

## 多臂赌博机赋能的IRS辅助MISO系统传输方案设计

宋云超, 李浩, 梁汇彬, 汪革

(南京邮电大学电子与光学工程学院、柔性电子(未来技术)学院, 江苏 南京 210023)

**摘要:** 智能反射面 (IRS, intelligent reflecting surface) 是一种能够提升毫米波通信性能并应对传输挑战的技术。然而, 传统方法在联合优化预编码和 IRS 相移时需要准确的信道状态信息, 导频消耗随着基站天线数和 IRS 元件数的增加而增大。为降低导频, 提出了一种基于多臂赌博机 (MAB, multi-armed bandit) 的传输设计方案。具体而言, 在 IRS 辅助单天线用户下行通信系统中, 通过双时间尺度的方法, 将 IRS 相移和预编码波束成形问题解耦。利用强化学习中的上下文 MAB 算法, 在每个相干时间块内根据上下文信息和奖励反馈来设计 IRS 相移矩阵。接着, 采用迫零方法设计预编码, 并引入贪婪法来选择每一时刻的动作, 以避免陷入局部最优。仿真结果表明, 与现有的算法相比, 所提算法能够实现更高的有效频谱效率。

**关键词:** 毫米波; 智能反射面; 多输入单输出; 上下文多臂赌博机

**中图分类号:** TN929.5

**文献标志码:** A

**doi:** 10.11959/j.issn.2096-3750.2025.00386

## A multi-armed bandits empowered transmission scheme for IRS-assisted MISO system

SONG Yunchao, LI Hao, LIANG Huibin, WANG Ge

College of Electronic and Optical Engineering and College of Flexible Electronics (Future Technology), Nanjing University of Posts and Telecommunications, Nanjing 210023, China

**Abstract:** Intelligent reflecting surface (IRS) is an effective technology that can enhance millimeter-wave communication performance and address transmission challenges. However, conventional methods require accurate channel state information when jointly optimizing precoding and IRS phase shifting, and pilot consumption increases with the number of base station antennas and IRS components. To reduce pilot overhead, a transmission design scheme based on multi-armed bandit (MAB) was proposed. Specifically, in IRS-assisted single antenna user downlink communication systems, the problem of IRS phase shift design and precoding beamforming was decoupled through a dual time scale approach. Using the contextual MAB algorithm in reinforcement learning, the IRS phase shift matrix was designed within each coherent time block based on contextual information and reward feedback. Next, the zero forcing method was used to design precoding, and the greedy method was introduced to select actions at each moment to avoid getting stuck in local optima. The simulation results show that the proposed algorithm can achieve higher effective spectral efficiency than the existing algorithms.

**Key words:** millimeter-wave, IRS, MISO, contextual MAB

### 0 引言

智能反射面 (IRS, intelligent reflecting surface)

被认为是一项较具潜力的6G技术, 它是由大量的反射元件组成的平面超表面, 通过无源反射元件将入射信号反射到特定的方向, 具有低功耗、减小用

收稿日期: 2023-08-25; 修回日期: 2024-01-23

通信作者: 宋云超, songyc@njupt.edu.cn

基金项目: 国家自然科学基金资助项目 (No. 62101282)

**Foundation Item:** The National Natural Science Foundation of China (No. 62101282)

户干扰等特点<sup>[1-3]</sup>。IRS可以极大地提高信号传输的性能，因此在无线通信领域被广泛研究<sup>[4-6]</sup>。

目前解决IRS相移矩阵和预编码的联合优化问题的方法主要包括两类，一类是通过传统优化方法求解。在假设信道已知的条件下，文献[7]通过拉格朗日对偶变换，对IRS相移和基站（BS, base station）预编码的联合优化问题进行解耦。利用分式规划方法对BS处的主动波束成形进行了优化，并使用3种低复杂度的算法对IRS的无源波束成形进行求解。文献[8]提出交替优化的方法，联合优化IRS相移和传输协方差矩阵，通过固定变量，每次只优化传输协方差矩阵或IRS的一个相移，通过不断迭代，最终找到局部最优解。文献[9]考虑在IRS辅助多个单天线用户的场景下，联合优化IRS相移和预编码波束成形。通过交替方向乘子法（ADMM, alternating direction method of multipliers）来解决优化问题。然而，上述IRS辅助通信系统的联合优化问题需要在信道状态信息（CSI, channel state information）已知的条件下完成，而CSI的获取在实践中具有一定的挑战性。特别地，为了减少硬件成本，IRS端通常不配备射频链，因此无法主动处理基带信号，并且通常无法直接分别估计BS到IRS和IRS到用户之间的信道<sup>[10]</sup>。为了减小CSI造成的导频开销，研究者们提出了另一类方法，即利用机器学习解决IRS相移矩阵和基站预编码的优化问题。在毫米波大规模多输入多输出系统中，文献[11]利用深度学习的方法解决预编码波束和IRS相移的选择问题，将每次选择的预编码和IRS相移视为深度神经网络（DNN, deep neural network）中的网络参数<sup>[12]</sup>，通过遍历找到有限码本的最优解。文献[13]提出了一个由双DNN组成的模型，该模型能够适用于码本空间大的场景。DNN通过给定特定场景下生成的一组数据集，训练得到对应的参数，从而实现用户接收频谱效率的优化。因此，从DNN获得的解在某些信道环境下可能无效，因为这些信道与训练集上统计的数据是不同的。此外，DNN的训练需要标签数据，且不适用于连续动作空间的场景。因此，在联合优化问题中，研究者们提出了使用强化学习的方法。与使用DNN不同，强化学习不需要标签数据，仅需通过与环境的交互来学习策略的选择，而且强化学习可适应不同环境，并且能够在不同的环境中学习到最佳策略。这些优点使强

化学习成为联合优化问题中一个非常有用的工具。文献[14]在联合优化基站和IRS反射元件的波束成形场景中，使用强化学习的方法来选择最优波束。多臂赌博机（MAB, multi-armed bandit）作为强化学习的重要分支，是一种无监督的学习方式，会根据实时反馈和观察结果进行自适应调整，且复杂度较低。利用MAB方法能够在减小导频开销的前提下，有效地处理毫米波通信的联合优化问题。因此，开展这方面的研究工作具有重要意义。

本文考虑了一个IRS辅助单天线用户通信的场景，通过对IRS、预编码的优化问题进行研究，在保证性能的前提下，降低了基站端的导频开销，提出了利用强化学习中上下文多臂赌博机（CMAB, contextual multi-armed bandit）的方法对IRS相移进行设计<sup>[15]</sup>。具体而言，本文的主要贡献如下。

1) 研究了IRS辅助的下行多输入单输出（MISO, multi-input single-output）系统的有效频谱效率最大化问题。利用双时间尺度的方法对IRS和预编码的联合优化问题进行解耦。在设计过程中，先利用CMAB对IRS的设计问题进行建模，该方法通过接收信号和信道参数更新IRS相移，不需要得到精确的信道信息，有效地降低了BS的导频开销，再利用迫零（ZF, zero forcing）准则求解预编码。

