

引用格式:方娟,刘珍珍,陈思琪,等. 基于深度强化学习的 IRS 辅助 NOMA 通信资源分配优化[J]. 北京工业大学学报, 2024, 50(8): 923-931.

FANG J, LIU Z Z, CHEN S Q, et al. IRS assisted NOMA communication resource allocation optimization based on deep reinforcement learning[J]. Journal of Beijing University of Technology, 2024, 50(8): 923-931. (in Chinese)

基于深度强化学习的 IRS 辅助 NOMA 通信资源分配优化

方娟,刘珍珍,陈思琪,李硕朋

(北京工业大学信息学部,北京 100124)

摘要:为了解决无法与边缘服务器建立直连通信链路的盲区边缘用户卸载任务的问题,设计了一个基于深度强化学习(deep reinforcement learning, DRL)的智能反射面(intelligent reflecting surface, IRS)辅助非正交多址(non-orthogonal multiple access, NOMA)通信的资源分配优化算法,以获得由系统和速率和能源效率(energy efficiency, EE)加权的最大系统收益,从而实现绿色高效通信。通过深度确定性策略梯度(deep deterministic policy gradient, DDPG)算法联合优化用户功率分配和 IRS 的反射相移矩阵。仿真结果表明,使用 DDPG 算法处理移动边缘计算(mobile edge computing, MEC)的通信资源分配优于其他几种对比实验算法。

关键词:非正交多址(non-orthogonal multiple access, NOMA);智能反射面(intelligent reflecting surface, IRS);深度确定性策略梯度(deep deterministic policy gradient, DDPG)算法;移动边缘计算(mobile edge computing, MEC);能源效率(energy efficiency, EE);系统收益

中图分类号:TN 929.5

文献标志码:A

文章编号:0254-0037(2024)08-0923-09

doi: 10.11936/bjtxb2023020013

IRS Assisted NOMA Communication Resource Allocation Optimization Based on Deep Reinforcement Learning

FANG Juan, LIU Zhenzhen, CHEN Siqi, LI Shuopeng

(Faculty of Information Technology, Beijing University of Technology, Beijing 100124, China)

Abstract: To solve the problem of blind spot edge user offloading tasks where direct communication links cannot be established with edge servers, an intelligent reflecting surface (IRS) assisted non-orthogonal multiple access (NOMA) communication resource allocation optimization algorithm based on deep reinforcement learning (DRL) was designed. The algorithm aims to obtain the maximum system benefit weighted by system sum rate and energy efficiency (EE) for green and efficient communication. The deep deterministic policy gradient (DDPG) algorithm was adopted to jointly optimize the power allocation and phase-shift matrix. The simulation results show that DDPG algorithm is superior to other comparative experimental algorithms in dealing with the communication resource allocation of mobile edge computing (MEC).

收稿日期:2023-02-15;修回日期:2023-06-27

基金项目:国家自然科学基金资助项目(61202076);北京市自然科学基金资助项目(4192007)

作者简介:方娟(1973—),女,教授,博士生导师,主要从事计算机体系结构、高性能计算、移动边缘计算方面的研究,
E-mail:fangjuan@bjut.edu.cn

Key words: non-orthogonal multiple access (NOMA); intelligent reflecting surface (IRS); deep deterministic policy gradient (DDPG) algorithm; mobile edge computing (MEC); energy efficiency (EE); system profits

近年来,随着物联网(Internet of things, IoT)的普及和无线蜂窝通信技术的飞速发展,移动设备的数量呈爆炸式增长。英国著名市场研究机构Juniper Search发布的白皮书显示:“到2024年,IOT连接总数将可能达到830亿”^[1]。在这样的发展趋势下,海量的移动设备产生的大规模连接将需要巨大的流量作为支撑,仅依靠云服务中心为所有网络设备提供计算、存储等服务,已经无法满足人们对于互联网的高效要求。移动边缘计算(mobile edge computing, MEC)可以将互联网中的网络资源、计算服务、存储能力迁移到用户的边缘^[2-4],不仅可以有效地减轻云计算中心的压力,而且利用其靠近用户端的物理优势,还能提供高可靠的连接,同时,MEC拥有对周围环境的感知能力,并且分布广泛的边缘服务器相比于云服务器对用户隐私数据也起到了更好的保护作用,满足了互联网行业在敏捷连接、实时业务以及隐私保护方面的关键需求^[5]。然而,随着人类社会越来越趋近于数字化生活,在频谱资源欠缺的巨大压力下,如何让MEC在海量的数据请求中做到瞬时响应、精准应答,仍然是当前MEC通信中急需解决的重要问题。

随着无线通信技术的快速迭代发展,正交频分多址(orthogonal frequency division multiple access, OFDMA)技术被广泛应用于长期演进技术(long term evolution, LTE)、LTE-advanced^[6]等第4代移动通信系统中^[7],但是OFDMA只对用户信号进行正交处理,面对大规模连接和高频谱效率的要求,无法进行频谱共享。与OFDMA不同,非正交多址(non-orthogonal multiple access, NOMA)具有共享频谱的特性,可以有效地解决频谱利用率低的问题^[8-10]。NOMA的核心思想是在同一时间、频率上给多个用户分配不同的功率,允许多个用户共享相同的子信道来进行数据传输^[11]。NOMA技术在使用过程中,首先,需要在发送端应用叠加编码(superposition coding, SC)技术进行子信道内的用户信号非正交处理^[12];其次,再对各子信道之间的信号进行正交处理;最后,在接收端利用串行连续干扰消除(successive interference cancellation, SIC)机制消除叠加用户信号产生的干扰^[13]。

