

doi: 10.14132/j.cnki.1673-5439.2025.05.004

基于RIS辅助的RSMA无蜂窝上行通信系统和速率优化算法

俞观晔¹, 于佳², 陈沛³, 杨仕钊⁴, 田雨露⁴, 朱晓荣⁴

1. 中国通信服务股份有限公司中浙信科技咨询有限公司, 浙江杭州 310014
2. 国网电力科学研究院有限公司, 江苏南京 210000
3. 国网天津市电力公司城南供电分公司, 天津 300000
4. 南京邮电大学通信与信息工程学院, 江苏南京 210003

摘要: 针对无蜂窝系统中通信链路易受环境影响、终端接入设备受限的问题, 引入智能反射面(Reconfigurable Intelligent Surface, RIS)和速率分割多址(Rate Splitting Multiple Access, RSMA)技术辅助通信, 提出了一种基于RIS辅助的RSMA无蜂窝上行通信系统和速率优化算法, 通过联合优化RIS的无源波束形成、接入点选择、子消息功率分配和解码顺序, 以实现系统和速率最大化。由于该优化问题为非凸的离散非线性规划, 考虑将其解耦成多个子问题进行交替优化, 首先采用基于信道增益的以用户为中心的匹配算法获取最优匹配方案, 接着采用连续凸近似(Successive Convex Approximation, SCA)算法解决子消息的功率分配问题, 通过黎曼共轭梯度法优化RIS的无源波束形成, 然后提出一种基于信道增益以及用户子消息分配比例的解码顺序算法。仿真结果表明, 与传统算法相比, 所提算法具有更优的和速率。

关键词: 智能反射面; 速率分割多址; 无蜂窝网络; 和速率

中图分类号: TN929.5 文献标志码: A 文章编号: 1673-5439(2025)05-0031-11

Sum-rate optimization for RIS-aided cell-free uplink communication systems with RSMA

YU Guanye¹, YU Jia², CHEN Pei³, YANG Shizhao⁴, TIAN Yulu⁴, ZHU Xiaorong⁴

1. Zhongzhixin Technology Consulting Co., Ltd., Hangzhou 310014, China
2. State Grid Electric Power Research Institute Co., Ltd., Nanjing 210000, China
3. State Grid Tianjin Electric Power Company Chengnan Power Supply Branch, Tianjin 300000, China
4. School of Communications and Information Engineering, Nanjing University of Posts and Telecommunications, Nanjing 210003, China

Abstract: To address the challenges of environmental sensitivity in communication links and limited terminal access in cell-free communication systems, this paper introduces reconfigurable intelligent surface (RIS) and rate splitting multiple access (RSMA) technologies to assist communication. A sum-rate optimization algorithm is proposed for RIS-aided cell-free uplink communication systems with RSMA. The algorithm aims to maximize the system sum-rate by jointly optimizing RIS passive beamforming, access point selection, power allocation of sub-messages, and decoding order. Given the non-convex nature of this dis-

收稿日期: 2024-11-18; 修回日期: 2024-12-16 本刊网址: <http://nyzr.njupt.edu.cn>

基金项目: 国家自然科学基金(92367102)资助项目

作者简介: 俞观晔, 男, 高级工程师; 朱晓荣(通信作者), 女, 博士, 教授, 博士生导师, xrzhu@njupt.edu.cn

引用本文: 俞观晔, 于佳, 陈沛, 等. 基于RIS辅助的RSMA无蜂窝上行通信系统和速率优化算法[J]. 南京邮电大学学报(自然科学版), 2025, 45(5): 31-41.

Citation: YU Guanye, YU Jia, CHEN Pei, et al. Sum-rate optimization for RIS-aided cell-free uplink communication systems with RSMA[J]. Journal of Nanjing University of Posts and Telecommunications (Natural Science Edition), 2025, 45(5): 31-41.

crete nonlinear programming problem, we decompose it into multiple subproblems and address them by alternatively using an optimization approach. First, a user-centric matching algorithm based on channel gain is used to obtain the optimal matching scheme. Second, the successive convex approximation (SCA) algorithm is used to solve the sub-message power distribution problem, and the Riemann conjugate gradient algorithm is utilized to solve the RIS passive beamforming problem. Third, a decoding order algorithm is developed based on channel gain and the sub-message allocation ratio. Simulation results demonstrate that the proposed algorithm achieves superior sum-rate performance compared to traditional algorithms.

Keywords: reconfigurable intelligent surface (RIS); rate splitting multiple access (RSMA); cell-free network; sum-rate

随着新兴物联网应用的发展,第六代移动通信网络被要求支持超可靠和超低延时服务的需求^[1]。无蜂窝(Cell-Free, CF)系统被认为是未来支持新兴应用的无线网络的潜在范例。在传统蜂窝网络中,小区间干扰(Inter-Cell Interference, ICI)会导致小区边缘用户服务质量下降,进而影响整个系统性能。CF网络架构被视为一种分布式系统的扩展实现,通过多个随机部署的接入点(Access Point, AP)共享同一资源块并协同服务用户,大大提高了边缘用户的性能,减少了ICI带来的负面影响,是未来6G网络中一种非常可观的物理层技术^[2-3]。然而,CF网络在实际部署中面临多用户动态接入、复杂传播环境以及干扰管理的挑战。

在未来的移动通信环境中,服务类型将更加多样化,一般来说,通信运营商需要为用户提供各种服务,既需要支持单播业务,如语音通信等,也需要支持多播业务,如直播业务和视频广播业务等,这就要求基站(Base Station, BS)支持多业务联合传输。为了支持多业务联合传输,速率分割多址技术(Rate Splitting Multiple Access, RSMA)被认为是一种极具前景的实现方法^[4]。与空分多址接入(Space Division Multiple Access, SDMA)依赖于将所有剩余干扰作为噪声处理、非正交多址接入(Non-Orthogonal Multiple Access, NOMA)依赖于完全解码干扰不同,RSMA可以对干扰进行部分解码,并将部分干扰作为噪声处理,凭借这种能力,RSMA在多用户部署中比SDMA或NOMA具有更高的频谱和能量效率^[5-6]。文献[7]提出了一种基于CF架构的联合多业务传输方案,基于以用户为中心的原则进行AP点的选择,并根据RSMA技术在每个接入点上形成广播、组播和单播消息,然后进行多业务信号的联合传输,提出了一种新的三步资源分配算法来求解和速率最大化的非凸问题。

近年来,可重构智能表面(Reconfigurable Intelligent Surface, RIS)作为未来6G中的一项关键技术,因其可以按需动态控制信道环境而备受学术界的关

注^[8-9]。RIS可以对信号传输链路进行智能的控制,尽管CF系统中用户与AP之间的直连链路通常较短,但在实际应用场景中,尤其是室内环境中,RIS的部署具有显著优势。例如,在室内环境下,用户与AP之间的视距传播可能因建筑物墙体、家具等障碍物阻塞,导致信道质量恶化。RIS的引入能够通过优化反射路径,重构信道环境,显著提升系统覆盖性能和通信可靠性。此外,RIS在多用户复杂场景中具有优化干扰管理和提升频谱效率的潜力。

目前已有大量研究将RIS与NOMA技术结合,在通信系统中发挥两者的优势^[10-11]。而RSMA是一种新型的、更通用的多址框架,构建基于RIS辅助的RSMA无线通信系统,是满足未来6G通信场景需求的一个极具潜力的方案^[12-14]。文献[15]研究了一种基于RSMA的RIS辅助认知无线网络,在多种约束条件下优化了主用户的安全能效,展现了在提高频谱效率方面的巨大优势。文献[16]研究了一种基于RIS辅助的RSMA通信网络,该网络涉及近端和蜂窝边缘用户,在BS功率约束、每个用户的服务质量(Quality of Service, QoS)约束和RIS的单位相移约束下,通过设计RIS的反射相移和BS的波束形成来最大化用户的加权和速率。文献[17]研究了在用户发射功率和无源波束形成约束下,单RIS辅助下的UL RSMA系统的和速率最大化问题,且验证了该系统在和速率吞吐量方面优于NOMA和OMA方案的系统。文献[18]进一步将场景复杂化为多RIS辅助的UL RSMA毫米波通信系统,为了有效缓解多用户场景下的高节点间干扰,提出了一种基于K-means聚类的低复杂度用户配对方案,然后开发了一种有效的低成本交替优化框架,实现和速率最大化。

考虑将RIS和RSMA技术结合用于上行CF网络中。目前还没有在无蜂窝网络中应用RIS和RSMA辅助上行通信的相关研究,本文的主要贡献如下:

(1) 将RSMA技术应用于以用户为中心的无蜂窝网络架构中,并引入RIS辅助信号传输,提出了一

种基于RIS辅助的RSMA无蜂窝上行通信系统和速率最大化算法研究。

(2) 提出了联合优化AP选择、用户子消息功率分配、RIS相移矩阵以及AP处消息解码顺序以实现上行通信系统和速率最大化问题。为了求解该问题,采用交替优化算法。首先,提出了基于用户中心准则的AP配对算法,获得用户最优基站选择;接着利用连续凸优化算法获得用户子消息功率最优分配,利用黎曼流形梯度算法获得RIS的最优相移。