2) 对于建立的MAB问题，提出CMAB对相移进行设计，即以一定概率选择最优解，剩余概率随机选择，从而进行探索和折中。对于设计过程中的优化问题，采用半正定松弛（SDR, semidefinite relaxation）方法<sup>[16-18]</sup>进行求解。对于预编码的求解，则需要发送导频，利用迫零方法使得接收信号最大化。本文所提的双时间尺度的算法，先通过统计信道状态信息（SCSI, statistical channel state information）求解IRS的相移，再利用瞬时信道状态信息（ICSI, instantaneous channel state information）求解预编码，所需要消耗的导频数与天线数相关，对于信道状态参数的设计，只需要消耗一个导频符号获得接收信号，与传统的信道估计方法相比<sup>[19]</sup>，有效地降低了训练开销，使每个相干时间块里传输数据的符号占比增加。

3) 仿真结果表明，在考虑BS端导频开销的条件下，本文所提算法的性能比传统方法的性能要好，并且所提算法为在线学习方法，能够适用于连续时间场景。

本文中,  $\mathbf{C}^{M \times N}$  表示  $M \times N$  维的复空间、 $\mathbf{X}^{-1}$ ,  $\mathbf{X}^H$  和  $\mathbf{X}^T$  分别表示矩阵的逆、共轭转置和转置,  $E$  表示数学期望,  $\text{diag}(\mathbf{c})$  表示对角元素为向量  $\mathbf{c}$  的对角矩阵,  $\otimes$  表示 Kronecker 积。

## 1 系统模型

系统模型如图1所示, 本文研究的是频分双工 (FDD, frequency-division duplex) MISO 系统<sup>[20-22]</sup>, BS 与用户之间的直接链路被障碍物阻挡, 因此 BS 仅通过 IRS 辅助发送信号给单天线用户<sup>[9]</sup>。BS 端采用均匀平面阵列 (UPA, uniform planar array), 且配备了  $N = N_x N_y$  根天线, 其中,  $N_x$  和  $N_y$  表示横轴和纵轴的天线数。IRS 端配备  $M = M_x M_y$  个无源反射元件, 其中, 横轴和纵轴无源反射元件的个数分别为  $M_x$  和  $M_y$ 。设 BS 到 IRS 的信道为  $\mathbf{H} \in \mathbf{C}^{M \times N}$ , IRS 到用户的信道为  $\mathbf{h}_r \in \mathbf{C}^{1 \times M}$ 。

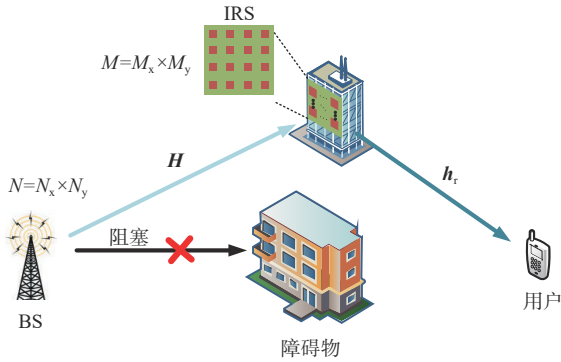


图1 系统模型

本文采用莱斯衰落信道模型, 其中,  $\bar{\mathbf{H}}$  表示直射路径的信道,  $\hat{\mathbf{H}}$  表示考虑扩散多路径信号的瑞利分布随机分量<sup>[23-25]</sup>。考虑用户缓慢移动的场景, 假设在  $T$  个时刻内 BS 到 IRS 的水平角、方位角和 IRS 到用户的角度不发生改变, 但在相干时间块间的信道增益和噪声会持续变化。因此, 在相干时间块  $t$  内, BS 到 IRS 的信道  $\mathbf{H}$  可以表示为

$$\mathbf{H}(t) = \sqrt{\frac{K}{K+1}} \bar{\mathbf{H}}(t) + \sqrt{\frac{1}{K+1}} \hat{\mathbf{H}}(t) \quad (1)$$

其中,  $K$  表示莱斯因子。由于 BS 端和 IRS 端都采用 UPA, 因此阵列响应向量可以表示为

$$\mathbf{a}_B(\eta_B, \gamma_B) = \begin{bmatrix} 1, \dots, e^{-j\frac{2\pi d}{\lambda}(N_x-1)\sin(\eta_B)\cos(\gamma_B)} \end{bmatrix}^T \otimes \begin{bmatrix} 1, \dots, e^{-j\frac{2\pi d}{\lambda}(N_y-1)\cos(\gamma_B)} \end{bmatrix}^T \quad (2)$$

$$\mathbf{a}_R(\eta_R, \gamma_R) = \begin{bmatrix} 1, \dots, e^{-j\frac{2\pi d}{\lambda}(M_x-1)\sin(\eta_R)\cos(\gamma_R)} \end{bmatrix}^T \otimes \begin{bmatrix} 1, \dots, e^{-j\frac{2\pi d}{\lambda}(M_y-1)\cos(\gamma_R)} \end{bmatrix}^T \quad (3)$$

其中,  $\eta_R$  和  $\gamma_R$  分别代表信号达到 IRS 的方位角和俯仰角,  $\mathbf{a}$  代表 UPA 的响应向量,  $\lambda$  为毫米波波长,  $d$  表示相邻天线或 IRS 元件之间的间距。为了简化问题, 假设 IRS 和 BS 的横轴和纵轴元件间距相同, 均为半波长  $d = \lambda/2$ 。因此, 在相干时间块  $t$  内, BS 到 IRS 的信道可以表示为

$$\bar{\mathbf{H}}(t) = \alpha(t) \mathbf{a}_R(\eta_R, \gamma_R) \mathbf{a}_B^H(\eta_B, \gamma_B) \quad (4)$$

其中,  $\alpha(t)$  表示路径增益。同理, IRS 到用户的信道可以表示为

$$\mathbf{h}_r(t) = \alpha_s(t) \mathbf{a}_R^H(\eta_B, \gamma_B) \quad (5)$$

基于上述信道模型, 用户在相干时间块  $t$  内的接收信号为

$$\mathbf{y}(t) = \sqrt{P} \hat{\mathbf{h}}(t) \mathbf{f}(t) s(t) + \mathbf{n}(t) \quad (6)$$

其中,  $\hat{\mathbf{h}}(t) = \mathbf{h}_r(t) \text{diag}(\boldsymbol{\theta}(t)) \mathbf{H}(t)$  为级联信道,  $P$  为 BS 端发射总功率,  $\boldsymbol{\theta}(t) = [e^{j\beta_1(t)}, \dots, e^{j\beta_N(t)}]^T \in \mathbf{C}^{M \times 1}$  表示 IRS 无源元件的反射角度, 且  $\beta_i \in [0, 2\pi)$ ,  $i = 1, \dots, N$  表示第  $i$  个反射元件的反射角度,  $\mathbf{f} \in \mathbf{C}^{N \times 1}$  表示预编码向量, 且  $\|\mathbf{f}\|_2^2 = 1$ ,  $s(t)$  表示在相干时间块  $t$  内发送的信号,  $\mathbf{n} \sim \text{CN}(0, \sigma^2)$  是加性高斯白噪声。本文的目标是通过优化 IRS 相移和数字预编码向量来最大化系统有效频谱效率。在连续时间系统中, 为便于比较不同信道状态信息下 IRS 和预编码的优化设计, 本文考虑了下行链路的导频开销  $\mu(t)$  和衰落相干时间块  $T_d$ , 定义了类似文献[26]的标准化的频谱效率和, 其形式为

$$R_{\text{avg}}(t) = \frac{1}{T} \sum_{t=1}^T \left(1 - \frac{\mu(t)}{T_d}\right) R(t) \quad (7)$$

