

基于非稳态 MAB 的低地球轨道卫星系统异构用户接入方法

林敏¹, 孙昊¹, 赵柏¹, 郭焱¹, 杨绿溪²

(1. 南京邮电大学通信与信息工程学院, 江苏 南京 210003; 2. 东南大学信息科学与工程学院, 江苏 南京 211189)

摘要: 针对吞吐量敏感用户和时延敏感用户同时接入低地球轨道卫星系统的场景, 提出了一种基于非稳态多臂赌博机 (MAB) 的非正交多址接入 (NOMA) 接入方法, 以提高系统的频谱效率和异构用户的接入灵活性。首先, 在构建联合效用函数以同时考虑时延敏感用户信息年龄 (AoI) 和吞吐量敏感用户传输速率需求的基础上, 建立以联合效用函数最大化为目标, 以速率门限和数据包缓存时间为约束条件的优化问题。由于传统的凸优化方法无法进行求解, 在采用波束成形抑制组间干扰的基础上, 提出了基于非稳态 MAB 的动态接入方案, 实现异构用户的智能配对与调度, 以及可靠接入。最后, 计算机仿真验证了所提方法的可靠性和优越性, 并进一步分析了折扣因子和数据包缓存门限值对系统性能的影响。

关键词: 低地球轨道卫星系统; 非正交多址接入; 信息年龄; 非稳态多臂赌博机

中图分类号: TN92

文献标志码: A

DOI: 10.11959/j.issn.1000-436x.2025051

Non-stationary MAB based access method for heterogeneous users in LEO satellite systems

LIN Min¹, SUN Hao¹, ZHAO Bai¹, GUO Yan¹, YANG Lyuxi²

1. School of Communications and Information Engineering, Nanjing University of Posts and Telecommunications, Nanjing 210003, China

2. School of Information Science and Engineering, Southeast University, Nanjing 211189, China

Abstract: Provided that both throughput-sensitive and delay-sensitive users access the low earth orbit (LEO) satellite system simultaneously, a non-orthogonal multiple access (NOMA) access method using non-stationary multi-armed bandit (MAB) was proposed to enhance the system spectral efficiency and access flexibility for the heterogeneous users. Firstly, a utility function was established to jointly consider the age of information (AoI) of delay-sensitive users and the throughput requirements of throughput-sensitive users. Then, an optimization problem was formulated to maximize the joint utility function, while satisfying the constraints of rate thresholds and packet buffering time. Since the traditional convex optimization methods couldn't solve this problem, by using the beamforming to suppress the inter-group interference, a dynamic access scheme based on the non-stationary MAB was proposed to realize the intelligent pairing, scheduling, and reliable access for the heterogeneous users. Finally, computer simulations demonstrate the reliability and superiority of the proposed method, and reveal the impact of discount factor and packet buffer threshold on the system performance.

Keywords: LEO satellite system, non-orthogonal multiple access, age of information, non-stationary MAB

0 引言

众所周知, 卫星通信因其覆盖范围广、传输速率高、不受地理条件限制等优点, 被认为是 6G 移

动通信系统的重要组成部分。根据轨道高度的不同, 人造卫星可以分为地球静止轨道 (GEO, geostationary earth orbit) 卫星、中地球轨道 (MEO,

收稿日期: 2024-10-23; 修回日期: 2025-03-07

基金项目: 国家自然科学基金资助项目 (No.62471255)

Foundation Item: The National Natural Science Foundation of China (No.62471255)

medium earth orbit) 卫星和低地球轨道 (LEO, low earth orbit) 卫星。虽然 GEO 卫星与 MEO 卫星在覆盖范围上具有较大优势, 但是其传输时延和传输损耗均相对较高。相比之下, LEO 卫星具有较低的传输时延和较小的传输损耗, 并且轨道资源更加丰富, 有利于组建下一代卫星互联网, 推动 6G 网络的建设, 因而近些年得到了业内的广泛研究和关注^[1-3]。

传统的卫星通信系统通常采用正交多址接入 (OMA, orthogonal multiple access) 技术, 能够在频域、时域或者码域满足正交的条件下为波束覆盖范围内的用户提供通信服务。然而随着接入用户数量的持续增加, 基于 OMA 技术的卫星系统容量已经无法满足未来 6G 网络的需求。鉴于此, 非正交多址接入 (NOMA, non-orthogonal multiple access) 技术应运而生, 该技术允许多个用户信号在同一个时频资源块上叠加传输, 能够有效提高无线通信网络的频谱利用效率并降低接入时延, 受到了学术界的广泛关注和研究^[4-7]。在 NOMA 传输系统中, 发送端对信号进行叠加编码, 而接收端则使用连续干扰消除 (SIC, successive interference cancellation) 技术进行多用户解码。为了保证准确的 SIC, 组内用户之间需要保持一定的信道增益差, 因此设计合理的用户调度方案以完成 NOMA 用户分组配对已经成为 NOMA 传输系统中重要的研究课题之一^[8-10]。一方面, 在吞吐量敏感的无线传输场景下, 相关文献以用户传输速率相关的服务质量 (QoS, quality of service) 为准则进行用户调度方案设计。文献[8]研究了下行 NOMA 系统中基于压缩感知的用户配对与功率分配问题, 建立以系统总速率最大化为优化目标的优化问题, 得到低复杂度的用户配对和功率分配结果, 然而该方法需要基于完美信道状态信息 (CSI, channel state information) 实现静态用户调度, 并不能很好地适应仅已知用户非完美 CSI 的场景。文献[9]则在非完美 CSI 情况下, 基于机器学习方法完成用户调度并实现传输速率最大化, 但是该算法未考虑其他类型用户业务需求。文献[10]针对蜂窝网络上行 NOMA 传输场景, 基于多臂赌博机 (MAB, multi-armed bandit) 模型提出相应的用户调度算法以满足速率和时延等多种业务需求, 但该算法仅可在用户业务需求稳态情况下完成用户调度, 并不能很好地适用于用户业务需求动态变化的场景。