考虑到NOMA的高频谱利用率的特性,文

献[14]提出了一种边缘计算感知的NOMA技术,通过优化用户集群、计算资源和通信资源分配,最大限度地降低了MEC的能耗。文献[15]通过用户、助手和无线接入基本三节点模型组成的新的边缘计算协作范式,利用NOMA在整个资源块上的并行传输来实现计算卸载。文献[16-17]研究了基于NOMA技术的MEC系统和速率最大化问题。文献[18]设计了一种用于多址多任务计算卸载、NOMA传输和计算资源分配的联合优化方案,并提出一种基于深度强化学习(deep reinforcement learning, DRL)的在线算法来解决时变信道下的最优卸载问题,目标是在系统延迟限制下最小化IoT设备的总能耗。

针对MEC和NOMA技术的结合,学者们分别从和速率、能耗与计算卸载等方面进行了研究,并且已经取得了很多有价值的研究成果,但是,仍然存在一些不足之处。在上述研究中,文献[14-17]使用的是传统算法,需要设定初始值,环境一旦改变,就需要重新求解,不利于应用到实际问题中;文献[18]考虑的是单一目标优化,并且模拟环境是在任何情况下移动设备都是直连边缘服务器,对于用户设备无法与边缘服务器建立直连链路的情况,并没有进行研究讨论。

在MEC的实际通信过程当中,由于城市中高层建筑物众多,很容易导致部分边缘用户与边缘服务器之间存在不可移除的障碍物,导致无法进行直连通信,从而无法发挥MEC的最大优势。新兴的智能反射面(intelligent reflecting surface, IRS)技术可以有效地解决此问题。IRS是一种可重构的电磁超导材料,IRS由大量被动反射元件组成,通过调整其元件的相移、振幅控制入射信号,从而重新配置无线传播环境^[19-21],可以使期望信号得到增强,干扰信号得到抑制^[22-23]。特别是,与有源中继/波束形成等传统技术相比,IRS不仅以全双工和无噪声的方式反射信号,不会产生自干扰,并且通过使用轻质无源元件,大大节省了能耗和硬件以及部署成本^[24],因此,IRS被认为是一种具有成本效益、节能和可持续发展的第6代移动通信网络技术^[25]。

文献[26-27]利用IRS帮助求解NOMA通信中最大化系统和速率问题。文献[28]研究了具有联合检测的双小区中多用户通信的最小化传输功率问

题。文献[29]设计了一个IRS辅助设备到设备(device to device, D2D)的卸载系统,提出了一种混合整数随机逐次凸逼近算法来解决波束形成和资源分配的联合设计问题。文献[30-31]针对IRS辅助NOMA系统的能量效率(energy efficiency, EE)进行了研究。文献[32]通过利用交替优化算法将多变量耦合的非凸优化问题分解成多个凸优化子问题进行求解,目的是为了最大化IRS辅助的NOMA网络鲁棒性和能效。

目前,针对MEC中处于盲区的边缘用户直连通信链路被中断,无法顺利卸载计算密集型和延迟敏感任务的问题,还没有行之有效的解决方案。另外,MEC通信的频谱资源有限,而NOMA技术具有高频谱的特性,能够解决多用户的资源共享问题,但是,关于NOMA-MEC通信系统中低延迟和低能耗多目标的研究还不够,因此,为了充分发挥MEC近距离分布式部署、资源贴近网络边缘的特点,实现MEC的通信全覆盖,在节约资源的前提下达到高效通信的目标,本文针对处于盲区的边缘用户的通信问题展开深入研究,并提出了相应的解决方案。

本文在MEC场景中,设计了一种基于DRL的IRS辅助NOMA通信资源分配优化方案,利用IRS的反射信号的特性,解决盲区边缘用户卸载任务通信链路中断问题。首先,建立通信模型和目标模型;其次,对目标模型中系统和速率与能效的影响因素进行分析;然后,将二者进行加权计算得到系统收益,通过提取影响系统收益的关键因素,发现决定系统收益最终取值的主要是用户发射功率和IRS的元件相移这2个参数。考虑到用户功率和IRS的相移都是属于连续数据,直接作为离散数据进行量化容易产生误差,并且现实通信环境复杂多变,因此,采用DRL中的深度确定性策略梯度(deep deterministic policy gradient, DDPG)算法,同时,对用户功率和IRS的相移进行联合优化,最终实现最大化系统收益的目标。

1 系统模型

本文研究了IRS辅助的基于NOMA-MEC通信系统的资源分配优化问题。假设通信场景如图1所示,在小区附近60 m处有一个基站,基站处部署了一台边缘服务器,边缘服务器为小区里的边缘用户提供计算和存储服务,基站和边缘用户都使用单一天线。另外,小区里楼宇太过密集,导致部分处于角

落里的边缘用户无法与边缘服务器建立直连通信链路,引起通信受阻问题,严重影响用户体验。因此,针对该问题,本文设计的通信模型中添加了IRS辅助,处于盲区的边缘用户可以依靠安装在临近建筑物表面上的IRS进行辅助通信,从而顺利地将任务卸载到边缘服务器上,IRS的元件数量为 N ,每个单独的元件用 n 表示,则 $n \in \{1, 2, \dots, N\}$, Φ 表示IRS反射对角矩阵,其表达式为 $\Phi = \text{diag}(\beta_1 e^{j\theta_1}, \beta_2 e^{j\theta_2}, \dots, \beta_N e^{j\theta_N})$ 。其中: $\text{diag}(\cdot)$ 表示对角函数; θ_n 表示IRS上第 n 个被动发射元件的相移, $\theta_n \in [0, 2\pi)$, β_n 表示IRS的振幅; $\beta_n \in [0, 1]$,由IRS的反射矩阵可以得知,其反射效果关键取决于相移和振幅这2个参数。本文为了达到IRS最理想的波束形成效果,设定所有信号经IRS元件反射后振幅不变,即采用固定的振幅 $\beta_n = 1$,因此,在本文中IRS的反射效率主要由 θ_n 决定。