其中,  $R(t)$  代表相干时间块  $t$  的频谱效率, 记为  $R(t) = \text{lb}(1 + \text{SNR}(t))$ ,  $\text{SNR}(t) = P |\hat{\mathbf{h}}(t) \mathbf{f}(t)|^2 / \sigma^2(t)$  表示信噪比。

在 IRS 辅助无线通信系统中, 传统方法为了获取 ICSI 需要花费大量的导频开销进行信道估计, 因此用于传输数据的时间将会减少, 从而降低了有效频谱效率。一些研究利用 SCSi 进行相位偏移和波束成形的设计, 由于 SCSi 在较长一段时间内保持不变, 获取 SCSi 所需的导频开销在总符号数中的占比远远小于获取 ICSI 导频开销的占比, 因此

提升了系统的有效频谱效率<sup>[27-28]</sup>。然而，当用户移动速度较快时，SCSI变化的速度也会加快，此时获取SCSI所需的导频开销将在总符号数中占据较大的比例，从而降低了有效频谱效率<sup>[29-30]</sup>。为解决这个问题，本文采用了CMAB方法来设计IRS的相位偏移，该方法无须获取CSI，通过训练获得信道状态参数并选择相移，从而提高系统的有效频谱效率。

## 2 基于CMAB的双时间尺度传输方案设计

本文所提的传输方案中，IRS相移和BS预编码向量分别是在SCSI和ICSI保持不变的情况下设计的，因此称为双时间尺度的传输方案设计。首先，本节介绍如何对预编码和IRS的联合优化问题进行解耦，然后，重点介绍如何利用CMAB方法设计IRS的相移。

### 2.1 问题表述

根据第1节的系统模型可知，本文旨在最大化系统的有效频谱效率。由式(7)可知，系统的有效频谱效率与接收信号能量有关，通过最大化接收信号能量来最大化系统的有效频谱效率，优化问题描述为

$$\begin{aligned}
 P1: \max_{\text{diag}(\theta), f} & \sum_{t=1}^T |\hat{h}(t) f(t)|^2 \\
 \text{s.t.} & |\theta_i| = 1, i = 1, \dots, M
 \end{aligned} \tag{8}$$

获取ICSI需要的导频数与IRS的元件数和BS的天线数相关，这导致较大的导频和时间成本。为

降低获取ICSI的成本，本文采用基于MAB的方法。在每个相干块中，通过预先设计好的IRS相位和迭代优化来降低导频开销。接着，利用ZF准则解决预编码问题。这种方法将导频数减少到与基站天线数相关，从而在一定程度上降低了导频开销。

相干时间块内的通信框架如图2所示，本文主要研究下行信道的传输设计方案，但也涉及上行反馈。具体的设计方案如下：首先初始化IRS相移，在下行传输中利用ICSI同时发送天线数个正交导频进行信道估计，图2中黑点符号用于发送导频。估计的信道考虑了已设计好的IRS相移。然后，通过上行反馈获取相应的信道，在BS端通过ZF准则设计预编码。为了收集上下文信息用于IRS相移的设计，还需要通过上行反馈发送一个导频符号获取本相干时间块内用户的 $y(t)$ ，剩余符号则用来做数据传输。在接下来在每个相干时间块内，先通过CMAB算法求解IRS相移，其他步骤如上。当用户的移动距离过大时，重新更新SCSI。

图2给出了本文的设计方案和传统的利用信道估计的传输方案。由于本文采用的是莱斯信道，无法利用信道的低秩特性，传统方法所需的导频符号数量是 $N \times M$ ，远大于本文所提方案的 $N + 1$ ，因此能够传输数据的符号数量相对较少。与其他文献类似，本文的IRS相移设计、上行反馈和预编码不占用下行符号。因此一个相干时间块内所有符号仅用来进行导频和有用信号的传输。

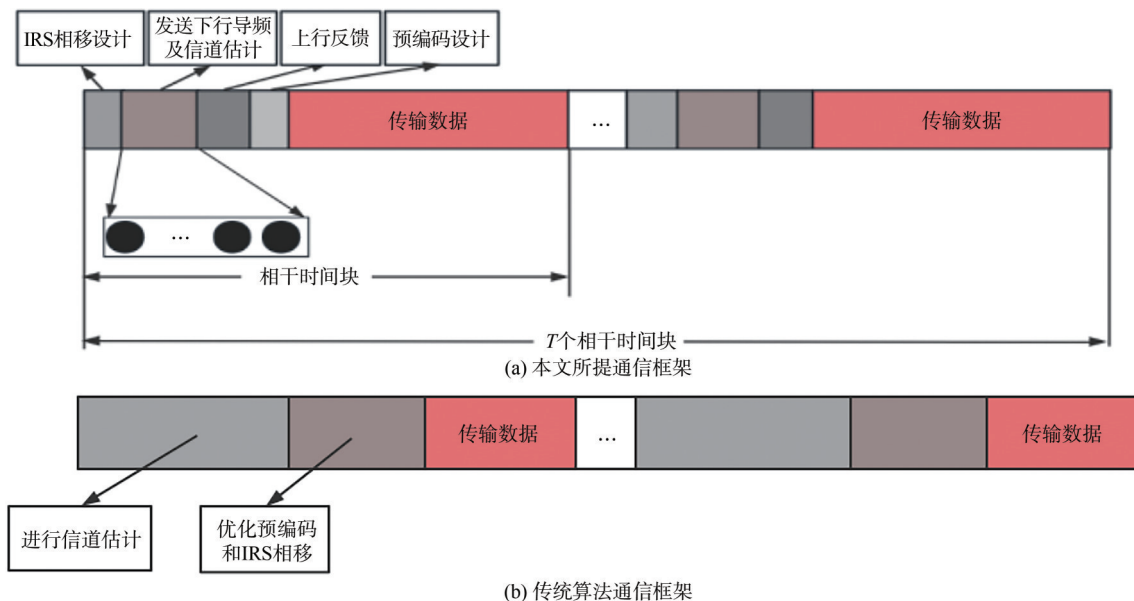


图2 相干时间块内的通信框架

在本文考虑的双时间尺度传输框架中, 假设统计 CSI 在较长时间段  $T$  内保持不变, 而瞬时 CSI 在信道相干时间内保持不变。通过发送导频, 用户在第  $t$  个相干时间接收到的信号表示为

$$\mathbf{y}^H(t) = \boldsymbol{\theta}^H(t) \text{diag}(\mathbf{h}_r(t)) \mathbf{H}(t) \mathbf{X}^H + \mathbf{n}^H(t) \quad (9)$$