另一方面, 为了满足下一代移动通信系统的实时性需求, 在时延敏感无线传输场景下, 如何进行优化调度方案设计也是移动通信系统设计中的另一个技术难题。为有效刻画信息的时效性, 文献[11]提出一种新的通信网络性能度量指标——信息年龄 (AoI, age of information), 其具体定义为数据包从产生到被成功接收所经历的总时间, 因此较低的 AoI 往往意味着较高的信息新鲜度。基于这一性能指标, 研究人员提出了多种不同算法以实现系统 AoI 最小化, 从而提高时延敏感型用户的服务质量。文献[12]研究了多信道场景中系统时效性最优问题, 分别提出了最小化峰值 AoI 和最小化平均 AoI 的调度方案, 实现了系统平均 AoI 的最小化。文献[13]研究了无线传感器网络中信息年龄最小化问题, 基于集中式训练分布式执行的多智能体近端策略优化算法, 有效降低了系统信息年龄。文献[14]研究了基于 AoI 最小化的无人机数据采集轨迹规划问题, 并提出了基于深度强化学习方法的轨迹优化算法, 显著降低了无人机采集数据的 AoI。虽然上述文献分别对吞吐量敏感型^[8-10]和时延敏感型^[12-14]业务传输场景展开深入研究, 但是下一代无线网络中需要同时满足不同业务类型的用户需求, 因此亟须从多个维度来综合提升无线通信系统的性能。

在这种情况下, 本文面向 LEO 卫星系统异构用户接入场景, 提出了一种基于非稳态 MAB 的 NOMA 高谱效灵活接入策略。首先利用动态 AoI 演进模型来刻画时延敏感用户的信息新鲜度, 采用信息传输速率表征吞吐量敏感用户的 QoS 需求, 构建异构用户共存下的联合效用函数。接下来, 以此为基础, 建立以联合效用函数最大化为目标、以速率门限和数据包缓存时间为约束条件的优化问题。由于时延敏感用户 AoI 及信道状态的随机性, 该优化问题无法采用传统的优化方法直接进行求解。因此, 本文通过部分迫零 (ZF, zero forcing) 波束成形方案消除吞吐量敏感用户间干扰, 并进一步提出基于非稳态 MAB 的时延敏感用户和吞吐量敏感用户的智能配对与动态调度方案, 保证异构用户的可靠接入, 且实现传输速率和 AoI 的联合优化。最后, 计算机仿真结果表明, 与现有的方案相比, 本文提出的方法可以提高系统的性能和异构用户的接入灵活性。

1 系统模型

多波束卫星多址接入系统中的上行 NOMA 接入模型如图 1 所示, 包含一颗 LEO 卫星, 以及覆盖范围内的 K 个吞吐量敏感用户和 N 个时延敏感用户。以吞吐量敏感用户为中心, 将地面用户分成 K 个组, 使每组内始终有一个能够与 LEO 卫星保持通信的吞吐量敏感用户, 以及多个需要进行调度和激活的时延敏感用户。记第 k 组中的时延敏感用户为 i_k , $i_k \in \{1, 2, \dots, S_k\}$, $k \in \{1, 2, \dots, K\}$, 其中 S_k 为第 k 组的时延敏感用户总数, 满足 $S_1 + S_2 + \dots + S_K = N$ 。在每个信号传输时隙, 卫星系统采用快速上行授权的接入方式, 将上行授权直接分配给激活时延敏感用户, 且激活时延敏感用户与组内吞吐量敏感用户进行智能配对, 并采用上行 NOMA 技术共享同一时频资源块接入卫星网络。此外, LEO 卫星平台配置 $L = L_x \times L_y$ 个阵元的均匀平面阵 (UPA, uniform planar array), 它能够以比较紧凑的结构获得较高的增益, 而地面用户则配备定向天线, 且吞吐量敏感用户的天线增益通常高于时延敏感用户的天线增益, 从而保证它们能实现 NOMA 传输。接下来, 介绍用户激活模型、信道与信号模型。

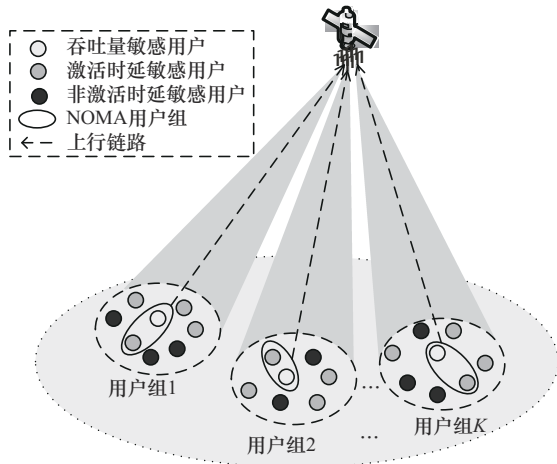


图 1 多波束卫星多址接入系统中的上行 NOMA 接入模型

1.1 基于流量预测的用户激活模型

在该多波束卫星多址接入系统中, 假设时延敏感用户首次接入卫星网络的激活概率在一定时间 T_A 内服从 Beta 分布, 其概率密度函数为^[15]

$$f(x) = \frac{x^{\alpha-1}(T_A - x)^{\beta-1}}{T_A^{\alpha+\beta-1}B(\alpha, \beta)} \quad (1)$$

其中, $B(\alpha, \beta)$ 是参数为 α, β 的 Beta 函数。将 T_A 分为 N_A 个时隙, 用户将集中在某个时段激活, 达到模拟大量用户短时间内集中接入的场景。在用户完成首次接入后, 将在后续 N_A 个时隙内重新激活, 其中激活的含义是时延敏感用户已准备好待发送的数据包。

卫星端使用基于隐马尔可夫模型的前向预测算法^[16], 以预测每个时隙的激活时延敏感用户的激活状态。将激活的用户列表表示为 $S_{a,k}$, 同时给出每个时延敏感用户的激活概率 P_{i_k} 。由于流量预测并不能完全正确地给出各时隙真实的时延敏感用户激活状态, 需要考虑算法预测误差对用户调度的影响。当发生预测错误时, 会错判用户的真实激活状态, 因此, 将预测误差 $e_p \in (0, 1)$ 定义为预测错误用户数与所有用户数的比值, 满足 $e_p \sim \mathcal{N}(\bar{e}_p, \sigma_p^2)$ 。本文后续将根据各时隙的预测激活列表对时延敏感用户进行调度, 并考虑算法预测误差对调度结果的影响。

1.2 信道与信号模型

本文考虑地面用户上行接入 LEO 卫星的通信场景, 综合考虑路径损耗、天线增益和小尺度衰落的影响, 卫星上行链路的信道矢量可以建模为

$$\mathbf{h} = \sqrt{G_r G_t} \ell \mathbf{g} \cdot \mathbf{a}(\theta_s, \varphi_s) \quad (2)$$

