

面向用户多样化业务需求的多波束卫星系统动态资源分配算法

柴 蓉, 刘 磊, 梁承超, 陈前斌
(重庆邮电大学通信与信息工程学院, 重庆 400065)

摘 要: 多波束卫星通信系统由于其高吞吐量和高资源利用率而受到广泛关注. 已有研究主要考虑多波束卫星通信系统的信道或功率分配问题, 但较少考虑用户分组和动态资源分配策略的联合优化设计, 导致系统性能受限. 此外, 现有研究往往假设固定的波束覆盖半径, 忽略了波束覆盖半径可变性对波束覆盖性能提升的影响. 本文研究了多波束卫星通信系统中的用户分组和资源分配问题, 提出了一种两阶段资源管理方案. 针对动态和多样化的用户需求, 首先设计一种基于 Voronoi 图的迭代用户分组算法以实现分组之间的负载均衡, 然后将子信道和功率分配问题建模为系统平均效用函数最大化问题. 为解决该问题, 将每个波束视为一个智能体, 采用一种基于多智能体深度 Q 网络 (Deep Q Network, DQN) 的算法来确定子信道和功率分配策略. 仿真结果表明, 与 K -均值用户分组方案相比, 本文所提出的基于 Voronoi 图的迭代用户分组算法对应的用户组负载差异值可降低 49.2%, 体现了本文所提算法在实现用户组间负载均衡方面的优势. 此外, 本文所提两阶段资源管理方案与现有文献中所提算法相比, 系统所提供容量与用户需求差值可降低 83.43%, 体现了本文所提算法在实现系统资源高效利用及用户服务需求保障方面的性能优势.

关键词: 多波束卫星; 用户分组; 子信道分配; 功率分配; 多智能体 DQN; 负载均衡

基金项目: 国家自然科学基金 (No.62271097, No.62071078)

中图分类号: TN927.23

文献标识码: A

文章编号: 0372-2112(2024)07-2438-11

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

DOI: 10.12263/DZXB.20240107

Diverse User Service Requirement-Oriented Dynamic Resource Allocation Algorithm for Multi-Beam Satellite Systems

CHAI Rong, LIU Lei, LIANG Cheng-chao, CHEN Qian-bin

(School of Communications and Information Engineering, Chongqing University of Posts and Telecommunications, Chongqing 400065, China)

Abstract: Multi-beam satellite communication systems have received considerable attention due to their high throughput and resource utilization. Existing research considers the channel or power allocation problems in multi-beam satellite communication systems but rarely addresses the joint optimization design of user grouping and dynamic resource allocation strategies, which limits system performance. Furthermore, current studies often assume a fixed beam coverage radius, overlooking the impact of variable beam coverage radius on improving beam coverage performance. In this paper, we study the problem of user grouping and resource allocation in multi-beam satellite communication systems, and propose a two-stage resource management scheme. Addressing the dynamic and diverse user service requirements, we first design a Voronoi diagram-based iterative user grouping algorithm to achieve load balancing among user groups. Then, we formulate the subchannel and power allocation problem as a system average utility function maximization problem. To solve the problem, we regard each satellite beam as an agent, and propose a multi-agent deep Q network (DQN)-based algorithm to determine the subchannel and power allocation strategy. Simulation results demonstrate that the iterative user grouping algorithm based on Voronoi diagram proposed in this paper reduces the discrepancy in user group loads by 49.2% compared to the K -means user grouping scheme, highlighting the advantage of the proposed algorithm in achieving load balancing among user groups. Furthermore, the two-stage resource management scheme presented in this paper, when compared to algorithm proposed in existing literature, reduces the gap between system capacity and user demand by 83.43%, showcasing the performance advantage of the proposed algorithm in efficiently utilizing system resources and ensuring user service demands.

Key words: multi-beam satellite; user grouping; subchannel allocation; power allocation; multi-agent deep Q network; load balancing

Foundation Item(s): National Natural Science Foundation of China (No.62271097, No.62071078)

1 引言

随着地面通信系统的快速发展,第五代移动通信技术(5th Generation mobile communication system, 5G)已获得广泛应用.然而,由于一些地区的复杂地形和高昂的设施建设成本,蜂窝通信系统的全球覆盖问题仍然难以解决.卫星通信系统作为地面通信系统的有力补充,具有覆盖范围广、系统容量大、不受地理条件限制及抗干扰能力强等特点^[1,2].为了满足日益增长的卫星通信需求,可应用多波束卫星通信系统.通过采用能量集中的多个高功率点波束,多波束卫星系统有望显著提升系统吞吐量并提高资源利用率^[3].在多波束卫星通信系统中,用户分组和资源分配策略是影响用户服务质量及系统性能的重要问题.

参考文献[4~7]研究了卫星系统中的用户分组问题.文献[4,5]以系统信道容量最大化为目标分别提出了用户分组方案.文献[6]联合设计了用户分组策略和最优波束宽度,以实现卫星的平均数据速率最大化并保证用户的服务质量.文献[7]中作者基于最大化系统能量效率原则,设计了一种卫星用户选择和地面用户分组策略,并提出一种基于能量效率的星地非正交多址网络功率分配算法.尽管已有文献设计卫星通信系统的用户分组方案^[4~7],现有研究通常基于用户位置及用户之间相互距离设计分组策略,较少考虑用户业务类型、请求速率等因素,可能导致组间负载差异较大、系统资源利用率下降及资源分配性能受限.

文献[8~10]研究了多波束卫星通信系统的频率或功率资源分配问题.文献[8]通过优化相邻波束之间子信道的可变数量实现系统吞吐量优化.文献[9]提出了一种卫星通信系统时域及子信道分配方案,以提高传输效率,并以较低的复杂度满足所请求的业务.文献[10]中作者针对星地融合网络下行链路增强型移动宽带用户及高可靠低延迟通信用户数据传输场景,研究功率及时频资源调度问题.文献[11~14]研究了多波束卫星通信系统的联合波束照射、功率和频率资源分配问题.文献[11,12]研究了地面业务分布不均匀场景下的跳波束和带宽分配问题.文献[13]联合优化功率分配和波束调度,以减小用户的流量需求和提供的容量之间的差距.文献[14]联合考虑卫星发送功率、波束增益和波束照射位置,以实现流量容纳率最大化.尽管文献[8~14]研究卫星通信系统的资源分配问题,现有研究较少考虑用户分组和动态资源分配策略的联合优化设计,导致系统性能受限.此外,现有研究多考虑确定场

景的资源分配策略,未能充分考虑实际场景中业务随机到达及信道随机性导致的系统不确定性,导致所提算法难以适用于实际动态随机场景.

文献[15~17]采用深度强化学习算法确定卫星通信系统中的资源分配策略.文献[15]提出了一种基于多智能体深度确定性策略梯度方法,联合优化设计功率控制及缓存部署策略,实现系统能量效率的优化.文献[16]中,为最大限度减小干扰,作者提出了一种多智能体深度强化学习算法来建立多波束上行链路信道分配策略.文献[17]提出了一种基于双延迟深度确定性策略梯度的卫星用户联合子信道及功率分配算法,实现用户传输速率最大化.尽管文献[15~17]考虑系统随机特性设计卫星网络设计资源分配算法,但主要考虑面向小区/用户组的资源分配策略,较少考虑面向用户的细粒度资源分配策略;部分研究仅考虑静态或准静态场景的子信道或功率分配方案,难以实现系统长期性能提升及多维资源的联合优化设计.