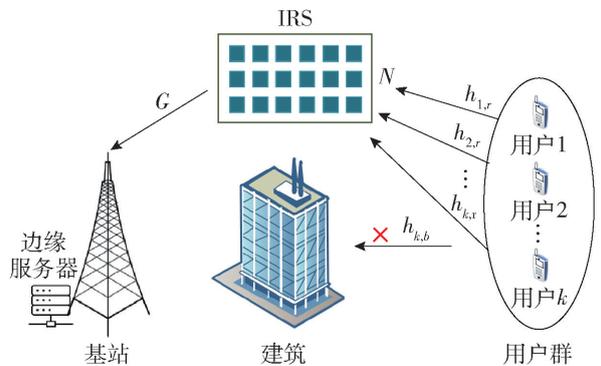


图1 IRS辅助NOMA-MEC系统通信链路

Fig. 1 IRS-assisted NOMA-MEC system communication link

假设每个边缘用户每次都只卸载一个计算密集或延迟敏感的任务,卸载的任务用 k 表示,需要卸载任务的用户数为 K ,则 $k \in \{1, 2, \dots, K\}$,边缘用户的卸载功率记作 P_k ,所有用户使用的总功率为 P_{sum} 。边缘用户与IRS之间的动态信道增益为 $h_{k,r} = g_{k,r} \cdot \text{PL}^{-1}(d_{k,r})$ 。其中: $g_{k,r} \sim \text{CN}(0, \sigma^2)$ 是服从小尺度瑞利衰落的信道增益,平均数为0,方差为 σ^2 ; $\text{PL}^{-1}(\cdot)$ 是路径损失函数; $d_{k,r}$ 表示边缘用户到IRS的距离。IRS与基站之间的信道增益表示为

$$G = g_{k,r} \cdot \text{PL}^{-1}(d_{k,r}) \cdot \Phi_T^{1/2}$$

式中: $g_{k,r}$ 为IRS与基站天线之间的信道增益; $d_{k,r}$ 为IRS元件到基站天线的距离; Φ_T 为IRS的反射协方差矩阵。本文假设盲区的边缘用户卸载任务时的信道增益满足 $|h_{1,r} \Phi G| \leq \dots \leq |h_{K,r} \Phi G|$,并且用户的信道增益排序都始终保持不变。边缘用户 k 的速率表示为

$$R_k = B_k \ln \left(1 + \frac{|h_{k,r} \Phi G|^2 P_k}{\sum_{i=1}^{k-1} |h_{i,r} \Phi G|^2 P_i + \sigma^2} \right) \quad (1)$$

式中 B_k 为用户 k 的传输带宽。根据 NOMA 技术协议,通过 SC 技术将多个用户叠加在同一个时间、频谱资源块上,然后在接收端使用 SIC 来消除叠加用户的干扰,SIC 遵循信号的信道质量越好,越能被优先解码的原则,因此,与用户 k 共享相同信道的用户,在用户 k 之前解码的信号都不作为其干扰信号,在用户 k 之后解码的信号 $\sum_{i=1}^{k-1} |h_{i,r} \Phi G|^2 P_i$ 才作为用户 k 的干扰信号。 σ^2 表示信号传输过程中系统产生的白色高斯加性噪声。

根据式(1)可以得出所有用户的和速率,即

$$R_{\text{sum}} = \sum_{k=1}^K R_k \quad (2)$$

评价通信质量的优劣不仅依靠系统和速率,还有另一个参考指标,即能效。本文模型中所有盲区的边缘用户在卸载计算密集型或延迟敏感任务时的能效可表示为

$$E_{\text{sum}} = \frac{R_{\text{sum}}}{P_{\text{sum}} + P_c} \quad (3)$$

式中 P_c 为通信过程中系统固定损耗功率。

本文不仅要考虑通信效率,而且还要节约资源,进行绿色通信,需要同时实现最大化系统和速率和最大化系统能效,因此,本文将用户和速率与系统能效进行加权求和,表示为系统收益 S^p ,其中 λ_1 和 λ_2 分别表示和速率和能效的权重,并且满足 $0 \leq \lambda_1, \lambda_2 \leq 1$ 。于是本文目标转化为求解最大系统收益,其值表示为

$$\max_{\theta_n, P_k} S^p = \lambda_1 R_{\text{sum}} + \lambda_2 E_{\text{sum}} \quad (4)$$

$$\text{s. t.} \quad P_k \leq P_{\text{max}} \quad (4a)$$

$$R_k \geq R_{\text{min}} \quad (4b)$$

$$\theta_n \in [0, 2\pi) \quad (4c)$$

$$\lambda_1 + \lambda_2 = 1 \quad (4d)$$

约束(4a)表示在资源有限的情况下,每个边缘用户卸载任务时功率不能超过设定的最大功率值,需要尽可能地节约资源,同时,不会因单个移动设备功率值过高而对其他移动设备的任务卸载产生过大干扰;约束(4b)表示每个用户的传输速度不能小于最低速率,保证用户的通信服务质量,避免信号质量极差导致任务卸载过程发生错误,以此确保通信的可靠性和稳定性;约束(4c)表示 IRS 的每个元件相移的取值范围;约束(4d)表示限制用户和速率和能