其中, $\ell = \left(\frac{c}{4\pi f d}\right)^2$ 表示卫星与用户之间的自由空间损耗路损, c , f 和 d 分别表示光速、载波频率和用户到卫星的距离, g 为信道小尺度衰落, G_r 表示卫星端接收天线增益, G_t 表示用户发射天线增益, $\mathbf{a}(\theta_s, \varphi_s) \in \mathbb{C}^{L_x L_y \times 1}$ 表示卫星配备 UPA 的阵列导向矢量, 即

$$\begin{aligned} \mathbf{a}(\theta_s, \varphi_s) &= \mathbf{a}_x(\theta_s, \varphi_s) \cdot \mathbf{a}_y^T(\theta_s, \varphi_s) = \\ & \left[1, e^{j\frac{2\pi}{\lambda} l_{2,1} \cdot r}, \dots, e^{j\frac{2\pi}{\lambda} l_{L_x,1} \cdot r} \right] \cdot \left[1, e^{j\frac{2\pi}{\lambda} l_{1,1} \cdot r}, \dots, e^{j\frac{2\pi}{\lambda} l_{1,L_y} \cdot r} \right]^T \end{aligned} \quad (3)$$

其中, θ_s 和 φ_s 分别表示卫星用户相对于卫星天线阵列的仰角和方位角, $\mathbf{l}_{l_x, l_y} = (d_x(l_x - 1), d_y(l_y - 1), 0)$, d_x 和 d_y 分别表示 UPA 水平方向和垂直方向的阵元间隔, $\mathbf{r} = (\cos \theta_s \cos \varphi_s, \cos \theta_s \sin \varphi_s, 0)$ 表示 X 轴和 Y 轴方向天线阵元间的距离, λ 为载波波长。

假设在 t 时隙, 调度第 k 个组的时延敏感用户 i_k 与该组中的吞吐量敏感用户进行配对, 并采用 NOMA 技术共享同一时频资源实现上行接入, 那么卫星接收到来自第 k 组的信号可以表示为

$$y_k(t) = \sqrt{P_k} \mathbf{w}_k^H \mathbf{h}_k x_k(t) + \sqrt{P_{i_k,k}} \mathbf{w}_k^H \mathbf{h}_{i_k,k} x_{i_k,k}(t) + I_k(t) + \mathbf{w}_k^H \mathbf{n}_k(t), k \in \{1, \dots, K\}, i_k \in \{1, 2, \dots, S_{a,k}\} \quad (4)$$

其中, $x_k(t)$ 和 $x_{i_k,k}(t)$ 分别表示第 k 组中吞吐量敏感用户和时延敏感用户的发送信号, 且满足 $E[|x_k(t)|^2] = E[|x_{i_k,k}(t)|^2] = 1$; \mathbf{n}_k 是均值为 0、方差为 σ_k^2 的加性白高斯噪声 (AWGN, additive white gaussian noise); $\mathbf{w}_k \in \mathbb{C}^{L_x L_y \times 1}$ 为对应第 k 组用户的卫星接收波束成形权矢量; $\mathbf{h}_k \in \mathbb{C}^{L_x L_y \times 1}$ 和 $\mathbf{h}_{i_k,k} \in \mathbb{C}^{L_x L_y \times 1}$ 分别表示第 k 组中吞吐量敏感用户和时延敏感用户到卫星的信道矢量; P_k 和 $P_{i_k,k}$ 分别为第 k 个 NOMA 组中吞吐量敏感用户和时延敏感用户的发射功率; $I_k(t) = \sum_{j=1, j \neq k}^K \mathbf{w}_k^H \mathbf{h}_j x_j(t) + \mathbf{w}_k^H \mathbf{h}_{i_j} x_{i_j}(t)$ 表示来自其他组用户的同频干扰。

为了更高效地解码多个用户的信号, 卫星采用了 SIC 技术, 其基本原理以及解码流程如下。接收端首先解码信道增益较高的用户, 并在解码过程中直接将信道增益较低的用户视作干扰。接着, 将已解码的强用户信号从接收到的混合信号中去除, 这样就能有效消除强用户对其他弱用户的干扰。考虑到吞吐量敏感用户通常配备较高增益的天线, 而时延敏感用户因为尺寸限制, 往往仅配备增益较低的微型天线, 因此可假设吞吐量敏感用户为强用户, 并将时延敏感用户视作干扰, 卫星首先对吞吐量敏感用户进行解码。在吞吐量敏感用户的信号被成功解码后, 卫星采用 SIC 技术将解码成功的吞吐量敏感用户信号从接收的总信号中消除。随后, 卫星继续解码时延敏感用户的信号, 此时, 时延敏感用户的信号不再受到吞吐量敏感用户的干扰。由此, 可以得到吞吐量敏感用户和时延敏感用户的信干噪比 (SINR, signal-to-interference-noise ratio) 分别为

$$\gamma_k = \frac{P_k |\mathbf{h}_k \mathbf{w}_k|^2}{P_{i_k} |\mathbf{h}_{i_k}^H \mathbf{w}_k|^2 + \sum_{j=1, j \neq k}^K P_j |\mathbf{h}_j^H \mathbf{w}_k|^2 + P_{i_j} |\mathbf{h}_{i_j}^H \mathbf{w}_k|^2 + \sigma_k^2} \quad k = 1, 2, \dots, K, \forall i_k \in S_{a,k}$$

$$\gamma_{i_k,k} = \frac{P_{i_k,k} |\mathbf{h}_{i_k,k} \mathbf{w}_k|^2}{\sum_{j=1, j \neq k}^K P_j |\mathbf{h}_j^H \mathbf{w}_k|^2 + P_{i_j} |\mathbf{h}_{i_j}^H \mathbf{w}_k|^2 + \sigma_k^2}$$

$$k = 1, 2, \dots, K, \forall i_k \in S_{a,k} \quad (5)$$

根据式(5), 可得两类用户的可达速率分别为

$$R_k = \text{lb}(1 + \gamma_k), k = 1, 2, \dots, K, \forall i_k \in S_{a,k}$$

$$R_{i_k,k} = \text{lb}(1 + \gamma_{i_k,k}), k = 1, 2, \dots, K, \forall i_k \in S_{a,k} \quad (6)$$