其中,  $\mathbf{n}^H(t) \in \mathbf{C}^{1 \times N_p}$  为噪声向量,  $\mathbf{X}^H \in \mathbf{C}^{M \times N_p}$  为传输的导频符号矩阵,  $N_p$  为导频数量且  $\mathbf{X}^H$  的第  $i$  行表示 BS 第  $i$  根天线发送的导频符号。  $\mathbf{X}^H$  的列向量相互正交且满足  $\mathbf{X}^H \mathbf{X} = \mathbf{I}$ 。为准确估计信道, 导频数量应不少于基站天线数, 所以本文假设  $N_p = N$ 。用户在接收到导频信号后对级联信道的  $\tilde{\mathbf{h}}(t)$  进行估计, 估计值可以表示为

$$\tilde{\mathbf{h}}(t) = \mathbf{y}^H(t) \mathbf{X} \quad (10)$$

然后, 用户将级联信道的估计值  $\tilde{\mathbf{h}}(t)$  反馈给 BS, BS 通过 ZF 准则<sup>[31]</sup>设计波束成形向量  $\mathbf{f}(t)$ , 并用于数据传输, 其中

$$\mathbf{f}(t) = (\tilde{\mathbf{h}}(t))^\text{H} (\tilde{\mathbf{h}}(t) \tilde{\mathbf{h}}^\text{H}(t))^{-1} \quad (11)$$

由于  $\mathbf{h}_r(t) \text{diag}(\boldsymbol{\theta}(t)) \mathbf{H}(t) = \boldsymbol{\theta}^H(t) \text{diag}(\mathbf{h}_r(t)) \mathbf{H}(t)$ , 则优化问题 P1 可转化为

$$\begin{aligned} \text{P2: } \max_{\boldsymbol{\theta}^H, \mathbf{f}} & \quad \left| \boldsymbol{\theta}^H(t) \mathbf{G}(t) \mathbf{f}(t) \right|^2 \\ \text{s.t. } & \quad |\theta_i| = 1, i = 1, \dots, M \end{aligned} \quad (12)$$

其中,  $\mathbf{G}(t) = \text{diag}(\mathbf{h}_r(t)) \mathbf{H}(t)$ 。因此, 优化问题 P2 通过解耦可以改写为

$$\begin{aligned} \text{P3: } \max_{\boldsymbol{\theta}^H} & \quad \left\| \boldsymbol{\theta}^H(t) \mathbf{G}(t) \right\|_2^2 \\ \text{s.t. } & \quad |\theta_i| = 1, i = 1, \dots, M \end{aligned} \quad (13)$$

对于优化问题 P3, 需要对  $\mathbf{G}(t)$  进行估计。考虑到 BS 天线数的增加可能会给准确获取 ICSI 带来挑战, 在每个相干块中进行 ICSI 估计会导致过高的导频开销。因此, 在 FDD 模式下, 本文利用长时间内 CSI 保持不变的特点和信道的统计特性, 以降低导频开销。接下来重点介绍 IRS 相移的设计。

## 2.2 CMAB 的模型

本文系统中, 假设 IRS 元件的反射角度是连续的, 每个元素都可以取  $[0, 2\pi)$  中的任意值, 在离散化的相干时间块  $t$  中, 代理 (agent) 作为 MISO 系统的控制器 (在本文中指 BS, 通过现场可编程门阵列控制器控制 IRS 相移), 首先将 IRS 相移看作是选择动作  $\mathbf{a}_t$ , 通过与环境  $\mathbf{s}_t$  交互后反馈得到一个奖励  $r_t$ , 并进行到下一个时刻  $t+1$ 。强化学习系统框架如图 3 所示, 下面详细介绍每一个参数的含义。

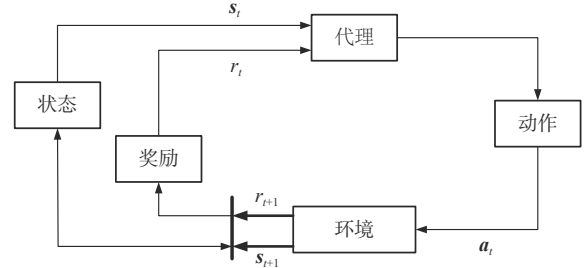


图3 强化学习系统框架

**状态 (state):** 在本文的信道模型中, 每个时间块的信道增益和噪声干扰是一直变换的, 将每个时间相干块中 BS 到 IRS 的信道  $\mathbf{H}(t)$  和 IRS 到用户的信道  $\mathbf{h}_r(t)$  ( $t = 1, \dots, T$ ) 以及噪声  $\mathbf{n}(t)$  看作为环境, 可表示为  $\mathbf{s}_t = [\mathbf{h}_r^\text{H}(t), \mathbf{H}(t), \mathbf{n}(t)]$ ,  $\mathbf{s}$  为信道信息的集合。

**动作 (action):** 设  $\boldsymbol{\theta}^H(t) \in \mathbf{C}^{1 \times M}$  是在时刻  $t$  选择的 IRS 相移, 该相移是从包含  $K$  个动作的码本空间里选择的。因为系统考虑的是连续时间场景, 所以  $K \rightarrow \infty$ , agent 在每个时间块中通过算法设计 IRS 的相移  $\boldsymbol{\theta}^H$ ,  $\mathbf{a}_t = [\boldsymbol{\theta}^H(t)]$ , 在每一回合中, 奖励最大的动作记为  $\arg \max \mathbf{a}_t$ 。

**奖励 (reward):** agent 根据选择的动作  $\mathbf{a}_t$ , 与环境  $\mathbf{s}_t = [\mathbf{h}_r^\text{H}(t), \mathbf{H}(t), \mathbf{n}(t)]$  交互后获得一个奖励, 该奖励用来衡量此刻 agent 选择动作的回报  $r_t$ 。

每个时间块  $t$  代表一个相干时间, 首先随机设计一个  $\boldsymbol{\theta}^H(t)$ , 然后利用环境  $\mathbf{h}_r(t)$  和  $\mathbf{H}(t)$ , agent 得到一个回报  $r_t$ , 接着利用前面相干时间块的上下文信息以及 SDR 的方法, 更新下一时刻的动作  $\boldsymbol{\theta}^H(t+1)$ , 并且将相干时间块  $t$  的信道状态信息保留, 在新的相干时间块  $t+1$  中执行动作  $\mathbf{a}_{t+1}$ , 并获得奖励  $r_{t+1}$ , 以此类推。在此场景中, 每次 IRS 的相移都通过之前保留的信息利用半正定松弛计算得到。由于用户在缓慢移动, 得到的上下文信息越多, 求解的 IRS 相移就越准确。

本文寻找到每一轮的最优臂  $\mathbf{a}_t^*$ , 简单来讲, 根据前  $t-1$  个时刻的动作  $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_{t-1}$  以及环境的相关信息, 推导出相干时间块  $t$  的动作  $\mathbf{a}_t$ , 这说明相干时间块  $t$  的动作依赖于前面时刻的动作信息和状态信息。目标是设计  $\mathbf{a}_t$  从而最小化遗憾, 遗憾指所选择的臂与每一时刻选择最优臂之间的差值。假设已知各时刻动作  $\mathbf{a}_t$ , 遗憾可以定义为

$$\text{Reg}(\mathbf{s}_t, \mathbf{a}_t) = \sum_{i=1}^T \mathbb{I}^r r_{\mathbf{a}_i} - \mathbb{E} \left( \sum_{i=1}^T r_{\mathbf{a}_i} \right) \quad (14)$$

