R_{m,s}(t)}+\dfrac{\gamma_{m,k}(t)D_m(t)}{f_{m,k}}+\tau_k^{\text{BS}}+\overline{\tau}_k^{\text{BS}}\leqslant\mathcal{T},\ \forall m,k,t$

C8：$\dfrac{D_{m,s}(t)}{R_{m,s}(t)}+\tau_{m,s}^{\text{LEO}}+T_{\text{LEO}}\leqslant\mathcal{T},\ \forall m,k,t$

C9：Eq.(7),Eq.(8),Eq.(34),Eq.(35)

C10：Eq.(16),Eq.(17),Eq.(20),

Eq.(21),Eq.(24)

式中：C1 为任务分配约束；C2 和 C3 表示无人机在每个时隙只能选择一个无人机；C4 为无人机与基站的带宽分配比例；C5 为无人机的速度约束；C6 表示计算频率资源约束；C6～C8 为任务时间约束，T_{LEO} 为无人机与 LEO 卫星的传播时延；C9 为无人机飞行约束和电池能量约束；C10 表示任务卸载和计算的排队延迟约束。

2 算法

问题 P1 中长期排队延迟约束与短期决策的耦合性、任务的随机性、信道状态和任务缓冲队列的动态性使得无人机的任务卸载和资源分配十分具有挑战性。本文首先采用虚拟队列[24]，将长期延迟约束转换为队列稳定性约束，其次采用 Lyapunov 优化算法将原始问题转化为确定的时隙问题[27-30]，最后通过 DRL 算法将无人机建模为智能体，利用神经网络来近似值函数动作，从而降低状态空间维数的同时学习最优策略。

2.1 Lyapunov 优化算法

通过引入虚拟队列，将长期排队延迟约束转化为稳定性队列，因此任务卸载队列、本地计算队列和基站计算队列有如下表示：

$$U_m^{\text{BS}}(t+1)=\left\{U_m^{\text{BS}}(t)+\frac{Q_m^{\text{BS}}(t)}{\overline{C}_k(t)}-\tau_{k,\max}^{\text{BS}}\right\}^+ \tag{36}$$

$$U_m^{\text{LEO}}(t+1)=\left\{U_m^{\text{LEO}}(t)+\frac{Q_m^{\text{LEO}}(t)}{\overline{C}(t)}-\tau_{\max}^{\text{LEO}}\right\}^+ \tag{37}$$

$$G_m^{\text{local}}(t+1)=\left\{G_m^{\text{local}}(t)+\frac{Q_m^{\text{local}}(t)}{\overline{C}_m(t)}-\tau_{m,\max}^{\text{local}}\right\}^+ \tag{38}$$

$$G_m^{\text{BS}}(t+1)=\left\{G_m^{\text{BS}}(t)+\frac{Q_m^{\text{BS}}(t)}{\overline{C}_k(t)}-\tau_{k,\max}^{\text{BS}}\right\}^+ \tag{39}$$

为求解问题 P1，采用 Lyapunov 优化算法进行求解，Lyapunov 函数可表示如下：

$$L(X(t))=\frac{1}{2}\left\{\sum_{m=1}^{M}\left[Q_m^{\text{local}}(t)^2+Q_m^{\text{UAV}}(t)^2+Q_m^{\text{LEO}}(t)^2+\sum_{k=1}^{K}S_{m,k}^{\text{BS}}(t)^2+E_m(t)^2\right]\right\} \tag{40}$$

Lyapunov 漂移函数的定义是相邻时隙函数变化的期望，表示如下：

$$\Delta L(X(t))=\mathbb{E}[L(X(t+1))-L(X(t))\,|\,X(t)] \tag{41}$$

式中：$\Delta L(X(t))$的大小对队列稳定性至关重要。

为了确保队列稳定性的同时最小化无人机总能耗，考虑采用 Lyapunov 漂移加罚，其定义为 Lyapunov 漂移加上平均能耗的期望值乘以权重参数，表示如下：

$$\Delta_V L(X(t))=\Delta L(X(t))+V\mathbb{E}[\varepsilon(t)\,|\,X(t)] \tag{42}$$

式中：非负权重参数 V 权衡了“惩罚最小化”和“队列稳定性”，即最小化能耗和减少队列长度。

因此，尽可能地最小化 $\Delta_V L(X(t))$的上界可以实现最优的无人机能耗。在所有可能的 $\varepsilon(t)=\sum_{t=1}^{T}\sum_{m=1}^{M}E_m^{\text{cons}}(t)$和 $V\geqslant 0$ 的情况下，漂移加罚的上界可表示如下：

