



研究与开发

基于深度强化学习的毫米波大规模 MIMO 系统资源联合优化

刘庆利^{1,2}, 李晓宇^{1,2}, 李蕊^{1,2}

(1. 大连大学信息工程学院, 辽宁 大连 116622;

2. 大连大学通信与网络重点实验室, 辽宁 大连 116622)

摘要: 针对毫米波大规模多输入多输出 (multiple-input multiple-output, MIMO) 系统中无线资源受限、功耗巨大且能量效率与系统容量相互制约而导致吞吐量和能量效率低的问题, 提出一种基于深度强化学习 (deep reinforcement learning, DRL) 的资源联合优化方法。该方法采用三阶段策略, 首先, 构建射频波束成形器, 通过少量的射频链降低硬件成本和总功耗; 其次, 利用有效信道状态信息设计基带预编码器; 最后, 设计应用双层 DRL 架构实现动态离散带宽和连续功率资源的分配。实验结果表明, 与单级全数字预编码和混合预编码的均等资源分配方法及基于粒子群优化的资源分配算法相比, 所提的联合优化方法显著提高了系统的吞吐量和能量效率。

关键词: 毫米波大规模 MIMO 系统; 资源分配; 联合优化

中图分类号: TN929.5

文献标志码: A

doi: 10.11959/j.issn.1000-0801.2024217

Deep reinforcement learning-based resource joint optimization for millimeter-wave massive MIMO systems

LIU Qingli^{1,2}, LI Xiaoyu^{1,2}, LI Rui^{1,2}

1. School of Information Engineering, Dalian University, Dalian 116622, China

2. Key Laboratory of Communication and Network, Dalian University, Dalian 116622, China

Abstract: Aiming at the problem of low throughput and energy efficiency caused by limited wireless resources, huge power consumption, and mutual constraints between energy efficiency and system capacity in millimeter-wave large-scale multiple-input multiple-output systems, a resource co-optimization method based on deep reinforcement learning was proposed. The method was adopted in a three-stage strategy, firstly, an RF beamformer was constructed to reduce the hardware cost and total power consumption through a small number of RF chains; secondly, a baseband precoder was designed using the effective channel state information; and finally, a two-tier deep reinforcement learning architecture was designed and applied to realize dynamic discrete bandwidth and continuous power resource allocation.

收稿日期: 2024-06-06; 修回日期: 2024-09-30

基金项目: 国家自然科学基金资助项目 (No.61931004)

Foundation Item: The National Natural Science Foundation of China (No.61931004)



tion. Experimental results show that the proposed joint optimization method significantly improves the throughput and energy efficiency of the system compared with the single-stage all-digital precoding and hybrid precoding equal resource allocation methods and the particle swarm optimization-based resource allocation algorithm.

Key words: millimeter-wave massive MIMO system, resource allocation, joint optimization

0 引言

毫米波通信是一种高频率的无线通信技术, 具有更高的带宽和更短的传输距离, 与传统的无线通信技术相比, 其优势明显。随着用户对多样化业务需求的急剧增加, 毫米波通信将在5G及未来的无线通信系统中受到越来越多的关注。然而, 毫米波通信也面临一些挑战, 其较短的波长和较高的大气吸收, 会导致严重的路径损耗^[1]。为了减轻毫米波通信中的路径损耗带来的影响, 需要形成高增益的定向波束, 这可以通过引入大规模多输入多输出 (multiple-input multiple-output, MIMO) 技术来实现^[2]。为了充分获得大规模MIMO技术提供的增益, 每根天线都需要使用专用的射频 (radio frequency, RF) 链, 然而, 由于空间受限, 毫米波难以实现大量的RF链部署, 这会导致成本高昂且系统功耗巨大, 在整个系统总功耗中占比高达70%。为了降低毫米波大规模MIMO系统的能耗, 业内提出了混合预编码 (hybrid precoding, HP) 方法, 需要的射频链路数量更少, 系统复杂度以及能耗更低。HP被认为是大规模MIMO系统中非常有前景的技术^[3]。两级HP架构通过有限数量的RF链互连RF和基带级^[4-7]。因此, 与单级全数字预编码 (fully digital precoding, FDP) 相比, HP显著提高了能量效率。另一方面, 宽带信号传输被认为可以实现极高的数据速率, 特别是在毫米波通信中。为了解决频率选择性问题, 正交频分复用 (orthogonal frequency division multiplexing, OFDM) 技术被广泛用于宽带信号传输^[8]。文献[9-11]针对大规模MIMO系统研究了基于OFDM的HP方

案, 为所有子载波设计一个RF波束成形器, 而基带预编码器的数量应与子载波的数量相同。在评估资源分配方案的有效性时, 系统容量作为传统性能指标已有大量研究支持, 同时, 由于绿色通信的推动, 能量效率的重要性日益凸显。随着天线数量的增加, 系统容量显著提升, 但能量效率达到一定程度可能会出现下降, 这两者之间相互制约。因此毫米波大规模MIMO系统在以系统容量及能量效率为性能指标的资源联合优化方面仍然值得深入研究。

文献[12]研究了大规模MIMO下行系统的功率分配问题, 以效益最大化为目标, 提出了一种低复杂度的功率分配算法, 该方法考虑了大规模MIMO系统的能耗过大问题, 提出能效优化目标减少系统功耗, 但是该方法依赖于构建的空间相关信道模型, 这意味如果实际环境中的信道特性与模型假设有较大差异, 算法性能可能会显著下降。文献[13]研究毫米波大规模MIMO系统中功率资源分配问题, 以用户服务质量 (quality of service, QoS) 为约束条件, 旨在最大化能量效率。该研究提出了一种迭代优化求解法来制定功率分配策略。尽管此方法在提升系统能效方面表现出色, 但它局限于功率资源的分配, 缺乏数据驱动的特性, 因而计算过程相对复杂。文献[14]研究了基于迫零 (zero forcing, ZF) 预编码的下行链路大规模MIMO系统中的能量效率资源分配问题。首先, 假设信道状态信息 (channel state information, CSI) 是完美的, 推导了遍历和速率的近似表达式。其次, 提出一种以系统能效最大化为目标的能效资源分配算法。尽管所提算法接近穷举算法的性能, 但它仅限于功率资源的单一