其中,

$$\mathbb{I}^* = \begin{cases} \mathbb{I} = 1, & \mathbb{I} = \arg \max_{1, \dots, K} \mathbf{a}_t \\ \mathbb{I} = 0, & \text{其他} \end{cases} \quad (15)$$

### 2.3 基于CMAB的IRS相移设计

由于本文研究的是单天线用户场景,可以将 $(\boldsymbol{\theta})^H \mathbf{G} \mathbf{f}$ 在 $T$ 个时刻内看作是呈线性关系的。通过这种关系提取一个维数为 $d$ 的上下文系数向量 $\boldsymbol{\zeta}_{t,a}$ ,即场景中的信道参数, $\mathbf{x}_{t,a}^H$ 用来表示特征向量,即动作 $\mathbf{a}_t = \mathbf{x}_{t,a} = \boldsymbol{\theta}_{t,a}$ 。对于第 $t$ 回合,每个臂的期望收益为臂 $\mathbb{I}^{(2)}$ 的线性函数。

$$\mathbb{E}[r_{t,a} | \mathbf{x}_{t,a}] = \mathbf{x}_{t,a}^H \boldsymbol{\zeta}_{t,a} \quad (16)$$

根据上下文信息 $\boldsymbol{\zeta}_{t,a}$ 来更新策略,这将对接下来的决策 $\mathbf{x}_{t,a}^H$ 产生影响。将 $t \times d$ 维矩阵 $\mathbf{D}_{t,a}$ 用来表示前 $t$ 个时刻特征向量的集合, $\mathbf{D}_{t,a} = [\mathbf{x}_{1,a}^H, \mathbf{x}_{2,a}^H, \dots, \mathbf{x}_{t,a}^H]$ 。将 $\mathbf{c}_{t,a}$ 表示前 $t$ 次观测值的奖励,即观测到的接收信号的集合 $\mathbf{c}_{t,a} = [y(1), y(2), \dots, y(t)]^H$ ,对于 $\mathbf{D}_{t,a}$ 和 $\mathbf{c}_{t,a}$ 可以通过岭回归估计求解系数向量。

$$\hat{\boldsymbol{\zeta}}_{t,a} = \mathbf{A}_{t,a}^{-1} \mathbf{b}_{t,a} \quad (17)$$

$$\mathbf{A}_{t,a} = \mathbf{D}_{t,a}^H \mathbf{D}_{t,a} + \mathbf{I}_d \quad (18)$$

$$\mathbf{b}_{t,a} = \mathbf{D}_{t,a}^H \mathbf{c}_{t,a} \quad (19)$$

其中,矩阵 $\mathbf{A}_{t,a}$ 记录了观察到的上下文, $\mathbf{I}_d$ 是一个 $d$ 维的单位向量, $\mathbf{b}_{t,a}$ 是根据奖励反馈加权的相应响应向量。

为平衡探索与选择此刻最优臂之间的平衡,相干时间块 $t$ 内IRS的设计服从 $\epsilon$ -贪婪法,

$$\mathbf{a}_t = \begin{cases} \arg \max_{\mathbf{x}_{t,a}^H} |\mathbf{x}_{t,a}^H \boldsymbol{\zeta}_{t,a}|^2, & \beta \leq 1 - \epsilon \\ \text{random}, & \text{其他} \end{cases} \quad (20)$$

其中, $\beta$ 是 $[0, 1]$ 之间均匀分布的随机变量, $\epsilon$ 为探索因子。

最优动作的计算是根据以下步骤得到的,通过得到的上下文信息 $\hat{\boldsymbol{\zeta}}_{t,a}$ ,利用置信度上界方法,更新 $t+1$ 时的动作 $\mathbf{a}_{t+1}$ 。 $\hat{r}_{t,a} \leftarrow \mathbf{x}_{t,a}^T \hat{\boldsymbol{\zeta}}_{t,a} + \eta \sqrt{\mathbf{x}_{t,a}^T \mathbf{A}_{t,a}^{-1} \mathbf{x}_{t,a}}$ 表示系数向量 $\boldsymbol{\zeta}_a$ 和特征向量 $\mathbf{x}_{t,a}^H$ 所获得的最优动作的线性组合再加上奖励偏差, $\eta$ 控制探索还是选最优动作, $\eta$ 越小,选最优动作的比重就越大, $\eta$ 越大,探索的比重就增加。

由于本文考虑的场景是连续的,所以选取所有的臂,然后比较每一个reward,通过不断地迭代找到均值最大的那个臂选为最优臂是不现实的,因此,在所提方法中,每一个相干时间块 $t+1$ 都是根

据前面时刻得到的信道状态信息更新这一时刻的动作 $\mathbf{a}_{t+1} = \boldsymbol{\theta}_{t+1,a}$ ,更新的策略中IRS的求解通过传统的SDR方法,将问题转化为

$$\begin{aligned} \text{P4: } \max_{\boldsymbol{\theta}_{t+1,a}^H} & \left| \boldsymbol{\theta}_{t+1,a}^H \hat{\boldsymbol{\zeta}}_{t,a} \right|^2 \\ \text{s.t. } & \boldsymbol{\theta}_i = 1, i = 1, \dots, M \end{aligned} \quad (21)$$

由于 $\boldsymbol{\theta}_{t+1,a}^H$ 的每个元素受到恒模约束的影响,可以通过文献[11]给出的高斯随机化的方法求解,优化问题P4转化为

$$\begin{aligned} \text{P4: } \max_V & \left\| (\text{tr}(\mathbf{W}(t)\mathbf{V}(t))) \right\|_2 \\ \text{s.t. } & V_{i,i} = 1, i = 1, \dots, M \quad \mathbf{V} \succeq 0 \end{aligned} \quad (22)$$

其中, $\mathbf{W}(t) = \hat{\boldsymbol{\zeta}}_{t,a} \hat{\boldsymbol{\zeta}}_{t,a}^H$ , $\mathbf{V}(t) = \boldsymbol{\theta}_{t+1,a} \boldsymbol{\theta}_{t+1,a}^H$

求解出 $\mathbf{V}$ 之后,经过奇异值分解得到 $\mathbf{V} = \mathbf{U} \boldsymbol{\Sigma} \mathbf{U}^H$ ,获得向量 $\bar{\mathbf{v}} = \mathbf{U} \boldsymbol{\Sigma}^{1/2} \mathbf{r}$ ,其中, $\mathbf{r} \in \text{CN}(0, \mathbf{I}_N)$ ,通过元素迭代法选出最大的 $\bar{\mathbf{v}}$ ,从而求得 $\boldsymbol{\theta}_{t+1,a}^H$ 。

所提方法充分利用前面时刻的信道状态信息,通过不断优化IRS的解,最后在时间段 $T$ 内寻找到IRS的优化解。在每一个相干时间块 $t$ 内本文都是根据上一时刻的上下文信息先设计IRS的角度,并且每个相干时间块内IRS相移的设计只需要一个导频符号获取当前时刻用户端的接收信号 $\mathbf{y}(t)$ ,来更新下一时刻的动作。与信道估计相比,大大地减小了导频开销和信道估计的时间,提高了有效频谱效率。