计算机工程 www.ecice06.com

$$\Delta_V L(X(t)) \leqslant C+\Lambda(t)+V\varepsilon(t)=$$
$$C+\sum_{m=1}^{M}\mathbb{E}[Z(t)\mid X(t)]+V\mathbb{E}[\varepsilon(t)\mid X(t)]=C+\sum_{m=1}^{M}\mathbb{E}[Q_m^{\text{local}}(t)(\chi_m(t)D_m(t)-Y_m^{\text{local}}(t))\mid X(t)]+$$
$$\sum_{m=1}^{M}\mathbb{E}[Q_m^{\text{UAV}}(t)(\gamma_{m,k}(t)D_m(t)-Y_m^{\text{UAV}}(t))\mid X(t)]+$$
$$\sum_{m=1}^{M}\mathbb{E}[Q_m^{\text{LEO}}(t)((1-\chi_m(t)-\gamma_{m,k}(t))-Y_m^{\text{LEO}}(t))\mid X(t)]+$$
$$\sum_{m=1}^{M}\sum_{k=1}^{K}\mathbb{E}[S_{m,k}^{\text{BS}}(t)(A_{m,k}(t)Y_m^{\text{UAV}}(t)-Y_m^{\text{UAV}}(t))\mid X(t)]+\sum_{m=1}^{M}\mathbb{E}\left[U_m^{\text{BS}}(t)\left(\frac{Q_m^{\text{BS}}(t)}{\overline{C}_k(t)}-\tau_{k,\max}^{\text{BS}}\right)\mid X(t)\right]+$$
$$\sum_{m=1}^{M}\mathbb{E}\left[\sum_{m=1}^{M}U_m^{\text{LEO}}(t)\left(\frac{Q_m^{\text{LEO}}(t)}{\overline{C}(t)}-\tau_{\max}^{\text{LEO}}\right)\mid X(t)\right]+\sum_{m=1}^{M}\mathbb{E}\left[G_m^{\text{local}}(t)\left(\frac{Q_m^{\text{local}}(t)}{\overline{C}_m(t)}-\tau_{m,\max}^{\text{local}}\right)\mid X(t)\right]+$$
$$\sum_{m=1}^{M}\mathbb{E}\left[G_m^{\text{BS}}(t)\left(\frac{Q_m^{\text{BS}}(t)}{\overline{C}_k(t)}-\tau_{k,\max}^{\text{BS}}\right)\mid X(t)\right]+\sum_{m=1}^{M}\mathbb{E}[E_m^{\text{UAV}}(t)(E_m^{\text{cons}}(t)-E_m^{\text{In}}(t))\mid X(t)]+$$
$$V\mathbb{E}[\varepsilon(t)\mid X(t)] \tag{43}$$

式中：C 为常数。

因此，问题 P1 可以通过最小化每个时隙的漂移加罚上限来获得。重新用问题 P2 表示如下：

$$\min V\varepsilon(t)+\sum_{m=1}^{M}Q_m^{\text{local}}(t)\left(\chi_m(t)D_m(t)-\frac{\delta_t f_m}{\rho_m}\right)+$$
$$\sum_{m=1}^{M}Q_m^{\text{UAV}}(t)(\gamma_{m,k}(t)D_m(t)-Y_m^{\text{UAV}}(t))+$$
$$\sum_{m=1}^{M}Q_m^{\text{LEO}}(t)((1-\chi_m(t)-\gamma_{m,k}(t))-Y_m^{\text{LEO}}(t))+$$
$$\sum_{m=1}^{M}\sum_{k=1}^{K}S_m^{\text{UAV}}(t)\left(A_{m,k}(t)Y_m^{\text{UAV}}(t)-\frac{\delta_t f_{m,k}}{\rho_m}\right)+$$
$$\sum_{m=1}^{M}U_m^{\text{BS}}(t)\left(\frac{Q_m^{\text{BS}}(t)}{\overline{C}_k(t)}-\tau_{k,\max}^{\text{BS}}\right)+$$
$$\sum_{m=1}^{M}U_m^{\text{LEO}}(t)\left(\frac{Q_m^{\text{LEO}}(t)}{\overline{C}(t)}-\tau_{\max}^{\text{LEO}}\right)+$$
$$\sum_{m=1}^{M}G_m^{\text{local}}(t)\left(\frac{Q_m^{\text{local}}(t)}{\overline{C}_m(t)}-\tau_{m,\max}^{\text{local}}\right)+$$
$$\sum_{m=1}^{M}G_m^{\text{BS}}(t)\left(\frac{Q_m^{\text{BS}}(t)}{\overline{C}_k(t)}-\tau_{k,\max}^{\text{BS}}\right)+$$
$$\sum_{m=1}^{M}E_m^{\text{UAV}}(t)(E_m^{\text{cons}}(t)-E_m^{\text{In}}(t))$$
$$\text{s.t. C1}\sim\text{C10}$$

2.2 双延迟深度确定性策略梯度(TD3PG)算法

由于未知的动态网络环境以及高维动作空间中含有连续变量，因此需要一种能够感知网络参数变化并动态更新卸载决策使目标函数最小化的算法。本文提出了 TD3PG 算法来与环境交互，学习奖励最大化策略，使目标函数最优化。

