

# LEO-RAN 切片场景联合用户关联和动态资源分配算法

陈赓<sup>1</sup>, 邢治薇<sup>1</sup>, 沈斐<sup>2</sup>, 曾庆田<sup>1</sup>

(1. 山东科技大学电子信息工程学院, 山东 青岛 266590; 2. 中国科学院上海微系统与信息技术研究所, 上海 200050)

**摘要:** 为了解决 6G 天地一体化网络的资源高效利用问题, 提出了一种面向高密度低地球轨道卫星-无线接入网 (LEO-RAN) 切片场景的联合用户关联和动态资源分配算法。考虑不同切片的最小速率、最大时延及资源比例等约束, 以频谱效率 (SE) 和不同切片服务水平协议 (SLA) 满意率 (SSR) 的加权和作为优化目标, 建立用户关联和资源分配的联合优化问题。首先设计基于多智能体深度确定性策略梯度 (MADDPG) 的网络切片算法确定切片资源比例, 然后采用基于拉格朗日对偶的用户关联算法确定最优的用户关联策略, 最后通过轮询调度机制将资源分配给用户。仿真结果表明, 所提算法在满足不同切片差异化 SLA 的同时能够有效提高 SE。与基于 MADDPG-RA、MATD3-LG、MATD3-RA、MASAC-LG 和 MASAC-RA 算法相比, 所提算法系统效用分别提升了 2.0%、2.3%、5.7%、8.7% 和 9.4%。

**关键词:** LEO 卫星通信; 网络切片; 用户关联; 多智能体深度确定性策略梯度; 系统效用

**中图分类号:** TN929.5

**文献标志码:** A

**DOI:** 10.11959/j.issn.1000-436x.2024041

## Joint user association and dynamic resource allocation algorithm for LEO-RAN slicing scenarios

CHEN Geng<sup>1</sup>, XING Zhiwei<sup>1</sup>, SHEN Fei<sup>2</sup>, ZENG Qingtian<sup>1</sup>

1. College of Electronic and Information Engineering, Shandong University of Science and Technology, Qingdao 266590, China

2. Shanghai Institute of Microsystem and Information Technology, Chinese Academy of Sciences, Shanghai 200050, China

**Abstract:** A joint user association and dynamic resource allocation algorithm was proposed for the slicing scenario of ultra dense low earth orbit-radio access network (LEO-RAN) in order to address the efficient utilization of resources of the integrated terrestrial-satellite network for 6G. Considering the constraints of the minimum rate, maximum delay and resource proportion of different slices, a joint optimization problem of user association and resource allocation was established to maximize the weighted sum of the SE and the differentiated SLA of different slices as the optimization objective. A network slicing algorithm based on multi-agent deep deterministic policy gradient (MADDPG) was proposed to determine the proportion of slicing resources, a Lagrange dual based user association algorithm was proposed to determine the optimal user association policy and the resources were allocated to users by using the round-robin scheduling mechanism. The simulation results show that the proposed algorithm can effectively improve SE while satisfying the differentiated SLA of different slices. Compared with MADDPG-RA, MATD3-LG, MATD3-RA, MASAC-LG and MASAC-RA algorithms, the system utility of the proposed algorithm is improved by 2.0%, 2.3%, 5.7%, 8.7% and 9.4%, respectively.

**Keywords:** LEO satellite communication, network slicing, user association, MADDPG, system utility

**收稿日期:** 2023-10-30; **修回日期:** 2023-12-26

**通信作者:** 曾庆田, qtzeng@sdust.edu.cn

**基金项目:** 国家自然科学基金资助项目 (No.61701284); 山东省自然科学基金资助项目 (No.ZR2022MF226); 山东科技大学青年教师人才培养计划基金资助项目 (No.BJ20221101); 青岛市应用基础研究计划基金资助项目 (No.19-6-2-1-cg); 山东科技大学菁英计划基金资助项目 (No.skr21-3-B-048); 中国科学院百人计划基金资助项目 (No.E06BRA1001); 山东省泰山学者计划基金资助项目 (No.ts20190936)

**Foundation Items:** The National Natural Science Foundation of China (No.61701284), The Natural Science Foundation of Shandong Province (No.ZR2022MF226), The Talented Young Teachers Training Program of Shandong University of Science and Technology (No.BJ20221101), The Innovative Research Foundation of Qingdao (No.19-6-2-1-cg), The Elite Plan Project of Shandong University of Science and Technology (No.skr21-3-B-048), The Hundred Talent Program of Chinese Academy of Sciences (No.E06BRA1001), The Taishan Scholar Program of Shandong Province (No.ts20190936)

## 0 引言

6G 新型用户设备和服务种类的增加为人们的生活提供了极大的便利。然而,当大规模设备连接到网络时,会引发对高带宽和计算性能的需求<sup>[1]</sup>。这给地面移动通信服务网络带来了巨大的挑战,传统的解决方案是通过增加地面基站(BS, base station)的数量来提升地面网络的通信能力。然而,由于网络容量和覆盖范围的限制,尤其是在环境恶劣的地区(如山区和沙漠),仅依靠地面网络无法有效提供高数据速率和可靠性的无线接入服务<sup>[2]</sup>。

由于低轨道高度和较小的信号衰减,低地球轨道(LEO, low earth orbit)卫星成为提供高带宽互联网连接的一种补充方案<sup>[3]</sup>。目前,已经成功发射的LEO卫星星座包括 OneWeb 和 SpaceX 的星链。OneWeb 计划在 2027 年之前建立一个拥有 600 多颗低轨卫星的星座,提供全球高速互联网接入服务。而 SpaceX 星链在 2021 年中期已经发射了 1 600 多颗低轨卫星,并计划在不久的将来再次发射数千颗卫星。具有低成本、低时延、高吞吐量特点的 LEO 卫星可以与基站通信,并提供无缝通信服务,这进一步推动了地面-卫星网络(TSN, terrestrial-satellite network)的发展<sup>[4]</sup>。

传统的多层地面蜂窝架构中基础设施都是同质的,然而在 TSN 场景中存在垂直平面特征。该场景中除了能够进行数据包传输的地面基站之外,还部署了覆盖范围较广阔的 LEO 卫星<sup>[5-6]</sup>。这 2 种基站的共存导致了 LEO-无线接入网(RAN, radio access network)的异构性。研究者已经提出关于在 TSN 和 RAN 中的资源分配的研究工作。文献[5]研究 LEO 卫星的收益最大化问题,提出一种面向卫星的限制性三面匹配算法来解决用户、高空平台站(HAPS, high-altitude platform station)和卫星之间的匹配问题。针对 LEO 边缘辅助的多层边缘计算系统,文献[6]采用交替优化算法分解原始优化问题,然后采用低复杂度迭代算法求解每个子问题,所提的卸载方案能够实现较低的计算时延和能量消耗。文献[7]研究了超密集(UD, ultra-dense)-LEO 卫星辅助的 TSN 中的云和边缘计算,将能耗最小化问题分解为用户关联、任务调度及计算资源分配子问题。针对 TSN,文献[8]提出了层次马尔可夫决策过程框架,该框架使用 2 个独立的智能体处理 2 个时间维度的网络管理决策。文献[9]提出一种

双时间尺度算法来解决 RAN 切片缓存、频谱和功率资源分配优化问题,仿真结果表明,该算法能在在保证不同切片用户的服务质量(QoS, quality of service)需求的同时提高频谱利用率。文献[10]针对动态网络场景中切片资源需求不确定的情况下最大化网络基础设施提供商的总收益及提供满足切片资源需求概率保证的问题展开研究,提出了一种次优方法,按顺序解决多个切片资源需求问题。针对空-天-地车联网场景,文献[11]提出了一种深度强化学习(DRL, deep reinforcement learning)算法来求解计算卸载决策和资源分配的联合优化问题。

除了上述提到的 TSN 和 RAN 切片技术的独立研究,还有一些综合考虑了空天地网络、TSN 中的 RAN 切片和资源分配问题的相关研究。针对空-天-地车联网场景,文献[12]提出了一种在线控制框架,采用李雅普诺夫优化理论动态分配频谱资源以提供独立的车载服务。文献[13]研究了基于切片的空天地车载网络协同任务卸载框架,将 RAN 切片和任务卸载的系统决策建模为任务完成次数最大化问题并将其分解为 3 个子问题,采用传统优化算法和 DRL 算法来求解子问题。针对卫星-地面车联网,文献[14]采用 DRL 算法将频谱资源预先分配给不同的切片,然后使用匹配算法将切片中的资源调度给车辆,该方案在满足服务质量的前提下有效降低了系统成本。文献[15]提出了一种基于多智能体深度确定性策略梯度(MADDPG, multi-agent deep deterministic policy gradient)的分布式管理方案,该方案可以快速进行车辆关联决策并为车辆分配多维资源,实现较高 QoS。文献[16]设计了基于动态边缘计算环境的多智能体双时延深度确定性策略梯度(MATD3, multi-agent twin delayed deep deterministic policy gradient),以寻找高效的无人机运动、计算资源分配和通信资源分配。文献[17]采用多智能体柔性动作-评价(MASAC, multi-agent soft actor-critic)来解决联合用户关联、无人机轨迹规划和功率分配问题,以优化无人机之间的最小加权频谱效率。TSN 服务场景的多样性不断增加,意味着不同的服务需共享底层频谱资源,因此很难保证异构网络中不同应用的服务质量<sup>[18]</sup>。同时不同的服务流量特征存在差异性,可能会对彼此产生较大的影响,尤其是在网络资源有限的情况下<sup>[19]</sup>。综上所述,现有的资源分配方法无法直接适用于 LEO-RAN 场景。作为新一代无线网络技术,网络

切片是基于相同的物理网络基础设施和资源构建一组独立的虚拟逻辑子网络, 其中每个逻辑子网络都被定制为特定类型的服务<sup>[20-21]</sup>。网络切片的关键问题是如何灵活有效地根据用户需求分配网络资源以保证不同用户的 QoS 需求<sup>[22]</sup>。