效的权重之和为 1。

从式(4)可以看出:首先,本文需要解决的是一个非凸、非定常多项式 (non-deterministic, polynomial, NP) 难问题,并且系统收益受 IRS 的相移和用户的传输功率所影响;其次,IRS 的相移和传输功率都是属于连续值;然后,本文设定的场景是动态的实时的通信环境。考虑到利用目前现有的传统算法,通过固定初始环境的方式已经无法快速地解决问题,因此,本文使用 DRL 中的 DDPG 算法来帮助在线联合优化 IRS 的相移和边缘用户的传输功率。DRL 算法具有强大的环境感知能力和策略泛化能力,能够帮助求解出理想的 IRS 相移和通信功率值,使系统收益最大化。

2 基于 DDPG 算法的传输功率和 IRS 相移联合设计

DDPG 算法是用来处理连续数据的一种将神经网络和学习策略函数 Actor-学习动作价值函数 Critic 体系结构相结合的无模型学习技术。DDPG 算法适合高维复杂的行为问题,其主要优势在于同时使用了连续动作和状态空间。

DDPG 算法主要使用了 4 种神经网络,分别为权重 μ 的参与者网络 $\mu(s|\theta^\mu)$ 、目标参与者网络 $\mu'(s|\theta^{\mu'})$ 、评论家网络 $Q(s, a|\theta^Q)$ 和目标评论家网络 $Q'(s, a|\theta^{Q'})$ 。参与者网络和目标参与者网络属于学习策略网络,评论家网络和目标评论家网络属于动作价值评判网络。DDPG 算法模型如图 2 所示,其运行流程具体说明如下。

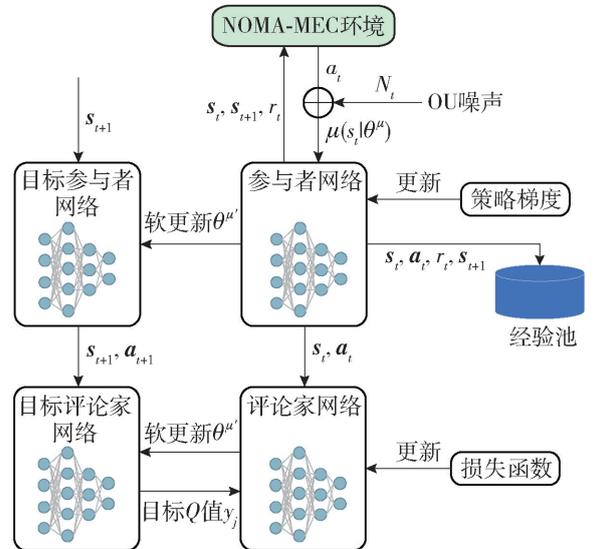


图 2 DDPG 算法模型

Fig. 2 DDPG algorithm model

在 t 时刻,参与者网络根据当前环境 s_t 生成一个确定性的动作 a_t ,为了使算法探索潜在更优策略,将其与环境进行深度交互,使策略的探索更加充分。在生成动作时,DDPG 算法引入了奥恩斯坦-乌伦贝克(Ornstein-Uhlenbeck,OU)随机过程产生的时序相关的噪声,作为动作生成随机噪声,因此,状态 s_t 产生的动作可表示为

$$\mathbf{a}_t = \boldsymbol{\mu}(s_t | \boldsymbol{\theta}^\mu) + N_t \quad (5)$$

式中 N_t 为 t 时刻的探索噪声。

在本文建立的 IRS 辅助的 NOMA-MEC 通信系统模型中,由于边缘用户通信的信道状态是复杂多变不可预测的,于是将边缘设备的动态信道状态和所有用户传输使用的总功率联合作为 DDPG 算法中参与者网络的环境状态,因此, t 时刻的环境状态 s_t 表示为

$$s_t = [h_{1,r}, h_{2,r}, \dots, h_{K,r}, P_{\text{sum}}] \quad (6)$$

将 IRS 元件的 θ_n 和边缘用户的 P_k 作为 DDPG 算法随环境生成的动作,表示为

$$\mathbf{a}_t = [\theta_1, \theta_2, \dots, \theta_N, P_1, P_2, \dots, P_k] \quad (7)$$

然后将 s_t 和 \mathbf{a}_t 送入评论家网络,该网络根据输入参数,输出一个动作价值函数值($s_t | \mathbf{a}_t$),表示对当前动作 \mathbf{a}_t 优劣的评判。

同时,NOMA-MEC 通信环境根据 DDPG 算法反馈的动作 \mathbf{a}_t 产生一个奖励值 r_t 和 $t+1$ 时刻的环境状态 s_{t+1} 。本文模型中设定系统收益 S^p 作为 DDPG 算法中的奖励值,因此,在 t 时刻的奖励值 r_t 可表示为

$$r_t = S_t^p \quad (8)$$

DDPG 算法吸收了深度 Q 网络(deep Q-network, DQN)算法的特性,同样具有目标网络和经验回放机制^[33-39]。首先,将每个 t 时刻产生的元组($s_t, \mathbf{a}_t, r_t, s_{t+1}$)存储到经验池 D 中;然后,从经验池取出固定数量 M 个样本数据($s_j, \mathbf{a}_j, r_j, s_{j+1}$)进行目标网络的训练。其中目标参与者网络利用 s_{j+1} 生成 \mathbf{a}_{j+1} ,将 \mathbf{a}_{j+1} 和 s_{j+1} 传出至目标评论家网络,用该网络计算目标 Q 值, Q 的表达式为

$$y_j = r_j + \gamma Q'(s_{t+1}, \boldsymbol{\mu}'(s_{t+1} | \boldsymbol{\theta}^{\mu'}) | \boldsymbol{\theta}^Q) \quad (9)$$