与传统的 DDPG 算法和双重深度 Q 网络(DDQN)算法相比，TD3PG 算法有两个评价网络，解决了 DDPG 算法和 DDQN 算法的过估计问题，同时通过目标策略平滑正则化来提高目标值的准确性，以保证网络训练更加稳定。此外，延迟策略技术的应用使 TD3PG 动作网络更新的频率低于评价网络，产生较小的方差评估值，从而获得更好的策略。

本文将优化目标表述为马尔可夫决策过程，用元组表示为(Z,a,R_e,Z')，其中，Z 为当前时隙的状态集，a 为有限的动作集，R_e 是智能体获得的奖励，Z'为下一时隙无人机的状态。

1)状态空间：状态空间包括无人机剩余能量、无人机位置、速度以及无人机到达的任务数量。系统状态空间可定义如下：

$$Z_m(t)\in\{q_m(t),E_m(t),V_m(t),D_m(t)\} \tag{44}$$

2)动作空间：智能体基于当前的状态决定执行的动作，其中包含卸载决策、充电决策、卸载任务量、计算资源、带宽资源比例、无人机飞行速度等。无人机的动作空间可表示如下：

$$a_m(t)\in\{A_{m,k}(t),A_{m,s}(t),\overline{A}_{m,k}(t),\chi_m(t),\gamma_{m,k}(t),f_m(t),\alpha_{m,k}(t),\upsilon_m(t)\} \tag{45}$$