值得注意的是, 不同用户组中吞吐量敏感用户和不同的时延敏感用户的配对, 以及不同组间的用户干扰都会显著影响所考虑系统的性能。为了实现性能的优化设计, 本文基于部分用户 ZF 波束成形, 提出了一种非稳态 MAB 用户接入方案。

2 基于非稳态 MAB 的异构用户接入策略

在无线通信场景中, 通常存在多种准则来衡量无线通信系统的性能, 其中系统传输速率和 AoI 分别从无线通信有效性和时效性 2 个方面描述系统的性能。因此, 本节综合考虑系统传输速率和 AoI 这 2 个性能准则建立联合优化问题, 保证吞吐量敏感用户和时延敏感用户同时接入卫星系统。根据时延敏感用户的激活模型及信道与信号模型的分析可知, 设计合理的用户接入策略是实现系统传输速率与 AoI 联合优化的关键。基于此, 本节提出了一种基于非稳态 MAB 的动态调度策略, 该策略能够灵活地调度时延敏感用户的配对接入, 从而在满足异构用户需求的同时, 优化系统整体性能。

2.1 优化问题的建立

在本文所研究的上行 NOMA 接入场景中, 由于吞吐量敏感用户先于时延敏感用户进行解码, 因此一旦吞吐量敏感用户解码失败, 时延敏感用户也会因 SIC 错误传播而无法成功解码, 从而导致系统传输性能的损失。因此, 为了保证吞吐量敏感用户用户的 QoS 需求, 其传输速率需要满足如下约束条件

$$R_k \geq R_{th} \quad (7)$$

其中, R_{th} 为预先定义的解码阈值。

另一方面, 考虑无线通信系统的信息时效性, 通常用 AoI 来描述状态更新中信息的时效性和新鲜度。用户的 AoI 定义为用户数据包从产生到被成功接收所经历的总时间。因此, AoI 越小, 信息新鲜度越高, 可以更好地保证信息的时效性。本文考虑

将时间划分为时隙，激活用户在每个时隙的开始被调度并发送数据包。若数据包成功传输，用户 AoI 重新归一，否则将随时隙递增。将用户激活状态记作 $A(t)$ ， $A(t) = 1$ 时表示 t 时刻该用户激活， $A(t) = 0$ 时表示用户未激活。不失一般性，本文假设时延敏感用户被调度时，只要满足与之配对的吞吐量敏感用户的传输速率大于阈值 R_{th} 的约束条件，就视为传输成功。将时延敏感用户的传输状态记为 $s_{k,i_k}(t)$ ， $k \in \{1, 2, \dots, K\}$ ， $i_k \in \{1, 2, \dots, S_k\}$ 。 $s_{k,i_k}(t) = 1$ 时表示 t 时隙成功传输， $s_{k,i_k}(t) = 0$ 时表示传输失败或未传输。同时，与文献[17]类似，本文考虑用户数据缓存时间存在上界 D_{th} ，如果超过该上界则丢弃该数据包，用户进行重新等待激活。将用户数据包缓存时间状态记为 $a(t)$ 。 $a(t) = 1$ 时表示 t 时隙数据包缓存时间未超过该上界， $a(t) = 0$ 时表示 t 时隙数据包缓存时间超过该上界。因此，激活状态时延敏感用户的 AoI 表达式为

$$\Delta_{k,i_k}(t+1) = \begin{cases} \min \{ D_{k,i_k}(t+1), D_{th} \}, & s_{k,i_k}(t) = 1 \\ \min \{ \Delta_{k,i_k}(t) + 1, D_{th} \}, & s_{k,i_k}(t) = 0 \end{cases} \quad (8)$$

其中， $D_{k,i}(t+1)$ 表示用户的本地信息年龄，定义为

$$D_{k,i_k}(t+1) = \begin{cases} \min \{ D_{k,i_k}(t) + 1, D_{th} \}, & a_{k,i_k}(t) = 1 \\ 1, & A_{k,i_k}(t+1) = 1 \end{cases} \quad (9)$$

本地信息年龄表示当时延敏感用户激活后且用户数据包缓存时间未达到上界时，在未被调度前所经历的时隙数。

为了保证吞吐量敏感用户 QoS 的前提下提高时延敏感用户的信息新鲜度，需要使每个时隙配对调度后的时延敏感用户 AoI 衰减量最大以及所调度的吞吐量敏感用户的传输速率最大。因此可以将两者结合得到效用函数表达式

$$U_k(t) = \delta_1 r_k^n(t) + \delta_2 \Delta_k^n(t), \quad k \in \{1, 2, \dots, K\} \quad (10)$$

其中， δ_1, δ_2 为权重因子，满足 $\delta_1 + \delta_2 = 1$ ； $r_k^n(t) = \frac{R_k(t)}{R_{max}}$ 为吞吐量敏感用户传输速率和的归一化表达式， R_{max} 为吞吐量敏感用户传输速率的最大值； $\Delta_k^n(t) = \frac{\Delta_k(t)}{D_{th}}$ 为时延敏感用户 AoI 归一化表达式，

$\Delta_k(t)$ 即 t 时隙与第 k 个吞吐量敏感用户配对的时延敏感用户 AoI。根据该效用函数，可以建立如下优化问题

$$\begin{aligned} \max_{\mathbf{w}_{k,i_k} \in \{1, 2, \dots, S_{a,k}\}} & \sum_{k=1}^K U_k(t) \\ \text{s.t.} & R_k(t) \geq R_{th} \\ & \Delta_{k,i_k}(t) \leq D_{th} \\ & \|\mathbf{w}_k\|^2 = 1 \end{aligned} \quad (11)$$

其中， R_{th} 和 D_{th} 分别是吞吐量敏感用户的最小速率和时延敏感用户的最大 AoI 限制，相应的 $r_k(t)$ 和 $\Delta_k(t)$ 为 t 时刻吞吐量敏感用户的传输速率和与之配对的时延敏感用户的 AoI。

由于优化问题式(11)中用户 AoI 和信道状态都是动态变化的，传统的凸优化方法难以求解。因此，在假设 LEO 卫星已经部署基于隐马尔可夫模型的前向预测算法的前提下，本文提出一种基于非稳态 MAB 模型的在线学习算法来解决上述优化问题。利用该求解方法，LEO 卫星能够在时延敏感用户 CSI 获取受限的条件下完成有效的动态调度，同时降低信令消耗和传输时延，大大提高通信效率。