针对现有研究存在的问题,本文研究多波束卫星系统中的用户分组、子信道和功率分配问题.考虑系统长期性能优化,将该问题建模为平均效用函数最大化问题.为求解所建模问题,提出了一种两阶段资源管理算法,首先设计基于Voronoi图的迭代用户分组算法,并基于最小圆覆盖对波束覆盖半径进行优化.进而提出基于多智能体深度 Q 网络(Deep Q Network, DQN)的动态资源分配算法,以确定联合资源分配策略.

2 系统模型

2.1 系统场景

如图1所示,本文考虑了一个由高轨道同步卫星、多个地面用户(Ground Users, GUs)和一个地面网关组成的多波束卫星通信系统.假设卫星配备信号发送设备,能够通过用户链路向GUs传输数据.为提高数据传输性能,卫星支持多波束传输,每个波束覆盖特定区域内的GUs.令 N 表示卫星波束的数量, M 表示卫星覆盖区域内的GU的数量.令用户 m 为 GU_m , c_m 表示 GU_m 的位置向量,可表示为 (x_m, y_m) , $1 \leq m \leq M$.

令 B_{tot} 表示卫星的总带宽,将总带宽划分为 K 个等长的子信道,每个子信道的带宽为 B_0 ,即 $B_0 = B_{tot}/K$,令 f_k 表示第 k 个子信道的载波频率.为避免波束内干扰,同一个波束覆盖的GU仅可占用不同的子信道.各波束之间采用全频率复用方案以提高频率资源利用率.将系统时间 T 划分为连续的等长时隙.不失一般性,假设

卫星到 GU 的信道特性可能随时间变化,但信道特性在
各时隙内保持不变.

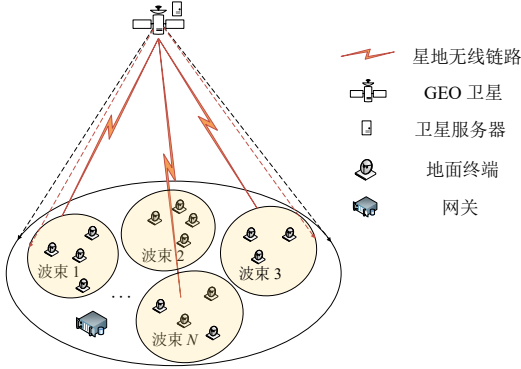


图1 系统模型图

2.2 通信业务建模

假设多波束卫星系统支持多种用户业务类型. 令 L 表示用户业务类型的数量, $r_{l,t}$ 表示 GU 在 t 时隙请求第 l 类服务时所需业务传输速率. 本文将 $r_{l,t}$ 建模为截断高斯分布随机变量, 其概率密度函数可表示为

$$f_{r_{l,t}}(x) = \frac{1}{\zeta} \exp\left[-\frac{(x-\mu_l)^2}{2\sigma_l^2}\right] \quad (1)$$

其中, μ_l 和 σ_l^2 分别表示 $r_{l,t}$ 的均值和方差, ζ 为归一化常数, 表示为

$$\zeta = \frac{\sqrt{2\pi}\sigma_l}{2} \left[\operatorname{erf}\left(\frac{\mu_l - R_l^{\min}}{\sqrt{2}\sigma_l}\right) - \operatorname{erf}\left(\frac{R_l^{\min} - \mu_l}{\sqrt{2}\sigma_l}\right) \right] \quad (2)$$

其中, $\operatorname{erf}(u) = \frac{2}{\sqrt{\pi}} \int_0^u \exp(-x^2) dx$ 为高斯误差函数, R_l^{\min} 表示第 l 类业务的最小需求速率.

令 $b_{m,l} \in \{0, 1\}$ 表示用户业务标识, 若 GU_m 需要第 l 类服务, 则 $b_{m,l} = 1$, 否则 $b_{m,l} = 0$. 假设每个 GU 仅选择一种类型的服务, 即 $\sum_{l=1}^L b_{m,l} = 1, 1 \leq m \leq M$.

2.3 通信信道建模

令 $h_{t,n,m,k}$ 表示 t 时隙卫星通过波束 n 占用子信道 k 传输数据至 GU_m 时对应用户链路的信道增益, 建模为

$$h_{t,n,m,k} = g_{t,n,m}^T g_m^R L_{t,m,k} L_t^r \quad (3)$$

其中, $g_{t,n,m}^T$ 表示发射天线增益, 建模为

$$g_{t,n,m}^T = g_n^{\max} \left(\frac{J_1(u_{t,n,m})}{2u_{t,n,m}} + 36 \frac{J_3(u_{t,n,m})}{(u_{t,n,m})^3} \right)^2 \quad (4)$$

其中, $J_1(\cdot)$ 和 $J_3(\cdot)$ 分别表示第一类一阶和三阶贝塞尔函数, g_n^{\max} 表示波束 n 的发射天线的最大增益, 建模为

$$g_n^{\max} = \eta \left(\frac{70\pi}{\theta_n} \right)^2 \quad (5)$$

其中, η 表示天线效率, θ_n 表示波束 n 的 3 dB 波束宽度对应的角度. 假设卫星可提供灵活的波束覆盖, 令 δ_n 表示波束 n 的覆盖半径. δ_n 的计算公式为

$$\delta_n = h_s \tan\left(\frac{\theta_n}{2}\right) \quad (6)$$

其中, h_s 表示卫星的高度.

式(4)中, $u_{t,n,m}$ 建模为

$$u_{t,n,m} = \frac{2.07123 \sin(\theta_{t,n,m})}{\sin(\theta_n)} \quad (7)$$

其中, $\theta_{t,n,m}$ 表示波束 n 的中心坐标与 GU_m 之间的离轴角. 令 (x_n^c, y_n^c) 表示波束 n 的中心坐标, 则 $\theta_{t,n,m}$ 表示为

$$\theta_{t,n,m} = \tan^{-1} \frac{\sqrt{(x_m - x_n^c)^2 + (y_m - y_n^c)^2}}{h_s} \quad (8)$$

式(3)中, g_m^R 表示 GU_m 的接收天线增益, $L_{t,m,k}$ 表示 t 时隙卫星占用子信道 k 传输数据至 GU_m 时对应链路的自由空间损耗, 建模为

$$L_{t,m,k} = \left(\frac{c}{4\pi d_{t,m} f_k} \right)^2 \quad (9)$$

其中, c 为光速, $d_{t,m}$ 为 t 时隙卫星与 GU_m 之间的距离. 式(3)中, L_t^r 表示用户链路的雨衰系数, 将其建模为对数正态分布, 其中, μ_t 和 σ_t^2 分别表示其均值和方差.

2.4 两阶段资源管理方案概述

本文假设卫星通信系统中的 GU 由卫星波束照射, 并为 GU 分配波束、子信道及功率资源策略. 为实现资源的高效利用, 需设计高效的波束及用户关联、子信道及功率资源分配策略. 具体而言, 根据给定的卫星波束半径及 GU 的地理位置, 可设计卫星波束照亮策略, 使得在一段时间内, 卫星的单个波束可照亮地面某个区域的若干个 GU; 而对于 GU 而言, 则需根据其不同业务需求分配相应的子信道和功率资源. 因此, 考虑到卫星波束照射的时间和区域尺度较大, 而用户的子信道及功率资源分配策略的时间和区域尺度较小, 本文提出一种两阶段资源管理方案.