3)奖励：奖励是执行当前动作后对智能体的反馈。奖励函数定义如下：

$$R_{e,m}[n]=-\Lambda(t)-V\varepsilon(t) \tag{46}$$

如图 2 所示，与传统的基于动作-评价的 DDPG 算法相比，TD3PG 架构包含一个动作网络和两个评价网络。

TD3PG 算法应用平滑技术对目标策略进行平滑值估计，该技术将削波随机噪声添加到目标动作中并进行小批量平均，可表示如下：

$$\varepsilon\sim\text{clip}(N(0,\bar{\omega}),-c,c) \tag{47}$$

根据 TD3PG 算法结构，动作网络参数由 θ_π 表示，两个评价网络分别由参数 θ_{Q1} 和 θ_{Q2} 表示。

计算机工程 www.ecice06.com

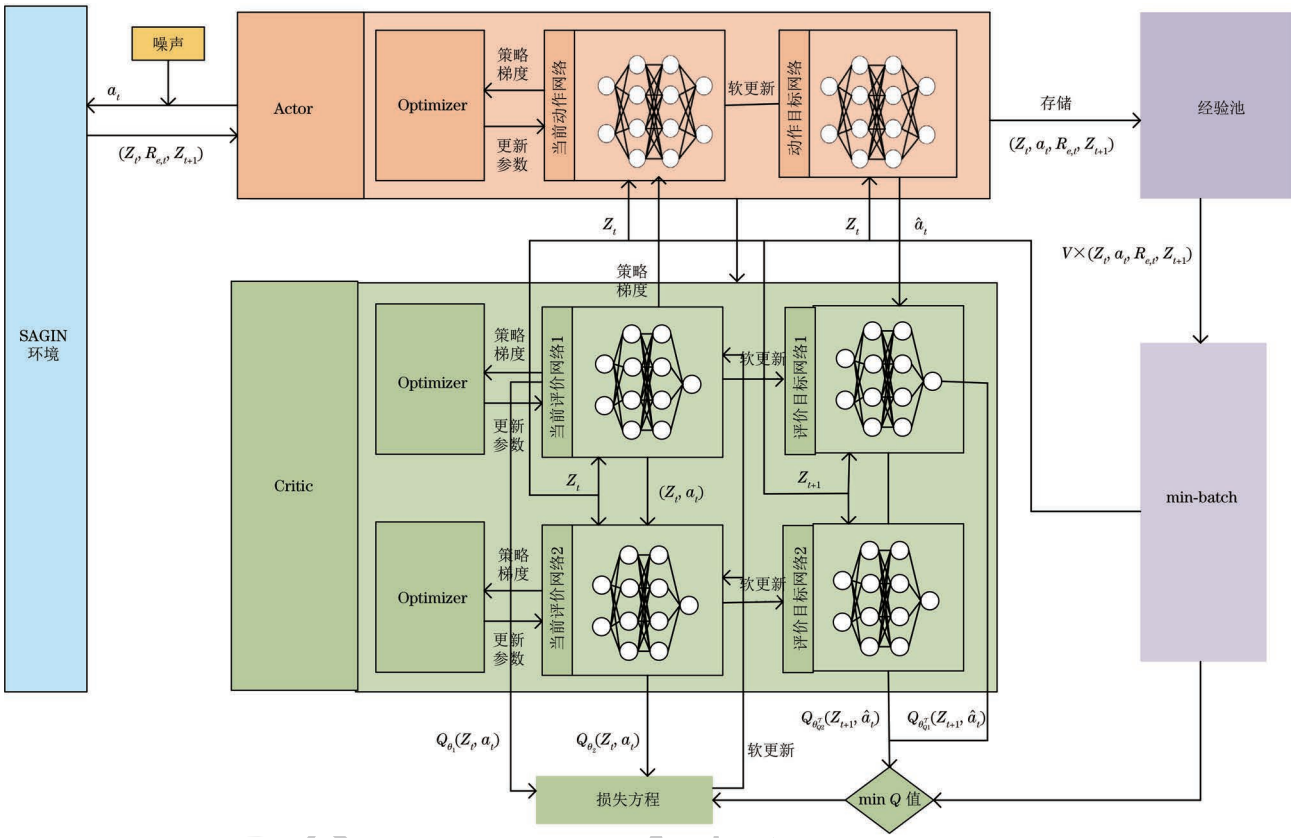


图 2 TD3PG 框架结构

Fig. 2 TD3PG framework structure

智能体首先观察当前状态 Z_t，并通过动作网络选择动作 a_t，执行动作 a_t 后智能体获得即时奖励 $R_{e,t}$ 和下一状态 Z_{t+1}，然后智能体选择使累积奖励最大化的策略，获得$(Z_t, a_t, R_{e,t}, Z_{t+1})$并将其存储在经验池中用于进一步采样。

针对算法中的目标值 V_t，可通过使用两个评论网络的最小 Q 值估计来计算，可表示如下：

$$V_t = R_{e,t} + \gamma \min_{i=1,2} Q_{\theta_{Qi}}(Z_t, (\pi(Z_t|\theta_\pi^T)+\varepsilon)|\theta_{Qi}^T) \tag{48}$$

式中：$\gamma \in [0,1]$为折扣因子。

两个评论网络的损失函数可定义如下：

$$L'(\theta_{Qi}) = \mathbb{E}[(V_t - Q_{\theta_{Qi}}(Z_t, a_t|\theta_{Qi}))^2], i=1,2 \tag{49}$$

损失函数的梯度可表示如下：

$$\nabla_{\theta_{Qi}} L'(\theta_{Qi}) = \mathbb{E}[(V_t - Q(Z_t, a_t|\theta_{Qi}))\nabla_{\theta_{Qi}} Q(Z_t, a_t|\theta_{Qi})], \quad i=1,2 \tag{50}$$

动作网络的策略梯度可表示如下：

$$\nabla_{\theta_\pi} J \approx \mathbb{E}[(\nabla_{a_t} Q(Z_t, a_t|\theta_{Q1})|_{a_t=\pi(Z_t|\theta_\pi)} \nabla_{\theta_\pi}\pi(Z_t|\theta_\pi))] \tag{51}$$

目标网络的更新可表示如下：

$$\theta_\pi^T \leftarrow \omega\theta_\pi + (1-\omega)\theta_\pi^T \tag{52}$$

$$\theta_{Qi}^T \leftarrow \omega\theta_{Qi} + (1-\omega)\theta_{Qi}^T, i=1,2 \tag{53}$$

式中：ω 是动作网络和评论网络的衰减率。

算法 1 TD3PG 算法

输入 最大训练集长度 N_{NUM}，折扣系数 γ

输出 损失函数、梯度

1. 初始化 $\theta_\pi^T \leftarrow \theta_\pi$、$\theta_{Q1}^T \leftarrow \theta_{Q1}$、$\theta_{Q2}^T \leftarrow \theta_{Q2}$ 以及经验池 $\mathcal{R}$
2. for Episode$\in\{1,2,\cdots,N_{NUM}\}$
3. 获取状态 Z
4. for t=1 to T
5. 添加探索噪声后采取行动 $a_t \sim \pi(Z_t|\theta_\pi) + N(0,\delta)$
6. 获得奖励 $R_{e,t}$ 并更新下一时刻状态 Z_{t+1}
7. 将$(Z_t, a_t, R_{e,t}, Z_{t+1})$存储在经验池中
8. 从经验池 $\mathcal{R}$ 中抽样$(Z_t, a_t, R_{e,t}, Z_{t+1})$用于训练
9. 通过 $\hat{a}_t \sim \pi(Z_{t+1}|\theta_\pi^T) + \varepsilon$ 计算下一个状态的贪婪动作
10. 计算目标值 V_t
11. 通过最小化损失函数方程更新评论者网络
12. 使用采样的策略梯度方程更新动作网络参数 θ_π
13. 更新目标网络
14. end for
15. end for

2.3 复杂度分析

本文的复杂度来源于 TD3PG 算法，其主要取决于神经网络的大小，假设动作网络有 $\bar{L}$ 层，其中第 $\bar{l}$ 层有 $n_{\bar{l}}$ 个神经元，因此动作网络的复杂度为 $O\left(\sum_{\bar{l}=2}^{\bar{L}-1}(n_{\bar{l}-1}n_{\bar{l}}+n_{\bar{l}}n_{\bar{l}+1})\right)$。同理，假设评价网络有 $\hat{L}$ 层，其中第 $\hat{l}$ 层有 $n_{\hat{l}}$ 个神经元，评价网络的复杂度为 $O\left(\sum_{\hat{l}=2}^{\hat{L}-1}(n_{\hat{l}-1}n_{\hat{l}}+n_{\hat{l}}n_{\hat{l}+1})\right)$。因此，本文算法的总复杂度为 $O\left(\sum_{\bar{l}=2}^{\bar{L}-1}(n_{\bar{l}-1}n_{\bar{l}}+n_{\bar{l}}n_{\bar{l}+1})+\sum_{\hat{l}=2}^{\hat{L}-1}(n_{\hat{l}-1}n_{\hat{l}}+n_{\hat{l}}n_{\hat{l}+1})\right)$。