(3) 引入了一种基于用户配对的算法,降低接收端执行串行干扰去除(Successive Interference Cancellation, SIC)的解码顺序复杂度。仿真结果表明,与传统的SDMA和NOMA多址技术相比,RSMA具有更高的系统和速率。

1 系统模型

如图1所示,考虑一个基于RIS辅助的具有时分双工的CF上行(Uplink, UL)RSMA通信系统,其中 K 个用户设备(User Equipment, UE)被 M 个AP所覆盖并与其进行上行通信,为了方便后续推导,假设UE和AP都配备了单天线。在AP与用户之间部署RIS以协助用户向AP传输数据,其中RIS含有 N 个反射单元。为了方便表述,定义 $\mathcal{M} = \{1, \dots, M\}$ 表示AP的集合, $\mathcal{K} = \{1, \dots, K\}$ 表示用户的集合,矩阵 $\mathbf{D} = \{d_{m,k}, \forall m \in \mathcal{M}, \forall k \in \mathcal{K}\}$ 表示用户和AP点之间的匹配关系,其中 $d_{m,k} = 1$ 表示第 k 个用户与第 m 个AP匹配成功并建立数据传输链路, \mathcal{M}_m 表示第 m 个AP接收与用户进行通信的集合。在导频训练阶段,用户通过上行链路发送被分配的导频序列给AP,在接收端AP利用接收到的导频信号进行信道估计,得到CSI。中央处理单元(Central Processing Unit, CPU)通过可靠的回程网络与所有AP进行通信,交换网络信息并执行统一调度。

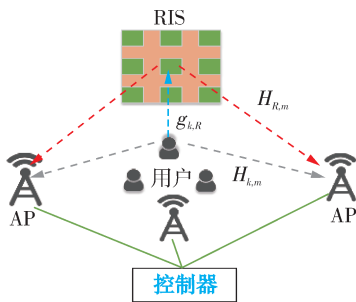


图1 基于RIS辅助的RSMA无蜂窝上行通信系统模型图

在该UL RSMA系统的数据传输阶段,考虑将每个用户的消息分成 J 个子消息,并在同一时间和频率资源内同时发送给AP^[19]。 $\mathcal{J} = \{1, \dots, J\}$ 表示每

个用户子消息的集合。 x_k 表示第 k 个用户发送的消息,即

$$x_k = \sum_{j \in \mathcal{J}} \sqrt{P_{kj}} s_{kj}, \quad \forall k \in \mathcal{K} \quad (1)$$

其中, s_{kj} 表示第 k 个用户的第 j 个子消息并且满足 $\mathbb{E}[|s_{kj}|^2] = 1$, P_{kj} 表示分配给第 k 个用户的第 j 个子消息的发射功率,每个用户的最大发射功率用 P_k^{\max} 表示,即

$$P_k = \sum_{j \in \mathcal{J}} P_{kj} \leq P_k^{\max}, \quad \forall k \in \mathcal{K} \quad (2)$$

AP通过两条信道来接收每个用户的信号,即UE-AP的直连信道以及UE-RIS-AP的级联信道。第 m 个AP点接收到用户的总信号可以表示为

$$y_m = \sum_{k=1}^K d_{m,k} (\mathbf{g}_{k,R}^n \Phi \mathbf{H}_{R,m} + \mathbf{H}_{k,m}) x_k + n = \sum_{k=1}^K d_{m,k} (\mathbf{g}_{k,R}^n \Phi \mathbf{H}_{R,m} + \mathbf{H}_{k,m}) \sum_{j \in \mathcal{J}} \sqrt{P_{kj}} s_{kj} + n \quad (3)$$

其中, $d_{m,k} \in \{0, 1\}$ 为一个0或1的变量,表示第 k 个用户与第 m 个AP点是否建立通信, $\mathbf{g}_{k,R} \in \mathbb{C}^{N \times 1}$ 表示第 k 个用户与RIS之间的信道增益, $\mathbf{H}_{R,m} \in \mathbb{C}^{N \times 1}$ 表示RIS与第 m 个AP点之间的信道增益, $\mathbf{H}_{k,m}$ 表示第 k 个用户与第 m 个AP点之间的信道增益。 $\Phi = \text{diag}(e^{j\phi_1}, \dots, e^{j\phi_N}) \in \mathbb{C}^{N \times N}$ 表示RIS被动波束形成的对角矩阵, ϕ_1, \dots, ϕ_N 分别表示RIS的 N 个反射单元的相移。 n 表示加性高斯白噪声,该噪声为服从均值为零、方差为 σ^2 的均匀复高斯随机变量,即 $n \sim \mathcal{CN}(0, \sigma_k^2)$ 。RIS与AP的位置是预先固定好的,这样做是为了确保RIS-UE与RIS-AP之间的LOS路径。信道模型采用准静态平坦衰落信道模型,假设在AP、RIS与UE之间的所有信道的CSI是完全已知的。

第 m 个AP利用SIC技术从接收到的消息 y_m 中解码所有用户的子消息,则总共有 $M = \sum_{k=1}^K d_{m,k} \times J$ 个子消息需要解码。假设第 m 个AP中子消息的解码顺序用集合 $\zeta_m = \{s_{kj}: k \in \mathcal{M}_m, j \in \mathcal{J}\}$ 表示,其中第一个元素先解码,接着第二个元素解码,以此类推。集合 ζ_m 属于大集合 Ω_m , Ω_m 表示第 m 个AP处子消息所有可能解码顺序的集合。设 ζ_m^{kj} 表示第 m 个AP处子消息 s_{kj} 的解码顺序,对于子消息 s_{kj} ,AP会将所有解码顺序比 s_{kj} 低的子消息剔除,并将剩余的子消息视为干扰。因此,RSMA系统中解码子消息 s_{kj} 的信噪比可以表示为

$$\gamma_m^{kj} = \frac{|\mathbf{g}_{k,R}^n \Phi \mathbf{H}_{R,m} + \mathbf{H}_{k,m}|^2 P_{kj}}{\sum_{(u,v) \in \mathcal{Q}_m^{kj}} |\mathbf{g}_{u,R}^n \Phi \mathbf{H}_{R,m} + \mathbf{H}_{u,m}|^2 P_{uv} + \sigma^2} \quad (4)$$

其中, \mathcal{Q}_m^{kj} 表示在第 m 个AP处解码顺序大于 s_{kj} 的所有子消息的集合,集合中的子消息全部在 s_{kj} 之后才

进行解码,即满足 $\mathcal{Q}_m^{bj} = \{(u, v): \zeta_m^{uv} > \zeta_m^{bj}\}$, 解码子消息 s_{kj} 可实现速率为

$$R_m^{kj} = \log_2(1 + \gamma_m^{kj}) = \log_2 \left(1 + \frac{|g_{k,R}^n \Phi H_{R,m} + H_{k,m}|^2 P_{kj}}{\sum_{(u,v) \in \mathcal{Q}_m^{bj}} |g_{u,R}^n \Phi H_{R,m} + H_{u,m}|^2 P_{uv} + \sigma^2} \right) \quad (5)$$

整个消息是通过 J 个子消息流传输的,因此可实现速率为

$$R_m^k = \sum_{j \in \mathcal{J}} R_m^{kj} \quad (6)$$

2 和速率最大化问题

在本文所建立的模型中,主要的研究目标是通过联合优化 AP 选择、用户子消息功率分配、RIS 无源波束形成和 AP 处解码顺序,使受 QoS 约束的基于 RIS 辅助的 RSMA 无蜂窝上行通信系统的和速率最大化。利用第 1 节的公式,优化问题可以表述为

$$\begin{aligned} \text{P1: } \max_{\mathbf{D}, \boldsymbol{\zeta}, \Phi, P} \quad & \sum_{m \in \mathcal{M}} \sum_{k \in \mathcal{K}} \sum_{j \in \mathcal{J}} d_{m,k} R_m^{kj} \\ \text{s. t. C1: } \quad & d_{m,k} \in \{0, 1\}, \quad \forall k \in \mathcal{K}, \quad \forall m \in \mathcal{M} \\ \text{C2: } \quad & |\Phi_{i,i}| = 1, \quad \forall i \in \{1, \dots, N\} \\ \text{C3: } \quad & \zeta_m \in \Omega_m, \quad \forall m \in \mathcal{M} \\ \text{C4: } \quad & R_m^k \geq R_k^{\min}, \quad \forall k \in \mathcal{K} \\ \text{C5: } \quad & R_{k1}: R_{k2}: \dots: R_{kj} = \\ & h_{k1}: h_{k2}: \dots: h_{kj}, \quad \forall k \in \mathcal{K} \\ \text{C6: } \quad & \sum_{j \in \mathcal{J}} P_{kj} \leq P_k^{\max}, \quad \forall k \in \mathcal{K} \end{aligned} \quad (7)$$