阶段一, 首先基于较大时间空间尺度设计用户分组及波束照射策略. 由于卫星可产生的波束资源数量及波束最大半径受限, 为保障卫星波束可为卫星覆盖区域内的所有用户提供服务, 本文提出一种基于 Voronoi 图的迭代用户分组算法将地理位置分布不均匀的用户划分为多个用户组, 并进而确定波束与用户组的关联策略. 阶段二, 基于所得到的用户分组及波束与用户的关联策略, 本文在较小时间尺度设计 GU 的子信道和功率分配策略. 首先定义系统平均效用函数, 并将

GU 的子信道和功率分配问题建模为系统平均效用函数最大化问题,进而将优化问题建模为 MDP 问题,并提出一种基于多智能体 DQN 的算法以确定子信道和功率分配策略. 本文所提两阶段资源管理方案的流程图如图 2 所示.

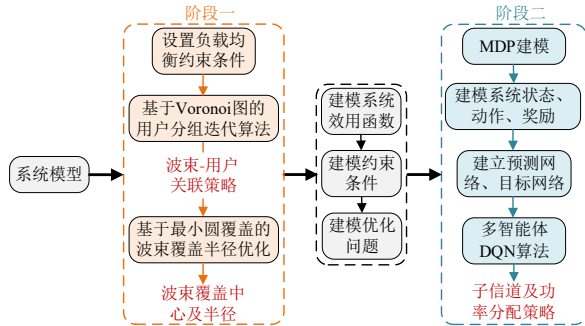


图2 两阶段资源管理方案流程图

3 基于 Voronoi 图的迭代用户分组算法

本节考虑用户分布和业务需求,提出一种基于 Voronoi 图的迭代用户分组算法.

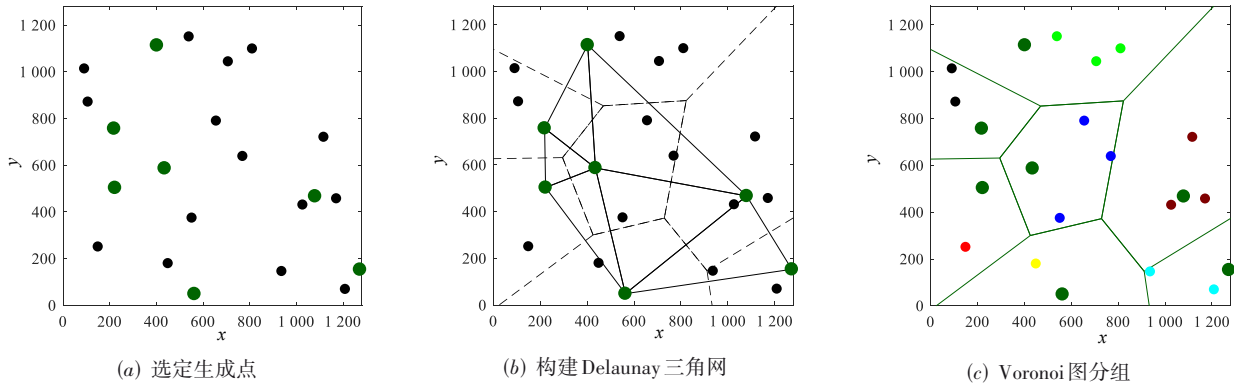


图3 Voronoi 图分组示意图

不失一般性,本文将 Voronoi 生成点的数量设置为卫星波束的数量 N . 令 $V = \{q_1, q_2, \dots, q_N\}$ 表示生成点的坐标向量集合, q_n 为第 n 个点的坐标向量,可表示为 (x_n^v, y_n^v) . 令 V_n 表示第 n 个生成点附近的用户坐标集合,其中各用户与 q_n 的距离较其他点更接近. 因此, V_n 可表示为

$$V_n = \{c_m \mid \|c_m - q_n\| \leq \|c_m - q_{n'}\|, \forall n' \neq n\} \quad (10)$$

显然,生成点的选择决定了经 Voronoi 图划分后各用户组的数量、大小、几何形状和位置,因此,通过设计 V 即可控制分区的所有属性,本文 3.3 小节对算法流程进行详述.

3.2 负载均衡约束条件

由于 GU 的业务需求为随机变量,为了评估每组的

3.1 Voronoi 图划分

由于 GU 所在区域由多个卫星波束覆盖,本文首先将 GU 划分为不同用户组,进而为各用户组分配子信道及功率资源. 本小节首先采用 Voronoi 图用户分组方法将卫星覆盖区域划分为多个区域. Voronoi 图是一组由连接两生成点线段的垂直平分线组成的连续多边形,每个连续多边形内的任一点到构成该多边形的生成点的距离小于到其他多边形生成点的距离. 基于 Voronoi 图分组方法的基本思想是生成一组有限点,其中,各点代表了其附近的一定子区域,进而依次将用户划分到距离其最近的子区域中. 以生成点数量为 7、用户数量为 15 为例,基于 Voronoi 图的区域划分及用户分组流程如图 3 所示. 首先,给定各用户及生成点的位置;进而连接各点以构成 Delaunay 三角网,并绘制各三角形中所有边的中垂线,其中 Delaunay 三角网是一系列相连但不重叠的三角形的集合,且这些三角形的外接圆不包含这个面域的其他任何点;最后,根据各中垂线及其交点将区域划分为 7 个子区域,每个子区域内的用户构成一个用户组.

流量需求,本节采用各类业务请求速率的平均值刻画表示各 GU 的需求. 令 R_n^i 表示第 n 个用户组中 GU 的平均业务需求之和, $1 \leq n \leq N$. R_n^i 可表示为

$$R_n^i = E \left(\sum_{m=1}^M \sum_{l=1}^L \alpha_{n,m} b_{m,l} r_{i,l} \right) = \sum_{m=1}^M \sum_{l=1}^L \alpha_{n,m} b_{m,l} \mu_l \quad (11)$$

其中, $\alpha_{n,m} \in \{0, 1\}$ 是波束照射变量,若 GU_{*m*} 由波束 n 提供服务,则 $\alpha_{n,m} = 1$,反之 $\alpha_{n,m} = 0$. 假设每个 GU 最多只能与一个波束相关联,即对于任意 m ,有 $\sum_{n=1}^N \alpha_{n,m} \leq 1$;且与每个波束相关联的 GU 的数量不能超过子信道的数量,即对于任意 n ,有 $\sum_{m=1}^M \alpha_{n,m} \leq K$.

为了满足组间负载均衡的目标,本节引入了负载均衡阈值,用 R_{th} 表示,并将负载均衡约束条件设置为

$$|R^{\max} - R^{\min}| \leq R_{th} \quad (12)$$

其中, $R^{\max} = \max\{R_1^r, R_2^r, \dots, R_N^r\}$, $R^{\min} = \min\{R_1^r, R_2^r, \dots, R_N^r\}$.

3.3 基于 Voronoi 图的用户分组迭代算法流程

本小节对所提出的基于 Voronoi 图的用户分组迭代算法的算法流程进行详述. 首先在 GU 所在区域内随机生成 N 个点, 并基于 Voronoi 图确定各点对应的用户区域, 即 $V_n, 1 \leq n \leq N$, 进而判断各组的平均需求是否满足式(12)中的约束条件. 若是, 则算法结束, 否则应用迭代方法更新生成点的位置, 重复这一过程, 直到满足负载均衡约束条件. 需指出的是, 本小节提出的基于 Voronoi 图的用户分组迭代算法本质上是一个通过移动生成点来实现负载均衡的迭代过程. 为实现生成点的高效移动及算法快速收敛, 本文采用类似分子扩散过程的负载梯度实现生成点的变更驱动.