此外, TSN 中多样化的节点增加了用户关联复杂性。具体来说, 传统地面基站提供最近的服务, 但其资源有限; LEO 网络提供较多的频谱资源, 但面临相对较长的传输时延。因此, 需要考虑用户关联策略, 以满足用户的时延和 QoS 要求。传统网络中的用户关联基于接收信号参考功率, 即最大信噪比 (MAX-SINR, max-signal to interference plus noise ratio), 但在异构网络中效果并不理想<sup>[23]</sup>。在 RAN 切片场景中, 用户关联问题涉及切片选择、基站关联和关联后的资源分配, 目标是在保证服务质量的同时提高资源利用率。文献[24]将 RAN 切片的设备关联问题建模为马尔可夫决策过程 (MDP, Markov decision process) 模型, 利用混合联邦强化学习方案来解决该问题, 仿真结果表明, 该方案在网络吞吐量和通信效率方面有显著的性能提升。针对差异化覆盖的 RAN 切片场景, 文献[25]提出了用户关联和资源分配的联合优化问题, 分别使用基于拉格朗日 (LG, Lagrange) 对偶法、拉格朗日乘子法和均匀分布法求解用户关联、切片资源分配和用户资源分配 3 个子问题, 仿真结果表明, 该算法能够有效提高负载均衡效用。

上述的研究工作很少考虑切片资源分配、用户关联以及 UD-LEO 卫星网络的联合应用, 因此, 在 UD-LEO 卫星网络补充地面基站通信覆盖范围的 RAN 切片场景中, 合理实现用户关联和切片资源分配, 在满足不同业务服务水平协议 (SLA, service level agreement) 的同时最大化系统频谱效率 (SE, spectrum efficiency) 仍然存在一定的挑战。针对以上挑战, 本文的主要研究工作如下。

1) 通过引入 UD-LEO 卫星网络作为 RAN 的一部分, 提出了一种面向高密度 LEO-RAN 切片场景的联合用户关联和动态资源分配的优化算法。考虑到不同的切片差异化的 SLA, 形成了在满足 SLA 的前提下提升系统 SE 的用户关联和资源分配的联合优化问题。

2) 将与优化问题中系统 SE 呈正相关的负载均衡效用函数作为用户关联问题的目标函数, 来构造

拉格朗日对偶问题, 并采用低复杂度的梯度投影法来求解此问题, 以获得最优的用户关联策略。

3) 由于环境的时变性和用户请求的动态性, 在联合用户关联和网络切片算法中, 每个基站和 LEO 卫星都充当一个智能体, 并与环境交互以获得经验来选择合适的切片策略, 实时调整切片资源策略。然后, 采用轮询调度将频谱资源调度到与基站或 LEO 卫星关联的用户。

4) 为验证所提出的基于 MADDPG-LG 的联合用户关联和网络切片算法的有效性, 与基于 MADDPG-随机关联 (RA, random association)<sup>[15]</sup>、MATD3-LG、MATD3-RA<sup>[16]</sup>、MASAC-LG 及 MASAC-RA<sup>[17]</sup>的联合用户关联和网络切片算法进行比较, 仿真结果表明, 所提算法能够在满足业务 SLA 的同时最大化 SE, 同时验证了基于拉格朗日对偶的用户关联算法的有效性。

## 1 系统模型和问题建模

### 1.1 系统模型

本文主要研究 LEO 卫星星座和多基站组成的 LEO-RAN 切片系统场景, 如图 1 所示, BS 和 LEO 卫星的集合分别表示为  $\mathcal{M} = \{1, \dots, m, \dots, M\}$  和  $\mathcal{L} = \{1, \dots, l, \dots, L\}$ 。

UD-LEO 卫星 (后文简称 LEO 卫星) 的部署确保了对所有 BS 和用户的无缝覆盖, 意味着每个 BS 和用户始终被至少一个 LEO 卫星覆盖。将 BS 和 LEO 卫星组成的下行传输系统视为卫星-地面基站 (STBS, satellite terrestrial base station), 系统中存在  $B = M + L$  个 STBS, 其集合表示为  $\mathcal{B} = \mathcal{M} \cup \mathcal{L}$ 。系统时隙的集合表示为  $\mathcal{T} = \{1, \dots, t, \dots, T\}$ , 固定的时间间隔为  $t_d$ 。现有切片集合表示为  $\mathcal{N} = \{1, \dots, n, \dots, N\}$ , 所提方案适用于任意数量的切片, 为了更好地说明框架, BS 和 LEO 卫星上部署 3 种相同类型的网络切片  $n \in \mathcal{N} = \{n_v, n_e, n_u\}$ , 即长期演进语音承载 (VoLTE, voice over long-term evolution)、增强型移动宽带 (eMBB, enhanced mobile broadband) 及超可靠低时延通信 (URLLC, ultra-reliable and low-latency communication) 切片。切片系统内用户集合表示为  $\mathcal{U} = \{1, \dots, u, \dots, U\}$ ,  $\mathcal{U}_n (\forall n \in \mathcal{N})$  表示切片  $n$  中用户集合,  $\mathcal{U}_b^n (\forall n \in \mathcal{N}, b \in \mathcal{B})$  表示切片  $n$  中与 STBS  $b$  关联的用户集合,  $u_b^n$  表示切片  $n$  中与 STBS  $b$  关联的用户。假设用户  $u$  只能与一个 STBS

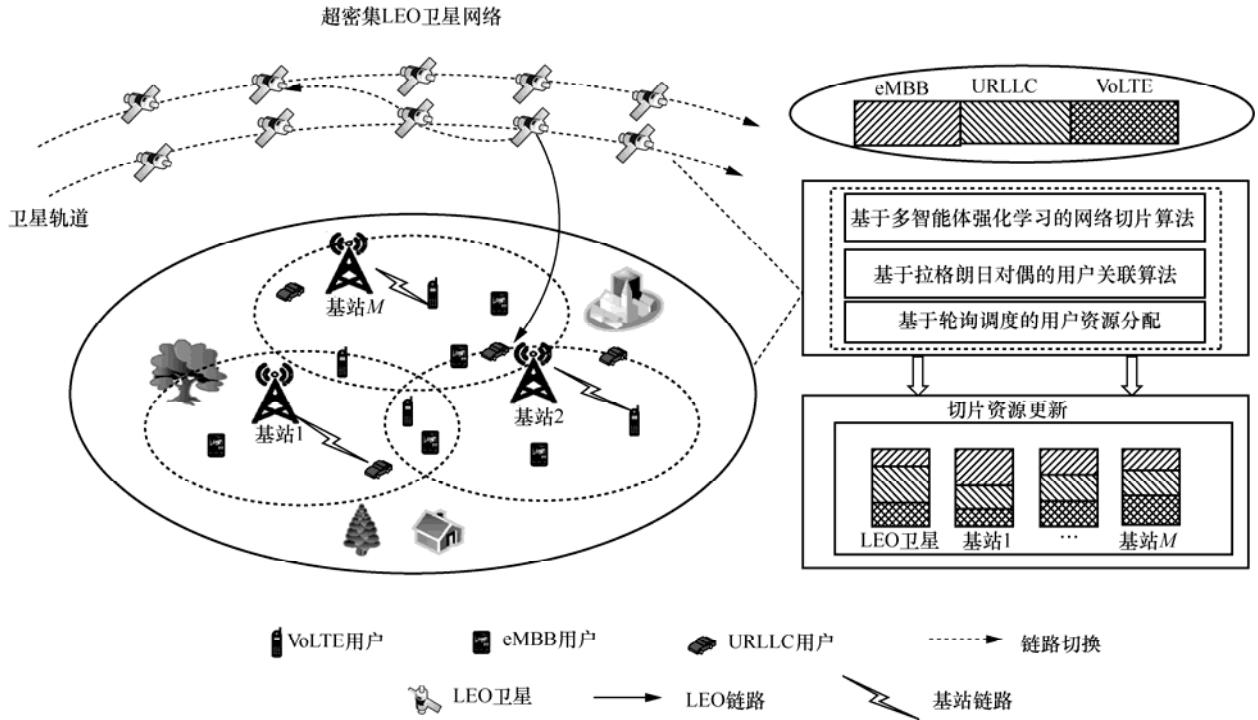


图 1 LEO-RAN 切片系统场景

关联并属于一个切片<sup>[26]</sup>。每个 STBS  $b$  拥有固定带宽，表示为  $W_B^{\max}$ ，每个资源块的带宽为  $W_{RB}$ 。在 UD-LEO 卫星网络中，当 LEO 卫星快速移动并离开当前目标区域时，切片操作可以在新到达的 LEO 卫星上继续进行以保持切片操作的连续性。这种动态切片策略使 UD-LEO 卫星网络能够灵活应对卫星位置的变化，持续覆盖地面目标区域以满足通信需求。为了更好地理解相关的数学公式和缩写，本文主要符号和变量如表 1 所示。

表 1 主要符号和变量

参数	含义	参数	含义
$\mathcal{M}$	基站集合	$f$	切片资源分配比例集合
$\mathcal{L}$	LEO 卫星集合	$z$	用户资源比例集合
$\mathcal{B}$	STBS 集合	$x$	用户关联集合
$\mathcal{T}$	时隙集合	$\Gamma_{ub}$	用户的可实现速率
$\mathcal{N}$	切片集合	$R_{u_b^n}$	用户的速率
$\mathcal{U}_n$	切片中用户的集合	$w$	负载均衡效用函数
$\mathcal{U}_b^n$	切片 $n$ 中与 STBS $b$ 关联的用户集合	$\lambda_{nb}$	STBS $b$ 中切片 $n$ 的拉格朗日乘子
$q_{u_b^n, t}^{\text{tran}}$	用户 $u_b^n$ 的传输队列	$k$	算法 1 中的迭代次数
$q_{u_b^n, i}^{\text{tran}}$	用户 $u_b^n$ 的传输时延	$\delta(k)$	动态步长

### 1.2 用户排队模型

假设每个用户  $u$  在 STBS  $b$  处都拥有一个数据队列以缓存到达的数据包，其中数据包根据先到先服务的原则进行传输。如果在当前时隙中传输的内容大于上一时隙队列积压的内容之和，则当前队列积压为空。相反，当传输速率不足时，队列积压将会随着时间的推移而增加，严重影响用户体验质量 (QoE, quality of experience)。在时隙  $t$ ，与 STBS  $b$  相关联的用户  $u_b^n$  的传输队列表示为  $q_{u_b^n, t}^{\text{tran}}$ 。因此，在时隙  $t+1$ ，用户  $u_b^n$  的传输队列为

$$q_{u_b^n, t+1}^{\text{tran}} = \max \left\{ q_{u_b^n, t}^{\text{tran}} - R_{u_b^n} t_d, 0 \right\} + A_{u_b^n, t}^{\text{tran}} \quad (1)$$