分配,并且迫零预编码的方式使得系统功耗过大。文献[15]指出,现有大规模 MIMO 系统通常采用单一的预编码方式,如单纯的数字预编码或模拟预编码方式,这不能很好地兼顾系统与个人的性能,所以为了在降低用户干扰的同时提升性能,需要采用混合预编码方式。文献[16]提出了一种基于群智能的功率分配方法,适用于大规模 MIMO 系统,并采用混合预编码架构降低系统功耗并提高能量效率。然而,该方法仅专注于功率资源的分配,未考虑系统中的其他资源。随着深度学习的蓬勃发展,相较于传统的非数据驱动技术,越来越多的学者开始投身于基于数据驱动的深度学习资源分配技术的研究中^[17-18]。然而,这些基于深度学习的技术通常需要大量标记数据集的支撑,这在一定程度上限制了其在实际应用中的普及和效果。为了应对资源管理的复杂挑战,研究者将目光投向了强化学习领域。将经典 Q 学习算法与深度神经网络结合,提出一种名为深度 Q 网络(deep Q-network, DQN)的深度强化学习(deep reinforcement learning, DRL)算法提升模型训练的稳定性和效率,在解决非凸优化问题方面展现出独特优势。传统优化方法往往难以处理高维非凸问题,容易陷入局部最优。而 DRL 通过探索-利用策略和神经网络的强大表达能力,能够在复杂的解空间中找到更优的解,尤其适合处理动态、非线性的系统优化问题。研究表明,DRL 方法在功率控制、频谱分配、数据缓存与卸载以及接入控制等关键问题展现出巨大潜力。文献[19]采用深度 Q 学习算法对目标网络中的功率资源进行高效分配,旨在实现每个用户请求满足的同时最小化总能耗。而文献[20]运用 DRL 算法,在异构网络的下行链路中进行用户调度和资源分配,以提高网络资源利用效率。文献[21]为实现系统吞吐量最大化的目标,巧妙运用多智能体 DRL 方法,对智能无线网络中的功率资源进行合理分配。在文献[22]中,作者借助 DRL 方法,

有效地分配通信网络中的频谱资源,以实现每个用户的效用值最大化。然而,这些方法主要聚焦于单一资源分配,如功率、频谱、子信道等,尚未全面考虑多种资源的联合分配策略。文献[23]提出了基于非正交多址系统的多资源分配方案,采用 DRL 方法,旨在最大化能效并对子信道和功率资源进行研究。尽管该方法在系统能效方面表现良好,但在系统吞吐量和多个用户干扰管理方面仍有改进空间。综上所述,尽管已经有大量研究将 DRL 应用到不同的领域中解决资源分配优化问题,但其应用于毫米波大规模 MIMO-OFDM 系统仍面临挑战,需要更多针对该特定系统的定制化优化方法。

针对上述问题,提出一种基于 DRL 算法的资源联合分配算法(resource allocation algorithm based on deep reinforcement learning, DRL-RA),以在毫米波大规模 MIMO-OFDM 系统中实现节能吞吐量最大化。本文的创新点和贡献如下。

(1) 双层 DRL 架构设计:上层采用 DQN 进行离散的带宽分配,下层采用近端策略优化(proximal policy optimization, PPO)算法进行连续的功率分配,实现了动态带宽和功率资源的联合优化。这种架构充分利用了 DQN 处理离散动作空间和探索-利用平衡的优势,以及 PPO 在连续动作空间中的稳定性和效率,能够有效处理高维度状态空间和混合(离散-连续)动作空间,这在传统方法中较难实现。

(2) RF 射频波束成形器构建:通过少量射频链设计射频波束成形器,显著降低硬件成本和总功耗。基于角度的混合预编码(angular based hybrid precoding, AB-HP)技术有效利用角度信息(angle of departure, AoD),进一步提升系统性能。

(3) 基带预编码器的优化:利用有效的信道状态信息为每个子载波构建基带预编码器,提升了系统的预编码效率和性能。

(4) 动态资源分配的智能化:通过所提出



的DRL架构,在用户之间有效地分配带宽和功率资源,实现吞吐量和能量效率的显著提升。实验结果表明,与传统的单级全数字预编码架构和混合预编码架构下的均等资源分配方法以及基于粒子群优化的资源分配算法相比,本文方法在系统吞吐量和能量效率方面具有明显的优势。

该方法适用于需要高能量效率和高系统吞吐量的毫米波大规模MIMO-OFDM系统,特别是在资源有限和硬件成本敏感的通信环境中,如下一代无线通信网络(5G及未来6G网络)、高速无线传输系统和大规模物联网(Internet of things, IoT)。

1 问题建模

大规模MIMO系统分为单用户MIMO和多用户MIMO,目前研究重点逐渐转向多用户MIMO,本文的研究在毫米波大规模多用户MIMO系统下展开。带有混合预编码的大规模MIMO系统如图1所示。

毫米波大规模MIMO系统通过 C 个子载波进行基于OFDM的下行传输。位于 G 组中的 K 个单天线用户由一个具有均匀矩形阵列(uniform rectangular array, URA)的基站与 $M=M_x \times M_y$ 个天线提供服务。基站利用 N_{RF} 个射频链连接基带级和RF级,每个子载波 i ($i=1,2,\dots,C$)的数据信号 $\mathbf{d}[i]$ 分别经过功率与带宽分配块 $\mathbf{P}[i]=$

$\text{diag}(\sqrt{p_{1,i}}, \sqrt{p_{2,i}}, \dots, \sqrt{p_{k,i}}) \in \mathbb{R}^{k \times k}$ 、 $\mathbf{Q}[i]=\text{diag}(\sqrt{q_{1,i}}, \sqrt{q_{2,i}}, \dots, \sqrt{q_{k,i}}) \in \mathbb{R}^{k \times k}$ 和数字基带预编码器 $\mathbf{B}[i]=[b_{1,j}, b_{2,j}, \dots, b_{k,j}] \in \mathbb{C}^{N_{\text{RF}} \times K}$,其中 $p_{k,i}$ 是子载波 i 处为第 k 个用户分配的功率; $q_{k,i}$ 是子载波 i 处为第 k 个用户分配的带宽。在应用快速傅里叶逆变换并添加循环前缀后,使用低成本移相器开发模拟RF波束成形器 $\mathbf{F} \in \mathbb{C}^{M \times N_{\text{RF}}}$ [24]。在子载波 i 处传输的数据信号为:

$$\mathbf{s}[i] = \mathbf{F}\mathbf{B}[i]\mathbf{P}[i]\mathbf{d}[i] \in \mathbb{C}^M \quad (1)$$

考虑URA结构和毫米波信道模型,第 k 个用户在子载波 i 处的信道向量定义为:

$$\mathbf{h}_k^T[i] = \sum_{l=1}^L T_{k,l,i}^\eta Z_{k,l,i} \boldsymbol{\phi}^T(\gamma_{x,k,l,i}, \gamma_{y,k,l,i}) e^{-j\frac{2\pi d_i}{C}} = \mathbf{Z}_k^T[i] \boldsymbol{\phi}_k[i] \quad (2)$$

其中, L 为路径数, $T_{k,l,i}$ 和 $Z_{k,l,i} \sim \text{CN}(0, \frac{1}{L})$ 分别是第 l 条路径在子载波 i 处的距离和复路径增益, η 是路径损耗指数, $\boldsymbol{\phi}() \in \mathbb{C}^M$ 是相位响应矢量, $\gamma_{x,k,l,i} = \sin(\theta_{k,l,i}) \cos(\psi_{k,l,i})$ 和 $\gamma_{y,k,l,i} = \sin(\theta_{k,l,i}) \sin(\psi_{k,l,i})$ 为反映相应路径的高度角AoD(elevation AoD, EAoD)与方位角AoD(azimuth AoD, AAoD)的系数。然后,将相位响应矢量表示为:

$$\boldsymbol{\phi}(\gamma_x, \gamma_y) = \left[1, e^{-j2\pi d_x \gamma_x}, e^{-j2\pi d_y \gamma_y}, \dots, e^{-j2\pi d(M_x-1)\gamma_x} \right]^T \otimes \left[1, e^{-j2\pi d_y \gamma_y}, e^{-j2\pi d_2 \gamma_y}, \dots, e^{-j2\pi d(M_y-1)\gamma_y} \right]^T \in \mathbb{C}^M \quad (3)$$