式中 γ 表示奖励折扣因子, $\gamma \in (0, 1)$ 。 γ 值越大,表示未来奖励对当前策略影响越大; γ 值越小,说明算法训练更关注当前的奖励对策略的影响。同时,为了更新评论家网络参数,使其能够计算出最优的动作价值函数,将目标评论家的目标 Q 值和当前评论家网络的动作价值函数做均方差,计算损失函数的值,通过最小化损失函数的值来更新评论家网络

参数,公式为

$$L = \frac{1}{M} \sum_j (y_j - Q(s_j, \mathbf{a}_j | \boldsymbol{\theta}^Q))^2 \quad (10)$$

本文设计的 DDPG 算法的主要目标就是要通过 DDPG 算法网络的学习,输出最优的确定性的动作,因此,算法通过链式法则和梯度上升的方式更新参与者网络参数,不断地进行网络策略的调整,更新法则表示为

$$\nabla_{\boldsymbol{\theta}^\mu} J = \frac{1}{M} \sum_i \nabla_a Q(s, \mathbf{a} | \boldsymbol{\theta}^Q) |_{s=s_i, \mathbf{a}=\boldsymbol{\mu}(s_i)} \nabla_{\boldsymbol{\theta}^\mu} \boldsymbol{\mu}(s | \boldsymbol{\theta}^\mu) |_{s_i} \quad (11)$$

目标参与者网络参数则是通过软更新方式定期从参与者网络参数中同步更新,更新逻辑表示为

$$\boldsymbol{\theta}^{\mu'} = \tau \boldsymbol{\theta}^\mu + (1 - \tau) \boldsymbol{\theta}^{\mu'} \quad (12)$$

式中 τ 表示软更新因子, $\tau \in (0, 1)$ 。

目标评论家网络参数同目标参与者网络参数更新策略一致,表示为

$$\boldsymbol{\theta}^{Q'} = \tau \boldsymbol{\theta}^Q + (1 - \tau) \boldsymbol{\theta}^{Q'} \quad (13)$$

通过大量的在线数据的训练,使 DDPG 算法模型的参数能够快速调整,从而输出最优动作,获取最大的奖励值,使本文求解的最大化系统收益问题得到解决。

3 仿真结果与分析

为了验证本文提出算法的合理性与有效性,以系统收益为实验的性能指标,将本文算法和其他算法进行了对比实验。实验中通信环境模型如图3所示。基站位于三维坐标的原点处,三维坐标的单位为 m,基站的位置为(0, 0, 10),小区中的盲区用户集中在 x 轴坐标(60, 0, 0)位置处,假设所有通信用户的坐标位置一致,IRS 的初始位置放置在坐标(20, 0, 10)处。

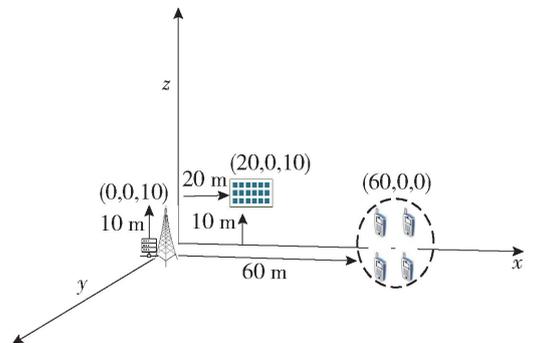


图3 IRS 支持的 NOMA-MEC 系统的实验模型

Fig. 3 IRS assisted NOMA-MEC system experimental model

在实验算法中,总系统带宽为 10 MHz,系统的

路径损耗为 $-170 + 20\lg d$, d 为系统模型中设备之间的物理距离,系统噪声为 0.01 dB,系统固定功率损耗为 0.1 dBm,系统和速率和能效的权重都设置为 0.5,同时实验中 IRS 的元件数量为 96,边缘用户数量为 4,激活函数采 Relu DDPG 算法的固定参数如表 1 所示。

表 1 DDPG 算法参数

Table 1 Parameters of DDPG algorithm

参数	值
经验池容量	10 000
训练集数	400
每个训练集迭代次数	100
抽样样本容量	32
参与者网络学习率	0.002
评论家网络学习率	0.004
奖励折扣因子	0.9
软更新因子	0.01
隐藏层第 1 层节点数	64
隐藏层第 2 层节点数	64

在训练了 400 次后 5 种不同算法计算出的系统收益结果如图 4 所示。图例中“DDPG 联合优化 NOMA 通信资源分配算法”表示在 IRS 辅助 NOMA-MEC 通信系统中,使用了 DDPG 算法优化用户传输功率和 IRS 相移的系统收益;“DDPG 单一优化 NOMA 通信功率算法”表示在 IRS 辅助 NOMA-MEC 通信系统中,使用 DDPG 算法优化用户传输功率,IRS 的相移矩阵遵循随机设置计算出的系统收益;“NOMA 通信资源随机分配算法”表示随机的用户功率和随机的 IRS 相移产生的系统收益;“无 IRS 辅助 NOMA 通信资源分配算法”表示在无 IRS 的基于 NOMA-MEC 通信系统中,使用随机的用户功率计算得到的系统收益;“OFDMA 通信资源随机分配算法”表示在 IRS 辅助基于 OFDMA 技术的 MEC 通信系统中,使用随机的用户功率和随机的 IRS 相移计算的系统收益。