其中， $R_{u_b^n}$  表示用户  $u_b^n$  的数据速率， $A_{u_b^n, t}^{\text{tran}}$  表示用户  $u_b^n$  在时隙  $t$  瞬时到达的数据包大小。

在系统模型中，一个完整数据包的处理时延为传输时延和排队时延之和。在时隙  $t$  生成的数据包为  $i$ 。因此，用户  $u_b^n$  传输数据包  $i$  的总时延为

$$q_{u_b^n, i}^{\text{tran}} = w_{u_b^n, i}^{\text{tran}} + \tau_{u_b^n, i}^{\text{tran}} \quad (2)$$

其中， $w_{u_b^n, i}^{\text{tran}}$  和  $\tau_{u_b^n, i}^{\text{tran}}$  分别表示数据包  $i$  的传输时延和排队时延。

### 1.3 通信模型

1) 地面基站-用户通信。 $h_{mu}^C$  为用户  $u$  和 BS  $m$  之间的信道, 且主要为视线线路 (LoS, line of sight), 表示为

$$h_{mu}^C = \xi d_{mu}^{-\alpha} \quad (3)$$

其中,  $d_{mu}$  为 BS  $m$  和用户  $u$  之间的距离;  $\xi$  为基准距离  $d_{mu} = 1 \text{ m}$  的单位信道增益;  $\alpha$  为路径损耗指数<sup>[6]</sup>。因此, 与基站  $m$  相关联的用户  $u$  在 C-band 波段的可实现速率为

$$\Gamma_{um} = \text{lb} \left( 1 + \frac{p_{mu} |h_{mu}^C|^2}{\sigma^2} \right) \quad (4)$$

其中,  $p_{mu}$  为 BS  $m$  的下行链路的传输功率,  $\sigma^2$  为噪声方差。

2) LEO 卫星-用户通信。对于 LEO 卫星-用户通信, 考虑了准静态衰落信道模型<sup>[6]</sup>。因此, 用户  $u$  和 LEO 卫星  $l$  之间的信道  $h_{lu}^{\text{Ka}}$  为

$$h_{lu}^{\text{Ka}} = \vartheta_{lu} \gamma_{lu} d_{lu}^{-\beta} \quad (5)$$

其中,  $\vartheta_{lu} \sim (0,1)$  为瑞利衰落的高斯变量,  $\gamma_{lu}$  为遵循对数-正态分布的阴影衰落,  $d_{lu}$  为 LEO 卫星  $l$  和用户  $u$  之间的距离,  $\beta$  为路径损耗指数<sup>[6]</sup>。

因此, 与 LEO 卫星  $l$  相关联的用户  $u$  在 Ka-band 波段上的可实现速率为

$$\Gamma_{ul} = \text{lb} \left( 1 + \frac{p_{lu} |h_{lu}^{\text{Ka}}|^2}{\tilde{\sigma}^2} \right) \quad (6)$$

其中,  $p_{lu}$  为 LEO 卫星  $l$  下行链路的传输功率,  $\tilde{\sigma}^2$  为噪声方差。

本文将  $\Gamma_{um}$  和  $\Gamma_{ul}$  在后续的计算表示中统称为  $\Gamma_{ub}$ 。

$f = \{f_{nb}\}, \forall n \in \mathcal{N}, b \in \mathcal{B}$  表示 STBS  $b$  的切片资源分配比例集合,  $f_{nb}$  表示 STBS  $b$  分配给切片  $n$  的资源比例, 其中  $0 \leq f_{nb} \leq 1$ ;  $z = \{z_{u_b^n}\}, \forall n \in \mathcal{N}, b \in \mathcal{B}, u_b^n \in \mathcal{U}_b^n$  表示用户资源比例集合,  $z_{u_b^n}$  表示与 STBS  $b$  关联的用户  $u$  由切片  $n$  分配的资源比例;  $x = \{x_{u_b^n}\}, \forall n \in \mathcal{N}, b \in \mathcal{B}, u_b^n \in \mathcal{U}_b^n$  表示用户关联集合,  $x_{u_b^n}$  表示用户  $u$  与 STBS  $b$  及切片  $n$  是否关联,

关联则  $x_{u_b^n} = 1$ , 否则  $x_{u_b^n} = 0$ 。

因此, 根据式(4)和式(6)得到的可实现速率, 与 STBS  $b$  相关联的切片  $n$  内的用户  $u_b^n$  实现的数据速率  $R_{u_b^n}$  表示为

$$R_{u_b^n} = \Gamma_{ub} z_{u_b^n} W_B^{\text{max}} \quad (7)$$

由此可得, 对于每个用户  $u \in \mathcal{U} = \{1, \dots, u, \dots, U\}$ , 总数据速率  $R_u$  取决于用户  $u$  与 STBS  $b$  关联的可实现速率  $\Gamma_{ub}$ 、切片  $n$  从 STBS  $b$  获得的资源比例  $f_{nb}$ , 以及切片  $n$  分配给用户  $u_b^n$  资源比例  $z_{u_b^n}$ 。由于每个用户  $u$  只能与一个 STBS 相关联并属于一个切片, 当满足  $x_{u_b^n} = 1$  时, 用户  $u$  可以实现的总数据速率为  $R_u = R_{u_b^n}$ 。

### 1.4 问题建模

在上述定义的 LEO-RAN 切片场景中, 重点寻找最优的用户关联和资源分配策略, 以满足在不同服务下用户的数据速率和时延方面的 SLA 要求。因此, 将系统效用定义为系统 SE 和不同切片服务水平协议满意率 (SSR, SLA satisfaction ratio) 的加权和, 并将其作为优化目标。

STBS  $b$  的 SE 为与其相关联的用户数据速率之和与总带宽  $W_B^{\text{max}}$  的比率, 表示为

$$\text{SE}_b = \frac{\sum_{n \in \mathcal{N}} \sum_{u_b^n \in \mathcal{U}_b^n} x_{u_b^n} R_{u_b^n}}{W_B^{\text{max}}} \quad (8)$$

系统 SE 表示为

$$\text{SE} = \frac{\sum_{b \in \mathcal{B}} \text{SE}_b}{B} \quad (9)$$

STBS  $b$  切片  $n$  的  $\text{SSR}_{n,b}$  表示成功传输的数据包与传输总的数据包个数的比值, 当速率和时延都满足所在切片  $n$  的性能指标时, 则该数据包被视为成功传输, 计算式为

$$\text{SSR}_{n,b} = \frac{\sum_{u_b^n \in \mathcal{U}_b^n} \sum_{q_{u_b^n} \in \mathcal{Q}_{n,b}} \mathbb{I} \left( R_{u_b^n} > \varphi_n^{\text{rate}}, \delta_{q_{u_b^n}} < \varphi_n^{\text{latency}} \right)}{\sum_{u_b^n \in \mathcal{U}_b^n} |\mathcal{Q}_{n,b}|} \quad (10)$$

其中,  $\mathcal{Q}_{n,b}$  为 STBS  $b$  的切片  $n$  所处理的数据包集

合,  $\varphi_n^{\text{rate}}$  和  $\varphi_n^{\text{latency}}$  分别表示切片  $n$  的速率和时延的性能指标。 $\mathbb{I}\left(R_{u_b^n} > \varphi_n^{\text{rate}}, \delta_{q_{u_b^n}} < \varphi_n^{\text{latency}}\right)$  是性能指标函数, 如果用户的速率和时延都满足所在切片  $n$  的性能指标, 则其函数值为 1, 否则为 0。

因此, 优化问题可以表示为

$$\begin{aligned} \mathbf{J}_1: \text{maximize}_{x, f, z} & \frac{\sum_{b \in \mathcal{B}} \alpha_b \text{SE}_b + \sum_{b \in \mathcal{B}} \sum_{n \in \mathcal{N}} \beta_n \text{SSR}_{n,b}}{B} \\ \text{s.t. } C_1: & \sum_{b \in \mathcal{B}} \sum_{n \in \mathcal{N}} x_{u_b^n} = 1, \forall u \in \mathcal{U} \\ C_2: & \sum_{n \in \mathcal{N}} f_{nb} = 1, \forall n \in \mathcal{N} \\ C_3: & f_{nb} \in [0, 1], \forall (n, b) \in \mathcal{N} \times \mathcal{B} \\ C_4: & \sum_{u \in \mathcal{U}} z_{u_b^n} \leq f_{nb}, \forall (n, b) \in \mathcal{N} \times \mathcal{B} \end{aligned} \quad (11)$$

其中,  $\alpha_b$  和  $\beta_n$  分别表示 STBS  $b$  实现的 SE 权重和切片  $n$  的 SSR 权重。约束条件  $C_1$  限制用户只能关联一个 STBS。约束条件  $C_2$  限制 STBS  $b$  上的所有切片获得的资源比例总和为 1。约束条件  $C_3$  表明 STBS  $b$  上每个切片获得的资源比例的取值范围。约束条件  $C_4$  表明切片  $n$  内与 STBS  $b$  相关联的用户的资源比例总和不能大于其切片  $n$  在 STBS  $b$  上获得的切片资源比例。

由于  $\mathbf{J}_1$  是一个联合用户关联和资源分配的优化问题, 每个用户获得的资源取决于用户关联和切片资源分配。这种耦合使优化问题难以解决, 因此本文首先确定用户资源比例集合  $z$ , 然后求解用户关联集合  $x$  和切片资源比例集合  $f$ 。

## 2 基于 MADDPG-LG 的联合用户关联和网络切片算法

### 2.1 用户资源分配

确定切片资源比例  $f$  之后, STBS 在同一切片内采用轮询调度将资源分配给用户。轮询调度把同一切片内所有用户的优先级都视为相等, 所有用户可以被周期性地调度资源, 能够保证同一切片内每个用户被调度的概率相同。