2.2 基于非稳态 MAB 的用户接入策略

为了消除各组间用户调度的相互影响，在仅已知吞吐量敏感用户的角度信息的前提下，采用部分 ZF 波束成形算法来消除吞吐量敏感用户间的干扰。由于时延敏感用户分布在吞吐量敏感用户周围，对吞吐量敏感用户 ZF 波束成形后，较小的组间干扰可以忽略，只需考虑组内用户的干扰。同时，为了使 LEO 卫星能够有效地完成上行传输用户的调度，本文提出基于非稳态 MAB 模型的求解算法，使 LEO 卫星在有限信息条件下实现对效用函数的学习，并根据被调度用户的反馈进行动态调整。

首先，LEO 卫星采用 ZF 波束成形技术对吞吐量敏感用户之间进行干扰消除，该算法的基本思想是使目标用户与其他用户之间的干扰为零。ZF 波束成形权矢量 $\mathbf{w}_{k,ZF}$ 设计为^[18]

$$\mathbf{w}_{k,ZF} = \frac{(\mathbf{I}_L - \mathbf{F}_e) \mathbf{a}(\theta_k, \varphi_k)}{\|(\mathbf{I}_L - \mathbf{F}_e) \mathbf{a}(\theta_k, \varphi_k)\|} \quad (12)$$

其中， $k \in \{1, 2, \dots, K\}$ ， \mathbf{I}_L 为 $L_x \times L_y$ 维单位矩阵， $\mathbf{F}_e = (\tilde{\mathbf{H}}_k)^H (\tilde{\mathbf{H}}_k (\tilde{\mathbf{H}}_k)^H)^{-1} \tilde{\mathbf{H}}_k$ ， $\tilde{\mathbf{H}}_k$ 为

$$\begin{aligned} \tilde{\mathbf{H}}_k &= [\mathbf{a}_1(\theta_1, \varphi_1), \dots, \mathbf{a}_{k-1}(\theta_{k-1}, \varphi_{k-1}), \\ &\mathbf{a}_{k+1}(\theta_{k+1}, \varphi_{k+1}), \dots, \mathbf{a}_K(\theta_K, \varphi_K)]^T \end{aligned} \quad (13)$$

进一步, 在使用ZF波束成形算法消除组间干扰后, 考虑到系统AoI的动态性及信道状态的时变性, 本文提出一种基于非稳态MAB模型的在线学习算法。在非稳态MAB模型^[19]的动作空间中包含多个摇臂, 决策者在每次决策时只能在可用摇臂中选取一个摇臂, 获得对应的奖励, 并根据奖励结果来优化后续的摇臂选取。因此, 在求解该非稳态MAB问题前, 应先确定该模型对应的动作空间和奖励设置。

1) 动作空间

在优化问题式(11)中, 将卫星作为决策者, 每个时隙中, 在各组中选择一个预测激活时延敏感用户与吞吐量敏感用户进行NOMA配对, 完成两用户NOMA的上行传输。对于卫星而言, 各组中所有预测激活的时延敏感用户就是其动作空间集合。

2) 奖励设置

第 k 组中吞吐量敏感用户对应的激活时延敏感用户预测列表为 $S_{a,k}$, $k \in \{1, 2, \dots, K\}$, LEO卫星在 t 时隙选择第 i 个摇臂, 根据优化问题式(11)可得该摇臂的瞬时奖励表达式为

$$\theta_i(t) = 1[R_i(t) \geq R_{th}]1[\Delta_i(t) \leq D_{th}]U_i(t) \quad (14)$$

其中, $1[\cdot]$ 为指示函数, 当满足括号中的条件时为1, 不满足时为0。2个指示函数分别表示优化问题式(11)中的速率和数据包缓存时间约束。

3) 非稳态MAB模型的求解

基于上文给出的动作空间和奖励设置, 采用折扣上界置信区间(D-UCB, discounted upper confidence bound)算法来进行非稳态MAB决策, 从而得到本文的NOMA用户调度方案。计算 t 时隙第 i 个摇臂的折扣期望奖励为

$$\bar{\theta}_i(t) = \frac{\sum_{s=1}^t \rho^{t-s} \theta_i(s)}{N_i(\rho, t)} \quad (15)$$

其中, $\rho \in (0, 1)$ 为设定好的折扣因子, 对应着历史奖励的利用程度; $N_i(\rho, t) = \sum_{s=1}^t \rho^{t-s} 1\{I_s = i\}$ 为摇臂 i 在 t 时隙的折扣选择总次数, $1\{I_s = i\}$ 为摇臂选择的指示函数, 当 s 时隙选择摇臂 i , 该表达式值为

1, 反之则为0。

根据式(16)可以得到各摇臂的索引值为

$$G_i(\rho, t) = \bar{\theta}_i(t) + 2\zeta \sqrt{\zeta \frac{\text{lb}n(\rho, t)}{N_i(\rho, t)}} \quad (16)$$

其中, $2\zeta \sqrt{\zeta \frac{\text{lb}n(\rho, t)}{N_i(\rho, t)}}$ 为折扣探索因子, ζ 和 ζ 均为

设定好的乘性因子, $n(\rho, t)$ 为动作空间中所有动作的累计折扣选择次数的总和。得到各摇臂的索引值后, 选择索引值最大的作为下一个时隙执行动作。如果存在首次激活的用户, LEO卫星会优先为其分配上行授权, 实现优先调度。基于部分ZF波束成形的D-UCB调度算法流程如算法1所示。

算法1 基于部分ZF波束成形的D-UCB调度算法

1) 根据式(12)得到各组的波束成形权矢量 $\mathbf{w}_{k,ZF}$, 初始化各组中时延敏感用户的折扣调度次数 $N_{k,i_k} = 0$, 总折扣调度次数 $n_{k,i_k} = 0$, 索引值 $G_{k,i_k} = \text{INF}$, 当前时隙数 $t = 0$, 设定总的时隙数为 T ;

2) 当各组中的时延敏感用户 $i_k \in S_{a,k}$, 根据 $I_{t,k} = \arg \max_{i_k \in S_{a,k}} G_{k,i_k}(\rho, t)$, 调度各组中的时延敏感用户, 更新 $N_{k,i_k} = \rho N_{k,i_k} + 1$, 根据式(14)得到对应的奖励;