CMAB算法如算法1所示,先初始化一个IRS的相移,然后初始化矩阵 $\mathbf{A}_{t,a}$ 和 $\mathbf{b}_{t,a}$ ,更新获得系数向量 $\hat{\boldsymbol{\zeta}}_{t,a}$ 和奖励 $r_t$ 。根据获得的系数向量 $\hat{\boldsymbol{\zeta}}_{t,a}$ ,并通过SDR方法,来更新下一个时刻的IRS相移 $\mathbf{x}_{t+1,a}^H$ ,当agent收集到足够多的上下文信息后,求解的IRS相移就会更准确。

#### 算法1 CMAB算法

输入  $\eta \in \mathbf{R}_+$

for  $t=1, 2, \dots, T$

if  $t=1$  then

$\mathbf{A}_{1,a} \leftarrow \mathbf{I}_d$  ( $d$ 维单位矩阵);

$\mathbf{b}_{1,a} \leftarrow \mathbf{0}_{d \times 1}$  ( $d$ 维零向量);

else

根据式(16)得到 $\hat{\boldsymbol{\zeta}}_{t,a}$ ,  $\hat{\boldsymbol{\zeta}}_{t,a} \leftarrow \mathbf{A}_{t,a}^{-1} \mathbf{b}_{t,a}$ ;

$\hat{r}_{t,a} \leftarrow \mathbf{x}_{t,a}^T \hat{\boldsymbol{\zeta}}_{t,a} + \eta \sqrt{\mathbf{x}_{t,a}^T \mathbf{A}_{t,a}^{-1} \mathbf{x}_{t,a}}$ ;

根据式(17)选择更新策略;

end if

if  $\beta \leq 1 - \epsilon$  then

根据式(19)更新下一时刻动作  $\mathbf{x}_{t+1,a}^H$ ;

else

$\mathbf{a}_t = \text{random}$ ;

$\mathbf{A}_{t,a} \leftarrow \mathbf{A}_{t,a} + \mathbf{x}_{t,a} \mathbf{x}_{t,a}^T$ ;

$\mathbf{b}_{t,a} \leftarrow \mathbf{b}_{t,a} + \hat{\mathbf{r}}_{t,a} \mathbf{x}_{t,a}$ ;

$t = t + 1$ ;

end if

end for

输出  $\mathbf{x}_{t+1,a}^H$

本文所用的ZF的复杂度为  $O(N^3)$ , CMAB算法的复杂度为  $O(M^{3.5} + TM^2)$ , 其中,  $T \gg M$  为随机的次数。ADMM算法的计算复杂度主要取决于天线数  $N$ 、IRS元件数  $M$  和迭代次数  $I$ 。该算法的总计算复杂度为  $O(INM^2T_p)$ , 其中,  $T_p$  指训练开销, 即导频符号周期的数量。相对于ADMM算法, CMAB算法具有更低的复杂度。

### 3 仿真分析

本节通过仿真分析比较了几种方法, 包括CMAB算法、ADMM算法<sup>[32-33]</sup>和随机选择IRS相移、预编码的方法, 以有效频谱效率为评价指标来体现算法的优劣。本文通过不同信噪比 (SNR, signal-to-noise ratio)、天线数量、IRS元件数和发送功率的有效频谱效率图来直观地体现不同方法的结果。

本文考虑的载波频率为28 GHz, 因此载波波长为10.7 mm。假设用户移动速度为1 m/s, 则由相干时间公式  $\lambda/2v$  可知, 相干时间为5.35 ms。使用5G NR帧结构, 子载波间隔为30 kHz, 则在固定时长为1 ms的子帧内包含两个时隙, 共28个符号。因此, 相干时间内可用的总符号数为  $28 \times 5.35 \approx 150$ 。BS配备  $N=16$  个天线, IRS配备的无源元件数  $M=64$ , 用户端是单天线用户。BS的发射功率设置为  $P=30$  dBm, 噪声功率  $\delta^2 = -50$  dBm, 探索因子  $\epsilon$  取0.1,  $\eta$  取0.9, IRS到用户的路径数相同, 都等于1, 复信道增益  $\alpha_r(t)$  和  $\alpha_s(t)$  都为服从独立同分布的复高斯随机变量分布  $CN(0, 1)$ , 设置的总相干时间块数量  $T=1000$ 。假设IRS的相移是连续的, 可以取  $[0, 2\pi)$ 。在ADMM算法下, 每个相干块内都需要知道ISCI, 因此需要发送导频进行信道估计, 然后利用传统的迭代方法得到对应的IRS相移和预编码, 在本文涉及的莱斯信道模型中, 无法利用低秩

特性降低信道估计的导频数, 所需要的导频符号与基站的天线数和IRS的元件数相关, 假设一个相干时间块内能够有效传输数据的时间占总时间的70%~90% (没有取定值是因为在实验图中IRS元件数和BS的天线数是发生变化的), 而CMAB算法是通过上下文信息得到下一个时刻的IRS角度, IRS的设计不需要进行信道估计。但在预编码设计时, 基站需要获取  $\hat{\mathbf{h}}(t)$ , 从而进行预编码设计。基站获取  $\hat{\mathbf{h}}(t)$  需要用户进行下行信道估计再进行上行反馈, 用户进行下行信道估计时需要基站发送导频, 而导频的数量与基站天线数相同。

莱斯信道下不同SNR的有效频谱效率如图4所示, 传统的ADMM算法是基于信道估计完成的, 需要的导频数、BS的天线数和IRS的元件数相关。在考虑导频开销的情况下, 随着信噪比的增加, 3种算法所获得的有效频谱效率都会逐步提高。CMAB算法在不同信噪比下的性能均优于ADMM算法, 原因是每个相干时间块  $t$  内ADMM算法都需要做一次信道估计, 通过发送导频估计信道减少了用户传输数据的时间, 而CMAB算法在每个时刻内获得一组数据, 并且通过前面时刻的数据设计IRS, 下行信道估计所需导频的数量与基站天线数相同, 减小了导频开销。本文所提方法不需要知道每个时刻完整的信道信息, 利用上下文信道信息, 通过不断地迭代, 最后得到优化解。

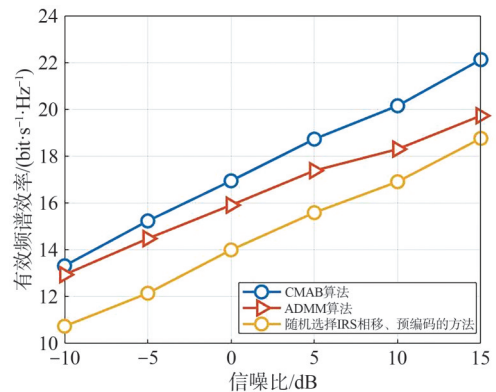


图4 莱斯信道下不同SNR的有效频谱效率

反射元件数  $M$  与有效频谱效率的关系如图5所示,  $M$  依次取16、25、36、49和64。可以看出, 随着反射元件的增加, CMAB算法、ADMM算法和随机选择IRS相移、预编码的方法获得的有效频谱效率逐步提升, 其中, CMAB算法获得的有效频谱效率最高, 明显高于ADMM算法和随

机选择IRS相移、预编码的方法,并且随着IRS元件数的增加,CMAB算法与ADMM算法的有效频谱效率的差值趋于平缓。图5结果说明了CMAB算法在优化IRS相移和预编码的问题上具有良好的性能。