其中, $\mathbf{P} = [\dots, P_{kj}, \dots]$ 代表用户的子消息功率分配, $\boldsymbol{\zeta} = [\dots, \zeta_m, \dots]$ 代表 AP 处消息解码顺序。 h_{kj} 是非负的,表示第 k 个用户对第 j 个子消息的分割比例并满足 $R_{kj} = h_{kj} R_k$, $\forall j \in \mathcal{J}$ 且 $\sum_{j \in \mathcal{J}} h_{kj} = 1$ 。约束 C1 表示

AP 与 UE 之间的匹配关系;约束 C2 表示 RIS 无源波束形成;约束 C3 表示子消息解码顺序约束;C4 表示用户最小速率需求;C5 表示用户子消息功率分配比例;C6 表示用户功率约束。

P1 中的目标函数由于变量的强耦合而非凸,很

$$\begin{aligned} \text{P2: } \max_P \quad & \sum_{m \in \mathcal{M}} \sum_{k \in \mathcal{K}} \sum_{j \in \mathcal{J}} d_{m,k} \log_2 \left(\frac{\sum_{(u,v) \in \mathcal{Q}_m^{bj+}} |g_{u,R}^n \Phi H_{R,m} + H_{u,m}|^2 P_{uv} + \sigma^2}{\sum_{(u,v) \in \mathcal{Q}_m^{bj}} |g_{u,R}^n \Phi H_{R,m} + H_{u,m}|^2 P_{uv} + \sigma^2} \right) \\ \text{s. t. C1: } \quad & \sum_{j \in \mathcal{J}} P_{kj} \leq P_k^{\max}, \quad \forall k \in \mathcal{K} \\ \text{C2: } \quad & \sum_{j=1}^{\mathcal{J}} \log_2 \left(1 + \frac{|g_{k,R}^n \Phi H_{R,m} + H_{k,m}|^2 P_{kj}}{\sum_{(u,v) \in \mathcal{Q}_m^{bj}} |g_{u,R}^n \Phi H_{R,m} + H_{u,m}|^2 P_{uv} + \sigma^2} \right) \geq R_k^{\min} \\ \text{C3: } \quad & R_{k1}: R_{k2}: \dots: R_{kj} = h_{k1}: h_{k2}: \dots: h_{kj}, \quad \forall k \in \mathcal{K} \end{aligned} \quad (8)$$

难转化为凸问题的一个重要原因就是信道增益依赖于相位矩阵 Φ 。此外,离散变量 $d_{m,k}$ 是一个 0-1 问题,也是非凸的,离散变量 ζ_m 使优化问题成为了极具挑战的混合整数非凸问题。

3 问题求解

由于 P1 的非凸性,需将原问题解耦为不同的子问题,首先解决 AP 与 UE 配对的问题,这里使用基于用户为中心的配对算法;其次针对给定的解码顺序分别解决用户功率分配问题和 RIS 无源波束形成问题;最后,采用 AO 算法迭代对这些子问题进行联合求解,给出优化问题的整体解,具体求解过程如下:

3.1 UE 与 AP 匹配问题求解

首先求解 UE 与 AP 匹配问题,确定优化问题 P1 中的变量 \mathbf{D} 。每个用户与不同的 AP 进行通信时信道链路具有不同的 CSI,因此合理有效的匹配方案对系统性能至关重要。令 $\mathbf{G} \in \mathbb{C}^{K \times M}$ 表示 UE 到 AP 间所有信道增益组成的矩阵,其中 $\mathbf{G}_{m,k} = g_{k,R}^n \Phi H_{R,m} + H_{k,m}$ 表示第 k 个用户与第 m 个 AP 点之间的信道增益。提出基于以用户为中心的匹配算法如算法 1 所示。

算法 1 基于 CSI 的 UE-AP 匹配算法

1. **输入:** 信道增益矩阵 \mathbf{G} , AP 集合用 $\mathcal{M} = \{1, \dots, M\}$ 表示; UE 集合用 $\mathcal{K} = \{1, \dots, K\}$ 表示, 阈值 $\varepsilon = 10^{-3}$;
2. **输出:** UE-AP 匹配矩阵 \mathbf{D} ;
3. for $k = 1$ to K do
4. 获得 $\mathbf{G}(:, k)$ 的最大值, 用户 k 选择信道增益最高的 AP 作为主 AP
5. for $m = 1$ to M do
6. 如果 $|\mathbf{G}(m, k) - \mathbf{G}(\text{主AP}, k)| \leq \varepsilon$, 退出, 并令 $\mathbf{D}(m, k) = 1$
7. 结束
8. 结束
9. 结束

3.2 用户子消息功率分配

对于给定的 UE-AP 匹配矩阵 \mathbf{D} 、RIS 反射相移矩阵 Φ 和子消息解码顺序 $\boldsymbol{\zeta}_m$, 和速率最大化问题 P1 可以改写为 P2, 即

其中, \mathcal{Q}_m^{kj+} 表示解码顺序大于等于 s_{kj} 的所有子消息的集合, 即 $\mathcal{Q}_m^{kj+} = \{(u, v): \zeta_{uv} \geq \zeta_{kj}\}$, C2 可以改写为

$$\begin{aligned} & |\mathbf{g}_{k,R}^n \Phi \mathbf{H}_{R,m} + H_{k,m}|^2 P_{kj} - (2^{(h_{kj} R_k^{\min})} - 1) \cdot \\ & \left(\sum_{(u,v) \in \mathcal{Q}_m^{kj}} |\mathbf{g}_{u,R}^n \Phi \mathbf{H}_{R,m} + H_{u,m}|^2 P_{uv} + \sigma^2 \right) \geq 0 \quad (9) \\ & \forall k \in \mathcal{K}, \forall j \in \mathcal{J} \end{aligned}$$

利用公平比例约束 C3 以及引入松弛变量 $\{R_k\}$, 优化问题 P2 可以简化为

$$\begin{aligned} \text{P3: } \max_{P,R} & \sum_{m \in \mathcal{M}} \sum_{k \in \mathcal{M}_m} d_{m,k} R_k \\ \text{s. t. C1: } & \sum_{j \in \mathcal{J}} P_{kj} \leq P_k^{\max}, \quad \forall k \in \mathcal{K} \\ \text{C2: } & |\mathbf{g}_{k,R}^n \Phi \mathbf{H}_{R,m} + H_{k,m}|^2 P_{kj} - (2^{(h_{kj} R_k^{\min})} - 1) \cdot \\ & \left(\sum_{(u,v) \in \mathcal{Q}_m^{kj}} |\mathbf{g}_{u,R}^n \Phi \mathbf{H}_{R,m} + H_{u,m}|^2 P_{uv} + \sigma^2 \right) \geq 0, \\ & \forall k \in \mathcal{K}, \forall j \in \mathcal{J} \end{aligned}$$

$$\begin{aligned} \text{C3: } \log_2 & \left(\frac{\sum_{(u,v) \in \mathcal{Q}_m^{kj+}} |\mathbf{g}_{u,R}^n \Phi \mathbf{H}_{R,m} + H_{u,m}|^2 P_{uv} + \sigma^2}{\sum_{(u,v) \in \mathcal{Q}_m^{kj}} |\mathbf{g}_{u,R}^n \Phi \mathbf{H}_{R,m} + H_{u,m}|^2 P_{uv} + \sigma^2} \right) \geq \\ & h_{kj} R_k, \quad \forall k \in \mathcal{K}, \forall j \in \mathcal{J} \quad (10) \end{aligned}$$

约束 C3 仍然是非凸的, 可以进一步将约束 C3 左侧写成凸差形式, 即

$$\begin{aligned} & \log_2 \left(\sum_{(u,v) \in \mathcal{Q}_m^{kj+}} |\mathbf{g}_{u,R}^n \Phi \mathbf{H}_{R,m} + H_{u,m}|^2 P_{uv} + \sigma^2 \right) - \\ & \log_2 \left(\sum_{(u,v) \in \mathcal{Q}_m^{kj}} |\mathbf{g}_{u,R}^n \Phi \mathbf{H}_{R,m} + H_{u,m}|^2 P_{uv} + \sigma^2 \right) \quad (11) \\ \text{令 } A_{kj} & = \log_2 \left(\sum_{(u,v) \in \mathcal{Q}_m^{kj}} |\mathbf{g}_{u,R}^n \Phi \mathbf{H}_{R,m} + H_{u,m}|^2 P_{uv} + \sigma^2 \right), \end{aligned}$$