因此, 用户的资源分配比例可以近似表示为

$$z_{u_b^n}^* = \frac{f_{nb}}{U_b^n} \quad (12)$$

其中,  $U_b^n = |\mathcal{U}_b^n|$  表示 STBS  $b$  切片  $n$  的用户数量。

根据用户资源分配策略得到式(12)所示的用户资源分配比例, 可将式(7)的切片  $n$  中与 STBS  $b$  关联的用户  $u_b^n$  的数据速率重新表示为

$$R_{u_b^n} = \frac{f_{nb} W_B^{\max} \Gamma_{ub}}{U_b^n} \quad (13)$$

### 2.2 基于拉格朗日对偶的用户关联算法

用户关联问题为切片内的用户与 STBS 之间关联的 0-1 二进制变量问题, 因此通过最大化负载均衡效用函数来求解用户关联问题<sup>[25]</sup>。负载均衡效用函数表示为

$$w = \sum_{b \in \mathcal{B}} \sum_{n \in \mathcal{N}} \sum_{u \in \mathcal{U}} x_{u_b^n} \log R_{u_b^n} \quad (14)$$

由于负载均衡效用函数与本文的优化问题式(11)中的 SE 成正相关, 因此求解用户关联问题实现负载均衡效用函数最大化, 即实现本文的优化问题式(11)中 SE 的最大化。

根据用户资源分配策略得到式(12)所示的用户资源分配比例, 可将式(14)转化为

$$w = \sum_{b \in \mathcal{B}} \sum_{n \in \mathcal{N}} \sum_{u \in \mathcal{U}} x_{u_b^n} \log \left( \frac{f_{nb} W_B^{\max} \Gamma_{ub}}{\sum_l x_{l_b^n}} \right) \quad (15)$$

假设切片资源比例集合  $f = \{f_{nb}\}, \forall n \in \mathcal{N}, b \in \mathcal{B}$  为固定值。由于每个切片内的用户集合  $\mathcal{U}_n$  在整个时间范围内被视为是固定的, 因此可以构建最大化每个切片的负载均衡效用函数的用户关联问题  $\mathbf{J}_2$ , 可以表示为

$$\begin{aligned} \mathbf{J}_2: \text{max}_x & \sum_{b \in \mathcal{B}} \sum_{u \in \mathcal{U}_n} x_{u_b^n} \log \frac{f_{nb} W_B^{\max} \Gamma_{ub}}{\sum_l x_{l_b^n}} \\ \text{s.t. } & \sum_{b \in \mathcal{B}} x_{u_b^n} = 1, \forall u \in \mathcal{U}_n \\ & x_{u_b^n} \in \{0, 1\}, \forall b, u \in \mathcal{B} \times \mathcal{U}_n \end{aligned} \quad (16)$$

式(16)中的优化变量均为 0-1 的离散变量, 如果采用全力搜索法, 算法复杂度较高。为了更加有效地求解此问题, 需要一种低复杂度的方法。因此松弛用户关联变量  $x_{u_b^n}$  并将其转化为连续变量, 即  $x_{u_b^n} \in [0, 1], \forall (b, u) \in \mathcal{B} \times \mathcal{U}_n$ , 式(16)可以转化为凸集

上的凸优化问题。

### 2.2.1 对偶分解

本节提出一种基于拉格朗日对偶分解算法<sup>[27]</sup>, 将式(16)的对偶问题解耦为2个子问题, 分别求解2个子问题。

引入一组新的变量, 即负载度量  $\sum_u x_{u_b}^n = U_b^n$ ,

式(16)等价形式表示为

$$\begin{aligned} \max_x \quad & \sum_b \sum_u x_{u_b}^n \log(f_{nb} W_B^{\max} \Gamma_{ub}) - \sum_b U_b^n \log(U_b^n) \\ \text{s.t.} \quad & \sum_{b \in \mathcal{B}} x_{u_b}^n = 1, \quad \forall u \in \mathcal{U}_n \\ & x_{u_b}^n \in [0, 1], \quad \forall n, u \in \mathcal{N} \times \mathcal{U}_n \\ & \sum_u x_{u_b}^n = U_b^n, \quad \forall b \in \mathcal{B} \\ & 0 \leq U_b^n \leq U_n, \quad \forall b \in \mathcal{B} \end{aligned} \quad (17)$$

其中, 引入冗余约束  $U_b^n \leq U_n$  来分析算法的收敛性,  $\sum_u x_{u_b}^n = U_b^n$  为对偶约束。因此, 采用拉格朗日对偶分解方法, 通过引入拉格朗日乘子  $\lambda_{nb}$  来松弛对偶约束, 拉格朗日对偶问题表示为

$$\mathbf{F}: \min_{\lambda} F(\lambda) = h_x(\lambda) + g_u(\lambda) \quad (18)$$

其中

$$\begin{aligned} h(\lambda) &= \max_x \sum_b \sum_u x_{u_b}^n (\log(f_{nb} W_B^{\max} \Gamma_{ub}) - \lambda_{nb}) \\ \text{s.t.} \quad & \sum_b x_{u_b}^n = 1 \\ & x_{u_b}^n \in [0, 1] \\ g(\lambda) &= \max_{U_b^n} \sum_b U_b^n (\lambda_{nb} - \log U_b^n) \\ \text{s.t.} \quad & U_b^n \leq U_n \end{aligned} \quad (19)$$

显然式(17)满足 Slater 条件<sup>[28]</sup>, 因此式(17)和式(18)之间的强对偶关系成立。求解对偶问题式(18)等价于求解原问题式(17)。  $x_{u_b}^n(\lambda)$  和  $U_{nb}(\lambda)$  分别表示2个子问题式(19)和式(20)的最大值。因此, 给定对偶最优值  $\lambda^*$ , 可以通过分别求解子问题式(19)和式(20)来获得原始最优解。

### 2.2.2 子问题求解和 $\lambda$ 更新

采用梯度投影法<sup>[29]</sup>来求解拉格朗日对偶问题式(18), 其中拉格朗日乘子  $\lambda$  的更新方向与  $\nabla F(\lambda)$

的方向相反。设  $k$  为当前迭代次数, 通过求解2个子问题式(19)和式(20)来计算对偶问题的梯度, 之后求解  $x_{u_b}^*(k)$  和  $U_b^{n*}(k)$  来更新  $\lambda$ 。

当迭代次数为  $k$  时, 通过求解子问题式(19), 每个用户  $u \in \mathcal{U}_n$  选择关联的 STBS, 表示为

$$b^* = \arg \max_b (\log(f_{nb} W_B^{\max} \Gamma_{ub}) - \lambda_{nb}(k)), u \in \mathcal{U}_n \quad (21)$$

其中, 对于用户选择关联的 STBS  $b^*$ ,  $x_{u_b}^n = 1$ ; 对于其他 STBS  $b \neq b^*$ ,  $x_{u_b}^n = 0$ 。因此, 当迭代次数为  $k$  时, 获得  $x_{u_b}^*(k)$ 。

当迭代次数为  $k$  时, 为了获得子问题式(20)的最大值, 在  $U_b^n \leq U_n$  的约束下, 每个 STBS  $b \in \mathcal{B}_n$  的梯度为0, 即

$$U_b^n(k) = \min\{U_n, e^{\lambda_{nb}(k)-1}\}, b \in \mathcal{B}_n \quad (22)$$

因此, 当迭代次数为  $k$  时, 获得  $U_b^{n*}(k)$ 。

根据式(21)和式(22)计算得到的  $x_{u_b}^*(k)$  和  $U_b^{n*}(k)$ , 拉格朗日乘子  $\lambda_{nb}$  通过以下方式更新

$$\lambda_{nb}(k+1) = \lambda_{nb}(k) - \delta(k)(U_b^{n*}(k) - \sum_u x_{u_b}^*(k)) \quad (23)$$

其中,  $\delta(k) > 0$  是一个动态更新的步长。

$\lambda_{nb}$  可被视为供求规律中的需求与供给之间的桥梁, 式(23)中的  $\sum_u x_{u_b}^n$  可以解释 STBS  $b$  的切片  $n$  的服务需求,  $U_b^n$  为 STBS  $b$  的切片  $n$  所能提供的服务。如果服务需求  $\sum_u x_{u_b}^n$  超过服务供应  $U_b^n$ , STBS  $b$  被视为超载,  $\lambda_{nb}$  的数值将减小; 如果服务需求  $\sum_u x_{u_b}^n$  小于服务供应  $U_b^n$ , 此时, STBS  $b$  被视为轻载,  $\lambda_{nb}$  的数值将增大。

### 2.2.3 动态步长和更新

求解式(16)的过程中, 需要确保收敛时间小于关联的周期, 因此, 本节提出一种迭代次数较小的步长更新方法。在迭代地执行 2.2.2 节所述步骤之后, 梯度算法将会收敛到一个接近最优的解  $\lambda^*$ 。步长  $\delta(k)$  更新规则表示为

$$\delta(k) = \gamma(k) \frac{F(\lambda(k)) - F(k)}{\|\partial F(\lambda(k))\|^2} \quad (24)$$

其中,  $F(k)$  是拉格朗日对偶问题式(18)最优值  $F^*$  的估计值,  $0 \leq \underline{\gamma} \leq \gamma(k) \leq \bar{\gamma} \leq 2$ ,  $\underline{\gamma}$  和  $\bar{\gamma}$  是标量<sup>[30]</sup>.  $F(k)$  的更新过程为

$$F(k) = \min_{0 \leq k' \leq k} F(\lambda(k')) - \varepsilon(k) \quad (25)$$

$\varepsilon(k)$  的更新过程为

$$\varepsilon(k+1) = \begin{cases} \rho\varepsilon(k), & F(\lambda(k+1)) \leq F(k) \\ \max\{\beta\varepsilon(k), \varepsilon\}, & \text{其他} \end{cases} \quad (26)$$

其中,  $\varepsilon$ 、 $\beta$  和  $\rho$  是固定的正数,  $\beta < 1$  和  $\rho > 1$ <sup>[29]</sup>.