由图 4 可知:使用了 IRS 辅助基于 NOMA-MEC 通信系统的系统收益明显高于无 IRS 辅助通信时的系统收益,这是由于添加 IRS 辅助通信,在用户与基站之间重新建立了虚拟直连链路,跨越了物理直连链路上的障碍物,明显减少了环境因素对通信网络造成的严重影响。其次,使用 DDPG 算法训练优化

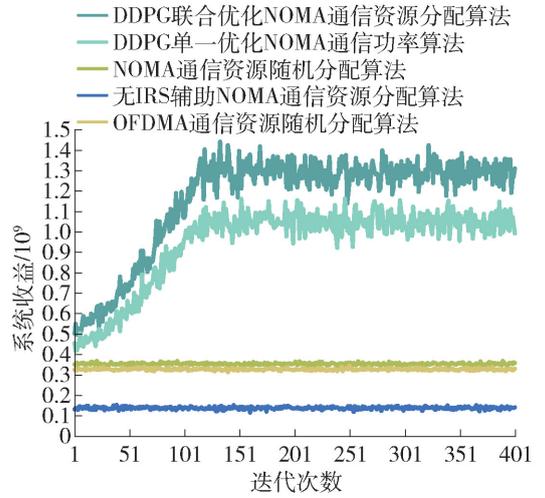


图 4 5 种实验迭代次数对比

Fig. 4 Comparison of five experimental under the iteration numbers

的通信能够获得比其他基准算法更高的系统收益。实验结果显示,“DDPG 联合优化 NOMA 通信资源分配算法”计算出的系统收益分别比“DDPG 单一优化 NOMA 通信功率算法”的结果高 22.1%,比“NOMA 通信资源随机分配算法”的结果提升 231.4%,比“OFDMA 通信资源随机分配算法”的结果增长 262.5%。这是因为 DDPG 算法将边缘用户的复合信道增益和总功率值作为输入参数,并以提升系统收益为结果导向,不断进行模型训练,从而调整策略,继而对 NOMA 通信资源的分配做出更好的选择。于此同时,在相同条件下,对比本文设计的算法得出的系统收益和“DDPG 单一优化 NOMA 通信功率算法”的结果可以发现,同时优化 2 个参数变量要优于单一功率参数的优化结果。

不同 IRS 元件数量下的 4 种算法的实验结果如图 5 所示。IRS 的反射效率取决于反射元件,反射元件数量的增加或减少直接影响通信的质量和稳定性。

由图 5 中的结果可以得出:首先,当 IRS 的元件数量逐渐从 48 个增加到 112 个时,每次增加 8 个元件,4 种算法计算出的系统收益都在增长,说明 IRS 元件数量越多,越能减少信号损耗,越有利于信号传输。其次,在不同 IRS 元件数量下,基于 DDPG 算法优化的 IRS 辅助基于 NOMA-MEC 通信的系统收益始终高于基准算法。随着 IRS 数量的增加,基于 DDPG 优化多参数算法的系统收益从 1.118×10^9 增长到 1.175×10^9 ;基于 DDPG 算法进行传输功率单一变量优化的系统收益从 0.923×10^9 增长到

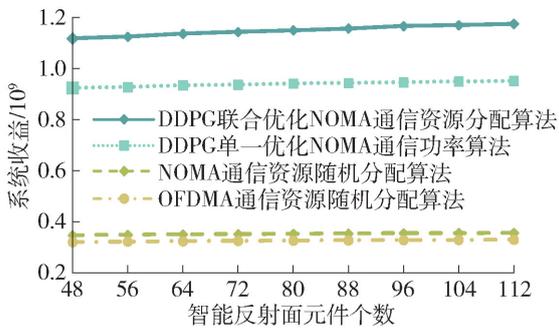


图5 IRS元件数量的对比实验结果

Fig. 5 Comparative experimental results of IRS elements numbers

0.951×10^9 ; 基于 NOMA-MEC 系统采用随机变量得到的系统收益从 0.346×10^9 增长到 0.355×10^9 ; 基于 OFDMA 技术的 MEC 系统采用随机参数的系统收益从 0.319×10^9 增长到 0.328×10^9 。对比结果表明,系统收益目标值受到 IRS 相移和传输功率这 2 个因素共同影响,因此,本文设计的算法采用了联合参数优化方法,将 IRS 相移和传输功率考虑在内,更好地提高了通信资源的利用率,从而验证了本文提出算法的可扩展性。

在基于 IRS 的不同物理位置下,4 种实验算法的对比结果如图 6 所示,IRS 的部署位置影响着信号的反射角度和覆盖范围。模拟实验中,边缘服务器和边缘用户的位置始终在上文设定的坐标位置处保持不变,而 IRS 的位置从 (10, 0, 10) 沿着 x 轴坐标向右移动,最终移动到坐标 (60, 0, 10) 处。