令 F_n 表示第 n 个生成点的驱动力, 定义为其他生成点在 q_n 处的合力, 即

$$F_n = \sum_{\substack{n'=1 \\ n' \neq n}}^N F_{n',n} \quad (13)$$

其中, $F_{n',n}$ 表示点 $q_{n'}$ 对点 q_n 的作用力, 建模为

$$F_{n',n} = \text{sign}(R_{n'}^r - R_n^r) \frac{|R_{n'}^r - R_n^r|^a}{(r_{n',n})^b} \tilde{r}_{n',n} \quad (14)$$

其中, a 和 b 均为用于调整负载差异影响的参数, $\text{sign}(\cdot)$ 表示符号函数, $r_{n',n} = \|q_{n'} - q_n\|$ 表示 $q_{n'}$ 和 q_n 之间的距离, $\tilde{r}_{n',n}$ 为单位向量, 建模为

$$\tilde{r}_{n',n} = \frac{r_{n',n}}{r_{n',n}} = \frac{q_{n'} - q_n}{\|q_{n'} - q_n\|} \quad (15)$$

设 $q_n^{(i)}$ 表示第 i 次迭代时第 n 个生成点的坐标, $F_n^{(i)}$ 表示第 i 次迭代时第 n 个生成点的驱动力. 根据动力学方程, 得到生成点的更新迭代方程:

$$q_n^{(i+1)} = q_n^{(i)} + \mu F_n^{(i)} \quad (16)$$

其中, μ 表示保证动力学演进稳定性的扩散或粘滞系数.

本文所提出的基于 Voronoi 图的用户分组迭代算法的步骤如下:

- (1) 令 $i=0$, 给定生成点集合 $\{q_1^{(i)}, q_2^{(i)}, \dots, q_N^{(i)}\}$;
- (2) 生成 Voronoi 图, 将 GU 划分至图中各关联区域;
- (3) 根据式(11)计算各组的流量负载;
- (4) 评估各组是否满足式(12)中的负载约束条件, 若是, 则算法终止, 输出用户分组策略;
- (5) 若否, 则根据式(13)计算各生成点的合力;
- (6) 根据式(16)更新生成点的位置, 得到生成点位置集合 $\{q_1^{(i+1)}, q_2^{(i+1)}, \dots, q_N^{(i+1)}\}$, 返回步骤(2).

应用本节提出的基于 Voronoi 图的迭代用户分组算法, 可将所有 GUs 划分为多个分组. 不失一般性, 假设波束 n 为第 n 组 GU 提供业务支持, 令 $\alpha_{n,m}^*$ 表示所得用户分组及波束照亮策略.

3.4 基于最小圆覆盖的波束覆盖半径优化

本文假设一个组由一个对应波束服务. 由于波束覆盖半径与波束发射天线的最大增益成反比, 为了提高信道增益, 波束应该覆盖具有最小可能半径的相关组中的 GUs. 本小节中, 采用最小圆覆盖方法来确定每个波束的中心坐标和覆盖半径.

最小圆覆盖法的主要思想是: 对于平面上的多个点, 给定每个点的坐标, 找到一个半径最小的圆来覆盖所有这些点. 假设给定组 n 的用户集, 用 V_n 表示. 基于最小圆覆盖确定波束 n 的中心坐标和覆盖半径的过程如下:

(1) 任意选择集合 V_n 中的两个点, 分别标记为 c_A 和 c_B ;

(2) 以连接两点 c_A 和 c_B 的线段作为直径, 线段的中点为圆心, 构造一个圆. 若所得到的圆形区域覆盖了 V_n 中的所有点, 则算法终止, 对应圆形区域的中心及半径即为波束 n 的中心点及覆盖半径; 否则, 选择圆外的一个点并将其标记为 c_C ;

(3) 若 c_A, c_B 和 c_C 形成的三角形为直角三角形或钝角三角形, 则重新选择圆外一点作为 c_C , 并判断所得三角形形状; 若所得三角形为锐角三角形, 则转至步骤(4);

(4) 由所得锐角三角形构造一个外接圆, 若该圆包含 V_n 中的所有点, 则算法结束, 对应圆形区域的中心及半径即为波束 n 的中心点及覆盖半径, 否则, 选择该外接圆外的一点, 将其标记为 c_D , 并选择 $\{c_A, c_B, c_C\}$ 中离 c_D 最远的点, 重新标记为 c_E . 在外接圆中, 确定过点 c_E 的圆的直径, 延长该直径后将平面分成两个半平面, 找到 $\{c_A, c_B, c_C\}$ 中与点 c_D 不在同一半平面中的点, 重新标记为 c_F , 令 $\{c_D, c_E, c_F\}$ 为 $\{c_A, c_B, c_C\}$, 转至步骤(3).

通过对每组使用最小圆覆盖方法, 可确定每个波束的中心坐标 (x_n^c, y_n^c) 和半径 δ_n . 需要注意的是, 波束覆盖半径不能小于相控阵天线可控范围的最小半径. 设 δ_0 表示波束的最小半径, 可表示为

$$\delta_0 = 0.443 \frac{c}{fD} h_s \quad (17)$$

其中, f 是卫星的载波频率, D 是卫星相控阵天线的直径. 若 $\delta_n \leq \delta_0$, 则设 $\delta_n = \delta_0$.

4 问题建模

为实现用户业务需求与系统业务提供之间的匹配, 本小节基于两者之间的差值建模系统效用函数, 进

而将系统长期动态资源分配问题建模为带约束的系统平均效用函数最大化问题.

4.1 系统效用函数建模

令 U 表示系统平均效用函数, 建模为

$$U = \frac{1}{T} \mathbb{E} \left(\sum_{t=1}^T U_t \right) \quad (18)$$

其中, U_t 表示 t 时隙系统效用函数, 建模为

$$U_t = \sum_{m=1}^M U_{t,m} \quad (19)$$

其中, $U_{t,m}$ 表示 t 时隙 GU_m 的效用函数, 建模为

$$U_{t,m} = \begin{cases} \lambda_1, & |R_{t,m} - R_{t,m}^r| \leq R_d \\ \varepsilon \lambda_1, & R_{t,m} - R_{t,m}^r > R_d \\ \lambda_2, & \text{其他} \end{cases} \quad (20)$$

其中, λ_1 为奖励因子, λ_2 为惩罚因子, ε 为奖励折扣因子, R_d 表示差异阈值. 若用户实际传输速率 $R_{t,m}$ 与用户需求速率 $R_{t,m}^r$ 之差的绝对值小于等于差异阈值 R_d , 也即系统提供容量与用户业务需求之间差异较小, 定义效用函数为奖励因子 λ_1 , 且令 $\lambda_1 > 0$; 若用户实际传输速率较用户需求速率高, 且两者之间差异超过阈值 R_d , 尽管此时系统传输速率可满足用户需求, 但存在系统资源浪费的情况, 故予以效用函数一定奖励折扣值, 也即 $\varepsilon \lambda_1$, 令 $0 < \varepsilon < 1$; 若用户实际传输速率远低于需求速率, 也即系统传输性能较差, 无法满足用户需求, 予以惩罚 λ_2 , 令 $\lambda_2 < 0$.