由于 A_{kj} 的非凸性, 约束 C3 仍是非凸的, 因此, 采用 SCA 方法将约束 C3 的非凸性转化为凸性。SCA 算法原理主要是利用前一次迭代的值, 将非凸函数逼近为等价的凸函数形式。令 $P_{kj}^{(a-1)}$ 表示第 $(a-1)$ 次迭代得到的功率分配值, 则在第 a 次迭代时, A_{kj} 可以近似为

$$\begin{aligned} A_{kj} & \leq \tilde{A}_{kj}^{(a)} \triangleq A_{kj}^{(a-1)} + \\ & \frac{|\mathbf{g}_{k,R}^n \Phi \mathbf{H}_{R,m} + H_{k,m}|^2 (P_{kj} - P_{kj}^{(a-1)})}{\sum_{(u,v) \in \mathcal{Q}_m^{kj}} |\mathbf{g}_{u,R}^n \Phi \mathbf{H}_{R,m} + H_{u,m}|^2 P_{uv}^{(a-1)} + \sigma^2} \quad (12) \end{aligned}$$

因此约束 C3 可以近似写成

$$\begin{aligned} & \log_2 \left(\sum_{(u,v) \in \mathcal{Q}_m^{kj+}} |\mathbf{g}_{u,R}^n \Phi \mathbf{H}_{R,m} + H_{u,m}|^2 P_{uv} + \sigma^2 \right) - \\ & \tilde{A}_{kj}^{(a)} \geq h_{kj} R_k, \quad \forall k \in \mathcal{K}, \forall j \in \mathcal{J} \quad (13) \end{aligned}$$

这样, 优化问题 P3 可以近似转换为优化问题 P4, 即

$$\begin{aligned} \text{P4: } \max_{P,R} & \sum_{m \in \mathcal{M}} \sum_{k \in \mathcal{M}_m} d_{m,k} R_k \\ \text{s. t. C1: } & \sum_{j \in \mathcal{J}} P_{kj} \leq P_k^{\max}, \quad \forall k \in \mathcal{K} \\ \text{C2: } & |\mathbf{g}_{k,R}^n \Phi \mathbf{H}_{R,m} + H_{k,m}|^2 P_{kj} - (2^{(h_{kj} R_k^{\min})} - 1) \cdot \\ & \left(\sum_{(u,v) \in \mathcal{Q}_m^{kj}} |\mathbf{g}_{u,R}^n \Phi \mathbf{H}_{R,m} + H_{u,m}|^2 P_{uv} + \sigma^2 \right) \geq 0, \\ & \forall k \in \mathcal{K}, \forall j \in \mathcal{J} \\ \text{C3: } \log_2 & \left(\sum_{(u,v) \in \mathcal{Q}_m^{kj+}} |\mathbf{g}_{u,R}^n \Phi \mathbf{H}_{R,m} + H_{u,m}|^2 P_{uv} + \sigma^2 \right) - \\ & \tilde{A}_{kj}^{(a)} \geq h_{kj} R_k, \quad \forall k \in \mathcal{K}, \forall j \in \mathcal{J} \quad (14) \end{aligned}$$

优化问题 P4 中目标函数是仿射的, 约束 C1 和 C2 是线性的, 约束 C3 的左侧是凹仿射函数的对数, 因此约束 C3 也是凸约束, 所以优化问题 P4 是关于变量 P 和 R 的凸优化问题, 可以使用标准的 CVX 工具箱进行求解。给定用户子消息的初始功率分配值 $\{P_{kj}^{(0)}\}$, 迭代求解优化问题 P4, 直到迭代次数达最大的迭代次数 A_{\max} 或求和速率 $\sum_{m \in \mathcal{M}} \sum_{k \in \mathcal{M}_m} d_{m,k} R_k$ 收敛, 算法 2 给出了求解的整个过程。

算法 2 功率分配的连续凸近似算法

1. **输入:** 给定 RIS 相移矩阵 Φ , 各 AP 处的解码顺序 ζ , 最大迭代次数 A_{\max} , 阈值 $\varepsilon = 10^{-3}$;

2. **初始化:** $\zeta = \zeta_0, \Phi = \Phi_0, R^{(-1)} = 0, a = 0, P_{kj}^{(0)} = \frac{P_k^{\max}}{J}, \forall k \in \mathcal{K}, \forall j \in \mathcal{J}$;

3. **while** $\frac{R^{(a)} - R^{(a-1)}}{R^{(a-1)}} \geq \varepsilon$ 或者 $a \leq A_{\max}$ **do**

4. 求解优化问题 P4 并计算系统和速率 $R^{(a)} = \sum_{m \in \mathcal{M}} \sum_{k \in \mathcal{M}_m} d_{m,k} R_k$;

5. $a = a + 1$;

6. **end while**;

7. **输出:** $\hat{R} = R^{(a)}, \hat{P} = P^{(a)}$;

3.3 RIS 被动波束形成

在给定的 UE-AP 匹配矩阵、用户功率分配以及解码顺序下, RIS 无源波束形成问题可以表述为

$$\begin{aligned} \text{P5: } \max_{\theta} & \sum_{m \in \mathcal{M}} \sum_{k \in \mathcal{M}_m} \sum_{j \in \mathcal{J}} d_{m,k} \\ & \log_2 \left(1 + \frac{|\mathbf{g}_{k,R}^n \text{diag}(\theta) \mathbf{H}_{R,m} + H_{k,m}|^2 P_{kj}}{\sum_{(u,v) \in \mathcal{Q}_m^{kj}} |\mathbf{g}_{u,R}^n \text{diag}(\theta) \mathbf{H}_{R,m} + H_{u,m}|^2 P_{uv} + \sigma^2} \right) \\ \text{s. t. C1: } & |\Phi_{i,i}| = 1, \quad \forall i \in \mathcal{N} \triangleq \{1, \dots, N\} \\ \text{C2: } & R_m^k \geq R_k^{\min} \\ \text{C3: } & R_{k_1} : R_{k_2} : \dots : R_{k_j} = \\ & h_{k_1} : h_{k_2} : \dots : h_{k_j}, \quad \forall k \in \mathcal{K}, j \in \mathcal{J} \quad (15) \end{aligned}$$

其中, $\boldsymbol{\theta} = [\theta_1, \dots, \theta_N]^T = \text{diag}(\boldsymbol{\Phi})$ 。然而, 由于目标函数的非凸性和 RIS 被动波束的单位模量约束 C1, 优化问题 P5 是非凸的。具有单位模约束的非凸优化问题的全局最优解通常是难以解决的, 为了有效地求解该优化问题, 提出了一种基于黎曼流形共轭梯度法获得局部最优解, 约束 C1 可以看作光滑的黎曼流形 V , 即

$$V = \{\boldsymbol{\theta} \in \mathbb{C}^{N \times 1}; |\theta_i| = 1, \forall i \in \mathcal{N}\} \quad (16)$$

所有与流形空间 V 在点 $\boldsymbol{\theta}$ 处相切的切向量的集合 $T_{\boldsymbol{\theta}}V$ 可以表示为

$$T_{\boldsymbol{\theta}}V = \{z \in \mathbb{C}^{N \times 1}; z^H \boldsymbol{\theta} = 0\} \quad (17)$$

特别地, 流形空间是一个拓扑空间, 类似于每个点上的欧几里得空间, 因此黎曼流形上的优化类似于欧几里得空间中的优化。为此, 提出了一种基于黎曼流形共轭梯度的 RIS 被动波束形成迭代算法进行局部求解问题 P5。首先从目标函数中确定欧几里得梯度, 再确定黎曼梯度。优化问题 P5 中的目标函数为

$$f_c(\boldsymbol{\theta}) = \sum_{m \in \mathcal{M}_k} \sum_{k \in \mathcal{M}_m} \sum_{j \in \mathcal{J}} d_{m,k} \log_2 \left[1 + \frac{|\mathbf{g}_{k,R}^H \text{diag}(\boldsymbol{\theta}) \mathbf{H}_{R,m} + H_{k,m}|^2 P_{kj}}{\sum_{(u,v) \in \mathcal{Q}_m^b} |\mathbf{g}_{u,R}^H \text{diag}(\boldsymbol{\theta}) \mathbf{H}_{R,m} + H_{u,m}|^2 P_{uv} + \sigma^2} \right] \quad (18)$$

其矢量的欧几里得梯度可表示为

$$\Delta f_c(\boldsymbol{\theta}) = \frac{\partial f_c}{\partial \boldsymbol{\theta}^*} = \sum_{m \in \mathcal{M}_k} \sum_{k \in \mathcal{M}_m} \sum_{j \in \mathcal{J}} d_{m,k} \left[\frac{\sum_{(u,v) \in \mathcal{Q}_m^b} P_{uv} (\mathbf{g}_{u,R} \odot \mathbf{H}_{R,m}) \odot (\mathbf{g}_{u,R} \odot \mathbf{H}_{R,m})^* \odot \boldsymbol{\theta}}{\sum_{(u,v) \in \mathcal{Q}_m^b} |\mathbf{g}_{u,R}^H \boldsymbol{\Phi} \mathbf{H}_{R,m} + H_{u,m}|^2 P_{uv} + \sigma^2} - \frac{\sum_{(u,v) \in \mathcal{Q}_m^b} P_{uv} (\mathbf{g}_{u,R} \odot \mathbf{H}_{R,m}) \odot (\mathbf{g}_{u,R} \odot \mathbf{H}_{R,m})^* \odot \boldsymbol{\theta}}{\sum_{(u,v) \in \mathcal{Q}_m^b} |\mathbf{g}_{u,R}^H \boldsymbol{\Phi} \mathbf{H}_{R,m} + H_{u,m}|^2 P_{uv} + \sigma^2} \right] \quad (19)$$