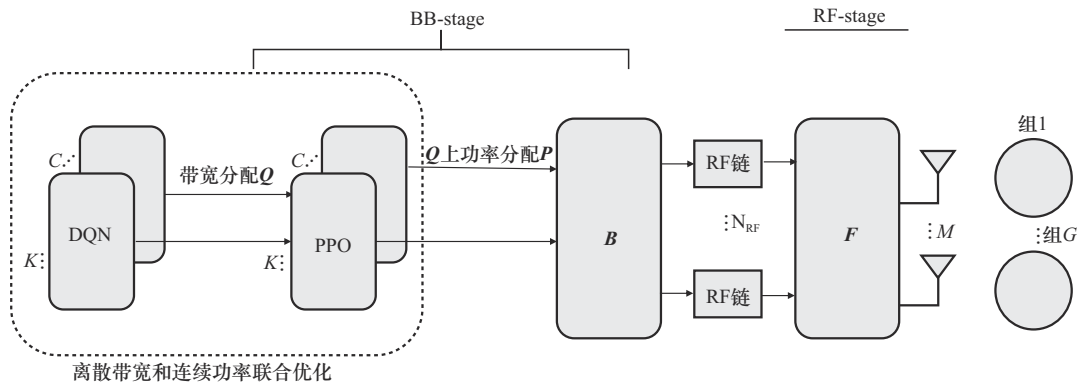


图1 带有混合预编码的大规模MIMO系统

其中, d 为由波长归一化的天线元件间距。

由式 (1) 和式 (2) 可知, 在子载波 i 处的第 k 个用户处接收到的信号为:

$$r_{k,i} = \mathbf{h}_k^T[i] \mathbf{s}[i] + \omega_{k,i} = \mathbf{h}_k^T[i] \mathbf{F} \mathbf{B}[i] \mathbf{P}[i] \mathbf{d}[i] + \omega_{k,i} = \sqrt{p_{k,i}} \mathbf{h}_k^T[i] \mathbf{F} \mathbf{b}_{k,i} d_{k,i} + \sum_{t \neq k}^K \sqrt{p_{t,i}} \mathbf{h}_k^T[i] \mathbf{F} \mathbf{b}_{t,i} d_{t,i} + \omega_{k,i} \quad (4)$$

其中, $\omega_{k,i} \sim \text{CN}(0, \sigma^2)$ 是高斯噪声。然后, 推导子载波 i 处的瞬时和速率 (即吞吐量) 表达式如下。

$$R_{\text{sum},i}(\mathbf{F}, \mathbf{B}[i], \mathbf{P}[i], \mathbf{Q}[i]) = \sum_{k=1}^K q_{k,i} \cdot \text{lb} \left(1 + \frac{p_{k,i} |\mathbf{h}_k^T[i] \mathbf{F} \mathbf{b}_{k,i}|^2}{\sum_{t \neq k}^K p_{t,i} |\mathbf{h}_k^T[i] \mathbf{F} \mathbf{b}_{t,i}|^2 + q_{k,i} \sigma^2} \right) \quad (5)$$

C 个子载波的总和速率为 $R_{\text{sum}} = \sum_{i=1}^C E\{R_{\text{sum},i}(\mathbf{F}, \mathbf{B}[i], \mathbf{P}[i], \mathbf{Q}[i])\}$, 每个子载波的平均和速率为 $\frac{R_{\text{sum}}}{C}$ 。因此, 将吞吐量最大化问题表述如下。

$$\left\{ \begin{array}{l} \max_{\mathbf{F}, \mathbf{B}[i], \mathbf{P}[i], \mathbf{Q}[i]} R_{\text{sum},i}(\mathbf{F}, \mathbf{B}[i], \mathbf{P}[i], \mathbf{Q}[i]) \\ \text{s.t. } E\left\{\|\mathbf{S}[i]\|_2^2\right\} = \\ \sum_{k=1}^K p_{k,i} \|\mathbf{F} \mathbf{b}_{k,i}\|_2^2 \leq \frac{P_T}{C}, p_{k,i} \geq 0, \forall k \\ \|\mathbf{F}\|_{m,n} = \frac{1}{\sqrt{M}}, \forall m, n \\ \sum_{k=1}^K q_{k,i} \leq \delta_i, \forall i \end{array} \right. \quad (6)$$

其中, $R_{\text{sum},i}(\mathbf{F}, \mathbf{B}[i], \mathbf{P}[i], \mathbf{Q}[i])$ 是式 (5) 中定义的子载波 i 处的总速率。为了确保载波不会溢出其物理容量, 限制了每个载波最大负载总带宽为 δ_i , 其中 3 个约束条件分别表示每个子载波的发射功率约束, 在射频级使用移相器而引起的恒定模量约束以及每个子载波的带宽约束是非凸优化问题, 因此首先通过 AB-HP 设计射频波束成形器和基带预编码器, 然后提出基于 DRL 方法的带宽与功率联合分配方法。

2 混合预编码

为毫米波大规模 MIMO-OFDM 系统开发基于角度的混合预编码体系结构, 以减少射频链的数量, 降低硬件成本和总功耗。

2.1 射频波束成形器

考虑位于 G 组中的 K 个用户, G 组在子载波 i 处的信道矩阵为:

$$\mathbf{H}_g[i] = [h_1[i], h_2[i], \dots, h_{K_g}[i]]^c = \mathbf{Z}_g[i] \boldsymbol{\phi}_g[i] \in \mathbb{C}^{K_g \times M} \quad (7)$$

其中, $\mathbf{Z}_g[i] = [z_1[i], z_2[i], \dots, z_{K_g}[i]]^T \in \mathbb{C}^{K_g \times L}$, $\boldsymbol{\phi}_g[i] \in \mathbb{C}^{L \times M}$, g 为组数 ($g \in [1, G]$), K_g 是第 g 组的用户数, $K = \sum_{g=1}^G K_g$, 将子载波 i 处的全尺寸信道矩阵定义为 $\mathbf{H}[i] = [\mathbf{H}_1^T[i], \mathbf{H}_2^T[i], \dots, \mathbf{H}_G^T[i]]^T \in \mathbb{C}^{L \times M}$ 。假设同一组中的用户经历了与文献[5-7]中相似的 AoD。因此, 为 G 组用户设计的射频波束成形器如下。

$$\mathbf{F} = [\mathbf{F}_1, \mathbf{F}_2, \dots, \mathbf{F}_G] \in \mathbb{C}^{M \times N_{\text{RF}}} \quad (8)$$

其中, $\mathbf{F}_g \in \mathbb{C}^{M \times N_{\text{RF}}}$ 是第 g 组的射频波束成形器, $N_{\text{RF}} = \sum_{g=1}^G N_{\text{RF},g}$ 。利用式 (7) 和式 (8) 可得到基带级的有效信道矩阵为:

$$\hat{\mathbf{H}}[i] = \mathbf{H}[i] \mathbf{F} = \begin{bmatrix} H_1[i] F_1 & H_1[i] F_2 & \dots & H_1[i] F_G \\ H_2[i] F_1 & H_2[i] F_2 & \dots & H_2[i] F_G \\ \vdots & \vdots & \ddots & \vdots \\ H_G[i] F_1 & H_G[i] F_2 & \dots & H_G[i] F_G \end{bmatrix} \in \mathbb{C}^{K \times N_{\text{RF}}} \quad (9)$$