因此, 在这个过程中期望达到一个比最优值小  $\varepsilon(k)$  的目标值  $F(k)$ 。当达到目标值  $F(k)$  时,  $\varepsilon(k)$  将保持不变 ( $\rho = 1$ ) 或变大 ( $\rho > 1$ )。在给定的迭代次数内, 如果没有达到目标值  $F(k)$ , 那么  $\varepsilon(k)$  会减小到阈值, 确保式(24)的步长  $\delta(k)$  不等于 0<sup>[30]</sup>, 由此可得定理 1。

**定理 1** 假设根据式(25)和式(26), 步长  $\delta(k)$  调整通过动态步长规则式(24)更新。如果  $F^* > -\infty$ , 其中  $F^*$  表示最优值, 则

$$\inf_k F(\lambda(k)) \leq F^* + \varepsilon \quad (27)$$

**证明** 函数  $F(\lambda)$  的导数为

$$\frac{\partial F}{\partial \lambda_{nb}}(\lambda) = U_b^n(\lambda) - \sum_u x_{u_b^n}(\lambda) \quad (28)$$

由于  $U_b^n = \sum_u x_{u_b^n} \leq U_n$ , 根据式(28), 当  $U_b^n$  和  $\sum_u x_{u_b^n}$  有界时, 对偶函数的  $\partial F$  也有界, 即

$$\sup_k \left\{ \left\| \frac{\partial F(\lambda(k))}{\partial \lambda} \right\| \right\} \leq \eta \quad (29)$$

其中,  $\eta$  是标量。因此, 拉格朗日对偶问题满足文献[30]中 6.3.6 命题的必要条件。证毕。

故本节方法能有效地逼近问题式(18)的最优解  $\lambda^*$ , 同时得到子问题式(19)和式(20)的最优解,  $x_{u_b^n}^*$  也是问题式(16)的期望解。

上述提出的基于拉格朗日对偶的用户关联算法给出接近最优用户关联  $x = \{x_{u_b^n}\}$  的求解步骤。

求解过程的重复次数与切片的数量有关, 如果有  $N$  个切片, 求解过程就执行  $N$  次。具体步骤如算法 1 所示。

**算法 1** 基于拉格朗日对偶的用户关联算法

**输入** 切片资源比例集合  $f$

**输出** 用户关联集合  $x^*$

- 1) 初始化  $k = 0$ ; 每个用户计算  $\Gamma_{ub}$
- 2) 重复计算步骤 3)至步骤 14)
- 3) for  $n = 1 : N$
- 4) for  $u = 1 : U_n$
- 5) 根据式(21), 用户选择关联的 STBS  $b^* \in \mathcal{B}_n$
- 6) end for
- 7) 更新用户关联集合  $x$
- 8) for  $b = 1 : B_n$
- 9) 根据式(22), 计算相应的  $U_b^{n*}(k)$
- 10) 根据式(23), 更新  $\lambda_{nb}(k)$
- 11) end for
- 12) end for
- 13)  $k \leftarrow k + 1$
- 14) 直至所有的  $\lambda_{nb}(k)$  收敛到  $\lambda_{nb}^*$
- 15) 获得最优的用户关联策略  $x^* = x$

### 2.3 基于多智能体强化学习的网络切片算法

#### 2.3.1 多智能体环境模型

在 LEO-RAN 切片的系统场景中, STBS 需要根据用户需求决定切片资源的分配以实现满足切片 SLA 的同时提高 SE。考虑到 STBS 资源切片的分配决策可能会影响环境状态, 同前一状态和前一动作共同作用于环境并获得新的随机状态。因此, 本文将建模为多智能体马尔可夫决策过程。在本文系统模型中, 对于 STBS  $b$  (智能体  $b$ ), 其状态空间  $\mathcal{S}_b$ 、动作空间  $\mathcal{A}_b$  及奖励函数  $\mathcal{R}_b$  定义如下。

状态空间  $\mathcal{S}_b$ 。STBS  $b$  观测到的状态为归一化的切片请求传输数据包数量。packet <sub>$n,b$</sub>  表示智能体  $b$  切片  $n$  请求传输的数据包总和, mean <sub>$n,b$</sub> 、std <sub>$n,b$</sub>  分别表示智能体  $b$  切片  $n$  传输数据包请求到达的间隔分布的平均值、标准差。因此状态表示为

$$\mathcal{S}_b = \{s_{n,b}\}, n \in \mathcal{N} = \{n_v, n_e, n_u\} \quad (30)$$

$$s_{n,b} = \frac{\text{packet}_{n,b} - \text{mean}_{n,b}}{\text{std}_{n,b}} \quad (31)$$

动作空间  $\mathcal{A}_b$ 。STBS  $b$  选取动作  $a_{n,b} \in \mathcal{A}_b$ , 表示智能体  $b$  根据当前状态  $s_{n,b} \in \mathcal{S}_b$  选择切片比例  $f_{nb}$ , 动作空间可以表示为

$$\mathcal{A}_b = \{a_{n,b}\}, n \in \mathcal{N} = \{n_v, n_e, n_u\} \quad (32)$$

奖励函数  $\mathcal{R}_b$ 。STBS  $b$  在与环境的交换中做出决策从而最大化奖励回报。eMBB、VoLTE 切片的 SLA 均设置为 0.95, URLLC 切片的 SLA 为 0.90, 系统 SE 的指标要求为 250 bit/(s·Hz)。同时引入分层即时奖励机制, 表示为

$$\mathcal{R}_b = f_2(\text{SSR}_{n_u,b}) \mathbb{I}(\text{SSR}_{n_v,b}, \text{SSR}_{n_e,b}) + f_1(1 - \mathbb{I}(\text{SSR}_{n_v,b}, \text{SSR}_{n_e,b})) \quad (33)$$

其中,  $\mathbb{I}(\text{SSR}_{n_v,b}, \text{SSR}_{n_e,b})$  表示 VoLTE 切片、eMBB 切片的 SSR, 如果 VoLTE 切片、eMBB 切片的 SSR 都分别满足切片的 SLA, 则其函数值为 1, 否则其函数值为 0。当 VoLTE 和 eMBB 切片的 SLA 被满足时, 奖励函数取决于  $f_2$ 。  $f_2$  可以表示为

$$f_2 = \begin{cases} 10(\text{SSR}_{n_u,b} - 0.7), & \text{SSR}_{n_u,b} < 0.9 \\ 4 + \max\{0, 10\alpha(\text{SE}_b - 250)\}, & \text{SSR}_{n_u,b} \geq 0.9 \end{cases} \quad (34)$$

当不满足时, 奖励函数只由  $f_1$  决定, 给予一个较大的负奖励, 表示为

$$f_1 = -5 \quad (35)$$

### 2.3.2 用于切片资源分配的 MADDPG 算法

MADDPG 算法是基于 Actor-Critic 的单智能体 DDPG 算法的扩展, 本文将应用到多智能体强化学习问题中, 其基础思想为: 集中式训练、分布式执行。MADDPG 中每个智能体的主网络包含 2 种网络: 1) 网络参数为  $\mu_b$  的 Actor, 用来做出符合当前环境和状态的决策; 2) 网络参数为  $\theta_b$  的 Critic, 用来评判策略网络输出动作的优劣。同时为了提高学习的稳定性引入目标网络结构: 1) 网络参数为  $\mu'_b$  的 Target Actor 网络; 2) 网络参数为  $\theta'_b$  的 Target Critic 网络。

为了防止  $Q$  值过拟合及增加算法的探索能力, Target Actor 网络中添加随机噪声以实现更加平滑的状态动作估计, 因此智能体  $b$  的输出动作值为  $a_b$  为

$$a_b = u'_b(s'_j) + \tilde{\epsilon} \quad (36)$$

其中,  $u_b$  为智能体  $b$  策略网络输出的策略;  $\tilde{\epsilon} \sim \text{clip}(N(0, \hat{\sigma}^2), -1, 1)$  为均值为 0、标准差为  $\hat{\sigma}^2$  并裁剪到  $[-1, 1]$  的随机噪声。

为了使训练过程更加稳定并提高采用效率, 每

个 STBS 会将记忆  $(s_b, a_b, r_b, s'_b)$  存储到回放缓冲区  $\mathcal{D}$ 。对于每个 STBS, 首先从回放缓冲区  $\mathcal{D}$  随机采样  $M_b$  小批量记忆  $(s_j, a_j, r_j, s'_j)$ , 然后将  $s_j$  输入 Actor 网络以生成策略  $\pi_b^\mu(s_j)$ 。因此, 每个 STBS 可以采用梯度策略更新 Actor 网络的参数, 表示为  $\nabla_{\mu_b} J(\mu_b) =$

$$\frac{1}{M_b} \sum_{j=1}^{M_b} \nabla_{\mu_b} \pi_b^\mu(s_b^j) \nabla_{a_b} Q_b^\theta(s_j, a_j^1, \dots, a_B^j) |_{a_b = \pi_b^\mu(s_b^j)} \quad (37)$$

然后, 根据当前策略  $\pi_b^\mu(s_j)$ , 通过 Critic 网络计算  $Q$  值  $Q_b^\theta(s_j, \pi_b^\mu(s_j))$  以最小化损失函数  $L(\theta_b)$ , 计算式为

$$L(\theta_b) = \frac{1}{M_b} \sum_{j=1}^{M_b} [y_j - Q_b^\theta(s_j, a_j)]^2 \quad (38)$$

其中, 目标值  $y_j$  的计算式为

$$y_j = r_j + \gamma Q_b^\theta(s'_j, a_j) \quad (39)$$

其中,  $\gamma$  为折扣因子。接下来, 每个 STBS 通过式(40)和式(41)来更新 Actor 和 Critic 网络参数。

$$\mu_b \leftarrow \mu_b - \lambda \nabla_{\mu_b} J(\mu_b) \quad (40)$$

$$\theta_b \leftarrow \theta_b - \lambda \nabla_{\theta_b} L(\theta_b) \quad (41)$$

Target 网络采用软更新的更新方式, 主网络参数用于目标网络的更新, 目标网络的更新幅度由目标网络更新率  $\tau$  决定,  $0 \leq \tau < 1$ , 更新方式为

$$\mu'_b \leftarrow \tau \mu_b + (1 - \tau) \mu'_b \quad (42)$$

$$\theta'_b \leftarrow \tau \theta_b + (1 - \tau) \theta'_b \quad (43)$$

基于 MADDPG 的网络切片算法如算法 2 所示。

#### 算法 2 基于 MADDPG 的网络切片算法

- 1) 初始化每个 STBS 的 Actor 网络以及 Target Actor
- 2) 初始化每个 STBS 的 Critic 网络以及 Target Critic
- 3) 初始化每个 STBS 的回放缓冲区  $\mathcal{D}_b$ , 初始化环境并获得状态  $s$
- 4) for step=1: Maxstep
- 5)     for  $b=1: B$
- 6)         根据状态  $s_b$  获得动作  $a_b$
- 7)         动作  $a_b$  经过 Softmax 函数进行调整
- 8)         该 STBS 执行动作并计算获得的奖励及下一状态