3 仿真与分析

本文采用 Python 进行仿真，考虑了一个 2 000 m×2 000 m 的区域，其中包含 2 个基站，2 架无人机和 1 颗 LEO 卫星，且每个计算任务的数据大小服从均匀分布。其他仿真参数设置如表 2 所示。

表 2 实验仿真参数

Table 2 Experimental simulation parameters

参数名	符号	参数值
飞行高度/m	H	100
最大飞行速度/(m·s^{-1})	$V_{\max}$	50
基站带宽/MHz	B	20
叶尖速度/(m·s^{-1})	H_{tip}	120
空气密度/(kg·m^{-3})	ρ_0	1.225
转子盘面积/m^2	$\bar{A}$	0.503
噪声功率谱密度/(dB·Hz^{-1})	σ	−170
平均悬停速度/(m·s^{-1})	v_0	4.03
阻力比	ψ	0.6
转子坚固度	$\bar{s}$	0.05
无人机飞行半径/m	r_{u}	600
基站半径/m	r_{BS}	300

为了证明本文算法的优势，采用以下对比算法：

1)无 WPT 技术的 TD3PG 算法。

2)DDPG 算法。

3)DDQN 算法。

4)固定带宽算法：在任意时隙，基站将根据连接的无人机数量均匀分配带宽。

5)贪婪优化算法：在任意时隙，每架无人机到达的任务按照本地计算∶基站计算∶低轨卫星上计算=0.1∶0.8∶0.1 进行卸载分配，基站和卫星将根据连接的无人机数量均匀分配计算资源。

6)卸载卫星算法：在任意时隙，每架无人机到达的任务选择本地处理和卸载到卫星。

7)卸载基站算法：在任意时隙，每架无人机到达的任务选择本地处理和卸载到基站。

8)随机算法：在任意时隙，每架无人机到达的任务完全本地处理或完全卸载到基站或完全卸载到卫星。

图 3 展示了 TD3PG 算法、DDPG 算法和 DDQN 算法在不同学习率组合情况下对网络累积奖励的影响，其中，α_π、α_μ 分别表示动作网络和评价网络的学习率。由图 3 可知，在初始时网络的累积奖励值随着不同的学习率组合而波动，但随着训练的增加，所有方案下的累积奖励都逐渐收敛，其中 TD3PG 算法的累积奖励最高，DDQN 算法的累积奖励最低，在 2 种学习率组合下 TD3PG 算法相对于 DDPG 算法的奖励分别提高了 13.6%、24.0%，相对于 DDQN 算法的累积奖励分别提高了 20.4%、12.7%，证明了 TD3PG 算法在任务卸载和计算方面的优越性。此外，在所有方案中，当学习率为 0.000 1 时累积奖励最高，这说明了低学习率有助于收敛到最优，因而网络的性能最好。

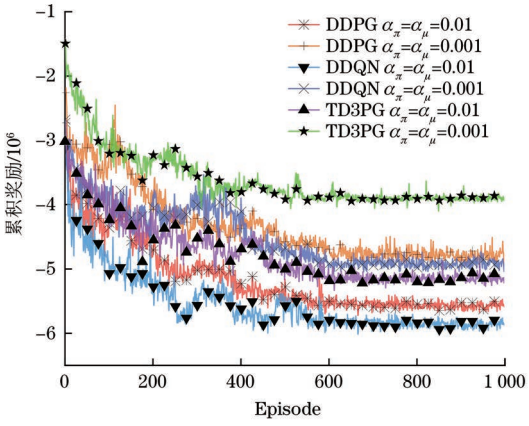


图 3 不同 Episode 的网络累积奖励

Fig.3 Cumulative network rewards of different Episodes

图 4 为不同时隙的队列延迟积压总和。由图 4 可知，随着时隙的增加，TP3PG 算法、无 WPT 的 TP3PG 算法、DDPG 算法和 DDQN 算法的延迟积压总和在逐渐增大后趋于稳定，达到动态平衡，同时 TP3PG 算法的队列积压总和相比无 WPT 的 TP3PG 算法、DDPG 算法和 DDQN 算法分别低了 5.8%、8.5%和 17.9%，证明了 TD3PG 算法和 WPT 技术能够有效改善网络性能。但是，其他算法的队列积压随时隙的增加而逐渐增大，其中：卸载基站算法的队列积压最高，这是由于偏远地区基站数量有限，使得网络的任务积压；卸载卫星算法由于卫星提供无时无刻的覆盖，因此积压相对较少；随机算法由于是单一卸载，因此任务积压相对较多；固定