其中, 对角块矩阵 $\mathbf{H}_g[i] \mathbf{F}_g = \mathbf{Z}_g[i] \boldsymbol{\phi}_g[i] \mathbf{F}_g \in \mathbb{C}^{K_g \times N_{\text{RF},g}}$ 是有效通道矩阵, 非对角块矩阵 $\mathbf{H}_t[i] \mathbf{F}_g = \mathbf{Z}_t[i] \boldsymbol{\phi}_t[i] \mathbf{F}_g \in \mathbb{C}^{K_t \times N_{\text{RF},g}}$ 是有效干扰通道矩阵, $\forall t \neq g$ 。因此, 射频波束成形器的设计目标是实现以下两个目标: 在期望的方向上最大化波束成形增益; 成功抑制用户组间的干扰。正如



文献[5]中所证明的, 这两个目标都是通过转向矢量 $\mathbf{e}(\gamma_x, \gamma_y) = \frac{1}{\sqrt{M}} \phi(\gamma_x, \gamma_y) \in \mathbb{C}^M$, 来实现射频波束

成形器 $\mathbf{F}_g$ 的角度构建, 其中 (γ_x, γ_y) 为角度对。为了用最少数量的角度对覆盖完整的高度角和方位角空间, M 个正交量化角度对定义为 $\lambda_u^x = -1 + \frac{2u-1}{M_x}$, $u=1, 2, \dots, M_x$; $\lambda_n^y = -1 + \frac{2n-1}{M_y}$, $n=1, 2, \dots, M_y$ 。

M_y 为 g 组构建射频波束成形器如下:

$$\mathbf{F}_g = [e(\lambda_{u1}^x, \lambda_{n1}^y), e(\lambda_{u2}^x, \lambda_{n2}^y), \dots, e(\lambda_{uN_{RF,g}}^x, \lambda_{nN_{RF,g}}^y)] \in \mathbb{C}^{M \times N_{RF,g}} \quad (10)$$

最后, 通过将式 (10) 代入式 (8), 推导出满足式 (6) 中给出的 CM 约束条件的完整射频预编码器 $\mathbf{F}$ 。

2.2 基带预编码器

为每个子载波开发了 C 个不同的基带预编码器, 主要目的是进一步抑制剩余的用户间干扰。利用简化尺寸的有效信道矩阵 $\hat{\mathbf{H}}[i]$, 对每个子载波采用正则化的零强迫技术。因此, 在子载波 i 处的基带预编码器被定义为:

$$\mathbf{B}[i] = [\hat{\mathbf{H}}[i]^H \mathbf{H}[i] + K\alpha \mathbf{I}]^{-1} (\hat{\mathbf{H}}[i])^H \in \mathbb{C}^{N_{RF} \times K} \quad (11)$$

其中, $\alpha = \frac{\sigma^2}{P^T}$ 为正则化参数。

3 基于 DRL 算法的资源联合优化

3.1 DRL 算法

推导出所有子载波的射频波束成形器 $\mathbf{F}$ 和基带预编码器 $\mathbf{B}[i]$ 的封闭解后, 式 (6) 中给出的吞吐量最大化问题变成了一个资源分配问题 (即在用户之间分配功率和带宽资源)。传统上, 此类问题通常采用凸优化算法求解。然而, 考虑毫米波大规模 MIMO 系统的复杂性和动态性, 凸优化方法面临着计算复杂度高、难以适应快速变化的信道条件和用户需求等挑战。而 DRL 结合了强

化学习和深度学习, 侧重于提升强化学习的决策能力。面对复杂环境, 传统决策方法难以应对, 而神经网络的强大拟合能力可以替代原有的决策功能。为此, 本文提出用 DRL 算法进行求解。设计一个双层的 DRL 架构, 其中包括 DQN 和 PPO 算法。上层 DQN 做离散的带宽分配, 下层 PPO 做连续的功率分配, 实现动态带宽和功率资源的联合分配优化。

DRL 算法包括 3 个要素: 状态、动作、奖励函数, 只有根据问题进行合理设置, 才能充分发挥 DRL 算法的优势, 达到预期的效果。本文将这三者设计如下。

(1) 状态集合 S

状态空间 S 表示系统当前的状态信息, 主要包括信道信息、每个用户采用的射频波束成形器以及基带预编码器, 将学习过程的第 t 个传输时间对应的状态记为:

$$s^t = \{\mathbf{H}_i^t[i], \mathbf{F}^t, \mathbf{B}^t[i]\} \quad (12)$$

其中, PPO 网络的输入状态集包含上层 DQN 的输出, 即当前时刻 DQN 在子载波 i 上为各用户分配的带宽 $\mathbf{Q}[i]$, 因此 PPO 的 t 时刻状态为:

$$s^t = \{\mathbf{H}_i^t[i], \mathbf{F}^t, \mathbf{B}^t[i], \mathbf{Q}^t[i]\} \quad (13)$$

(2) 动作集合 A

在毫米波大规模 MIMO 系统中, 动作空间 A 表示系统在时间步 t 所采取的动作, 包括带宽分配和功率分配。本文设计的双层 DRL 架构中, 上层 DQN 用于离散带宽分配 $\mathbf{Q}^t[i]$, 下层 PPO 用于连续功率分配 $\mathbf{P}^t[i]$ 。因此, 第 t 个传输时间对应的动作记为:

$$a^t = \{\mathbf{P}^t[i], \mathbf{Q}^t[i]\} \quad (14)$$

(3) 奖励设置 R

奖励函数用于评价带宽和功率分配的效果, 反映智能体与环境交互过程中采取动作之后环境给予智能体的反馈。该函数的设计直接关系智能

体长期的学习和收益。将式 (6) 设置为奖励函数, 旨在最大化系统容量, 将学习过程的第 t 个传输时间对应的奖励记为 r^t 。

$$r^t = R_{\text{sum}, i}(\mathbf{F}, \mathbf{B}[i], \mathbf{P}[i], \mathbf{Q}[i]) \quad (15)$$

在 DRL 中, 智能体通过与环境的互动, 不断调整策略以最大化累积奖励。智能体通过试错机制, 在环境中采取动作, 并根据环境反馈 (奖励和新状态) 更新策略。这个过程通过神经网络的强大拟合能力来实现, 代替了传统的决策功能。

给定时间 t , 智能体的目标是找到一个最优策略, 在给定的时间段后使体现的累积奖励最大化:

$$G^t = \sum_{v=t}^T \gamma^{v-t} r_v \quad (16)$$

其中, T 为事件的长度, r_v 为在时刻 v 获得的奖励, γ 为折扣因子。智能体会考虑当前时间的奖励和未来的奖励。

3.2 DQN+PPO 双层 DRL 架构

根据前文描述, 明确了优化目标, 将通过 DRL 架构实现动态带宽和功率资源的联合优化。具体来说, DQN 和 PPO 算法采用双层 DRL 架构, 如图 2 所示。DQN 用于离散的带宽分配, 而 PPO 算法用于连续的功率分配。以下是 DQN 和 PPO 算法的详细函数定义及其必要参数。