- 9) 记忆存储到回放缓冲区
- 10) end for
- 11) 对 STBS 分别从经验池中随机取样小批量记忆
- 12) 通过最小化式(38)中的损失函数,更新 Critic 网络的权值。
- 13) 通过式(40)和式(41)更新 Actor 网络和 Critic 网络参数
- 14) 通过式(42)和式(43)更新 Target Actor 网络和 Target Critic 网络参数
- 15) end for

### 3 仿真实验

#### 3.1 仿真场景与仿真参数设置

本节通过应用 Python3.8.6 和 PyTorch1.12.0 进行仿真,验证了所提 MADDPG-LG 算法在 LEO-RAN 场景下的网络切片性能。该场景包含一个卫星星座和 4 个基站协同为用户提供通信服务,采用卫星工具包(STK, satellite tool kit)来构建整个拓扑结构。卫星星座由 60×40 颗 LEO 卫星组成,轨道高度为 550 km,确保至少 3 颗卫星对目标区域的无缝覆盖<sup>[6]</sup>。由于卫星星座负责在全球范围内提供服务,因此只有一部分通信资源被分配给目标区域。将 Ka-band 波段卫星通信的可用带宽设置为 20 MHz, Ka 频段的具体数值为 30 GHz。4 个基站的带宽设置为 10 MHz。为了体现用户关联的负载均衡的效果,将目标区域划分为两部分,上半部分区域是用户密集的业务中心且用户分布在此部分,剩下的用户分布在下半部分区域部分。仿真参数设置如表 2 所示<sup>[31-32]</sup>。

MADDPG-LG 算法由两层全连接隐藏层组成,每层含有 64 个神经元; Actor 和 Critic 的学习率均设置为 0.001; 批大小和经验回放单元的大小分别设置

为 128 和 1 000 000; 折扣因子设置为 0.95。

#### 3.2 仿真结果分析

本节通过与 5 种基准算法的比较,验证所提基于 MADDPG-LG 的联合用户关联和网络切片算法的有效性。5 种基准算法分别为基于 MADDPG-RA、MATD3-LG、MATD3-RA、MASAC-LG 及 MASAC-RA 的联合用户关联和网络切片算法。

##### 3.2.1 超参数分析

###### 1) 不同学习速率 LR 的收敛性能

采用不同的 LR 对所提 MADDPG-LG 算法进行仿真,以确定适合该算法的 LR 的最优值,不同学习率的收敛性能如图 2 所示。从图 2 可以观察到,当 LR 设置为 0.001、0.005、0.000 1 以及 0.000 5 时,算法收敛。但当 LR 设置为 0.000 1 时,参数较小、更新步长较小,导致算法收敛速度较慢且收敛后的数值较低;另一方面,当 LR 设置为 0.005 时,由于 LR 过大,算法收敛速度快,但收敛后的数值较低。因此,为了保证训练更稳定而且具有较好的性能,最优 LR 设置为 0.001。

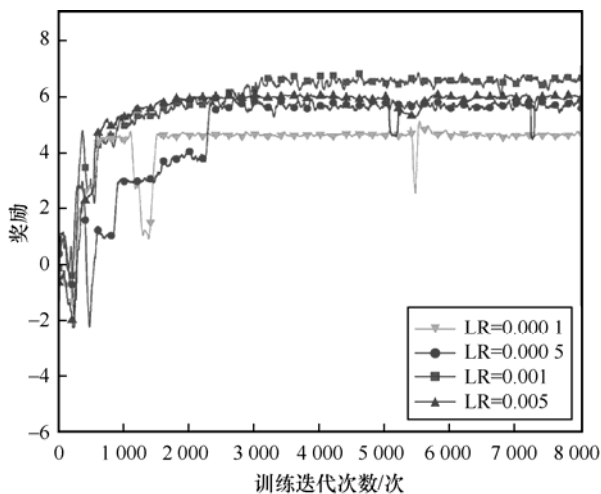


图 2 不同学习率的收敛性能

表 2 仿真参数设置

参数	VoLTE 参数值	eMBB 参数值	URLLC 参数值
轮询调度/ms	0.5	0.5	0.5
用户数量/个	333	667	1 000
数据包请求到达间隔分布	均匀分布 [min=0, max=160 ms]	截断帕累托分布 [指数分布=1.2, mean=6 ms, max=12.5 ms]	指数分布 [mean=180 ms]
数据包的大小分布	40 B	[截断帕累托分布=1.2, mean=100 B, max=250 B]	0.15 MB
切片要求	速率	100 Mbit/s	10 Mbit/s
	时延/ms	10	3

### 2) 不同用户数的系统效用对比

系统效用随用户数量的变化如图 3 所示。从图 3 可以看出，随着用户数量的增加，系统效用呈先上升后下降趋势。其原因在于当系统内用户数量小于 2 000 个时，带宽资源充足且系统效用取决于数据包数量，用户越多，请求的数据包越多，导致系统效用呈上升趋势；但当用户数量大于 2 000 个时，STBS 带宽资源有限并且用户传输数据包持续增多导致 SSR 降低，因此系统效用逐渐减小。

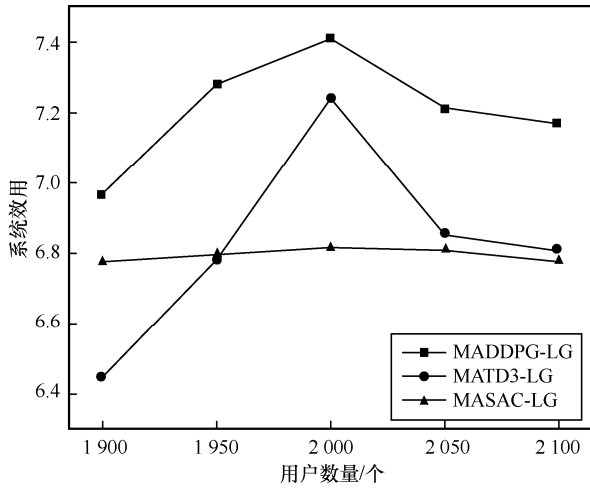


图 3 系统效用随用户数量的变化

根据对学习率和用户数的超参数分析，接下来在验证所提出的基于 MDDPG-LG 的联合用户关联和网络切片算法有效性的过程中，将用户数设置为 2 000 个，并将学习率设置为 0.001。

### 3.2.2 不同用户关联方式的对比

为了更好地评估所提算法的有效性，将其与 RA 算法和基于 MAX-SINR 的用户关联算法进行对比。不同算法下系统效用随训练迭代次数的变化如图 4 所示。从图 4 可以观察到，当用户关联方式为 LG 用户关联算法时，系统效用稳定在 7.41 左右；RA 算法最终能够达到收敛状态，系统效用只能稳定在 7.27 左右；基于 MAX-SINR 的用户关联算法达到收敛状态后，系统效用稳定在 7.13 左右。因此，LG 用户关联算法为最优的用户关联方式，相比于 RA 算法、基于 MAX-SINR 的用户关联算法分别提升了 1.9%、3.9%。

### 3.2.3 STBS 的性能分析

#### 1) STBS 的效用对比

本节通过分析 STBS 的效用，讨论了采用基于 MDDPG-LG 的联合用户关联和网络切片算法得到的每个智能体的性能。STBS 的效用随训练迭代

次数的变化如图 5 所示。从图 5 可以看到，LEO 卫星的效用大约在 3 000 次左右收敛，效用稳定在 7.81 左右；BS1 和 BS2 的效用大约在 1 000 次左右收敛，效用分别稳定在 7.89 和 7.81 左右；BS3 和 BS4 的效用略低，BS3 的效用最终收敛在 6.82 左右，BS4 的效用最终收敛在 6.73 左右。

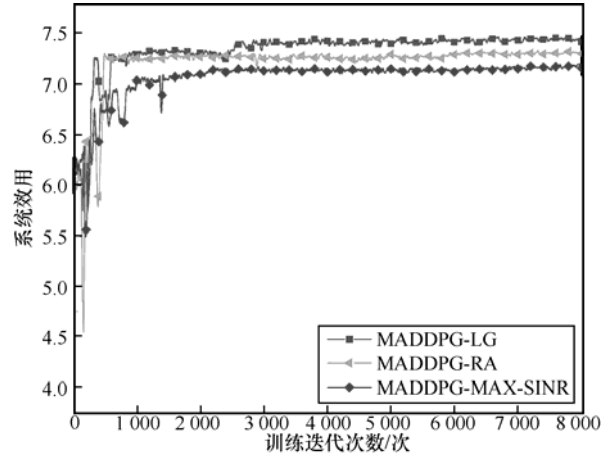


图 4 不同算法下系统效用随训练迭代次数的变化

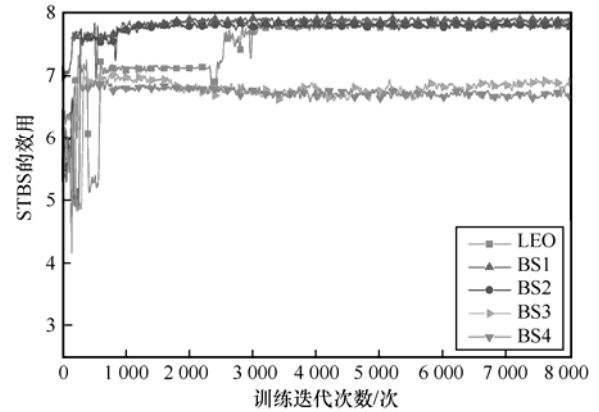


图 5 STBS 的效用随训练迭代次数的变化

#### 2) STBS 上切片 SSR 的对比

图 6~图 8 分别给出了 STBS 的 VoLTE、eMBB 和 URLLC 切片的 SSR 随训练迭代次数的变化。从图 6 可以看出，STBS 的 VoLTE 切片的 SSR 稳定在 0.998~1.000。从图 7 可以看出，LEO 卫星和基站上的 eMBB 切片的 SSR 稳定在 0.97 左右。随着迭代次数的增加，LEO 卫星上的 URLLC 切片的 SSR 稳定在 0.95 左右，同时基站上的 URLLC 切片的 SSR 稳定在 0.92 左右。相比于 VoLTE、eMBB 切片，URLLC 切片的 SSR 数值较低，这是因为 URLLC 切片的服务要求最高，用户对数据包时延的要求较高。