函数 $f_c(\boldsymbol{\theta})$ 在 $\boldsymbol{\theta}$ 点处的黎曼梯度是基于欧几里得梯度 $\Delta f_c(\boldsymbol{\theta})$ 在切空间 $T_{\boldsymbol{\theta}}V$ 上正交投影计算的, 即

$$\mathbf{g}_{\boldsymbol{\theta}} = P_{\boldsymbol{\theta}}(\Delta f_c(\boldsymbol{\theta})) = \Delta f_c(\boldsymbol{\theta}) - \text{Re}[\Delta f_c(\boldsymbol{\theta}) \odot \boldsymbol{\theta}^*] \odot \boldsymbol{\theta} \quad (20)$$

其中, $P_{\boldsymbol{\theta}}(z)$ 为变量 z 在切空间上的正交投影。

共轭梯度法迭代求出流形上的局部最优点。注意到在给定的搜索方向下, 缩回有助于限制流形空间 V 上的搜索迭代次数, 即

$$\text{Ret}_{\boldsymbol{\theta}}(\alpha \mathbf{d}) = \left[\frac{\theta_1 + \alpha d_1}{|\theta_1 + \alpha d_1|}, \dots, \frac{\theta_N + \alpha d_N}{|\theta_N + \alpha d_N|} \right]^T \quad (21)$$

其中, α 表示搜索的步长, $\mathbf{d} = [d_1, \dots, d_N]^T$ 表示搜索方向。步长 α 可以通过 Armijo 回溯搜索算法来获得, 第 $(b+1)$ 次迭代共轭梯度搜索方向的更新规则可以类似于在欧几里得空间的情况下获得, 即

$$\mathbf{d}^{(b+1)} = \mathbf{g}^{(b+1)} + \lambda^{(b+1)} P_{\boldsymbol{\theta}^{(b+1)}}(\mathbf{d}^{(b)}) = \mathbf{g}^{(b+1)} + \lambda^{(b+1)} (\mathbf{d}^{(b)} - \text{Re}[\mathbf{d}^{(b)} \odot \boldsymbol{\theta}^{(b+1)*}] \odot \boldsymbol{\theta}^{(b+1)}) \quad (22)$$

其中, $\mathbf{g}^{(b+1)}$ 表示黎曼梯度, $\lambda^{(b+1)}$ 为 Polak-Ribiere 参数, 可以由式(23)计算得到。

$$\lambda^{(b+1)} = \mathbf{g}^{(b+1)H} \left(\frac{\mathbf{g}^{(b+1)} - P_{\boldsymbol{\theta}^{(b+1)}}(\mathbf{g}^{(b)})}{\|P_{\boldsymbol{\theta}^{(b+1)}}(\mathbf{g}^{(b)})\|_F^2} \right) \quad (23)$$

RIS 相移的更新点可以通过收缩运算获得。基于黎曼流形梯度法的 RIS 无源波束形成的求解过程如算法 3 所示。

算法 3 RIS 相移的黎曼共轭梯度算法

1. 输入: 满足约束 C1 的任意 $\boldsymbol{\theta}^{(0)}$;
2. 初始化: $\mathbf{d}^{(0)} = P_{\boldsymbol{\theta}^{(0)}}(\Delta f_c(\boldsymbol{\theta})|_{\boldsymbol{\theta}=\boldsymbol{\theta}^{(0)}})$, $f_c(\boldsymbol{\theta}^{(-1)}) = 0$, $b = 0$, 阈值 $\varepsilon = 10^{-3}$;
3. while $\frac{f_c(\boldsymbol{\theta}^{(b)}) - f_c(\boldsymbol{\theta}^{(b-1)})}{f_c(\boldsymbol{\theta}^{(b-1)})} \geq \varepsilon$ do;
4. 选择 Armijo 回溯线搜索步长 $\alpha^{(b)}$;
5. 更新 RIS 相移: $\boldsymbol{\theta}^{(b+1)} = \text{Ret}_{\boldsymbol{\theta}^{(b)}}(\alpha \mathbf{d}) = \left[\frac{\theta_1 + \alpha d_1}{|\theta_1 + \alpha d_1|}, \dots, \frac{\theta_N + \alpha d_N}{|\theta_N + \alpha d_N|} \right]^T$;
6. 计算黎曼梯度 $\mathbf{g}^{(b+1)} = P_{\boldsymbol{\theta}^{(b+1)}}(\Delta f_c(\boldsymbol{\theta})|_{\boldsymbol{\theta}=\boldsymbol{\theta}^{(b+1)}}) = \Delta f_c(\boldsymbol{\theta})|_{\boldsymbol{\theta}=\boldsymbol{\theta}^{(b+1)}} - \text{Re}[\Delta f_c(\boldsymbol{\theta})|_{\boldsymbol{\theta}=\boldsymbol{\theta}^{(b+1)}} \odot \boldsymbol{\theta}^{(b+1)*}] \odot \boldsymbol{\theta}^{(b+1)}$;
7. 计算 Polak-Ribiere 参数 $\lambda^{(b+1)}$ 如式(22)所示;
8. 根据式(21)确定共轭方向;
9. $b = b + 1$;
10. end while;
11. 输出: $\hat{\boldsymbol{\theta}} = \boldsymbol{\theta}^{(b+1)}$ 。

3.4 解码顺序问题求解

在求解了 UE-AP 匹配问题并获得了匹配矩阵 \mathbf{D} 后, 每个 AP 匹配到的用户集合可以用 $\mathcal{M} = \{\mathcal{M}_1, \dots, \mathcal{M}_M\}$ 表示, 因此需要求解每个 AP 上的最优解码顺序。选择最优的解码顺序可以有效地消除其他用户的子报文产生的干扰, 进而提高系统的整体吞吐量。给定 UE-AP 匹配矩阵 \mathbf{D} 、用户子消息功率分配 \mathbf{P} 以及 RIS 被动波束 $\boldsymbol{\Phi}$, 优化问题 P1 可以写成

$$\begin{aligned}
\text{P6: } \max_{\zeta} & \sum_{m \in \mathcal{M}} \sum_{k \in \mathcal{M}_{-j}} \sum_{\mathcal{J}} d_{m,k} \\
& \log_2 \left(\frac{\sum_{(u,v) \in \mathcal{Q}_m^{j+}} |g_{u,R}^n \Phi H_{R,m} + H_{u,m}|^2 P_{uv} + \sigma^2}{\sum_{(u,v) \in \mathcal{Q}_m^j} |g_{u,R}^n \Phi H_{R,m} + H_{u,m}|^2 P_{uv} + \sigma^2} \right) \\
\text{s. t. } & \text{C1: } \zeta_m \in \Omega_m, \forall m \in \mathcal{M} \\
& \text{C2: } R_m^k \geq R_k^{\min} \\
& \text{C3: } R_{k_1}: R_{k_2}: \dots: R_{k_J} = h_{k_1}: h_{k_2}: \dots: h_{k_J}, \quad \forall k \in \mathcal{K}
\end{aligned} \quad (24)$$

优化问题 P6 由于非凸目标函数和离散变量的强耦合而难以求解。最优的解码顺序选择可以看作是难以处理的混合整数非线性规划问题,目前几乎还没有求解这类问题的标准方法,目前,求解用户的子报文的解码顺序的算法大都是根据用户的信道增益进行的。

可以使用基于穷举搜索算法来获取 UL-RSMA 的最优解码顺序,穷举搜索算法是最优的,但计算复杂度会随着子消息数量呈指数型增长,因此复杂度是极高的,所以在实际系统中并不会优先考虑。本文提出了一种基于 UE-AP 信道增益以及各个用户子消息分割比例的解码顺序,并在仿真部分证明了该算法的可行性以及优越性。

假设 s_i 和 s_j 分别为信道增益为 g_i 和 g_j 以及功率分配为 P_i 和 P_j 相邻的解码顺序上的两个子消息。假设解码顺序小于 s_i 和 s_j 的子消息的干扰上限为 τ , 即

$$|g_i|^2 P_i + |g_j|^2 P_j + n + \sigma^2 \leq \tau \quad (25)$$

其中, n 表示解码顺序高于 s_i 和 s_j 的其他子消息产生的干扰。现在有两种方案,即 (a) s_i 先解码; (b) s_j 先解码。对比两种方案,对于 (a) 方案需满足

$$\frac{|g_i|^2 P_i}{|g_j|^2 P_j + n + \sigma^2} \geq \gamma_i^{\min}, \quad \frac{|g_j|^2 P_j}{n + \sigma^2} \geq \gamma_j^{\min} \quad (26)$$