式(20)中, $R_{t,m}^r = \sum_{l=1}^L b_{m,l} r_{t,l}$ 表示 t 时隙 GU_m 的业务请求速率, $R_{t,m}$ 表示 t 时隙 GU_m 的实际传输速率, 建模为

$$R_{t,m} = \sum_{n=1}^N \sum_{k=1}^K \alpha_{n,m}^* \beta_{t,m,k} B_0 \log_2(1 + \gamma_{t,n,m,k}) \quad (21)$$

其中, $\beta_{t,m,k}$ 表示 t 时隙 GU_m 的子信道分配变量, 若 t 时隙 GU_m 占用第 k 个子信道, 则 $\beta_{t,m,k} = 1$, 否则 $\beta_{t,m,k} = 0$. $\gamma_{t,n,m,k}$ 为信干噪比 (Signal-to-Interference-plus-Noise Ratio, SINR), 建模为

$$\gamma_{t,n,m,k} = \frac{p_{t,n,m,k} h_{t,n,m,k}}{I_{t,n,m,k} + N_0 B_0} \quad (22)$$

其中, $p_{t,n,m,k}$ 表示 t 时隙波束 n 占用第 k 个子信道向 GU_m 发送数据时对应的发送功率, N_0 为噪声功率密度, $I_{t,n,m,k}$ 表示 t 时隙 GU_m 接收到的波束间干扰, 可建模为

$$I_{t,n,m,k} = \sum_{m'=1}^M \sum_{n'=1}^N \sum_{m''=1}^K \alpha_{n',m'}^* \beta_{t,m',k} p_{t,n',m',k} h_{t,n',m',k} \quad (23)$$

4.2 约束条件建模

本小节详细讨论优化问题约束条件.

(1) 子信道分配约束

每个用户在任一时隙最多只能占用一个子信

道, 即

$$\text{C1: } \sum_{k=1}^K \beta_{t,m,k} \leq 1, \quad \forall m \quad (24)$$

为防止过度的同频干扰, 不允许由同一波束服务的 GU 使用相同的子信道, 即

$$\text{C2: } \sum_{m=1}^M \alpha_{n,m} \beta_{t,m,k} \leq 1, \quad \forall k, n \quad (25)$$

(2) 功率分配约束

只有当 GU 占用了子信道时, 才能为其分配相应的功率资源, 即

$$\text{C3: } p_{t,n,m,k} > 0, \alpha_{n,m} \beta_{t,m,k} = 1 \quad (26)$$

设 P_n^{\max} 表示波束 n 的最大功率, 最大功率约束可表示为

$$\text{C4: } \sum_{m=1}^M \sum_{k=1}^K \alpha_{n,m} \beta_{t,m,k} p_{t,n,m,k} \leq P_n^{\max}, \quad \forall n \quad (27)$$

设 P_{tot} 表示卫星的总发射功率, 则所有波束使用的总功率资源不应超过 P_{tot} , 即

$$\text{C5: } \sum_{n=1}^N \sum_{m=1}^M \sum_{k=1}^K \alpha_{n,m} \beta_{t,m,k} p_{t,n,m,k} \leq P_{\text{tot}} \quad (28)$$

(3) 服务质量约束

每个用户的实际传输速率不能低于最小业务请求速率, 即

$$\text{C6: } R_{t,m} \geq \sum_{l=1}^L b_{m,l} R_l^{\min}, \quad \forall m \quad (29)$$

4.3 优化问题建模

本文综合考虑卫星通信系统资源状况、地面服务需求和系统资源共享, 将所考虑的多波束卫星系统的资源分配问题建模为带约束条件下系统效用函数最大化问题:

$$\begin{aligned} \max_{\beta_{t,m,k}, P_{t,n,m,k}} \quad & U \\ \text{s.t.} \quad & \text{C1} \sim \text{C6}. \end{aligned} \quad (30)$$

通过求解上述问题, 可获得联合子信道和功率分配策略.

5 优化问题求解

式(30)中的优化问题是一个动态混合整数非线性规划问题, 使用传统算法很难求解. 本节中, 将子信道和功率分配问题建模为马尔可夫决策过程 (Markov Decision Process, MDP), 并提出一种基于多智能体 DQN 的算法来确定资源分配策略.

5.1 MDP 建模

将所提出的问题建模为由三元组 $\langle S, A, R \rangle$ 表示的 MDP, 其中, S 表示系统状态集, A 为动作集, R 为系统奖励. MDP 的状态、动作和奖励函数具体如下:

(1) 状态: 令 $S_t = \{s_{t,1}, s_{t,2}, \dots, s_{t,N}\}$ 表示 t 时隙的全局

状态,其中, $s_{t,n}$ 表示 t 时隙波束 n 的状态. $s_{t,n}$ 包含波束 n 所服务的GUs的信道增益和通信需求,建模为

$$s_{t,n} = (\mathbf{h}_{t,n,m}, R_{t,m}^r | \alpha_{n,m} = 1, \forall m) \quad (31)$$

其中, $\mathbf{h}_{t,n,m}$ 表示波束 n 照射的GU m 的信道增益矢量,表示为

$$\mathbf{h}_{t,n,m} = (h_{t,n,m,1}, h_{t,n,m,2}, \dots, h_{t,n,m,K}) \quad (32)$$

(2)动作:令 $A_t = \{a_{t,1}, a_{t,2}, \dots, a_{t,N}\}$ 表示 t 时隙的全局动作,其中 $a_{t,n}$ 表示 t 时隙波束 n 的动作. $a_{t,n}$ 包含波束 n 所服务的GUs的子信道分配和功率分配变量,建模为

$$a_{t,n} = (\boldsymbol{\beta}_{t,m}, \mathbf{p}_{t,n,m} | \alpha_{n,m} = 1, \forall m) \quad (33)$$

其中, $\boldsymbol{\beta}_{t,m} = (\beta_{t,m,1}, \beta_{t,m,2}, \dots, \beta_{t,m,K})$ 表示GU m 在 t 时隙的子信道分配向量, $\mathbf{p}_{t,n,m} = (p_{t,n,m,1}, p_{t,n,m,2}, \dots, p_{t,n,m,K})$ 表示 t 时隙波束 n 向GU m 发送数据时的功率分配向量.由于 $p_{t,n,m,k}$ 是连续变量,产生的连续的动作空间不能直接应用于DQN.为解决该问题,利用离散化机制将连续的功率分配变量转换为离散的功率阶数.具体地,将卫星波束的发射功率均匀地划分为 J 个阶数,且设定只有特定的功率可以分配给GUs.设 p_j 表示第 j 阶的功率,建模为

$$p_j = \frac{P_{\text{tot}} j}{J} \quad (34)$$

令 $q_{t,m,j}$ 表示 t 时隙GU m 的功率阶数选择变量, $1 \leq j \leq J$.若 t 时隙卫星波束发送至GU m 的功率为 p_j ,则 $q_{t,m,j} = 1$,否则 $q_{t,m,j} = 0$.任一时隙一个用户仅能选择一个功率阶数.因此,可重写功率分配变量 $p_{t,n,m,k}$ 为

$$p_{t,n,m,k} = \sum_{j=1}^J \alpha_{n,m} \beta_{t,m,k} q_{t,m,j} p_j \quad (35)$$

相应地,动作 $a_{t,n}$ 可重写为

$$a_{t,n} = \{\boldsymbol{\beta}_{t,m}, \mathbf{p}_{t,m} | \alpha_{n,m} = 1, \forall m\} \quad (36)$$

其中, $\mathbf{p}_{t,m} = [q_{t,m,1} \ q_{t,m,2} \ \dots \ q_{t,m,J}]$ 表示GU m 在 t 时隙的离散功率阶数选择矩阵.

(3)奖励函数:单个智能体通过选择一个动作来获得相应的奖励.为有效地分配功率和频率资源以实现系统效用函数最大化,本文定义为 t 时隙的奖励函数为 $R_t = U_t$.

5.2 基于多智能体DQN的资源分配算法

本小节中,将 N 个波束视为 N 个智能体,智能体之间相互合作与竞争以最大限度地提高系统收益.每个智能体由预测网络和目标网络组成.预测网络用于评估不同的动作,其网络参数实时更新;目标网络用于辅助预测网络更新网络参数,其网络参数基于预测网络按照一定间隔进行更新.多智能体DQN的框架如图4所示.