3) 根据式(15)、式(16)更新各组内激活时延敏感用户的 $N_{k,i_k}, n_{k,i_k}, G_{k,i_k}$, 若 $G_{k,i_k} = \text{INF}$, 则保持不变, 同时设置 $t = t + 1$;

4) 若满足 $t > T$, 则迭代结束; 否则, 返回步骤2)。

3 仿真结果与分析

本节基于MATLAB 2018a仿真平台来验证所提方案的有效性和优越性。仿真场景参数设置如下: 吞吐量敏感用户数 $K = 4$, 各组内时延敏感用户数 $S_k = 50$, 且各时延敏感用户随机均匀分布在吞吐量敏感用户周围 $10 \text{ km} \times 10 \text{ km}$ 的范围内, 时延敏感用户激活概率 $P_a \in [0.5, 1]$; 配置在LEO卫星的天线数为 $L_x = L_y = 8$, 服务范围为半径为 500 km 的圆形区域。假设覆盖范围中心为坐标原点, 吞吐量敏感用户位置坐标分别为 $\{(-150 \text{ km}, 100 \text{ km}), (-125 \text{ km}, 150 \text{ km}), (-100 \text{ km}, -175 \text{ km}), (-100 \text{ km}, -175 \text{ km}), (-125 \text{ km}, 125 \text{ km})\}$, 其余仿真参数如表1所示。

表 1 仿真参数

参数	值
工作频率 f/GHz	20
卫星接收天线最大增益 G_r/dBi	30
仿真总时隙 T	8000
卫星高度 H/km	550
乘性系数 ζ, ξ	$\zeta = 2, \xi = 4$
效用函数权重因子 δ_1, δ_2	$\delta_1 = 0.4, \delta_2 = 0.6$
Beta 函数参数 α, β	$\alpha = 3, \beta = 4$
首次激活时隙长度 N_A	15

对于本文算法 (D-UCB 算法) 而言, 折扣因子 ρ 的选择会直接影响算法的性能。当 $\rho > 0$ 时, 随着 ρ 的增大, 算法对历史奖励的依赖程度增加, 反之则降低; 当 $\rho = 1$ 时, 算法则完全基于历史奖励进行决策, 退化为传统的置信上界 (UCB, upper confidence bound) 算法。本文选用 $\rho = 0$ 、 $\rho = 0.8$ 和 $\rho = 1$ 这 3 种不同的折扣因子, 比较了不同情况下每个时隙内吞吐量敏感用户的平均传输速率, 以评估不同折扣因子下系统频谱效率的表现。根据图 2 可知, 当 $\rho = 0$ 和 $\rho = 1$ 时, D-UCB 算法下吞吐量敏感用户的平均传输速率在前几个时隙中增长到一定数值后, 就会随时隙数的增加逐渐降低, 最终趋于平稳。与 $\rho = 1$ 相比, 当 $\rho = 0$ 时, 由于完全忽略历史奖励, 算法仅依据瞬时奖励做出决策, 导致平均传输速率下降更快, 系统性能较差。当 $\rho = 0.8$ 时, D-UCB 算法的平均传输速率保持稳定, 并优于随机调度和 ϵ 贪心算法。这表明, 在奖励非稳态的情况下, D-UCB 算法通过对历史奖励的选择性利用, 能够取得更好的性能表现。

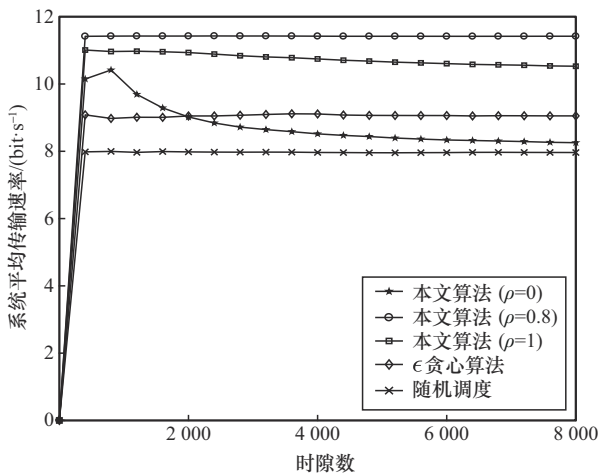


图 2 吞吐量敏感用户平均传输速率对比

数据包丢弃是由于时延敏感用户准备好传输的数据包等待时隙过长, 从而产生新的数据包替代旧的数据包, 丢弃旧数据包资源。图 3 给出了数据包缓存门限值小于用户数、等于用户数和大于用户数 3 种情况下, 不同算法的系统数据包丢弃情况。由图 3 可知, 在 $D_{th} = 30$ 、 $D_{th} = 50$ 、 $D_{th} = 70$ 这 3 种情况下, 本文算法的数据包丢弃数量都低于 UCB 算法、随机调度和 ϵ 贪心算法的系统数据包丢弃数量。当数据包缓存门限值 $D_{th} = 30$ 时, 由于用户数大于该门限值, 数据包丢弃的情况难以避免, 但本文算法的系统数据包丢弃数量仍低于其他 3 种算法。这是由于该 MAB 问题的奖励是非稳态分布的, D-UCB 算法通过对历史奖励的选择性利用, 可以更好地应对奖励的动态变化, 选择更优的用户进行调度。随着 D_{th} 的增加, 系统数据包丢弃数量都会随之下降, 这是因为随着 D_{th} 的增加, 同一时隙数内产生数据包数量减少并且数据包缓存时间更久, 造成的数据包丢弃数就会随之降低。