www.ecjs.com.cn 计算机工程

带宽和贪婪优化算法的任务和带宽为固定分配，导致基站的任务量过载以及无人机任务的积压，从而使整个网络的积压总和持续增加。

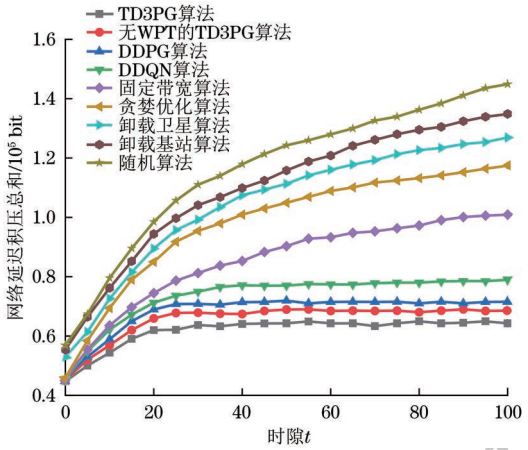


图 4 不同时隙的网络延迟积压总和

Fig. 4 Total network delay backlogs of different time slots

图 5 展示了不同算法下无人机最大计算资源与其总能耗的关系。由图 5 可知，各种算法的能耗随着无人机的最大计算资源的增加而增加，但总体而言，TD3PG 算法的无人机能耗明显小于其他对比算法，证明了 TD3PG 算法的优越性能，同时证明了 WPT 技术能够降低无人机网络的能耗。随着无人机最大计算资源的增大，DDPG 算法和 DDQN 算法的能耗逐渐大于 TD3PG 算法，这是因为 TD3PG 算法通过两个独立评论网络以及使用延迟学习，有效地提高了网络性能并缓解了 DDPG 算法的过估计，从而能够有效降低无人机网络的总能耗。此外，随机算法的能耗明显高于其他算法，这是因为任务卸载的随机性导致无人机的能耗增大。

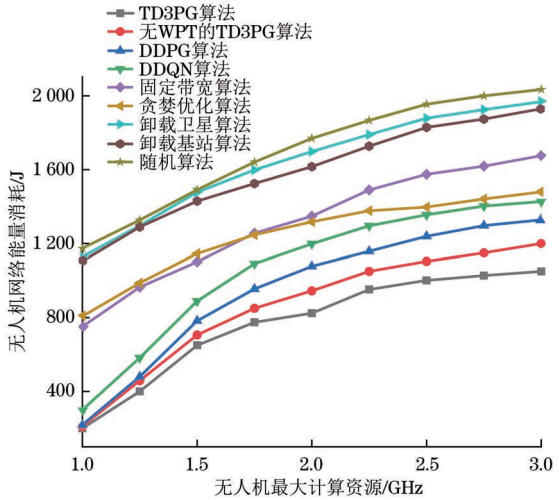


图 5 无人机最大计算能力下无人机网络能量消耗

Fig. 5 UAV network energy consumption under UAV maximum computing ability

到达的任务量与无人机网络的能耗关系如图 6 所示。由图 6 可知，无人机的能耗随着到达任务量的增加而增加，其中，TP3PG 算法的能量消耗最少，随机算法的能量消耗最多，卸载卫星算法和卸载基站算法次之，这是因为卸载的单一性和固定性不能适应网络的动态变化，导致网络能耗增大。此外，当卸载任务时，固定带宽算法中无人机因卸载能力受限导致所连接的基站无法在较短的时间内完成卸载，从而导致无人机的计算任务增多，无人机的能耗快速升高。贪婪优化算法将大量任务卸载到基站上，降低了本地处理的任务量，优化了本地计算资源分配和关联控制，因此无人机的能耗可以保持在较低水平。总体而言，本文提出的动态卸载方案优于其他对比方案。