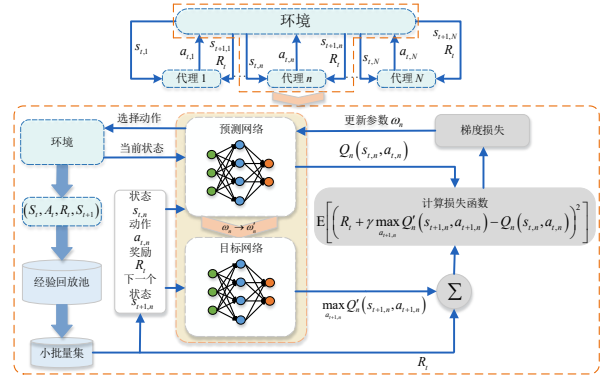


图4 MADQN框架示意图

本节所用算法的基本思想概括如下.首先,初始化所有智能体的各神经网络、经验池 D_b 和系统状态 S_t ;对于智能体 n ,给定状态 $s_{t,n}$,其预测网络通过应用 ϵ -贪婪策略选择动作 $a_{t,n}$;通过执行由各智能体动作组成的联合动作 A_t ,得到全局奖励 R_t ,系统转移至状态 s_{t+1} ,并将四元组 (S_t, A_t, R_t, S_{t+1}) 存储到经验回放池 D_b 中;当 D_b 中有足够的样本时,随机提取一些样本,并使用最小均方误差(Minimum Square Error, MSE)作为损失函数来更新预测网络参数.智能体 n 的损失函数可表示为

$$L(\omega_n) = E_{(S_t, A_t, R_t, S_{t+1}) \sim D_b} \left[\left(y_{t,n} - Q_n(s_{t,n}, a_{t,n} | \omega_n) \right)^2 \right] \quad (37)$$

其中, $Q_n(s_{t,n}, a_{t,n} | \omega_n)$ 表示智能体 n 的预测网络的 Q 值, ω_n 为智能体 n 的预测网络的参数.令 $y_{t,n}$ 表示用于更新预测网络的网络参数的标签值,定义为

$$y_{t,n} = R_t + \gamma \max_{a_{t+1,n}} Q'_n(s_{t+1,n}, a_{t+1,n} | \omega'_n) \quad (38)$$

其中, γ 为折扣因子, $Q'_n(s_{t,n}, a_{t,n} | \omega'_n)$ 表示智能体 n 的目标网络的 Q 值, ω'_n 是智能体 n 的目标网络的参数.通过迭代更新时间差来估计相应的动作值函数,表示为

$$Q_n^*(s_{t,n}, a_{t,n}) = Q_n(s_{t,n}, a_{t,n}) + \alpha \left[R_t + \gamma \max_{a_{t+1,n}} Q'_n(s_{t+1,n}, a_{t+1,n}) - Q_n(s_{t,n}, a_{t,n}) \right] \quad (39)$$

其中, α 为学习率.在目标网络滞后若干步后,将从预测网络中复制参数以更新目标网络,使训练过程更加稳定.当上述算法收敛之后,即可获得对应长期性能优化的子信道和功率分配策略.

本小节所提多智能体DQN算法的动作空间较大,导致收敛时间较长.为解决该问题,本文对功率阶数选择变量进行了预处理以减少动作空间和计算复杂度.功率阶数选择变量的预处理包括波束间和波束内两个阶段.在第一阶段,为单个卫星波束设定最大可用功率电平.根据用户分组结果和每个波束的业务需求均值占总需求的比例,为每个波束设置最大可用功率电平;

在第二阶段,为每个卫星波束内的 GU 设定最大可用功率电平. 为此,本文将两个 GUs 分配为一组,并为每组设定最大可用功率电平. 具体的用户配对规则为:根据业务需求对 GU 进行降序排序,并将业务需求高的 GU 与业务需求低的 GU 进行一一配对.

需说明的是,本文考虑到卫星波束照射的时间和区域尺度较大,而各用户的子信道及功率资源分配的时间和区域尺度较小,提出一种两阶段资源管理方案. 在第一阶段中,采用基于 Voronoi 图的迭代用户分组方法可得到波束-用户关联策略. 由于波束覆盖半径与波束发射天线的最大增益成反比,为基于波束最小可能的半径覆盖相关组中的用户,需优化波束设计中心及覆盖半径,以提高发射天线增益,本文提出基于最小圆覆盖的波束覆盖半径优化方案,该阶段未涉及波束边界干扰问题. 在算法第二阶段,基于给定的波束覆盖策略,联合设计用户子信道及功率分配策略. 由于占用相同子信道且距离相近的用户可能存在干扰问题,在本文提出的基于多智能体 DQN 的子信道及功率分配算法中,通过为相邻波束的边界用户分配不同子信道及设计功率控制策略可避免或减小干扰.

6 仿真结果分析

本小节对本文所提算法性能进行仿真评估. 仿真中,设定仿真区域的大小为 $1\ 280\ \text{km} \times 1\ 280\ \text{km}$, GU 随机分布于区域中,仿真中使用的参数如表 1 所示. 本文分别基于 MATLAB 和 Pycharm 软件评估所提出的用户分组和动态资源分配算法. 为实现所提出的基于多智能体 DQN 的算法框架,本文利用 Google TensorFlow-2.10.0 为各智能体建立了两个完全相同的全连接神经网络分别作为预测网络和目标网络. 每个神经网络由一个输入层、一个输出层和一个包含 1 000 个神经元的隐藏层组成. ϵ -贪婪策略中探索参数的初始值设置为 0.5,并逐渐递增至最大值 0.9. 折扣因子 γ 固定为 0.5,学习率 α 设置为 0.008. 经验回放池的容量设置为 10 000,小批次样本大小为 600. 设置激活单元为校正线性单元,并采用 RMSProp 优化器以最小化 MSE. 关于神经网络参数的设定,本文使用基于性能的验证方法微调学习率、探索率、折扣因子和样本大小. 最终的选择是由其他参考文献、经验观察以及在探索和开发、收敛速度和泛化能力之间实现最佳权衡而确定的.

算法仿真评估阶段,我们基于 MATLAB 软件随机产生 50 个非均匀分布的用户. 首先,采用所提出的基于 Voronoi 图的迭代用户分组算法确定波束-用户关联策略 $\alpha_{n,m}^*$;进而,基于波束-用户关联策略,采用最小圆覆盖方法,在满足最小波束半径的约束条件下,确定各波束照射的中心坐标 (x_n^c, y_n^c) 和覆盖半径 δ_n . 由于所设

计的用户资源分配算法需基于所设定的用户坐标位置,以及所确定的波束与用户关联策略、各波束中心位置及覆盖半径等数据,故由 MATLAB 生成多个 Excel 文件并将文件数据导入 Pycharm 软件中,作为多智能体 DQN 算法中的输入环境信息,用以计算信道增益、用户实际传输速率及确定对应各状态的资源分配动作. 通过智能体与环境的交互和智能体神经网络的更新,在多智能体 DQN 算法收敛后,即可得到用户子信道及功率分配策略.