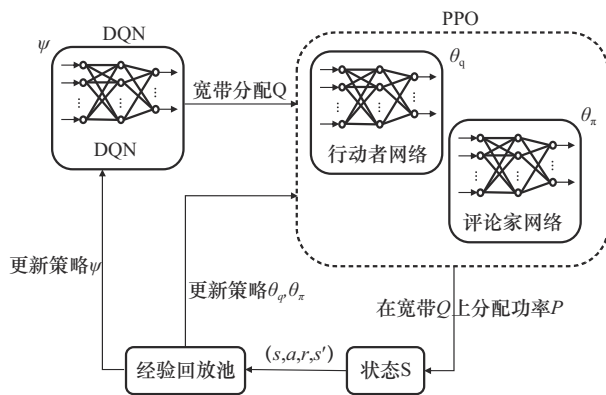


图2 DQN和PPO算法双层DRL架构

DQN算法能在探索和利用之间找到良好的平衡, 用于在给定状态下选择离散的带宽分配动作。根据定义的状态空间和动作空间, 通过贝尔曼方程和梯度下降方法进行网络参数更新。贝尔曼方程用于描述当前状态的价值函数与下一个状态的价值函数之间的关系。具体来说, 状态 s 下采取动作 a 的预期回报 $Q(s, a)$ 可以表示为:

$$Q(s, a) = E_s[r + \gamma \max_{a'} Q(s', a') | s, a] \quad (17)$$

其中, $Q(s, a)$ 是在状态 s 下采取动作 a 的预期回报, r 是即时奖励, γ 是折扣因子, 权衡当前奖励和未来奖励的重要性, s' 是下一个状态, a' 是在下一个状态 s' 下的动作选择策略。

为了优化 DQN 参数, 使用梯度下降方法最小化均方误差 (mean square error, MSE) 损失函数:

$$L(\psi) = E_{s, a, r, s'} \left\{ \left[Q(s, a; \psi) - (r + \gamma \max_{a'} Q(s', a'; \psi')) \right]^2 \right\} \quad (18)$$

其中, ψ 是网络参数, ψ' 是目标网络的参数, (s, a, r, s') 是从经验回放缓冲池中采样的经验数据。目标函数的更新 $\psi' \leftarrow \tau \psi + (1 - \tau) \psi'$ 是为了稳定训练过程, 防止频繁更新导致的不稳定性, τ 是更新率, 控制新旧网络参数的混合程度。

为了克服 DQN 在连续动作上的不足, 在下层使用 PPO 算法分配功率。除了能进行连续动作, PPO 算法通过限制策略更新的幅度 (使用剪裁策略比率) 来确保策略更新的稳定性, 也避免过大的策略更新可能导致的不稳定性问题。这种机制使得 PPO 算法在训练过程中更容易收敛到较好的策略, 尤其在复杂的非凸优化问题中表现优异。

PPO 算法同时训练一个行动者 (actor) 网络和一个评论家 (critic) 网络。行动者网络的任务是学习策略函数 $\pi_{\theta_q}(a|s)$, 其中 θ_q 是网络参数, 它



决定在状态 s 下采取动作 a 的概率分布。在行动者网络中，并不直接使用传统的贝尔曼函数计算动作的长期回报。行动者网络通过梯度上升优化策略梯度，使得在当前策略下选择动作的期望奖励最大化。评论家网络的任务是学习值函数 $V_{\theta_{\pi}}(s)$ ，这里 θ_{π} 是网络的参数，用于估计状态 s 下的预期回报或价值。贝尔曼函数用于估计当前状态的值函数 $V_{\theta_{\pi}}(s) = Ea \sim \pi[Q(s, a)]$ 。

PPO 算法使用梯度下降更新策略网络的参数，通过最小化损失函数来优化策略。策略函数的损失函数使用了剪裁比率 $r_t(\theta)$ 和优势估计 $\hat{A}_t$ ，限制策略更新的幅度以确保稳定性：

$$L(\theta) = E_{s,a,r,s'}[\min(r_t(\theta)\hat{A}_t, \text{clip}(r_t(\theta), 1-\epsilon, 1+\epsilon)\hat{A}_t)] \quad (19)$$

其中， ϵ 是剪裁范围， $r_t(\theta) = \frac{\pi_{\theta_{\text{new}}}(a_t|s_t)}{\pi_{\theta_{\text{old}}}(a_t|s_t)}$ 是决策比率， $\hat{A}_t$ 是优势估计。同样，PPO 算法使用目标策略函数 $\theta' \leftarrow \tau\theta + (1-\tau)\theta'$ 定期更新目标策略网络参数。

DQN+PPO 算法联合带宽和功率优化分配如算法 1 所示。

算法 1 DQN+PPO 算法联合带宽和功率优化分配

初始化环境，包括射频波束成形矩阵 $\mathbf{F}$ 和基带预编码矩阵 $\mathbf{B}$

初始化超参数 total_episodes 、 ϵ # ϵ 为探索率

初始化 DQN 参数 ψ

初始化 行动者网络参数 θ_q

初始化 评论家网络参数 θ_{π}

for episode in range(total_episodes):

DQN 部分：确定带宽分配

while True:

根据探索率 ϵ_{dqn} 随机选择带宽或从 DQN 中选择最优带宽

if $\text{np.random.uniform}(0, 1) < \epsilon_{\text{dqn}}$:

带宽分配 $q = \text{env.sample_bandwidth}()$ # 随机选择带宽

else:

$q = \text{dqn.act}(s, \psi)$ # 根据 DQN 选择最优带宽

执行带宽分配，获取下一状态和奖励

$s', r = \text{env.step_bandwidth}(q)$

将状态转移存入经验回放池

$\text{replay_buffer.append}((s, a, r, s'))$

更新 DQN 参数

$\psi = \text{optimize_dqn}(\text{replay_buffer}, \psi)$

$s = s'$ # 更新状态

if $\text{env.done_bandwidth}()$: # 如果带宽分配阶段结束

break

PPO 部分：确定功率分配

while True:

if $\text{np.random.uniform}(0, 1) < \epsilon_{\text{ppo}}$: # 以 ϵ 的概率随机选择动作进行探索

功率分配 $p = \text{env.action_space.sample}()$

else:

$p = \text{np.argmax}(Q(s, a; \theta_q))$ # 根据行动者网络选择最优功率

$s', r = \text{env.step}(p, q)$ # 执行动作，获取下一状态和奖励

$\text{replay_buffer.append}((s, a, r, s'))$ # 将状态转移加入经验回放池

更新 PPO 网络参数

计算优势函数 $\hat{A}_t = Q(s, a) - V(s)$

计算旧策略动作概率 $\pi_{\theta_{\text{old}}}(a_t|s_t)$

计算新策略动作概率 $\pi_{\theta_{\text{new}}}(a_t|s_t)$

计算概率比 $r_t(\theta) = \frac{\pi_{\theta_{\text{new}}}(a_t|s_t)}{\pi_{\theta_{\text{old}}}(a_t|s_t)}$

if $r < 1$:

增大学习率

$\text{actor.increase_lr}()$

$\text{critic.increase_lr}()$ # 如果新策略比旧策略差，则增大学习率

计算损失 $L(\theta) = E_{s,a,r,s'}[\min(r_t(\theta)\hat{A}_t, \text{clip}(r_t(\theta), 1-\varepsilon, 1+\varepsilon)\hat{A}_t)]$

$\theta_\pi = \text{optimize}(L, \theta_\pi)$ # 更新行动者网络参数

$\theta_q = \text{optimize}(L, \theta_q)$ # 更新评论家网络参数

$s = s'$ # 转移到下一个状态

if done: # 如果当前回合结束

break

周期性记录训练过程相关指标

if episode % log_interval == 0:

log_stats(episode, replay_buffer)

4 实验仿真与分析

本文采用典型的毫米波大规模 MIMO 系统作为仿真场景。系统包括一个基站和多个用户终端。基站配备有大规模天线阵列，可以同时服务多个用户。系统采用 OFDM 调制，信道模型采用标准的毫米波信道模型。网络架构构建了双层的 DRL 算法架构，包括 DQN 和 PPO 算法两种。DQN 包括一个输入层、两个隐藏层和一个输出层。PPO 算法包括一个行动者网络和一个评论家网络，两个网络都有一个输入层、两个隐藏层和一个输出层。

为了验证所提出的基于 DRL 的资源联合优化方法的性能，进行了详细的仿真实验。将本文提出的算法与均等资源分配算法（equal RA，EQ-RA）、基于粒子群优化的资源分配算法（particle swarm optimization based RA，PSO-RA）^[7] 进行比较，其中，EQ-RA 算法中功率与带宽资源在用户之间平均分配。在 EQ-RA 方案中，研究了 FDP 和 HP 之间的性能比较。网络参数设置和算法参数设置见表 1、表 2。表 1 列出了所需网络仿真参数^[25]，考虑了配备有 $M=256$ 个天线单元的基站。

以总和速率 R_{sum} 与总功耗 P_{total} 之比来评估毫米波大规模 MIMO-OFDM 系统的能量效率^[5-6]：

表 1 网络仿真参数设置

仿真参数	参数取值
天线数量 M	$16 \times 16 = 256$
小区半径/m	100
用户与基站的水平距离/m	10 ~ 90
小组个数	$G = 3$
每个小组的用户个数	$K_g = \frac{K}{G}$
EAoD 相关参数	$\theta_g = 60^\circ, \delta_g^\theta = 15^\circ$
AAoD 相关参数	$\psi_g = 60^\circ + 120^\circ(g-1), \delta_g^\psi = 11^\circ$
路径损失指数 η	3.76
噪声 PSD/(dBm·Hz) ⁻¹	-174
信道带宽/kHz	10
路径数量 L	10
天线间距（波长） d	0.5
网络实现/次	1 000

表 2 算法参数设置

仿真参数	参数取值
DQN 学习率	0.001
PPO 学习率	0.000 3
广义优势估计系数	0.95
DQN 隐藏层	(128, 128)
PPO 隐藏层	(64, 64)
熵系数	0.2
折扣因子	0.99
探索策略	ε -greedy 策略 ($\varepsilon=0.999\ 5$)
小批量样本大小	64
激活函数	ReLU

$$\beta = \frac{R_{\text{sum}}}{P_{\text{total}}} = \frac{\sum_{i=1}^C E\{R_{\text{sum},i}(\mathbf{F}, \mathbf{B}[i], \mathbf{P}[i], \mathbf{Q}[i])\}}{P_T + N_{\text{RF}}P_{\text{RF}} + N_{\text{PS}}P_{\text{PS}}} \quad (20)$$

其中，系统功耗 P_{total} 由传输功率 P_T 、每个射频链（移相器）的功耗 $P_{\text{RF}}(P_{\text{PS}})$ 和射频链（移相器）的数量 $N_{\text{RF}}(N_{\text{PS}})$ 等几部分构成。假设 $P_{\text{RF}} = 250\text{ mW}$, $P_{\text{PS}} = 1\text{ mW}$ ^[9]，对于 FDP 结构，射频链路数量等于天线数（ $N_{\text{RF}} = M = 256$ ），并且不需要任何移相器（即 $N_{\text{PS}} = 0$ ）。而对于所提出的 AB-



HP 结构, 射频链路数量 $N_{RF}=36$, 移相器数量 $N_{PS}=N_{RF} \times M=9\,216$ 个。因此, 相比 FDP 结构, AB-HP 结构通过减少射频链的数量, 在满足 $M=256$ 根天线的前提下, RF 链的数量从 256 条减少到 36 条 (即硬件成本/复杂性和信道状态信息 (channel state information, CSI) 开销大小减少 87%), 有效降低了系统的总功耗, 这与本文提出的基于 DRL 的资源联合优化方法目标是一致的。

图 3 展示了各个算法在不同子载波数量上吞吐量对比, 设置 $P_T=40$ dBm、用户数 $K=15$ 。图 3 中的数值结果表明, 算法能随着子载波数量提升而提升系统吞吐量, 其中本文算法的提升效果最明显。与 EQ-RA 相比, 本文算法对 C 分别为 8、16、32、64 时子载波的吞吐量提高了 18.9%、41.4%、42.9% 和 45.58%。与 PSO-RA 算法相比, 本文所提算法在吞吐量方面也有所提升。从整体趋势来看, 本文算法的吞吐量曲线始终高于其他算法, 并且随着子载波数量的增加, 性能差距越来越明显。采用 EQ-RA 的 HP 吞吐量相较于 FDP 确实有所牺牲, 但是本文的双层 DRL 架构能根据系统状态和需求, 自适应地进行带宽和功率的联合优化, 弥补了 HP 在吞吐量方面的不足, 能更好地利用频谱资源提升系统性能, 在大规模多载波系统中的优势尤为突出。

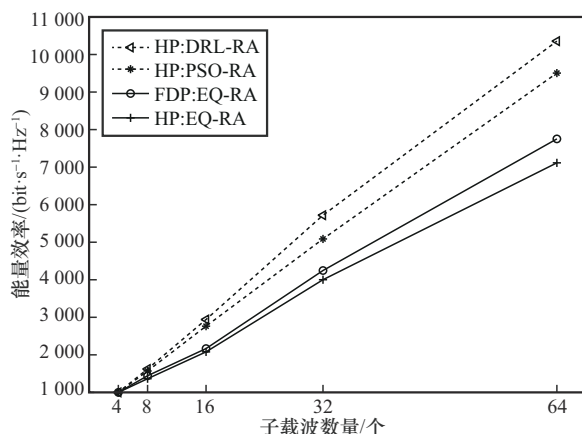


图3 各个算法在不同子载波数量上的系统吞吐量对比

图 4 展示了各个算法的系统能量效率对比, 同样设置 $P_T=40$ dBm、用户数 $K=15$ 。从数值结果来看, 所有算法的能量效率均随着子载波数量的增加而上升, 本文算法在各种数设置下都表现出最高的能效。值得注意的是, 与 FDP 方案相比, HP 架构可以显著提高能量效率, 这是因为采用 HP 架构可以通过减少射频链数量降低硬件成本, 提高能量效率。在子载波数 M 分别为 4、16、64 时本文算法与 EQ-RA 算法相比, 能效提高了 125%、269% 和 272%; 与 PSO-RA 算法相比, 能效提高了 3.8%、9.4% 和 10.3%。随着子载波数量的增加, 本文算法和 EQ-RA、PSO-RA 算法之间的性能差距不断扩大, 尤其是 EQ-RA, 显然这种均等分配方式和基于智能算法的能效优化效果远不如本文所设计的 DRL-RA 算法, 混合预编码架构相比全数字预编码架构表现出更高的能量效率。