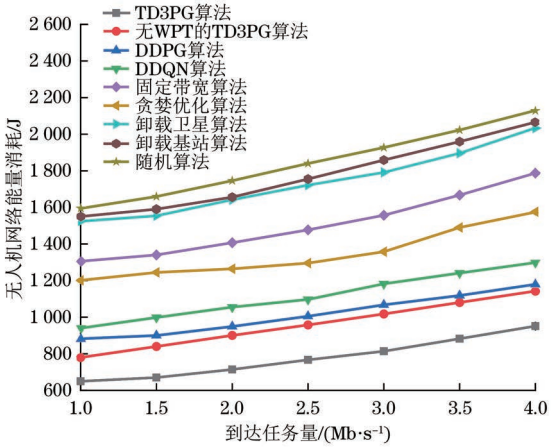


图 6 到达任务量对无人机网络能耗的影响

Fig. 6 Effect of task arrivals on UAV network energy consumption

不同算法下无人机的任务排队延迟如图 7 所示。数值结果表明，TD3PG 算法、无 WPT 的 TD3PG算法、DDPG算法和DDQN算法的排队延

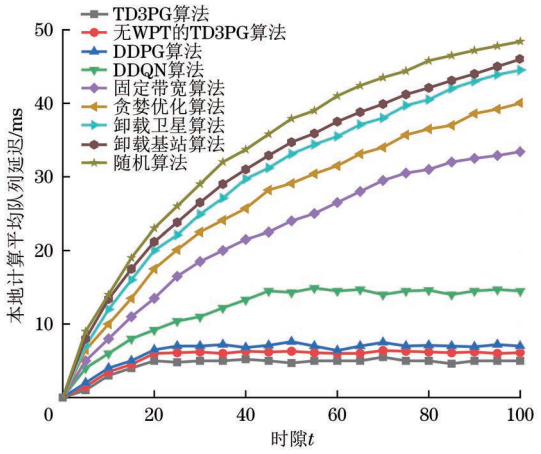


图 7 不同时隙下本地计算的平均排队延迟

Fig. 7 Average queue delay of local computing at different time slots

迟先增加后趋于稳定，证明了网络的稳定性以及较好的排队延迟性能，同时证明了 WPT 技术的优越性。此外，其他算法的排队延迟随着时间的增加逐渐增大，其中：随机算法因卸载单一，平均队列延迟最大；卸载基站算法和贪婪优化算法因基站数量有限，排队延迟相对较大；卸载卫星算法因卫星提供无时无刻的覆盖，延迟得到一定的缓解；固定带宽算法受到卸载传输的限制，本地的任务量相对增加，因而具有一定的延迟。

4 结束语

本文为缓解偏远地区 IoT 设备面对计算密集型任务的压力，提出一种 SAG-MEC 网络下无人机动态任务计算框架，采用部分卸载模式通过优化无人机动态关联、任务分配、计算资源分配和带宽分配最小化无人机能耗，同时采用新型 WPT 充电技术来维持网络的可持续性。结合该优化问题长期队列延迟约束与短期决策的耦合性、任务的随机性、信道状态和任务缓冲队列的动态性，提出基于 Lyapunov 优化的 TD3PG 算法，在排队延迟的约束下，首先在获得网络稳定性的同时引入虚拟队列将长期队列约束转化为稳定性约束，其次利用 Lyapunov 算法将目标函数进行转化，最后通过协调无人机利用 TD3PG 算法来学习最优卸载资源分配策略以增加系统的回报，从而最小化无人机网络的能耗。仿真结果表明，该算法在保持队列稳定性的同时能够最大限度地降低能耗和队列延迟。未来，本团队将进一步研究 SAGIN 中无人机与多个 LEO 卫星切换的动态任务卸载问题。

参考文献

[1] BEDI G, VENAYAGAMOORTHY G K, SINGH R, et al. Review of Internet of Things (IoT) in electric power and energy systems[J]. IEEE Internet of Things Journal, 2018, 5(2): 847-870.

[2] WANG H C, WANG J L, DING G R, et al. Robust spectrum sharing in air-ground integrated networks: opportunities and challenges [J]. IEEE Wireless Communications, 2020, 27(3): 148-155.

[3] SAHOO B P S, PUTHAL D, SHARMA P K. Toward advanced UAV communications: properties, research challenges, and future potential[J]. IEEE Internet of Things Magazine, 2022, 5(1): 154-159.

[4] LI Z D, WANG Y, LIU M, et al. Energy efficient resource allocation for UAV-assisted space-air-ground Internet of remote things networks[J]. IEEE Access, 2019, 7: 145348-145362.

[5] AI B, MOLISCH A F, RUPP M, et al. 5G key technologies for smart railways [J]. Proceedings of the IEEE, 2020, 108(6): 856-893.

[6] JIA Z Y, SHENG M, LI J D, et al. LEO-satellite-assisted UAV: joint trajectory and data collection for Internet of remote things in 6G aerial access networks[J]. IEEE Internet of Things Journal, 2020, 8(12): 9814-9826.

[7] 刘亮，毛武平，李汶蔚，等. 空天地一体化边缘计算网络中基于博弈论的任务卸载策略[J]. 计算机工程，2025，51(2)：238-249.
LIU L, MAO W P, LI W W, et al. Task offloading strategy based on game theory in the space-air-ground integrated edge computing networks [J]. Computer Engineering, 2025, 51(2): 238-249. (in Chinese)

[8] TANG Q Q, LI B. Overview of mobile edge computing in space-air-ground integrated network [J]. Radio Communications Technology, 2021, 47(1): 25-35.

[9] CUI H X, ZHANG J, GENG Y H, et al. Space-Air-Ground Integrated Network (SAGIN) for 6G: requirements, architecture and challenges [J]. China Communications, 2022, 19(2): 90-108.

[10] JI B F, WANG Y N, SONG K, et al. A survey of computational intelligence for 6G: key technologies, applications and trends[J]. IEEE Transactions on Industrial Informatics, 2021, 17(10): 7145-7154.

[11] LIU J J, SHI Y P, FADLULLAH Z M, et al. Space-air-ground integrated network: a survey [J]. IEEE Communications Surveys & Tutorials, 2018, 20(4): 2714-2741.