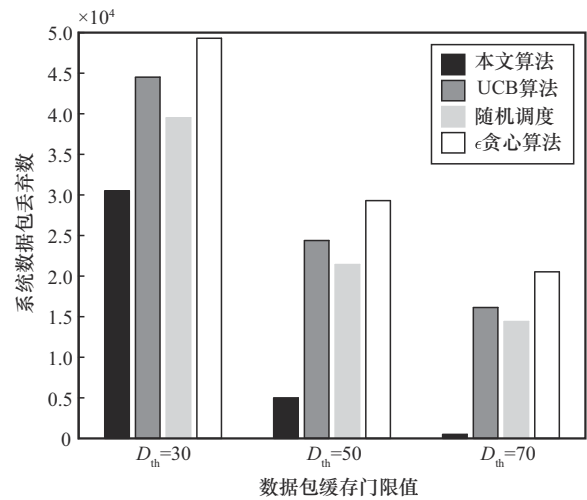


图 3 不同数据包缓存门限值下系统数据包丢弃数对比

图 4 给出了在数据包缓存门限值 D_{th} 小于用户数、等于用户数和大于用户数 3 种情况下, 采用不同接入算法后的系统平均 AoI。由于系统 AoI 收敛较快, 为方便展示, 本文只给出前 2 000 时隙数下不同算法的系统平均 AoI。由图 4 可知, 在 3 种不同门限值情况下, 系统平均 AoI 收敛后, 本文算法的系统 AoI 均低于 UCB 算法、随机调度和 ϵ 贪心算法的系统 AoI, 且随着门限值的增加, 本文算法的性能更优。在 $D_{th} = 50$ 时, 前 400 时隙随机调度的系统平均 AoI 略低于本文算法, 这是由于本文算法

需要根据历史奖励进行学习,在调度初期并不能实现最精准的调度。同时,由于此时门限值较高,时延敏感用户激活状态保持更久,前期的平均AoI较高。随着时隙的增加,本文算法能有效地根据历史奖励进行决策,实现更优的调度结果。

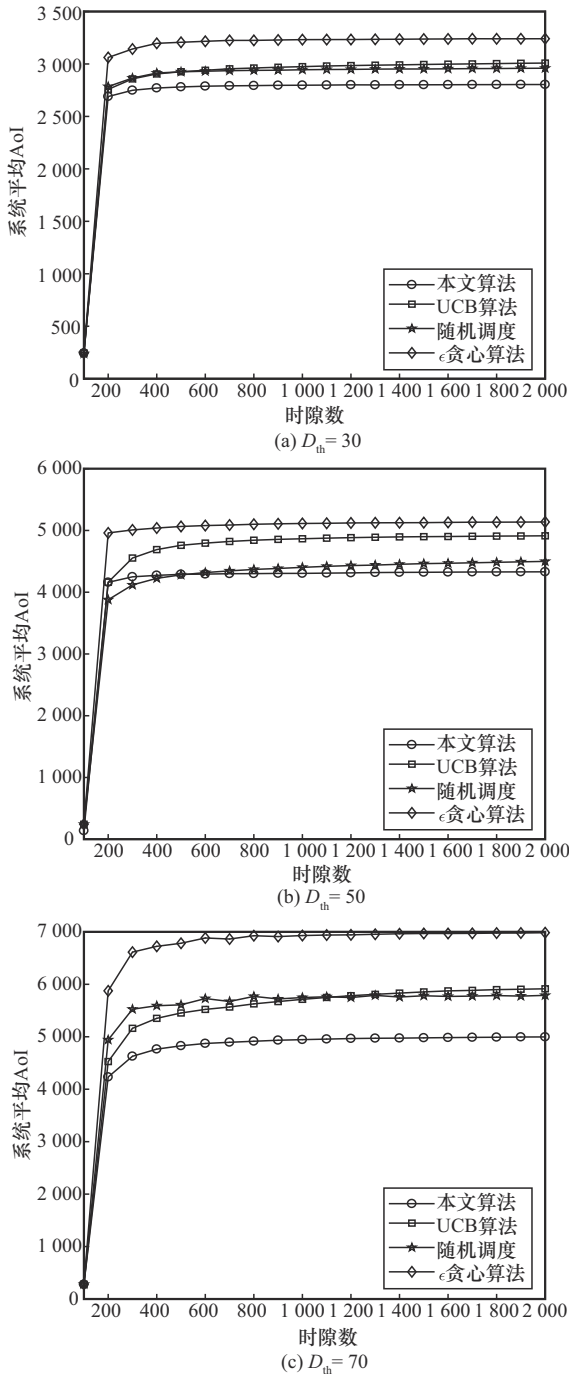


图 4 不同数据包缓存阈值下系统平均AoI对比

图5给出了不同预测误差概率下系统调度资源浪费数对比。调度资源浪费指的是当LEO卫星预测用户

处于激活状态时,分配其上行接入授权,从而调度完成上行NOMA传输。但是由于预测误差,该用户并未真正激活,从而调度失败,未完成NOMA上行传输。由图5可知,随着预测模型误差概率的增加,系统的调度资源浪费次数也随之增加。这是因为当预测模型误差概率变大时,各时隙预测激活用户列表错误也随之增加,导致在调度过程中调度非激活用户的概率变大,从而造成更多的调度资源浪费。

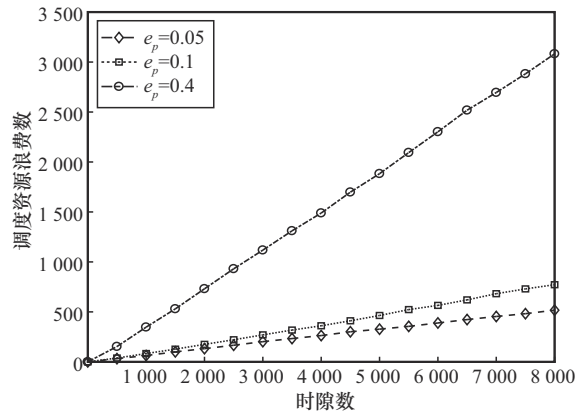


图 5 不同预测误差概率下系统调度资源浪费数对比

4 结束语

本文研究了LEO卫星系统异构用户上行NOMA接入策略,旨在最小化时延敏感用户的AoI,并保证提升吞吐量敏感用户的传输速率。在同时考虑时延敏感用户信息年龄和吞吐量敏感用户传输速率需求的基础上,建立联合效用函数最大化为目标,且用户的速率门限和数据包缓存时间满足要求为约束条件的优化问题。由于传统的优化方法无法对该问题直接进行求解,在通过部分ZF波束成形技术消除组间干扰的基础上,提出了基于MAB的动态调度算法,实现了时延敏感与吞吐量敏感用户的智能配对,以保证异构用户的可靠性接入。仿真结果表明,本文的D-UCB算法相比其他现有方案更适合该场景,在保证系统QoS的前提下,显著降低了数据包丢弃率与系统AoI,尤其在数据包缓存门限值大于用户数时表现尤其优越。

参考文献:

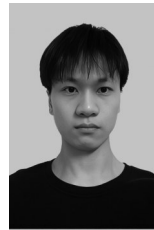
[1] LE T T T, HASSAN N U, CHEN X M, et al. A survey on random access protocols in direct-access LEO satellite-based IoT communication[J]. IEEE Communications Surveys & Tutorials, 2025, 27(1): 426-462.
 [2] 林敏,李菲,王子宁,等.智能反射面辅助的太赫兹卫星通信鲁棒安全

- 波束成形算法[J]. 电子学报, 2023, 51(10): 2715-2723.
- LIN M, LI F, WANG Z N, et al. Robust secure beamforming algorithm for IRS-aided terahertz satellite communication systems[J]. Acta Electronica Sinica, 2023, 51(10): 2715-2723.
- [3] 林敏, 阚鹏程, 赵柏, 等. 基于非稳态 MAB 的 LEO 卫星跳波束时隙分配算法[J]. 通信学报, 2023, 44(8): 134-143.
- LIN M, KAN P C, ZHAO B, et al. Time-slot allocation algorithm for LEO satellite beam hopping based on non-stationary MAB[J]. Journal on Communications, 2023, 44(8): 134-143.
- [4] KHAVARI-MOGHADDAM S, FARAHMAND S, RAZAVIZADEH S M, et al. Optimum solutions for weighted sum-rate of NOMA and TDMA in wireless-powered IoT networks[J]. IEEE Internet of Things Journal, 2024, 11(2): 3302-3315.
- [5] 杨龙, 李函洋, 任萌萌, 等. 面向缓存网络内容分发的非正交传输策略研究[J]. 通信学报, 2023, 44(10): 124-136.
- YANG L, LI H Y, REN M M, et al. Research on non-orthogonal transmission strategy for content delivery in caching networks[J]. Journal on Communications, 2023, 44(10): 124-136.
- [6] 林敏, 郭怀波, 刘笑宇, 等. IRS 辅助的星地融合认知网络中断性能分析[J]. 通信学报, 2023, 44(7): 64-75.
- LIN M, GUO H B, LIU X Y, et al. Outage performance analysis of IRS-aided cognitive satellite-terrestrial network[J]. Journal on Communications, 2023, 44(7): 64-75.
- [7] KONG H C, LIN M, WANG Z N, et al. Performance analysis for rate splitting uplink NOMA transmission in high throughput satellite systems[J]. IEEE Wireless Communications Letters, 2022, 11(4): 816-820.
- [8] MAKITA T, MUTA O. QoS-aware low-complexity user pairing based on compressed sensing in downlink NOMA[C]//Proceedings of the 2021 IEEE 94th Vehicular Technology Conference (VTC2021-Fall). Piscataway: IEEE Press, 2021: 1-5.
- [9] WANG W X, YU W J, FOH C H, et al. User scheduling in NOMA random access using contextual multi-armed bandits[C]//Proceedings of the 2022 IEEE Globecom Workshops (GC Wkshps). Piscataway: IEEE Press, 2022: 112-117.
- [10] TANAB M E, HAMOUDA W. Fast-grant learning-based approach for machine-type communications with NOMA[C]//Proceedings of the IEEE International Conference on Communications. Piscataway: IEEE Press, 2021: 1-6.
- [11] KAUL S, YATES R, GRUTESER M. Real-time status: how often should one update? [C]//Proceedings of the IEEE INFOCOM. Piscataway: IEEE Press, 2012: 2731-2735.
- [12] SOMBABU B, MOHARIR S. Age-of-information based scheduling for multi-channel systems[J]. IEEE Transactions on Wireless Communications, 2020, 19(7): 4439-4448.
- [13] 胡昊南, 韩铭, 李文鹏, 等. 面向无线传感器网络信息年龄的多无人机轨迹优化算法[J]. 电子与信息学报, 2024, 46(4): 1222-1230.
- HU H N, HAN M, LI W P, et al. Multi-unmanned aerial vehicles trajectory optimization for age of information minimization in wireless sensor networks[J]. Journal of Electronics & Information Technology, 2024, 46(4): 1222-1230.
- [14] ZHOU C H, HE H L, YANG P, et al. Deep RL-based trajectory planning for AoI minimization in UAV-assisted IoT[C]//Proceedings of the 2019 11th International Conference on Wireless Communications and Signal Processing (WCSP). Piscataway: IEEE Press, 2019: 1-6.
- [15] WEI C H, BIANCHI G, CHENG R G. Modeling and analysis of random access channels with bursty arrivals in OFDMA wireless networks[J]. IEEE transactions on wireless communications, 2014, 14(4): 1940-1953.
- [16] SHEHAB M, HAGELSKJÆR A K, KALØR A E, et al. Traffic prediction based fast uplink grant for massive IoT[C]//Proceedings of the 2020 IEEE 31st Annual International Symposium on Personal, Indoor and Mobile Radio Communications. Piscataway: IEEE Press, 2020: 1-6.
- [17] YAN B W, WANG X J, XU C, et al. Age-based scheduling in Internet of things with bursty traffic over time varying channels[C]//Proceedings of the 2021 Computing, Communications and IoT Applications (ComComAp). Piscataway: IEEE Press, 2021: 242-247.
- [18] ZHAO B, LIN M, CHENG M, et al. Robust downlink transmission design in IRS-assisted cognitive satellite and terrestrial networks[J]. IEEE Journal on Selected Areas in Communications, 2023, 41(8): 2514-2529.
- [19] GARIVIER A, MOULINES E. On upper-confidence bound policies for switching bandit problems[C]//International Conference on Algorithmic Learning Theory. Berlin: Springer, 2011: 174-188.

[作者简介]



林敏 (1972–), 男, 浙江台州人, 博士, 南京邮电大学教授、博士生导师, 主要研究方向为无线通信系统、智能信号处理、天线新技术等。



孙昊 (2000–), 男, 安徽铜陵人, 南京邮电大学硕士生, 主要研究方向为卫星通信、无线通信、人工智能。



赵柏 (1995–), 男, 江苏泗阳人, 南京邮电大学博士生, 主要研究方向为无线通信中的智能信号处理技术、人工智能、混合多址接入技术等。



郭焱 (2000–), 女, 山西寿阳人, 南京邮电大学博士生, 主要研究方向为无线通信、非正交多址接入等。



杨绿溪 (1964–), 男, 安徽桐城人, 博士, 东南大学教授、博士生导师, 主要研究方向为无线通信中的信号处理、MIMO 通信、智能无线通信和统计信号处理等。