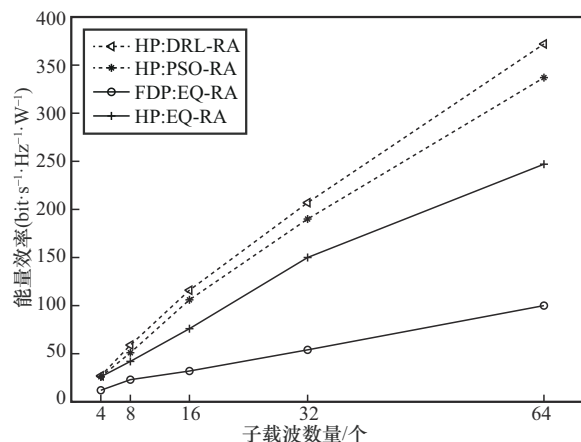


图4 各个算法的系统能量效率对比

此外, 为了进一步说明本文所提 DRL-RA 算法对于提升系统吞吐量方面的优势, 将 $P_T=20$ dBm、40 dBm、60 dBm, $K=15$ 下 DRL-RA 算法与其他算法的和速率增益比进行比较。将和速率增益比定义为 $\frac{R_{\text{sum, DRL-RA}}}{R_{\text{sum, 其他算法}}}$, 不同功率下各个算法在子载波数分别为 1、2、4、8、16、32、64 时的和速率增益比如图 5 所示。

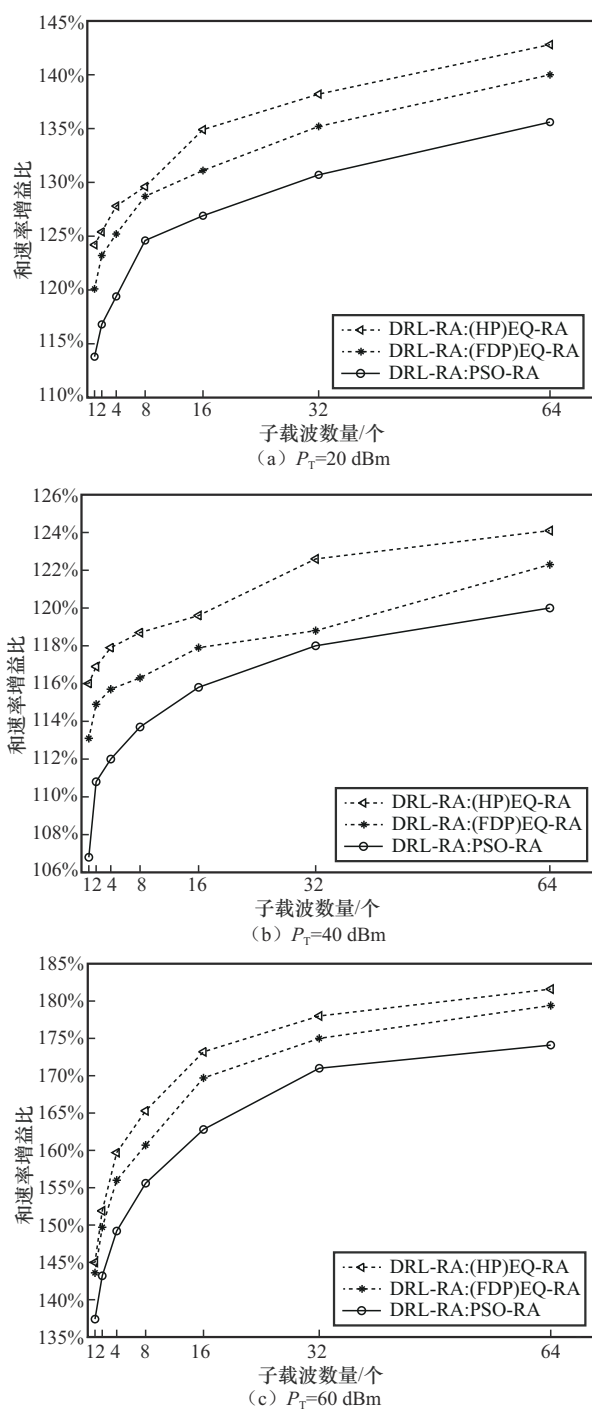


图5 不同功率下各个算法的和速率增益比

由于DRL-R A算法的优越性,和速率增益比均大于1。图5数据表明随着子载波数量的增多,DRL-R A算法与其他算法的和速率增益比也在不断增加,其中DRL-R A算法与采用HP架构的EQ-R A算法的和速率增益比最高。且随着发射功率

的增加,和速率增益比有所降低。所以这一事实的重要性在低发射功率的情况下显得更为突出。

表3是不同算法的时间复杂度对比。其中 M 是发射天线数, K 是用户数, N_{RF} 是射频链路数(通常 $N_{\text{RF}} \ll M$), T 是迭代次数, P 是粒子数, C 是子载波数, E 是训练的回合数, T 是每个回合的时间步数, H_1 和 H_2 分别是两个隐藏层的节点数, x 是环境模拟的复杂度,可以近似为 $O(MK + K \times C)$ 。

表3 不同算法时间复杂度对比

算法	时间复杂度
FDP: EQ-R A	$O(MK)$
HP: EQ-R A	$O(N_{\text{RF}} \times K + M \times N_{\text{RF}})$
HP: PSO-R A	$O(T \times P \times (MK + K \times C))$
HP: DPL-R A (训练)	$O(E \times T \times (H_1 + H_2 + x))$
HP: DPL-R A (推理)	$O(H_1 + H_2)$

从算法复杂度的角度来看,DRL-R A算法相比其他方法也展现出优势。首先,FDP: EQ-R A和HP: EQ-R A的时间复杂度相对较低,但它们是静态分配方法,无法适应动态变化的信道条件和用户需求。HP: EQ-R A由于使用了混合预编码,复杂度比FDP: EQ-R A稍高,但在实际系统中可能更节能。PSO-R A的时间复杂度随着迭代次数、粒子数和系统规模(天线数、用户数、子载波数)的增加而显著增加。尽管它找到较优解,但在大规模系统或需要快速决策的场景中不太适用。DRL-R A在训练阶段的时间复杂度较高,但这是离线进行的。训练完成后,其推理阶段的复杂度相对较低,主要取决于神经网络的规模。

DRL-R A的优势在于它能够学习复杂的环境动态和长期策略。虽然训练阶段耗时较长,但一旦训练完成,它能快速适应变化的环境,这使其更适合应用于实际的大规模无线通信系统。这种时间复杂度的改善有利于提高系统的响应速度和部署效率,在实时性要求高的场景中尤为重要。