[12] MAO B M, TANG F X, KAWAMOTO Y, et al. Optimizing computation offloading in satellite-UAV-served 6G IoT: a deep learning approach[J]. IEEE Network, 2021, 35(4): 102-108.

[13] 李晓青，贺占权，周卫彤. 支持 MEC 的天地一体化网络下任务卸载和资源分配[J]. 计算机应用与软件，2023，40(2)：130-137.
LI X Q, HE Z Q, ZHOU W T. Task offloading and resource allocation for the MEC enabled integrated satellite-terrestrial network[J]. Computer Applications and Software, 2023, 40(2): 130-137. (in Chinese)

[14] LIU J Y, ZHAO X W, QIN P, et al. Joint dynamic task offloading and resource scheduling for WPT enabled space-air-ground power Internet of Things[J]. IEEE Transactions on Network Science and Engineering, 2021, 9(2): 660-677.

[15] TANG Q Q, FEI Z S, LI B, et al. Computation offloading in LEO satellite networks with hybrid cloud and edge computing [J]. IEEE Internet of Things Journal, 2021, 8(11): 9164-9176.

[16] ZHANG S B, LIU A J, HAN C, et al. Multiagent reinforcement learning-based orbital edge offloading in SAGIN supporting Internet of remote things [J]. IEEE Internet of Things Journal, 2023, 10(23): 20472-20483.

[17] CHEN Y L, AI B, NIU Y, et al. Energy-constrained computation offloading in space-air-ground integrated networks using distributionally robust optimization[J]. IEEE Transactions on Vehicular Technology, 2021, 70(11): 12113-12125.

[18] ZHOU C H, WU W, HE H L, et al. Deep reinforcement learning for delay-oriented IoT task scheduling in SAGIN[J]. IEEE Transactions on Wireless Communications, 2021, 20(2): 911-925.

[19] ZHOU C H, WU W, HE H L, et al. Delay-aware IoT task scheduling in space-air-ground integrated network [C]// Proceedings of the IEEE Global Communications Conference. Washington D.C., USA: IEEE Press, 2019: 1-9.

[20] 李斌，刘文帅，费泽松. 面向空天地异构网络的边缘计算部分任务卸载策略[J]. 电子与信息学报，2022，44(9)：3091-3098.
LI B, LIU W S, FEI Z S. Partial computation offloading for mobile edge computing in space-air-ground integrated network [J]. Journal of Electronics & Information

www.ecice06.com

Technology, 2022, 44(9): 3091-3098. (in Chinese)
[21] LIU Y D, LI B, YANG Y X, et al. Joint delay and completion rate optimization for dependent task offloading in space-air-ground integrated network[C]//Proceedings of the 9th International Conference on Computer and Communications (ICCC). Washington D. C., USA: IEEE Press, 2023: 327-331.
[22] ZHENG K C, JIANG G D, LIU X Y, et al. DRL-based offloading for computation delay minimization in wireless-powered multi-access edge computing[J]. IEEE Transactions on Communications, 2023, 71(3): 1755-1770.
[23] WANG Y H, LI B, HE J H, et al. Joint latency-oriented, energy consumption, and carbon emission for a space-air-ground integrated network with newly designed power technology[J]. Electronics, 2023, 12(17): 3537.
[24] QIN P, FU Y, XIE Y B, et al. Multi-agent learning-based optimal task offloading and UAV trajectory planning for AGIN-power IoT [J]. IEEE Transactions on Communications, 2023, 71(7): 4005-4017.
[25] LITTLE J D C. A proof for the queuing formula: $L=\lambda W$ [J]. Operations Research, 1961, 9(3): 383-387.
[26] ZHANG Z B, LI X H, AN J P, et al. Model-free attitude control of spacecraft based on PID-guide TD3 algorithm[J]. International Journal of Aerospace Engineering, 2020(1): 8874619.
[27] KUMAR A S, ZHAO L, FERNANDO X. Task offloading and resource allocation in vehicular networks: a Lyapunov-based deep reinforcement learning approach [J]. IEEE Transactions on Vehicular Technology, 2023, 72(10): 13360-13373.
[28] MA G F, WANG X W, HU M J, et al. DRL-based computation offloading with queue stability for vehicular-cloud-assisted mobile edge computing systems[J]. IEEE Transactions on Intelligent Vehicles, 2023, 8(4): 2797-2809.
[29] WU Q Q, ZENG Y, ZHANG R. Joint trajectory and communication design for multi-UAV enabled wireless networks [J]. IEEE Transactions on Wireless Communications, 2018, 17(3): 2109-2121.
[30] DENG D, LI J X, JHAVERI R H, et al. Reinforcement-learning-based optimization on energy efficiency in UAV networks for IoT[J]. IEEE Internet of Things Journal, 2023, 10(3): 2767-2775.

文字编辑 陆燕菲
栏目编辑 宋 圆

计算机工程 www.ecice06.com