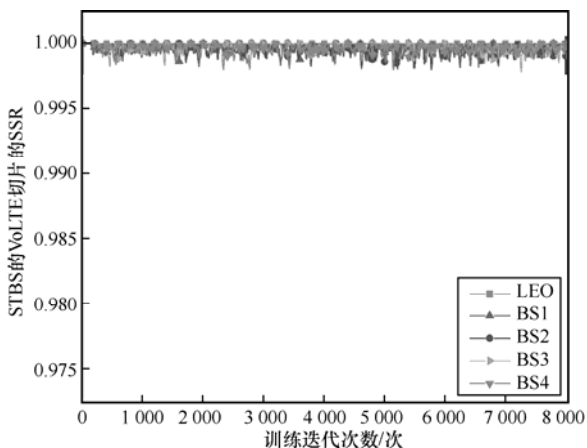


图6 STBS的VoLTE切片的SSR随训练迭代次数的变化

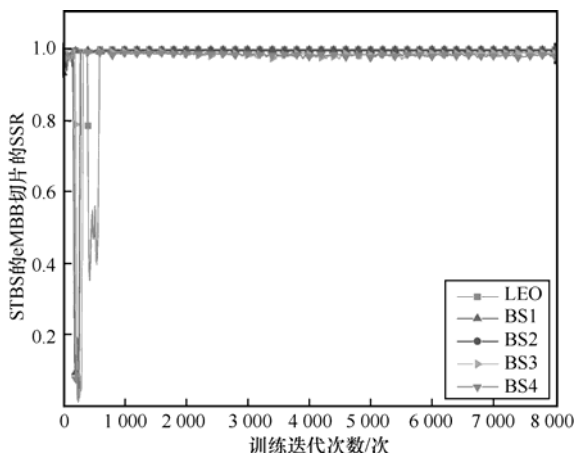


图7 STBS的eMBB切片的SSR随训练迭代次数的变化

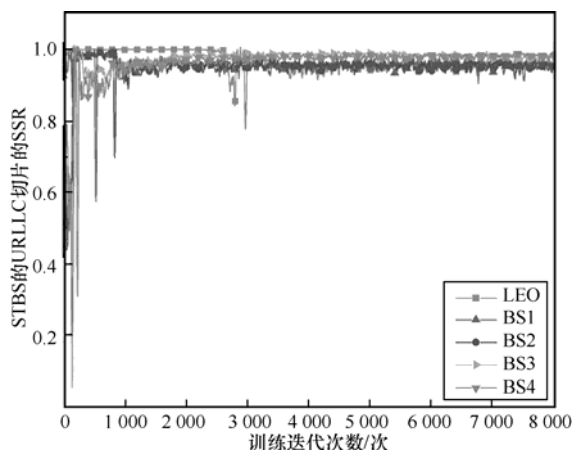


图8 STBS的URLLC切片的SSR随训练迭代次数的变化

### 3.2.4 性能对比

为了更好地评估所提算法在系统SE和切片SSR加权和优化问题的性能,将所提算法与其他几种基准算法进行了比较。

#### 1) 系统效用的对比

系统效用随训练迭代次数的变化如图9所示。

$SE_b$ 的权重 $\alpha_b$ 设置为0.01,  $SSR_{n,b}$ 的系数 $\beta_n$ 设置为[1,1,3],共训练迭代8000次。从图9可以观察到,所提算法在大约3000次收敛,系统效用稳定在7.41左右。MADDPG-RA算法最终能够达到收敛状态,但是系统效用只能稳定在7.27左右,整体系统效用较差。这是因为该场景下的区域分为用户密集的业务中心和不密集的业务中心,RA算法导致一些STBS超载,而另一些STBS轻载,资源利用率较低。MATD3-LG算法的系统效用比MATD3-RA算法高3.6%左右,其原因在于LG用户关联算法使用户关联更加合理。MASAC-LG算法的系统效用相比MASAC-RA算法的系统效用提升不明显,且整体的系统效用较低。整个用户关联和资源分配的过程中,所提算法能够在有限的训练迭代次数下找到合理用户关联和资源分配策略,更适用于解决LEO-RAN切片场景下的优化系统SE和切片SSR的优化问题。所提出的MADDPG-LG算法与MADDPG-RA、MATD3-LG、MATD3-RA、MASAC-LG及MASAC-RA算法相比,系统效用分别提升了2.0%、2.3%、5.7%、8.7%和9.4%。

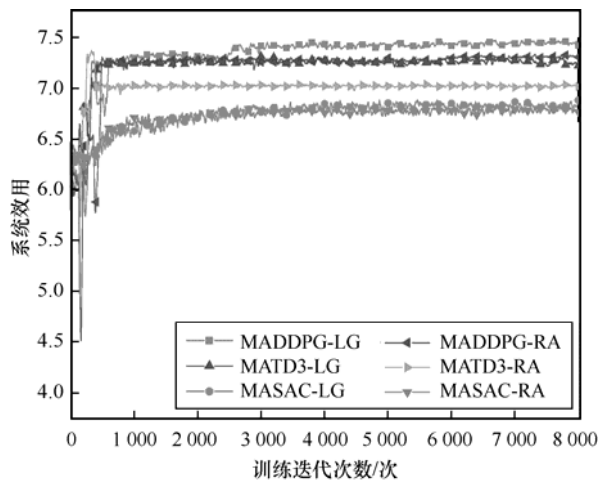


图9 系统效用随训练迭代次数的变化

#### 2) 切片SSR的对比

图10~图12分别给出了VoLTE、eMBB和URLLC切片的SSR随训练迭代次数的变化。从图10可以看出,VoLTE切片的SSR稳定在0.999~1.000。从图11可以看出,eMBB切片的SSR都较为稳定,所提MADDPG-LG算法、MADDPG-RA、MATD3-LG、MASAC-LG和MASAC-RA算法基本能够稳定在0.98,甚至能达到1。由于URLLC切片的服务要求最高,URLLC用户对数据包时延

的要求较高，随着迭代次数的增加，所提算法能够稳定在 0.98 左右。

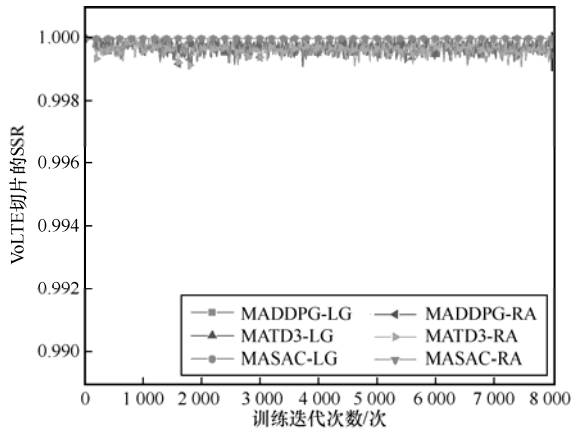


图 10 VoLTE 切片的 SSR 随训练迭代次数的变化

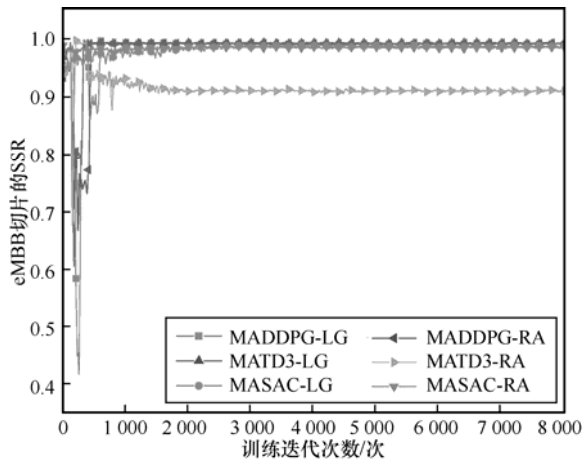


图 11 eMBB 切片的 SSR 随训练迭代次数的变化

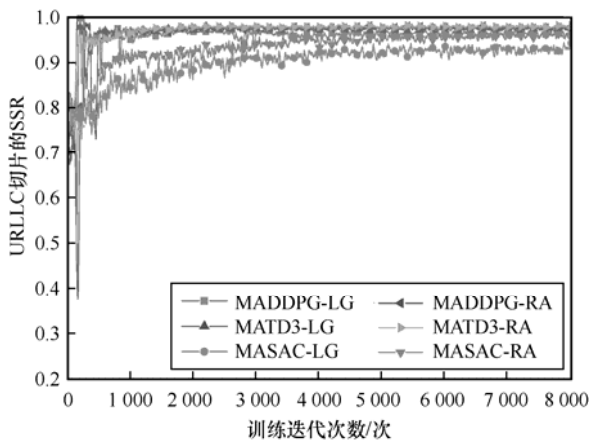


图 12 URLLC 切片的 SSR 随训练迭代次数的变化

### 3) 系统 SE 的对比

系统 SE 随训练迭代次数的变化如图 13 所示。从图 13 可以看出，存在个别高于收敛值的异常点，这是因为在该状态下没有找到合适的分配策略，某

个切片被分配过多的资源，而其他切片的资源较少，导致分配资源较少的切片 SSR 很大程度上不能满足服务指标要求，所以此次奖励很小。所提出的 MADDPG-LG 算法的系统 SE 稳定在 253 bit/(s·Hz) 左右，相较于 MATD3-LG 算法的 232 bit/(s·Hz) 以及 MASAC-LG 算法的 205 bit/(s·Hz) 都有一定程度的提升。

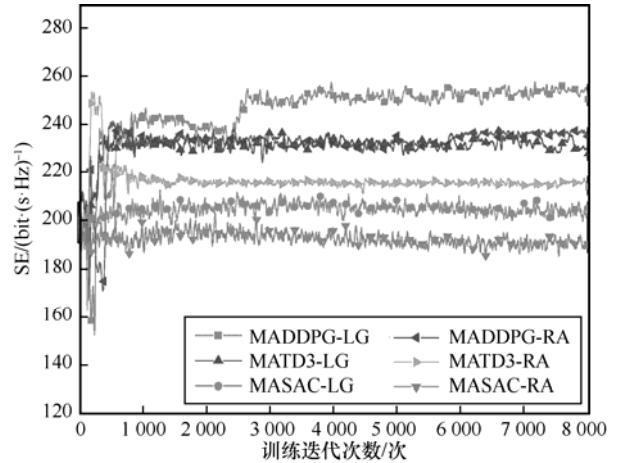


图 13 系统 SE 随训练迭代次数的变化

### 4) 不同基站带宽的系统效用对比

系统效用随着基站带宽的变化如图 14 所示。从图 14 可以看出，随着基站带宽的增加，系统效用呈上升趋势。相比其他基准算法，所提算法系统效用上升趋势具有明显的优势。但是随着基站带宽的持续增加，系统效用的增长速度减慢。