表 1 仿真参数设置

仿真参数	数值
卫星高度(h_s)	35 786 km
载波频率(f)	20 GHz
卫星可用总带宽(B_{tot})	200 MHz
卫星总发射功率(P_{tot})	150 W
波束数量(N)	7
地面用户数量(M)	50
子信道数量(K)	10
地面业务类型数量(L)	4
接收天线增益(g_m^R)	52 dBi
卫星天线效率(η)	0.5
卫星天线直径(D)	2.5 m
噪声功率谱密度(N_0)	-174 dBm/Hz
奖励/惩罚因子(λ_1, λ_2)	1, -1
奖励折扣因子(c)	0.5
调整参数(a, b)	1, 2
扩散/粘滞系数(μ)	0.1
离散功率阶数(J)	150

图 5 显示了三种不同用户分组方法的结果,即固定波束宽度、基于 K -均值和基于 Voronoi 图的迭代用户分组. 从图中可以看出,不规则的波束布局和可变的波束宽度可以提供灵活的波束-用户映射. 图 6 显示了上述不同用户分组方案下各组的负载. 以各用户组通信需求的最大值和最小值的差作为负载均衡性能度量,以衡量各分组方案的负载差异. 由图中结果计算可知,固定波束宽度、基于 K -均值和基于 Voronoi 图的迭代用户分组算法对应的用户组负载差异分别为:893.3 Mbps、503.4 Mbps、272.2 Mbps,即与基于 K -均值的用户分组方案相比,基于 Voronoi 图的迭代用户分组负载差异值可降低 45.9%. 这证明了本文所提出的用户分组算法在实现用户组间负载均衡方面的有效性.

图 7 绘制了不同算法下系统效用函数与算法步数的关系图,包括本文提出的基于多智能体 DQN 的算法和文献[9]中提出的算法. 从图中可以看出,两种算法在若干训练步数之后达到收敛,证明了算法的有效性. 同时,本文所提算法在训练 3 500 步左右即达到收敛,

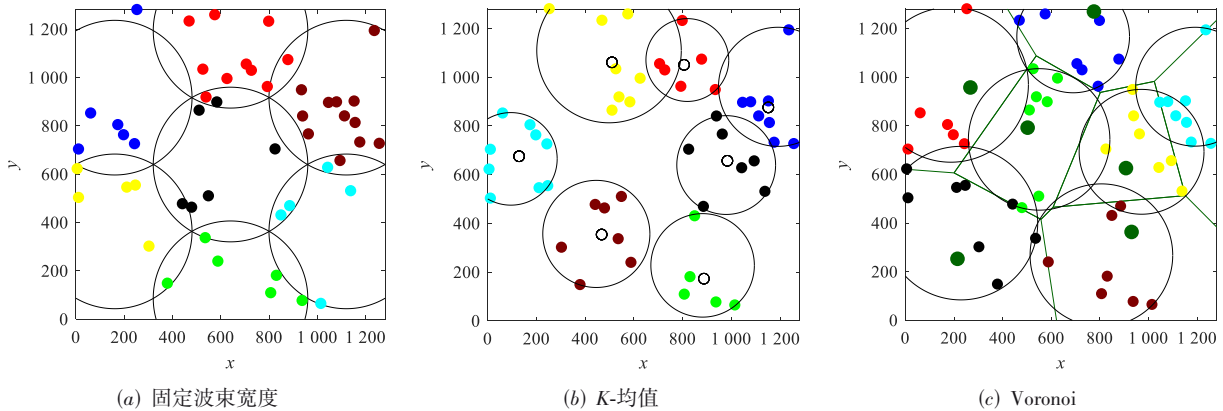


图5 不同用户分组方案结果示意图

而文献[9]中所提算法在训练5 000步之后趋于收敛,表明本文算法具有更快的收敛速度.作为对比,本文绘制了对应于不同学习率的效用函数,从图中可以看出,学习率的降低导致算法收敛速度变慢.将本文提出的算法与文献[9]中提出的算法进行比较,可以看出本文提出的算法收敛后具有更高的系统效用.这是因为文献[9]中的算法主要考虑了子信道分配,仅采用相同发送功率及固定覆盖半径的波束照射策略.

看出,文献[9]中算法的资源分配结果存在提供容量过高导致资源浪费和提供容量过低导致无法满足服务质量的情况.为评估系统提供容量与用户需求之间的匹配程度,将系统提供容量与用户需求之间差值的绝对值作为性能度量.通过计算可知,采用文献[9]中算法,系统所提供的容量与用户需求差值的绝对值之和为427.4 Mbps,而采用本文所提算法得到的此数值为70.8 Mbps.与文献[9]所提算法相比,本文所提出的算法对应的系统提供容量与用户需求差值可降低83.43%.因此,本文所提出的算法能够更好地实现系统提供容量与用户业务需求之间的匹配.

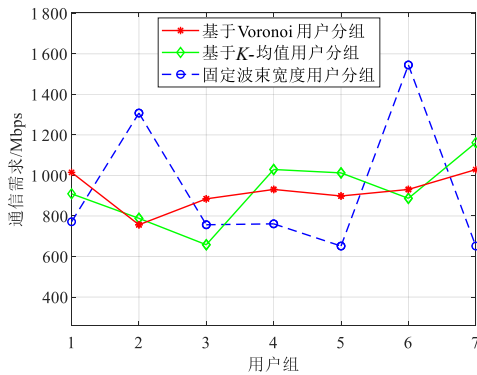


图6 各用户组的流量需求示意图

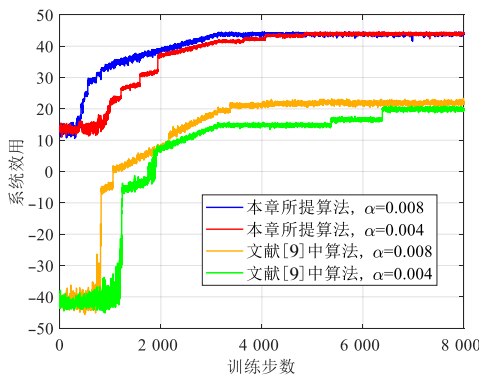


图7 系统效用与训练步数关系图

图8分别绘制了本文和文献[9]中提出的算法得到的各波束的平均通信需求和提供的容量.从图中可以

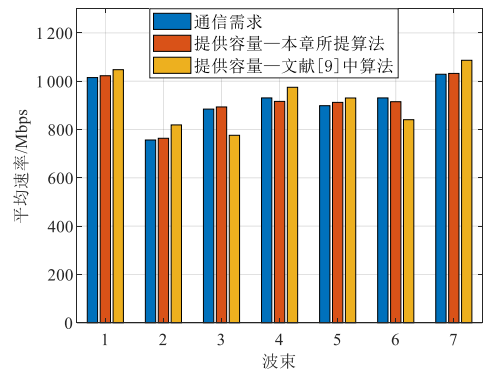


图8 各波束通信需求与提供容量对

图9描述了系统效用与卫星发射功率的关系.从图中可以看出,随着卫星发射功率的增加,系统效用先呈增长趋势,当卫星发射功率达到一定值后,系统效用减小.原因是当功率较小时,增加功率会提供更高的数据传输速率以减少提供容量与通信需求之间的差异.然而,当卫星发射功率达到一定值时,继续增加功率会导致过高的系统容量,造成提供容量与通信需求之间的差值增大.此外,由图可知,低噪声功率密度下系统具有更好的性能.