其中需满足 $0 < P_i, P_j \leq P_{\max}$, 根据式 (25), 功率下界可以表示为

$$\begin{aligned}
P_i & \geq \frac{\gamma_i^{\min} (\gamma_j^{\min} + 1) (n + \sigma^2)}{|g_i|^2}, \\
P_j & \geq \frac{\gamma_j^{\min} (n + \sigma^2)}{|g_j|^2}
\end{aligned} \quad (27)$$

相应地, 功率上界可以确定为

$$P_i \leq P_{\max}, P_j \leq \min \left\{ P_{\max}, \frac{|g_i|^2 P_{\max} - n - \sigma^2}{\gamma_i^{\min} |g_j|^2} \right\} \quad (28)$$

从式 (25~28) 可以得到, 两个子消息 s_i 和 s_j 的最大信号增强 $|g_i|^2 P_i + |g_j|^2 P_j$ 可以表示为

$$\begin{aligned}
\varepsilon & \leq \varepsilon_i \triangleq \min \left\{ |g_i|^2 P_{\max} \left(1 + \frac{1}{\gamma_i^{\min}} \right), \right. \\
& \left. (|g_i|^2 + |g_j|^2) P_{\max}, \tau - n - \sigma^2 \right\}
\end{aligned} \quad (29)$$

对于 (b) 方案: s_j 先解码时, 与前面的情形类似, 两个子消息 s_i 和 s_j 的最大信号增强 $|g_i|^2 P_i + |g_j|^2 P_j$ 可以表示为

$$\begin{aligned}
\varepsilon & \leq \varepsilon_j \triangleq \min \left\{ |g_j|^2 P_{\max} \left(1 + \frac{1}{\gamma_j^{\min}} \right), \right. \\
& \left. (|g_i|^2 + |g_j|^2) P_{\max}, \tau - n - \sigma^2 \right\}
\end{aligned} \quad (30)$$

通过比较式 (29) 和式 (30) 可知, 如果 $|g_i|^2 \left(1 + \frac{1}{\gamma_i^{\min}} \right) > |g_j|^2 \left(1 + \frac{1}{\gamma_j^{\min}} \right)$, 则有 $\varepsilon_i > \varepsilon_j$, 此时方案 (a) 的解码顺序可以提供最大信号增益; 方案 (b) 同理。这证明了解码顺序的最优策略是先解码信道增益高的子消息。综上所述, 我们可以得出上行 RSMA 的最优解码顺序为递减顺序: $V_{k_j} = |g_j|^2 \left(1 + \frac{1}{\gamma_j^{\min}} \right) = |g_j|^2 P_{\max} \left(1 + \frac{1}{2^{h_j R_k^{\min}}} \right)$, 因此设计了一种基于 UE-AP 信道增益 $|g_j|^2$ 以及各个用户子消息分割比例 h_{k_j} 的解码顺序。

3.5 优化问题整体求解

综合上述算法, 提出了一种基于 AO 算法的联合求解方案, 交替优化 UE-AP 的匹配矩阵、用户的最佳功率分配和 RIS 无源波束形成, 其中最佳的 UE-AP 匹配矩阵根据算法 1 获得, 用户功率分配子问题和 RIS 无源波束形成子问题采用最优解码顺序算法考虑的译码顺序交替迭代求解, 使用算法 3 获取 RIS 的无源波束形成, 接着将获得的 RIS 无源波束矩阵用于求解算法 2 中的用户功率分配, 在下次迭代过程中, 利用上一次迭代过程获得的值进行计算, 直到达到收敛, 详细的过程如算法 4 所示。

算法 4 基于交替优化算法的整体求解

1. 输入: 初始化 D_0, P_0, Φ_0 , 最大迭代次数 d^{\max} ;
2. 初始化: $P^{(0)} = P_0, R^{(-1)} = 0$, 阈值 $\varepsilon = 10^{-3}$;
3. while $\frac{R^{(c)} - R^{(c-1)}}{R^{(c-1)}} \geq \varepsilon$ 或者 $c \leq d^{\max}$ do;
4. 根据算法 1 设置 UE-AP 匹配矩阵 $D^{(c+1)}$;
5. 根据最优解码顺序算法获得 ζ ;
6. 给定功率分配 $P^{(c)}$, 执行算法 3 获得 $\Phi^{(c+1)}$;
7. 利用 $\Phi^{(c+1)}$, 执行算法 2 获得系统和速率 $R^{(c+1)}$ 和用户功率分配 $P^{(c+1)}$;
8. $c = c + 1$;
9. end while;
10. 输出: $\hat{R} = R^{(c)}$;

在基于CSI的UE-AP配对算法中,两层for循环包含 MK 个变量。基于SCA的功率分配问题P2中存在 $2KJ$ 个变量以及 $3KJ + K$ 个约束条件,考虑到最坏情况,让SCA算法收敛到最大迭代次数 A^{\max} ,因此,基于SCA的用户功率分配算法复杂度为 $O(A^{\max}(KJ + K)^2(3KJ + K))$,基于黎曼流形梯度的RIS无源波束设计的最坏复杂度为 $O((2N)^{1.5})^{[20]}$,假设统一解在最大次 d^{\max} 迭代中收敛,则该算法的总体复杂度为 $O(d^{\max}(MK + A^{\max}(KJ + K)^2(3KJ + K) + (2N)^{1.5}))$ 。因此,算法的复杂度主要取决于UE、AP个数以及子消息的数量,RIS元素的数量一般比较少。最优解码顺序仅在 J 个顺序中选择,因此最优解码顺序选择不会显著改变算法的计算复杂度。

4 仿真分析

在本节通过蒙特卡罗仿真,验证所提出的基于RIS辅助的RSMA无蜂窝上行通信系统以及所提出的求解方案具有更好的系统性能,仿真结果是经过500次蒙特卡罗模拟得到的平均值。考虑一个 500×500 平方米区域作为研究场景,该区域内随机分布6个用户,15个配有单天线的AP为用户提供服务,用户高度设为1.65 m,AP高度设为15 m,每个用户的总发射功率为 $P^{\max} = 500$ mW。RIS预先部署在20 m高的建筑物上,假设该建筑物位于原点,即RIS坐标为 $[0, 0, 20]$ m,且RIS包含 $N = 20$ 个反射单元。路径损耗模型表示为 $L(d) = L_0(dd_0)^{-\gamma}$,式中 d 表示通信设备之间的距离, $L_0 = -10$ dB表示参考距离 $d_0 = 1$ m处的路径损耗, $\gamma = 2.2$ 表示路径损耗指数。为了考虑小尺度衰落,假设所有信道都是均匀分布的,包含LOS和NLOS分量。因此,信道模型表示为

$$G = \sqrt{K_f/(K_f + 1)}G_L + \sqrt{1/(K_f + 1)}G_N \quad (31)$$

其中, $k_f = 10$ 表示Rician因子, $G_L = 1$ 表示LOS分量, G_N 表示参数 $L = 1$ 服从Rayleigh分布的NLOS分量,其中 G 中的每个元素与距离相关路径损失模型的平方根相乘。噪声功率 σ^2 设置为 -100 dBm,每个用户的信号被分成两个子消息,即 $J = 2$,最大迭代次数 $d^{\max} = 10$,子报文优先级索引设置为 $h_{kj} = 0.5, \forall k \in \mathcal{K}, \forall j \in \mathcal{J}$ 。RIS初始相移为随机值,初始发射功率设置为 $P_{kj} = P^{\max}/J, \forall k \in \mathcal{K}, \forall j \in \mathcal{J}$,用户的最低速率需求设置为 $r_k^{\min} = 0.5$ bit/s/Hz, $\forall k \in \mathcal{K}$ 。假设同一个AP服务下的UE在同一时间/频带下工作,带宽为1 Hz。提出的算法3获得解的方案称为RIS-RSMA,为了公平比对性能基准,分别用算法3

类似的方案求解基于RIS辅助的SDMA无蜂窝上行通信系统和基于RIS辅助的NOMA无蜂窝上行通信系统的等效和速率最大化问题,分别为RIS-SDMA和RIS-NOMA。

如图2和3所示,对比分析了基于RIS辅助的RSMA、NOMA和SDMA系统中的SCA的功率分配(SCA-RIS-RSMA)算法和黎曼流形梯度法的RIS无源波束设计(RCG-RIS-RSMA)算法在各自相对应系统中的收敛行为,两种算法的最大迭代次数分别用 A_{SCA}^{\max} 和 A_{RCG}^{\max} 表示。从图中可以看出,算法2基本在5~6次迭代内收敛,算法3在15~20次迭代内收敛,这表明提出的基于SCA的功率分配和基于黎曼流形梯度算法的RIS无源波束设计的快速收敛性。首先,为了分析固定RIS相移下功率分配算法的优越性,将本文所设计的算法与穷举法进行了对比,搜索步长设定为 $\delta_p = 5$ mW,此时穷举法需要进行 $\left(\frac{P^{\max}}{\delta_p}\right)^{KJ}$ 次搜索,从图2中可以看出,具有较小步长的穷举法所带来的性能确实会好一点,但随着用户数和子消息数的增长,其带来的复杂度会呈指数型增长。其次我们分析固定功率分配下的RIS相移设