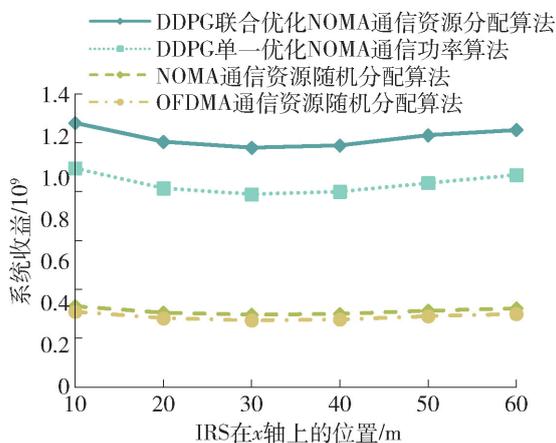


图6 IRS在不同位置下的对比实验结果

Fig. 6 Comparative experimental results of IRS at different positions

从图 6 中可以看出:首先,当 IRS 从靠近边缘服务器的位置逐渐远离边缘服务器,再到逐渐靠近用户设备端,期间 4 种算法的系统收益都是先逐渐减

小,再逐渐增大,这是因为通信距离越远,通信过程中受到的干扰噪声就越多,信号损失就越大,这说明在 IRS 实际应用中,IRS 的部署位置应该选在临近边缘服务器处或者靠近用户端,这样能够获得较好的信道增益。其次,当 IRS 处于不同位置时,本文提出的基于 DDPG 优化算法计算得到的系统收益始终高于其他对比算法,原因是本文设计的 DDPG 联合优化算法就是在基于动态通信环境下进行设计建模和控制的,而 IRS 位置的移动对边缘用户的通信环境会产生影响,但对算法训练不会产生影响,因此,说明本文提出的算法具有非常明显的优势,能够显著改善 MEC 的通信效率和质量。

4 结论

1) 为了充分发挥 MEC 低延迟、强感知的特性,实现 MEC 的通信全覆盖目标,解决处于盲区的边缘用户快速卸载计算密集型和延迟敏感任务的通信问题,本文设计了一个 IRS 辅助基于 NOMA-MEC 上行通信链路场景,并建立了一个三边模型,分别由用户端、IRS、边缘服务器组成。然后,提出利用 DRL 中的 DDPG 算法进行联合优化用户传输功率和 IRS 相移,最终得到了最优的系统收益,为实现边缘用户的顺利通信、提高 MEC 的吞吐量和能源效率提供了有效的解决方案。

2) 本文设计的算法模型,不仅考虑了用户的和速率,同时,还兼顾了能量效率,在动态通信环境中,进行了多目标联合优化,而且算法的自适应能力非常强大,不仅缓解了频谱资源紧张问题,提高了通信效率,而且响应了国家的节能减排计划,实现了绿色通信。

3) 本文针对提出的算法模型,将其与目前流行的通信技术进行了多维度的深入比较。实验结果表明,本文提出的算法在多方面都明显优于其他基准算法,具有实际效益。

参考文献:

- [1] ROTHMULLER M, BARKER S. IoT-the Internet of transformation 2020 [R]. Hampshire, UK: Juniper Research, 2020.
- [2] MAO Y Y, YOU C S, ZHANG J, et al. A survey on mobile edge computing: the communication perspective [J]. IEEE Communications Surveys & Tutorials, 2017, 19(4): 2322-2358.
- [3] DING Z G, FAN P Z, POOR H V. Impact of non-orthogonal multiple access on the offloading of mobile edge

- computing [J]. *IEEE Transactions on Communications*, 2019, 67(1): 375-390.
- [4] HU F, CHEN B. Channel coding scheme for relay edge computing wireless networks via homomorphic encryption and NOMA [J]. *IEEE Transactions on Cognitive Communications and Networking*, 2020, 6(4): 1180-1192.
- [5] CAO K Y, LIU Y F, MENG G J, et al. An overview on edge computing research [J]. *IEEE Access*, 2020, 8: 85714-85728.
- [6] GHOSH A, RATASUK R, MONDAL B, et al. LTE-advanced: next-generation wireless broadband technology [J]. *IEEE Wireless Communications*, 2010, 17(3): 10-22.
- [7] DI B Y, SONG L Y, LI Y H. Sub-channel assignment, power allocation, and user scheduling for non-orthogonal multiple access networks [J]. *IEEE Transactions on Wireless Communications*, 2016, 15(11): 7686-7698.
- [8] SAAD W, BENNIS M, CHEN M Z. A vision of 6G wireless systems: applications, trends, technologies, and open research problems [J]. *IEEE Network*, 2020, 34(3): 134-142.
- [9] QIAN L P, SHI B H, WU Y, et al. NOMA-enabled mobile edge computing for Internet of things via joint communication and computation resource allocations [J]. *IEEE Internet of Things Journal*, 2020, 7(1): 718-733.
- [10] DING Z G, SCHOBBER R, POOR H V. No-pain No-gain: DRL assisted optimization in energy-constrained CR-NOMA networks [J]. *IEEE Transactions on Communications*, 2021, 69(9): 5917-5932.
- [11] XIAO Z Y, ZHU L P, CHOI J H, et al. Joint power allocation and beamforming for non-orthogonal multiple access (NOMA) in 5G millimeter wave communications [J]. *IEEE Transactions on Wireless Communications*, 2018, 17(5): 2961-2974.
- [12] HE M L, LI Y, WANG X F, et al. NOMA resource allocation method in IoV based on prioritized DQN-DDPG network [J]. *EURASIP Journal on Advances in Signal Processing*, 2021, 2021: 120.
- [13] LIU Y W, QIN Z J, ELKASHLAN M, et al. Non-orthogonal multiple access for 5G and beyond [J]. *Proceedings of the IEEE*, 2017, 105(2): 2347-2381.
- [14] KIANI A, ANSARI N. Edge computing aware NOMA for 5G networks [J]. *IEEE Internet of Things Journal*, 2018, 5(2): 1299-1306.
- [15] LIU Y. Exploiting NOMA for cooperative edge computing [J]. *IEEE Wireless Communications*, 2019, 26(5): 99-103.
- [16] WANG F, XU J, DING Z G. Multi-antenna NOMA for computation offloading in multiuser mobile edge computing systems [J]. *IEEE Transactions on Communications*, 2019, 67(3): 2450-2463.
- [17] HANIF M F, DING Z G, RATNARAJAH T, et al. A minorization-maximization method for optimizing sum rate in the downlink of non-orthogonal multiple access systems [J]. *IEEE Transactions on Signal Processing*, 2016, 64(1): 76-88.
- [18] QIAN L P, WU Y, JIANG F L, et al. NOMA assisted multi-task multi-access mobile edge computing via deep reinforcement learning for industrial internet of things [J]. *IEEE Transactions on Industrial Informatics*, 2021, 17(8): 5688-5698.
- [19] XU Z G, LIU J X, ZOU J W, et al. Energy-efficient design for IRS-assisted NOMA-based mobile edge computing [J]. *IEEE Communications Letters*, 2022, 26(7): 1618-1622.
- [20] WU Q Q, ZHANG R. Towards smart and reconfigurable environment: intelligent reflecting surface aided wireless network [J]. *IEEE Communications Magazine*, 2020, 58(1): 106-112.
- [21] ZHENG B X, YOU C S, ZHANG R. Double-IRS assisted multi-user MIMO: cooperative passive beamforming design [J]. *IEEE Transactions on Wireless Communications*, 2021, 20(7): 4513-4526.
- [22] LIU Y W, ZHANG S W, MU X D, et al. Evolution of NOMA toward next generation multiple access (NGMA) for 6G [J]. *IEEE Journal on Selected Areas in Communications*, 2022, 40(4): 1037-1071.
- [23] WU Q Q, ZHOU X B, SCHOBBER R. IRS-assisted wireless powered NOMA: do we really need different phase shifts in DL and UL? [J]. *IEEE Wireless Communications Letters*, 2021, 10(7): 1493-1497.
- [24] ZHENG B X, WU Q Q, ZHANG R. Intelligent reflecting surface-assisted multiple access with user pairing: NOMA or OMA? [J]. *IEEE Communications Letters*, 2020, 24(4): 753-757.
- [25] HUANG C W, MO R H, YUEN C. Reconfigurable intelligent surface assisted multiuser MISO systems exploiting deep reinforcement learning [J]. *IEEE Journal on Selected Areas in Communications*, 2020, 38(8): 1839-1850.
- [26] SHEHAB M, CIFTLER B S, KHATTAB T, et al. Deep reinforcement learning powered IRS-assisted downlink NOMA [J]. *IEEE Open Journal of the Communications Society*, 2022, 3: 729-739.
- [27] WANG H, LIU C, SHI Z, et al. Power minimization for two-cell IRS-aided NOMA systems with joint detection