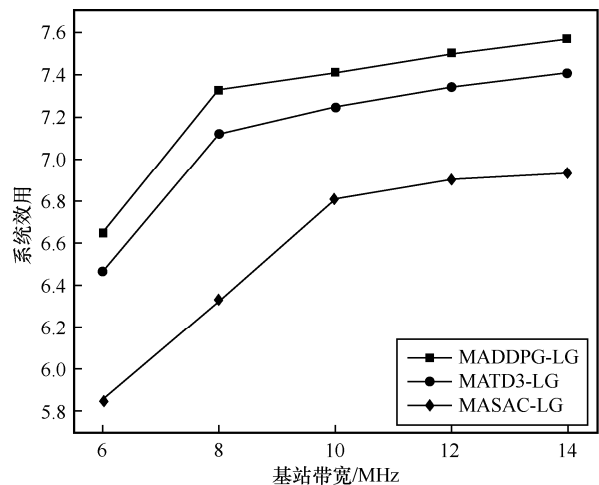


图 14 系统效用随着基站带宽的变化

## 4 结束语

本文面向 UD-LEO 卫星网络辅助的 TSN 切片

场景中的联合用户关联和动态资源分配的问题进行了研究。以最大化 SE 和不同切片 SSR 的加权和为系统效用并将其作为优化目标,提出了一种基于 MADDPG-LG 的联合用户关联和网络切片算法。该算法分别通过基于 MADDPG 的网络切片算法确定资源切片比例,然后基于拉格朗日对偶的用户关联算法确定最优的用户关联策略,最后采用轮询调度将资源分配给用户。将基于 MADDPG-LG 的联合用户关联和网络切片算法与 MADDPG-RA、MATD3-LG、MATD3-RA、MASAC-LG 和 MASAC-RA 算法进行了对比,仿真结果表明,所提算法在满足切片 SLA、系统 SE 以及系统效用等方面具有更好的性能,同时验证了基于拉格朗日对偶的用户关联算法的有效性。

### 参考文献:

- [1] NGUYEN D C, DING M, PATHIRANA P N, et al. 6G Internet of things: a comprehensive survey[J]. *IEEE Internet of Things Journal*, 2022, 9(1): 359-383.
- [2] ZHANG J K, CHEN T H, ZHONG S D, et al. Aeronautical ad-hoc networking for the Internet-above-the-clouds[J]. *Proceedings of the IEEE*, 2019, 107(5): 868-911.
- [3] ZHANG Z J, ZHANG W Y, TSENG F H. Satellite mobile edge computing: improving QoS of high-speed satellite-terrestrial networks using edge computing techniques[J]. *IEEE Network*, 2019, 33(1): 70-76.
- [4] WANG Y T, SU Z, NI J B, et al. Blockchain-empowered space-air-ground integrated networks: opportunities, challenges, and solutions[J]. *IEEE Communications Surveys & Tutorials*, 2022, 24(1): 160-209.
- [5] JIA Z Y, SHENG M, LI J D, et al. Joint HAP access and LEO satellite backhaul in 6G: matching game-based approaches[J]. *IEEE Journal on Selected Areas in Communications*, 2021, 39(4): 1147-1159.
- [6] CAO X L, YANG B, SHEN Y L, et al. Edge-assisted multi-layer offloading optimization of LEO satellite-terrestrial integrated networks[J]. *IEEE Journal on Selected Areas in Communications*, 2023, 41(2): 381-398.
- [7] TANG Z X, YU K, YANG G N, et al. New bridge to cloud: an ultra-dense LEO assisted green computation offloading approach[J]. *IEEE Transactions on Green Communications and Networking*, 2023, 7(2): 552-564.
- [8] HAN D R, YE Q, PENG H X, et al. Two-timescale learning-based task offloading for remote IoT in integrated satellite-terrestrial networks[J]. *IEEE Internet of Things Journal*, 2023, 10(12): 10131-10145.
- [9] 元伟敬, 宋清洋, 郭磊. 面向软件定义多模态车联网的双时间尺度 RAN 切片资源分配[J]. *通信学报*, 2022, 43(4): 60-70.
- QI W J, SONG Q Y, GUO L. Dual time scale resource allocation for RAN slicing in software-defined oriented polymorphic IoV[J]. *Journal on Communications*, 2022, 43(4): 60-70.
- [10] LUU Q T, KERBOEUF S, KIEFFER M. Uncertainty-aware resource provisioning for network slicing[J]. *IEEE Transactions on Network and Service Management*, 2021, 18(1): 79-93.
- [11] 曾锋, 张政, 陈志刚. 基于深度强化学习的计算卸载与资源分配策略[J]. *通信学报*, 2023, 44(7): 124-135.
- ZENG F, ZHANG Z, CHEN Z G. Computation offloading and resource allocation strategy based on deep reinforcement learning[J]. *Journal on Communications*, 2023, 44(7): 124-135.
- [12] LYU F, YANG P, WU H Q, et al. Service-oriented dynamic resource slicing and optimization for space-air-ground integrated vehicular networks[J]. *IEEE Transactions on Intelligent Transportation Systems*, 2022, 23(7): 7469-7483.
- [13] SHEN H, TIAN Y B, WANG T J, et al. Slicing-based task offloading in space-air-ground integrated vehicular networks[J]. *IEEE Transactions on Mobile Computing*, 2023: doi.org/10.1109/TMC.2023.3283.852.
- [14] WU H Q, CHEN J Y, ZHOU C H, et al. Learning-based joint resource slicing and scheduling in space-terrestrial integrated vehicular networks[J]. *Journal of Communications and Information Networks*, 2021, 6(3): 208-223.
- [15] PENG H X, SHEN X M. Multi-agent reinforcement learning based resource management in MEC-and UAV-assisted vehicular networks[J]. *IEEE Journal on Selected Areas in Communications*, 2021, 39(1): 131-141.
- [16] ZHAO N, YE Z Y, PEI Y Y, et al. Multi-agent deep reinforcement learning for task offloading in UAV-assisted mobile edge computing[J]. *IEEE Transactions on Wireless Communications*, 2022, 21(9): 6949-6960.
- [17] QIN Y H, ZHANG Z S, LI X L, et al. Deep reinforcement learning based resource allocation and trajectory planning in integrated sensing and communications UAV network[J]. *IEEE Transactions on Wireless Communications*, 2023, 22(11): 8158-8169.
- [18] FU S, GAO J, ZHAO L. Collaborative multi-resource allocation in terrestrial-satellite network towards 6G[J]. *IEEE Transactions on Wireless Communications*, 2021, 20(11): 7057-7071.
- [19] JIA Z Y, SHENG M, LI J D, et al. Toward data collection and transmission in 6G space-air-ground integrated networks: cooperative HAP and LEO satellite schemes[J]. *IEEE Internet of Things Journal*, 2022, 9(13): 10516-10528.
- [20] CAO H T, DU J B, ZHAO H T, et al. Toward tailored resource allocation of slices in 6G networks with softwarization and virtualization[J]. *IEEE Internet of Things Journal*, 2022, 9(9): 6623-6637.
- [21] AFOLABI I, TALEB T, SAMDANIS K, et al. Network slicing and softwarization: a survey on principles, enabling technologies, and solutions[J]. *IEEE Communications Surveys & Tutorials*, 2018, 20(3): 2429-2453.

- [22] HUA Y X, LI R P, ZHAO Z F, et al. GAN-powered deep distributional reinforcement learning for resource management in network slicing[J]. IEEE Journal on Selected Areas in Communications, 2020, 38(2): 334-349.
- [23] GE X, LI X H, JIN H, et al. Joint user association and user scheduling for load balancing in heterogeneous networks[J]. IEEE Transactions on Wireless Communications, 2018, 17(5): 3211-3225.
- [24] LIU Y J, FENG G, SUN Y, et al. Device association for RAN slicing based on hybrid federated deep reinforcement learning[J]. IEEE Transactions on Vehicular Technology, 2020, 69(12): 15731-15745.
- [25] YE Y K, ZHANG T K, YANG L W. Joint user association and resource allocation for load balancing in RAN slicing[J]. Physical Communication, 2021, 49: 101459.
- [26] WANG Z Y, WEI Y F, YU F R, et al. Utility optimization for resource allocation in multi-access edge network slicing: a twin-actor deep deterministic policy gradient approach[J]. IEEE Transactions on Wireless Communications, 2022, 21(8): 5842-5856.
- [27] LOW S H, LAPSLEY D E. Optimization flow control-I: basic algorithm and convergence[J]. IEEE/ACM Transactions on Networking, 1999, 7(6): 861-874.
- [28] BOYD S P, VANDENBERGHE L. Convex optimization[M]. Cambridge: Cambridge University Press, 2004.
- [29] BERTSEKAS D, TSITSIKLIS J. Parallel and distributed computation: numerical methods[M]. New Jersey: Prentice Hall, 1989.
- [30] BERTSEKAS D. Convex optimization theory[M]. Massachusetts: Athena Scientific, 2009.
- [31] 3GPP. Further advancements for E-UTRA physical layer aspects:

TR36. 814 (V9.0.0) [S]. 2010.

- [32] 3GPP. Service requirements for 5G system (Release 15): TS 22.261 V18.3.0[S]. 2021.

#### [作者简介]



陈赓(1984-), 男, 山东潍坊人, 博士, 山东科技大学副教授、硕士生导师, 主要研究方向为异构网络、泛在网络和软件定义移动网络方面的无线资源管理和优化算法。



邢治薇(1996-), 女, 山东日照人, 山东科技大学硕士生, 主要研究方向为网络切片、资源分配。

沈斐(1983-), 女, 江苏南京人, 博士, 中国科学院上海微系统与信息技术研究所研究员、博士生导师, 主要研究方向为无线通信、边缘计算和雾计算的资源优化。

曾庆田(1976-), 男, 山东潍坊人, 博士, 山东科技大学教授、博士生导师, 主要研究方向为 Petri 网、过程挖掘和知识管理。