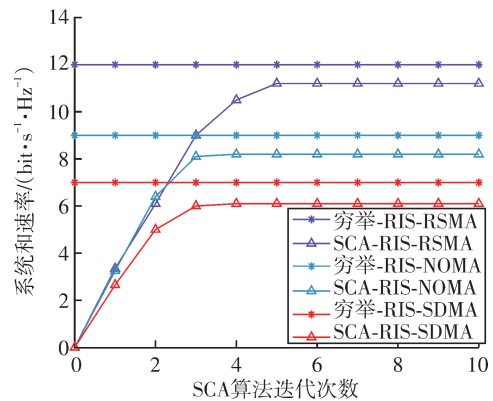


图2 基于SCA的子消息功率分配迭代图

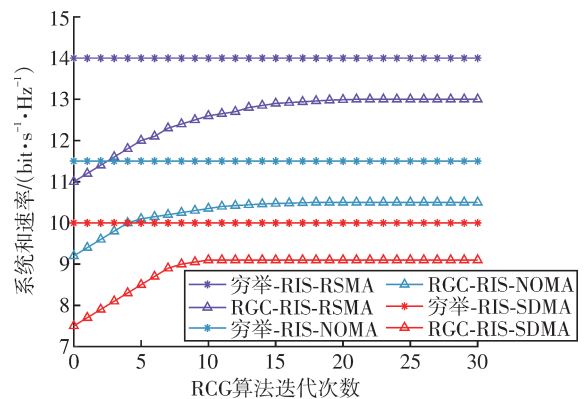


图3 基于黎曼流形梯度法的RIS相移迭代图

计,同样与穷举法进行对比,搜索步长设定为 $\delta_\theta = \frac{\pi}{180}$,此时穷举法需要进行 $\left(\frac{2\pi}{\delta_\theta}\right)^N$ 次搜索,从图3中可以看出,具有较小步长的穷举法所带来的性能确实会好一点,但随着RIS反射单元的增大,其带来的复杂度会呈指数型增长。从图中还能看出,基于RIS辅助的RSMA迭代收敛次数高于NOMA和SDMA,这是因为RSMA机制会将用户消息进一步拆分为子消息,这样会导致接收端解码更多的消息数,所以复杂度会提高,但与此同时,其可以提高系统的鲁棒性,进而提升系统性能。还检查了所提出的基于交替迭代框架的统一求解方案的收敛性,如图4所示,其中作为对比的穷举方案设置搜索步长分别为 $\delta_p = 1 \text{ mW}$ 、 $\delta_\theta = \frac{\pi}{180}$,尽管所提出的联合求解方案得到的性能略低于穷举法,但具有计算复杂度较低的优点,且在3~4次迭代内实现了收敛。

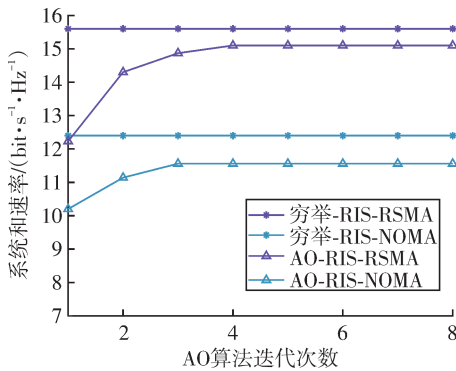


图4 基于交替迭代算法的整体求解迭代图

图5展示了所提出的RIS-RSMA方案随着用户最大发射功率 P^{\max} 改变的性能变化。显而易见,增大用户的最大传输功率会提升系统的速率吞吐量,但随着用户发射功率的继续增大,速率会趋于缓慢上升并达到饱和状态,这是因为功率达到一定量后,此时在信道容量受限的情况下,即使发射功率再增大,也无法再显著提升用户速率。在RSMA中,用户消息被分为多个子消息,用户可以更好地利用功率资源,AP解码子消息的基数变大了,系统的整体鲁棒性变得更好,因此RSMA系统比等效的NOMA和SDMA系统具有更好的性能。还加入了采用随机解码方案(RIS-RSMA-RD)的性能曲线作为对比,相较本文提出的算法,非理想的SIC会导致系统和速率下降,但依然优于NOMA。与此同时,将不同多址技术下有无RIS辅助的无蜂窝系统性能做了

对比,仿真曲线可以看出,RIS辅助的系统具有更大的速率吞吐量,因为RIS可以改善信道传输质量。

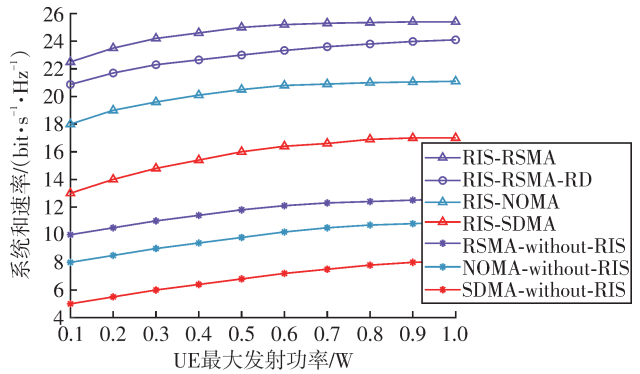


图5 和速率与UE最大发射功率 P^{\max} 关系图

图6展示了提出的RIS系统在不同的RIS反射元件数下的性能表现。RIS的加入使得用户和AP之间的信道增益得到了提升,因此系统和速率得到了提升,RIS相移器的增多可以将用户发射的更多信号功率反射给AP,进而获得更大的功率增益。除此之外,RIS提供了更高的资源分配灵活性和更强的波束形成增益,从而提高了通信吞吐量。还比较了随机相移的结果,最优无源波束形成比随机相移方案提供了更大的速率吞吐量,提高了近10%~20%。

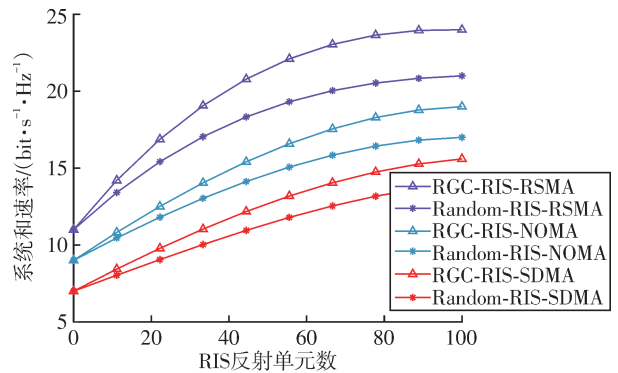


图6 和速率与RIS发射单元数的关系图

图7展示了所提出的RIS-RSMA系统和速率与用户子消息数的性能关系。NOMA可以被认为是RSMA的特殊情况,其中RSMA会将每个用户的功率按一定的比例分配给多个子消息,从而达到最大速率区域,进而实现更大的频谱效率,而NOMA只将用户功率分配给完整的消息。增加每个用户的子消息数会产生更大的有效增量,因此可以实现更大的系统和速率。结合表1和图8可以看出,当 $h_1 > h_2$ 时, ζ^4 方案更佳;当 $h_1 < h_2$ 时, ζ^2 方案更佳,这表明基于用户信道增益降序和子消息比例升序的解码

顺序的频谱效率更接近基于穷举搜索得到的最优解,即图 8 中的 ζ^5 。与此同时,当 $h_1 = h_2$ 时, ζ^1 和 ζ^2 获得相等的性能时,等比例分割子消息比不同比例分割子消息可以获得更好的性能。因此,可以得出结论,等比例分割消息可以提高 RSMA 上行通信系统的整体性能。

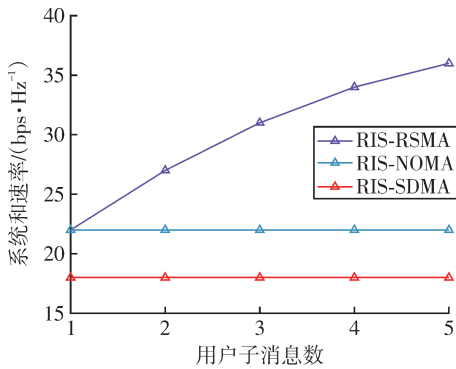


图 7 和速率与用户子消息数关系图

表 1 四种解码方案说明

解码方案	信道增益	比例指数	描述
ζ^1	上升	上升	$\zeta_{11} < \zeta_{12} < \dots < \zeta_{1K} < \dots < \zeta_{K1} < \zeta_{K2} < \dots < \zeta_{KK}$
ζ^2	上升	下降	$\zeta_{K1} < \zeta_{K2} < \dots < \zeta_{KK} < \dots < \zeta_{11} < \zeta_{12} < \dots < \zeta_{1K}$
ζ^3	下降	上升	$\zeta_{1K} < \dots < \zeta_{12} < \zeta_{11} < \dots < \zeta_{KK} < \dots < \zeta_{K2} < \zeta_{K1}$
ζ^4	下降	下降	$\zeta_{KK} < \dots < \zeta_{K2} < \zeta_{K1} < \dots < \zeta_{1K} < \dots < \zeta_{12} < \zeta_{11}$