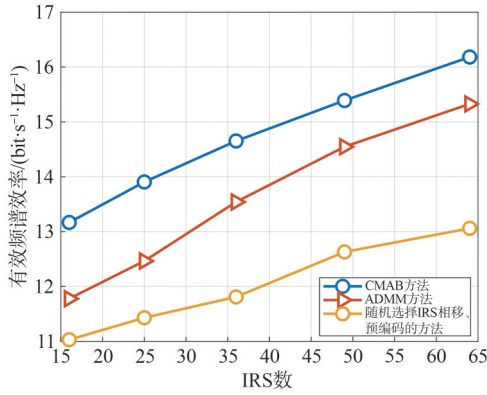


图5 反射元件数 $M$ 与有效频谱效率的关系

BS的不同天线数 $N$ 与有效频谱效率的关系如图6所示, $N$ 依次取2、4、9、16和25,IRS元件数为49。与ADMM算法相比,CMAB算法的性能更好。随着BS天线数 $N$ 的增多,系统得到的有效频谱效率也逐渐增多。

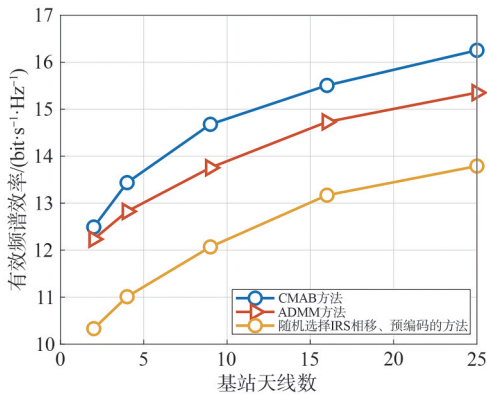


图6 BS的不同天线数 $N$ 与有效频谱效率的关系

发送功率 $P$ 与有效频谱效率的关系如图7所示, $M=36$ , $P$ 依次取10 dBm、20 dBm、30 dBm、40 dBm和50 dBm。随着发送功率的增加,CMAB算法得到的有效频谱效率一直高于ADMM算法和随机选择IRS相移、预编码的方法的有效频谱效率,并且与ADMM算法得到的有效频谱效率之间的差距趋于平缓。

CMAB算法在不同时刻的有效频谱效率如图8所示,其中,天线数 $N=16$ ,IRS元件数 $M=36$ ,发送功率 $P=30$  dBm。一开始由于探索获得的有效频谱

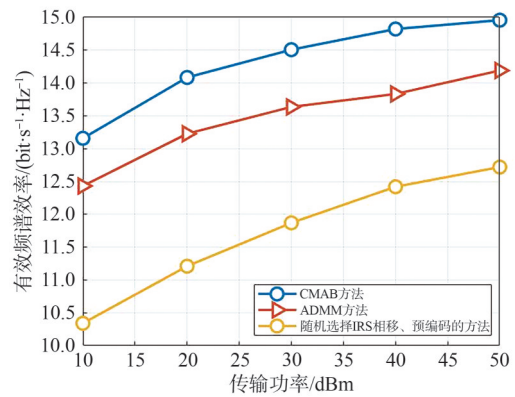


图7 发送功率 $P$ 与有效频谱效率的关系

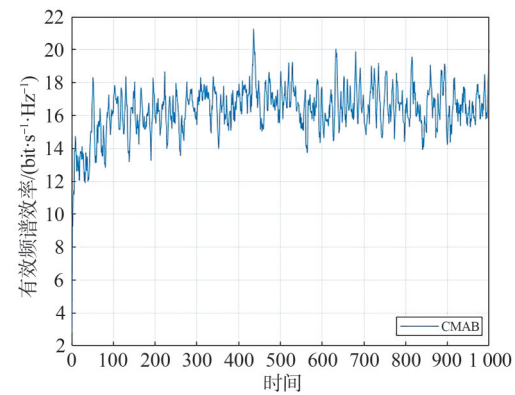


图8 CMAB算法在不同时刻的有效频谱效率

效率相对较差,随着探索次数的增加,获得的有效频谱效率也随之稳定下来。图8中曲线的波动是因为每个相干时间块内信道增益是变化的,所以每次得到的有效频谱效率也是随之改变的。

## 4 结束语

本文考虑的是单天线用户利用IRS辅助通信的场景,通过优化BS的预编码和IRS无源元件的相移矩阵,来实现有效频谱效率的最大化。具体而言,首先,通过目标函数将预编码和无源IRS的相移矩阵进行解耦。采用双时间尺度的方法,将IRS相移的设计问题转化为MAB问题,并利用上下文信息不断进行臂的选择,最终获得最优解。然后,通过ZF准则优化预编码。为了评估CMAB算法的性能,本文将其与ADMM算法、随机选择IRS相移、预编码的方法进行了对比。仿真实验证明了在处理连续场景下的IRS相移设计时,CMAB算法具有较高的频谱效率。研究结果表明了CMAB算法在IRS相移设计中的有效性,可利用上下文信息对IRS相移进行设计。接下来,将研究拓展到多用户问题,并考虑用户之间的平衡关系,为每个用户设