5 结束语

本文提出了一种基于DRL的资源分配方法，用于实现毫米波大规模MIMO-OFDM系统中的节能吞吐量最大化。该方法利用基于角度的混合预编码方案，有效减少了射频链的数量和CSI开销。AB-HP方案先为所有子载波设计RF波束成形器，再通过有效缩减CSI为每个子载波构建不同的基带预编码器。最后，设计了双层DRL架构，上层采用DQN进行离散带宽分配，下层使用近端策略优化算法进行连续功率分配，有效处理了高维状态空间和混合动作空间的问题。仿真实验验证了所提方法在吞吐量和能量效率方面的显著优势。毫米波通信在实际应用中可能会面临更复杂多变的环境，例如移动通信、异构网络和动态干扰等。未来的研究可以探索该方法在更复杂场景下的适应性，以及如何提高系统的鲁棒性和可靠性。

参考文献:

- [1] ASHRAF S, SHEIKH J A, ASHRAF A, et al. 5G millimeter wave technology: an overview[M]//SHEIKH J A, KHAN T, KANAUIA B K, eds. Energy Systems in Electrical Engineering. Singapore: Springer Nature Singapore, 2024: 97-112.
- [2] ELIJAH O, ABDUL RAHIM S K, NEW W K, et al. Intelligent massive MIMO systems for beyond 5G networks: an overview and future trends[J]. IEEE Access, 2022, 10: 102532-102563.
- [3] AHMED I, KHAMMARI H, SHAHID A, et al. A survey on hybrid beamforming techniques in 5G: architecture and system model perspectives[J]. IEEE Communications Surveys & Tutorials, 2018, 20(4): 3060-3097.
- [4] KOC A, LE-NGOC T. Full-duplex mmWave massive MIMO systems: a joint hybrid precoding/combining and self-interference cancellation design[J]. IEEE Open Journal of the Communications Society, 2021(2): 754-774.
- [5] KOC A, MASMOUDI A, LE-NGOC T. 3D angular-based hybrid precoding and user grouping for uniform rectangular arrays in massive MU-MIMO systems[J]. IEEE Access, 2020, 8: 84689-84712.
- [6] MAHMOOD M, KOC A, LE-NGOC T. Energy-efficient MU-massive-MIMO hybrid precoder design: low-resolution phase shifters and digital-to-analog converters for 2D antenna array structures[J]. IEEE Open Journal of the Communications Society, 2021, 2: 1842-1861.
- [7] ZHENG W Y, KOC A, LE-NGOC T. Sub-connected hybrid precoding architectures in massive MIMO systems[C]//Proceedings of the GLOBECOM 2020-2020 IEEE Global Communications Conference. Piscataway: IEEE Press, 2020: 1-6.
- [8] HARKAT H, MONTEIRO P, GAMEIRO A, et al. A survey on MIMO-OFDM systems: Review of recent trends[J]. Signals, 2022, 3(2): 359-395.
- [9] BISHE F, KOC A, LE-NGOC T. Intelligent subcarrier allocation in hybrid beamforming multi-user mMIMO-OFDM systems[C]//Proceedings of the 2023 IEEE 97th Vehicular Technology Conference (VTC2023-Spring). Piscataway: IEEE Press, 2023: 1-5.
- [10] YILDIRIM I, KOC A, BASAR E, et al. RIS-Aided Angular-Based Hybrid Beamforming Design in mmWave Massive MIMO Systems[C]//Proceedings of the GLOBECOM 2022-2022 IEEE Global Communications Conference. Piscataway: IEEE Press, 2022: 5267-5272.
- [11] CHOI J, OH H. Hybrid precoding and combining based on rate balancing for wideband mmWave multiuser MIMO systems[J]. IEEE Access, 2023, 11: 12987-12999.
- [12] KAUR D, GUPTA G, JHA V. A game theoretic bandwidth allocation scheme towards improving the fairness of XG-PON systems[C]//Proceedings of the 2021 International Conference on Information and Communication Technology Convergence (ICTC). Piscataway: IEEE Press, 2021: 921-926.
- [13] ZENG M, HAO W M, DOBRE O A, et al. Energy-efficient power allocation in uplink mmWave massive MIMO with NOMA[J]. IEEE Transactions on Vehicular Technology, 2019, 68(3): 3000-3004.
- [14] PENG J B, HU Y, ZHANG F, et al. Energy efficiency resource allocation in downlink massive MIMO system with ZF precoding[C]//Proceedings of the International Conference on Frontiers of Electronics, Information and Computation Technologies. New York: ACM, 2021.
- [15] SUN Q, ZHAO H, WANG J, et al. Deep learning-based joint CSI feedback and hybrid precoding in FDD mmwave massive MIMO systems[J]. Entropy, 2022, 24(4): 441.
- [16] KOC A, LE-NGOC T. Swarm intelligence based power allocation in hybrid millimeter-wave massive MIMO systems[C]//IEEE Wireless Communications and Networking Conference (WCNC 2021). IEEE, 2021.
- [17] ZHU X Y, KOC A, MORAWSKI R, et al. A deep learning and

- geospatial data-based channel estimation technique for hybrid massive MIMO systems[J]. IEEE Access, 2021, 9: 145115-145132.
- [18] MAHMOOD M, KOC A, MORAWSKI R, et al. Achieving capacity gains in practical full-duplex massive MIMO systems: a multi-objective optimization approach using hybrid beamforming[J]. IEEE Open Journal of the Communications Society, 2024, 5:2268-2286.
- [19] XU Z Y, WANG Y Z, TANG J, et al. A deep reinforcement learning based framework for power-efficient resource allocation in cloud RANs[C]//Proceedings of the 2017 IEEE International Conference on Communications (ICC). Piscataway: IEEE Press, 2017: 1-6.
- [20] ZHAO N, LIANG Y C, NIYATO D, et al. Deep reinforcement learning for user association and resource allocation in heterogeneous cellular networks[J]. IEEE Transactions on Wireless Communications, 2019, 18(11): 5141-5152.
- [21] NASIR Y S, GUO D N. Multi-agent deep reinforcement learning for dynamic power allocation in wireless networks[J]. IEEE Journal on Selected Areas in Communications, 2019, 37(10): 2239-2250.
- [22] NAPARSTEK O, COHEN K. Deep multi-user reinforcement learning for distributed dynamic spectrum access[J]. IEEE Transactions on Wireless Communications, 2019, 18(1): 310-323.
- [23] WANG X M, ZHANG Y H, SHEN R J, et al. DRL-based energy-efficient resource allocation frameworks for uplink NOMA systems[J]. IEEE Internet of Things Journal, 2020, 7(8): 7279-7294.
- [24] DU J B, XU W, ZHAO C M, et al. Weighted spectral efficiency optimization for hybrid beamforming in multiuser massive MIMO-OFDM systems[J]. IEEE Transactions on Vehicular Technology, 2019, 68(10): 9698-9712.
- [25] 3GPP. 5G: study on channel model for frequencies from 0.5 to 100 GHz: TR 38.901[S]. 2020.

[作者简介]



刘庆利 (1981-), 男, 大连大学信息工程学院、通信与网络重点实验室教授, 主要研究方向为无人机系统、网络通信等。



李晓宇 (1998-), 女, 大连大学信息工程学院、通信与网络重点实验室硕士生, 主要研究方向为网络通信。



李蕊 (1997-), 女, 大连大学信息工程学院、通信与网络重点实验室硕士生, 主要研究方向为网络通信。