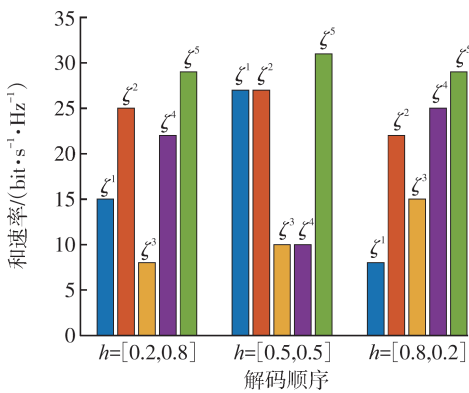


图 8 和速率与子消息解码顺序关系图

图 9 显示了非完美 CSI 条件下 RSMA 和 NOMA 的系统性能对比,揭示了不同噪声功率对系统和速率的影响。随着噪声功率的增加, RIS-RSMA 的性能下降较 RIS-NOMA 更快,主要是由于信道估计误差放大了子消息功率分配和干扰处理的不确定性。在高噪声环境下,信噪比的降低使得干扰的影响相

较噪声变得次要, RSMA 通过子消息分割和部分干扰解码的优势被削弱,系统的性能更接近传统的完全解码干扰的机制,即 NOMA。因此, RSMA 在高噪声条件下表现出与 NOMA 接近的和速率,但仍然略优于后者。

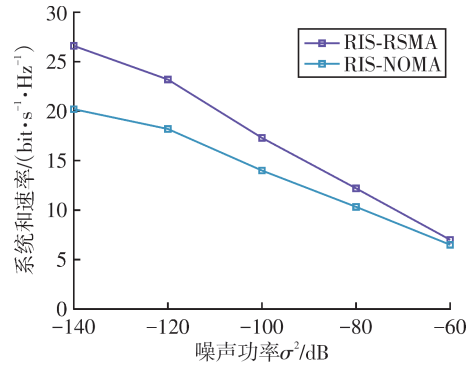


图 9 非完美 CSI 条件下性能对比图

5 结束语

为了提高上行通信系统的性能,研究了基于 RIS 辅助的 RSMA 无蜂窝上行通信系统和速率最大化问题。具体地,为了最大化系统和速率,建立了一个关于用户与 AP 匹配选择、子消息解码顺序、RIS 无源波束形成以及各 UL 用户的子消息功率分配的联合优化问题。由于所提问题的非凸性以及是离散非线性规划,原问题是一个非凸优化问题,因此需要将该问题解耦成不同的子问题。首先是基于信道增益的 UE-AP 匹配问题求解,其次采用 SCA 方法求解用户子消息功率分配问题,再者采用黎曼共轭梯度算法求解 RIS 无源波束形成问题,接着提出一种基于信道增益以及用户子消息分配比例的解码顺序优化算法,最终采用交替优化算法获得全局最优解。此外还分析了算法的复杂度,仿真结果表明,与 SDMA 和 NOMA 相比,所提出的 RSMA 系统具有更高的和速率。

参考文献:

- [1] WANG C X, YOU X H, GAO X Q, et al. On the road to 6G: visions, requirements, key technologies, and testbeds [J]. IEEE Communications Surveys & Tutorials, 2023, 25(2): 905-974.
- [2] NGO H Q, ASHIKHMIN A, YANG H, et al. Cell-free massive MIMO versus small cells [J]. IEEE Transactions on Wireless Communications, 2017, 16(3): 1834-1850.
- [3] INTERDONATO G, NGO H Q, LARSSON E G, et al. How much do downlink pilots improve cell-free massive

- MIMO? [C]//2016 IEEE Global Communications Conference (GLOBECOM). 2016: 1-7.
- [4] MAO Y J, DIZDAR O, CLERCKX B, et al. Rate-splitting multiple access: fundamentals, survey, and future research trends [J]. *IEEE Communications Surveys & Tutorials*, 2022, 24(4): 2073-2126.
- [5] MAO Y J, CLERCKX B, LI V O K. Rate-splitting multiple access for downlink communication systems: bridging, generalizing, and outperforming SDMA and NOMA [J]. *EURASIP Journal on Wireless Communications and Networking*, 2018, 2018(1): 133.
- [6] CLERCKX B, MAO Y J, SCHOBER R, et al. Rate-splitting unifying SDMA, OMA, NOMA, and multicasting in MISO broadcast channel: a simple two-user rate analysis [J]. *IEEE Wireless Communications Letters*, 2020, 9(3): 349-353.
- [7] WANG W J, LI L H, DENG G, et al. A joint multiservice transmission scheme for RSMA-aided cell-free mMIMO system [J]. *IEEE Communications Letters*, 2023, 27(2): 591-594.
- [8] WU Q Q, ZHENG B X, YOU C S, et al. Intelligent surfaces empowered wireless network: recent advances and the road to 6G [J]. *Proceedings of the IEEE*, 2024, 112(7): 724-763.
- [9] YILDIRIM I, UYRUS A, BASAR E. Modeling and analysis of reconfigurable intelligent surfaces for indoor and outdoor applications in future wireless networks [J]. *IEEE Transactions on Communications*, 2020, 69(2): 1290-1301.
- [10] 王强, 王鸿. 智能反射面辅助的下行NOMA系统和速率最大化研究 [J]. *南京邮电大学学报(自然科学版)*, 2022, 42(1): 23-29.
WANG Qiang, WANG Hong. On sum rate maximization for IRS-aided downlink NOMA systems [J]. *Journal of Nanjing University of Posts and Telecommunications (Natural Science Edition)*, 2022, 42(1): 23-29. (in Chinese)
- [11] 李志成, 王鸿. 智能反射面辅助的非正交多址系统协作多点传输方案研究 [J]. *南京邮电大学学报(自然科学版)*, 2024, 44(2): 45-52.
LI Zhicheng, WANG Hong. Cooperative transmission design for IRS-aided NOMA systems [J]. *Journal of Nanjing University of Posts and Telecommunications (Natural Science Edition)*, 2024, 44(2): 45-52. (in Chinese)
- [12] BANSAL A, SINGH K, LI C P. Analysis of hierarchical rate splitting for intelligent reflecting surfaces-aided downlink multiuser MISO communications [C] //IEEE Open Journal of the Communications Society. 2021: 785-798.
- [13] FU H, FENG S L, KWAN NG D W. Resource allocation design for IRS-aided downlink MU-MISO RSMA systems [C] //IEEE International Conference on Communications Workshops (ICC Workshops). 2021: 1-6.
- [14] LI H Y, MAO Y J, DIZDAR O, et al. Rate-splitting multiple access for 6G: part III: interplay with reconfigurable intelligent surfaces [J]. *IEEE Communications Letters*, 2022, 26(10): 2242-2246.
- [15] 吴雪雯, 马竞宵, 薛小平, 等. 基于RSMA的IRS辅助认知无线网络安全波束成形方案 [J]. *南京邮电大学学报(自然科学版)*, 2024, 44(6): 25-34.
WU Xuewen, MA Jingxiao, XUE Xiaoping, et al. A secure beamforming method for IRS-assisted cognitive radio networks based on RSMA [J]. *Journal of Nanjing University of Posts and Telecommunications (Natural Science Edition)*, 2024, 44(6): 25-34. (in Chinese)
- [16] JOLLY A, SINGH K, BISWAS S. RSMA for IRS aided 6G communication systems: joint active and passive beamforming design [C] //2021 IEEE International Conference on Advanced Networks and Telecommunications Systems (ANTS). 2021: 7-12.
- [17] KATWE M, SINGH K, CLERCKX B, et al. Rate splitting multiple access for sum-rate maximization in IRS aided uplink communications [J]. *IEEE Transactions on Wireless Communications*, 2022, 22(4): 2246-2261.
- [18] KATWE M, SINGH K, CLERCKX B, et al. Rate-splitting multiple access and dynamic user clustering for sum-rate maximization in multiple RISs-aided uplink mmWave system [J]. *IEEE Transactions on Communications*, 2022, 70(11): 7365-7383.
- [19] YANG Z H, CHEN M Z, SAAD W, et al. Sum-rate maximization of uplink rate splitting multiple access (RSMA) communication [J]. *IEEE Transactions on Mobile Computing*, 2022, 21(7): 2596-2609.
- [20] GUO H Y, LIANG Y C, CHEN J, et al. Weighted sum-rate maximization for intelligent reflecting surface enhanced wireless networks [C] //2019 IEEE Global Communications Conference (GLOBECOM). 2019: 1-6.

(责任编辑:李小溪)