定一个能够正常通信的下界阈值，以实现有效频谱效率和的最大化。

### 参考文献:

- [1] 王长江, 傅友华. IRS辅助毫米波MIMO系统波束成形优化的低复杂度方案[J]. 信号处理, 2022, 38(9): 1965-1974.  
WANG C J, FU Y H. A low-complexity scheme for beamforming optimization of IRS-assisted mmWave MIMO system[J]. Journal of Signal Processing, 2022, 38(9): 1965-1974.
- [2] 范敏, 邵思源, 贺超, 等. 下一代毫米波无线局域网: 愿景与关键使能技术[J]. 物联网学报, 2023, 7(3): 1-14.  
FAN M, SHAO S Y, HE C, et al. Next-generation mmWave WLAN: vision and key enabling technologies[J]. Chinese Journal on Internet of Things, 2023, 7(3): 1-14.
- [3] 潘旭敏, 邓凯元, 陈前斌. 超大规模MIMO-OTFS系统上行辅助的下行信道估计方法[J]. 物联网学报, 2023, 7(4): 28-38.  
PU X M, DENG K Y, CHEN Q B. Uplink assisted downlink channel estimation method of extra-large scale MIMO-OTFS system[J]. Chinese Journal on Internet of Things, 2023, 7(4): 28-38.
- [4] 江玉涵, 邹玉龙, 郑宝玉. 新一代超密集异构蜂窝无线网络研究[J]. 信号处理, 2020, 36(2): 159-167.  
JIANG Y H, ZOU Y L, ZHENG B Y. Research on new generation of ultra-dense heterogeneous cellular wireless networks[J]. Journal of Signal Processing, 2020, 36(2): 159-167.
- [5] ABID W, HAJJAJ M, MEJRI A, et al. Weighted Kalman based detection for uplink MIMO systems[C]//Proceedings of the 2020 International Conference on Software, Telecommunications and Computer Networks (SoftCOM). Piscataway: IEEE Press, 2020: 1-6.
- [6] PAHLAVAN K, KRISHNAMURTHY P. Evolution and impact of Wi-Fi technology and applications: a historical perspective[J]. International Journal of Wireless Information Networks, 2021, 28(1): 3-19.
- [7] GUO H Y, LIANG Y C, CHEN J, et al. Weighted sum-rate maximization for intelligent reflecting surface enhanced wireless networks[C]//Proceedings of the 2019 IEEE Global Communications Conference (GLOBECOM). Piscataway: IEEE Press, 2019: 1-6.
- [8] ZHANG S W, ZHANG R. Capacity characterization for intelligent reflecting surface aided MIMO communication[J]. IEEE Journal on Selected Areas in Communications, 2020, 38(8): 1823-1838.
- [9] LI Y Q, JIANG M, ZHANG Q, et al. Joint beamforming design in multi-cluster MISO NOMA reconfigurable intelligent surface-aided downlink communication networks[J]. IEEE Transactions on Communications, 2021, 69(1): 664-674.
- [10] 傅友华, 陈栋. 混合智能反射表面结构辅助的毫米波通信信道估计[J]. 通信学报, 2021, 42(10): 189-196.  
FU Y H, CHEN D. Channel estimation for hybrid intelligent reflecting surface structure assisted mmWave communications[J]. Journal on Communications, 2021, 42(10): 189-196.
- [11] GAO J B, ZHONG C J, LI G Y, et al. Online deep neural network for optimization in wireless communications[J]. IEEE Wireless Communications Letters, 2022, 11(5): 933-937.
- [12] 李国权, 徐永海, 林金朝, 等. 基于Res-DNN的端到端MIMO系统信号检测算法[J]. 物联网学报, 2022, 6(1): 65-72.  
LI G Q, XU Y H, LIN J Z, et al. Res-DNN based signal detection algorithm for end-to-end MIMO systems[J]. Chinese Journal on Internet of Things, 2022, 6(1): 65-72.
- [13] LI K Y, HUANG C, GONG Y, et al. Double deep learning for joint phase-shift and beamforming based on cascaded channels in RIS-assisted MIMO networks[J]. IEEE Wireless Communications Letters, 2023, 12(4): 659-663.
- [14] LEE Y, LEE J H, KO Y C. Beamforming optimization for IRS-assisted mmWave V2I communication systems via reinforcement learning[J]. IEEE Access, 2022, 10: 60521-60533.
- [15] XU C, SI Y Y, ZHU L H, et al. Dynamic data transaction in crowd-sensing based on multi-armed bandits and shapley value[J]. IEEE Transactions on Sustainable Computing, 2022, 7(3): 609-618.
- [16] LUO Z Q, MA W K, SO A M C, et al. Semidefinite relaxation of quadratic optimization problems[J]. IEEE Signal Processing Magazine, 2010, 27(3): 20-34.
- [17] JALDEN J, OTTERSTEN B. The diversity order of the semidefinite relaxation detector[J]. IEEE Transactions on Information Theory, 2008, 54(4): 1406-1422.
- [18] WU Q Q, ZHANG R. Intelligent reflecting surface enhanced wireless network via joint active and passive beamforming[J]. IEEE Transactions on Wireless Communications, 2019, 18(11): 5394-5409.
- [19] HAN Y, JIN S, WEN C K, et al. Channel estimation for extremely large-scale massive MIMO systems[J]. IEEE Wireless Communications Letters, 2020, 9(5): 633-637.
- [20] QIN Z A, YIN H F, CAO Y D, et al. A partial reciprocity-based channel prediction framework for FDD massive MIMO with high mobility[J]. IEEE Transactions on Wireless Communications, 2022, 21(11): 9638-9652.
- [21] HAN Y, LIU Q, WEN C K, et al. Tracking FDD massive MIMO downlink channels by exploiting delay and angular reciprocity[J]. IEEE Journal of Selected Topics in Signal Processing, 2019, 13(5): 1062-1076.
- [22] HUGL K, KALLIOLA K, LAURILA J. Spatial reciprocity of uplink and downlink radio channels in FDD systems[C]//Proceedings of the Eur. Cooper. Sci. Technol. (EURO-COST). Piscataway: IEEE Press, 2002, 273(2): 066.
- [23] LI X, JIN S, SURAWEERA H A, et al. Statistical 3-D beamforming for large-scale MIMO downlink systems over Rician fading channels[J]. IEEE Transactions on Communications, 2016, 64(4): 1529-1543.
- [24] SALHAB A M, SAMUH M H. Accurate performance analysis of reconfigurable intelligent surfaces over Rician fading channels[J]. IEEE Wireless Communications Letters, 2021, 10(5): 1051-1055.

- [25] PENG Z J, WANG S X, PAN C H, et al. Multi-pair two-way massive MIMO DF relaying over rician fading channels under imperfect CSI[J]. IEEE Wireless Communications Letters, 2022, 11(2): 225-229.
- [26] KIM D, POOR H V, LEE N. Online learning to precode for FDD massive MIMO systems[C]//Proceedings of the 2020 IEEE Globecom Workshops (GC Wkshps). Piscataway: IEEE Press, 2020: 1-6.
- [27] JI S P, WANG Q S, WU S Y, et al. Deep learning based user grouping for FD-MIMO systems exploiting statistical channel state information[J]. China Communications, 2021, 18(7): 183-196.
- [28] ZHENG Z T, ZHU C Z, JIANG B, et al. Statistical channel state information acquisition for massive MIMO communications[C]//Proceedings of the 2015 International Conference on Wireless Communications & Signal Processing (WCSP). Piscataway: IEEE Press, 2015: 1-5.
- [29] SONG Y C, LIU C, LIU Y L, et al. Joint spatial division and multiplexing in massive MIMO: a neighbor-based approach[J]. IEEE Transactions on Wireless Communications, 2020, 19(11): 7392-7406.
- [30] SONG Y C, LIU C, WANG W, et al. Domain selective precoding in 3-D massive MIMO systems[J]. IEEE Journal of Selected Topics in Signal Processing, 2019, 13(5): 1103-1118.
- [31] SONG Y C, LIU C, WANG W, et al. Cluster-group-based two-stage beamforming for massive MIMO[J]. IEEE Transactions on Communications, 2022, 70(3): 1984-1998.
- [32] CHENG X R, LI L H, DU L T. ADMM-based channel estimation for mmWave massive MIMO systems[C]//Proceedings of the 2017 3rd IEEE International Conference on Computer and Communications (ICCC). Piscataway: IEEE Press, 2017: 152-157.
- [33] LIU H, ZHANG J Y, WU Q Q, et al. ADMM based channel estimation for RISs aided millimeter wave communications[J]. IEEE Communications Letters, 2021, 25(9): 2894-2898.

## [作者简介]



宋云超(1988-), 男, 博士, 南京邮电大学电子与光学工程学院、柔性电子(未来技术)学院副教授, 主要研究方向为大规模MIMO、智能反射面、强化学习。



李浩(1999-), 男, 南京邮电大学电子与光学工程学院、柔性电子(未来技术)学院硕士生, 主要研究方向为MIMO、多臂赌博机等。



梁汇彬(1998-), 男, 南京邮电大学电子与光学工程学院、柔性电子(未来技术)学院博士生, 主要研究方向为大规模MIMO通信、多臂赌博机、卫星通信等。



汪革(1994-), 女, 南京邮电大学电子与光学工程学院、柔性电子(未来技术)学院博士生, 主要研究方向为无人机通信、多臂赌博机等。