- [J]. IEEE Communications Letters, 2021, 25(5): 1635-1639.
- [28] LIU Y Z, HU Q Y, CAI Y L, et al. Latency minimization in intelligent reflecting surface assisted D2D offloading systems [J]. IEEE Communications Letters, 2021, 25(9): 3046-3050.
- [29] ZENG M, LI X W, LI G, et al. Sum rate maximization for IRS-assisted uplink NOMA [J]. IEEE Communications Letters, 2021, 25(1): 234-238.
- [30] LI G, ZENG M, MISHRA D, et al. Energy-efficient design for IRS-empowered uplink MIMO-NOMA systems [J]. IEEE Transactions on Vehicular Technology, 2022, 71(9): 9490-9500.
- [31] FANG F, XU Y Q, PHAM Q V, et al. Energy-efficient design of IRS-NOMA networks[J]. IEEE Transactions on Vehicular Technology, 2020, 69(11): 14088-14092.
- [32] 刘期烈, 辛雅楠, 高俊鹏, 等. 可重构智能表面辅助的非正交多址接入网络鲁棒能量效率资源分配算法 [J]. 电子与信息学报, 2022, 44(7): 2332-2341.
LIU Q L, XIN Y N, GAO J P, et al. Robust energy efficiency resource allocation algorithm in reconfigurable intelligent surface-assisted non-orthogonal multiple access networks [J]. Journal of Electronics & Information Technology, 2022, 44(7): 2332-2341. (in Chinese)
- [33] 贺梦利. 基于 DRL 的 NOMA 系统用户分组及功率分配算法研究[D]. 哈尔滨: 黑龙江大学, 2022.
HE M L. Research on user grouping and power allocation algorithm of NOMA system based on DRL[D]. Harbin; Heilongjiang University, 2022. (in Chinese)
- [34] LIU Q, WU J, HU L, et al. Optimal energy efficiency used DDPG in IRS-NOMA wireless communications[J]. Symmetry, 2022, 14(5): 1018-1027.
- [35] CHEN J, XING H L, XIAO Z W, et al. A DRL agent for jointly optimizing computation offloading and resource allocation in MEC[J]. IEEE Internet of Things Journal, 2021, 8(24): 17508-17524.
- [36] SHI Z Y, XIE X Z, LU H B, et al. Deep reinforcement learning-based multidimensional resource management for energy harvesting cognitive NOMA communications[J]. IEEE Transactions on Communications, 2022, 70(5): 3110-3125.
- [37] 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.
- [38] XU J P, AI B, CHEN L Y, et al. Deep reinforcement learning for computation and communication resource allocation in multiaccess MEC assisted railway IoT networks [J]. IEEE Transactions on Intelligent Transportation Systems, 2022, 23(12): 23797-23808.
- [39] CHEN J, GUO L, JIA J, et al. Resource allocation for IRS assisted SGF NOMA transmission: a MADRL approach [J]. IEEE Journal on Selected Areas in Communications, 2022, 40(4): 1302-1316.

(责任编辑 梁 洁)