图10绘制了系统总带宽为195 MHz情况下不同资

源方案对应的系统效用与卫星发射功率之间的关系图。由图可见,随着卫星发射功率的增加,各类资源管理方案对应的系统效用随之增加。这是因为卫星发射功率的增加导致系统数据传输速率增加,从而可提升系统效用。对比不同用户分组方法的系统效用可知,采用基于 Voronoi 图的迭代用户分组方法相比基于 K -均值的用户分组方法在不同卫星发射功率下均可以达到更高的系统效用。这是因为基于 K -均值的用户分组方法仅考虑用户位置,而本文所提出的基于 Voronoi 图的迭代用户分组方法兼顾用户间的距离与多样化的速率需求,可较好地实现波束间负载均衡及波束的灵活部署,从而可有效提升系统性能。对比不同子信道及功率分配算法可知,本文所提出的基于多智能体 DQN 的子信道及功率分配算法相比文献[9]所提出的方法可实现更高的系统效用。这是因为文献[9]所提出的方法仅考虑子信道及时域的资源分配策略,对卫星功率采用均分的方式,从而限制了资源分配的灵活性及系统效用的提升。

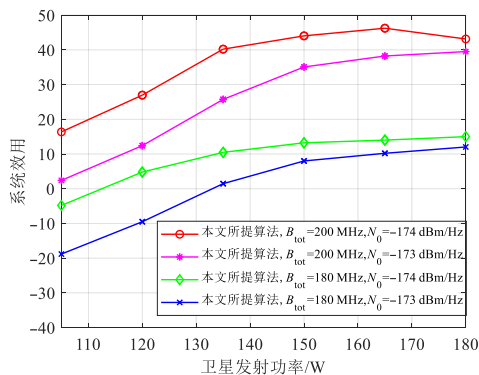
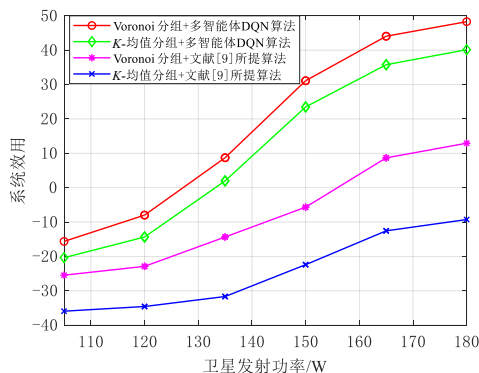


图9 系统效用与卫星发射功率关系图

图10 系统效用与卫星发射功率关系图($B_{\text{tot}}=195$ MHz)

7 结论

本文研究了多波束卫星通信系统中的用户分组和资源分配问题,提出了两阶段资源管理算法,包括基于 Voronoi 图的迭代用户分组算法和基于多智能体 DQN

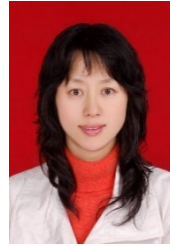
的子信道和功率分配算法。数值结果表明,用户分组算法能够有效地实现分组间的负载均衡;基于多智能体 DQN 的资源分配算法与现有算法相比,具有更快的收敛速度,且可实现系统所提供的容量与用户通信需求之间更好的匹配。

参考文献

- [1] KODHELI O, LAGUNAS E, MATURO N, et al. Satellite communications in the new space era: A survey and future challenges[J]. IEEE Communications Surveys & Tutorials, 2021, 23(1): 70-109.
- [2] ZHOU D, SHENG M, LI J D, et al. Aerospace integrated networks innovation for empowering 6G: A survey and future challenges[J]. IEEE Communications Surveys & Tutorials, 2023, 25(2): 975-1019.
- [3] KISSELEFF S, LAGUNAS E, ABDU T S, et al. Radio resource management techniques for multibeam satellite systems[J]. IEEE Communications Letters, 2021, 25(8): 2448-2452.
- [4] LI C, ZHU H P, TANG J Y, et al. User grouping in multiuser satellite MIMO downlink with fairness consideration[J]. IEEE Wireless Communications Letters, 2022, 11(8): 1575-1579.
- [5] DAI C Q, LI S P, WU J S, et al. Distributed user association with grouping in satellite—Terrestrial integrated networks[J]. IEEE Internet of Things Journal, 2022, 9(12): 10244-10256.
- [6] LIU B K, JIANG C X, KUANG L L, et al. Joint user grouping and beamwidth optimization for satellite multicast with phased array antennas[C]//GLOBECOM 2020 - 2020 IEEE Global Communications Conference. Piscataway: IEEE, 2020: 1-6.
- [7] 李信, 李勇军, 赵尚弘. 基于能量效率的星地 NOMA 网络功率分配算法[J]. 电子学报, 2023, 51(5): 1310-1318.
LI X, LI Y J, ZHAO S H. Power allocation in satellite-terrestrial noma network based on energy efficiency[J]. Acta Electronica Sinica, 2023, 51(5): 1310-1318. (in Chinese)
- [8] KAWAMOTO Y, KAMEI T, TAKAHASHI M, et al. Flexible resource allocation with inter-beam interference in satellite communication systems with a digital channelizer[J]. IEEE Transactions on Wireless Communications, 2020, 19(5): 2934-2945.
- [9] HU X, LIAO X L, LIU Z J, et al. Multi-agent deep reinforcement learning-based flexible satellite payload for mobile terminals[J]. IEEE Transactions on Vehicular Technology, 2020, 69(9): 9849-9865.

- [10] 柴蓉, 陈米铃, 李锦红. 基于效用优化的星地融合网络联合用户关联及资源块调度算法[J]. 电子学报, 2023, 51(12): 3483-3495.
CHAI R, CHEN M L, LI J H. Utility optimization-based joint user association and resource allocation algorithm for integrated satellite-terrestrial network[J]. Acta Electronica Sinica, 2023, 51(12): 3483-3495. (in Chinese)
- [11] LIN Z Y, NI Z Y, KUANG L L, et al. Dynamic beam pattern and bandwidth allocation based on multi-agent deep reinforcement learning for beam hopping satellite systems [J]. IEEE Transactions on Vehicular Technology, 2022, 71(4): 3917-3930.
- [12] XIAO A L, CHEN Z M, WU S, et al. Collaborative long-short term bandwidth allocation for satellite-terrestrial networks[J]. IEEE Communications Letters, 2022, 26(5): 1121-1125.
- [13] WANG A Y, LEI L, LAGUNAS E, et al. Joint optimization of beam-hopping design and NOMA-assisted transmission for flexible satellite systems[J]. IEEE Transactions on Wireless Communications, 2022, 21(10): 8846-8858.
- [14] TAKAHASHI M, KAWAMOTO Y, KATO N, et al. DBF-based fusion control of transmit power and beam directivity for flexible resource allocation in HTS communication system toward B5G[J]. IEEE Transactions on Wireless Communications, 2022, 21(1): 95-105.
- [15] LI X N, ZHANG H J, ZHOU H, et al. Multi-agent DRL for resource allocation and cache design in terrestrial-satellite networks[J]. IEEE Transactions on Wireless Communications, 2023, 22(8): 5031-5042.
- [16] CHO Y, YANG W, OH D, et al. Multi-agent deep reinforcement learning for interference-aware channel allocation in non-terrestrial networks[J]. IEEE Communications Letters, 2023, 27(3): 936-940.
- [17] DENG D H, WANG C W, PANG M L, et al. Dynamic resource allocation with deep reinforcement learning in multibeam satellite communication[J]. IEEE Wireless Communications Letters, 2023, 12(1): 75-79.

作者简介



柴蓉 女, 1974年出生. 现为重庆邮电大学教授、博士生导师, 研究方向为空天体一体化网络架构及关键技术、无线资源管理及移动性管理技术等.

E-mail: chairong@cqupt.edu.cn



刘磊 男, 1998年4月出生. 现为重庆邮电大学硕士研究生, 研究方向为卫星通信智能资源管理.

E-mail: s210131134@stu.cqupt.edu.cn



梁承超 男, 1988年出生. 现为重庆邮电大学教授、博士生导师, 研究方向为移动通信、无线网络、卫星互联网及优化理论等.

E-mail: liangcc@cqupt.edu.cn



陈前斌 男, 1967年出生. 现为重庆邮电大学教授, 研究方向为无线通信、通信网络理论多媒体技术等. 中国电子学会会员编号: E190000713S.

E-mail: chenqb@cqupt.edu.cn