

数字孪生使能的智能超表面边缘计算网络任务卸载

苏健 钱震 李斌*

(南京信息工程大学计算机与软件学院 南京 210044)

(南京信息工程大学江苏省大气环境与装备技术协同创新中心 南京 210044)

摘要: 针对新兴的计算密集型应用对移动用户高计算性能需求问题, 该文提出一种数字孪生(DT)结合智能反射面(RIS)辅助的移动边缘计算(MEC)任务卸载方案。首先, 在满足用户传输功率、用户和资源设备能耗、计算资源限制条件下, 通过联合优化用户卸载决策、用户传输功率、RIS相移、波束成形矢量、计算资源分配, 建立一个系统能耗最小化问题; 其次, 将该非凸组合优化问题分解为3个子问题, 使用深度双Q网络(DDQN)方法确定用户卸载策略; 然后对每个训练时间步进行一次求解, 基于交替迭代方法得到问题的优化解。仿真结果表明, 基于DDQN的算法训练速度较快, 有效降低了系统总能耗。

关键词: 智能超表面; 数字孪生; 移动边缘计算; 深度强化学习

中图分类号: TN929.5

文献标识码: A

文章编号: 1009-5896(2022)07-2416-09

DOI: [10.11999/JEIT220180](https://doi.org/10.11999/JEIT220180)

Digital Twin Empowered Task Offloading for RIS-Assisted Edge Computing Networks

SU Jian QIAN Zhen LI Bin

(School of Computer and Software, Nanjing University of Information Science and Technology,
Nanjing 210044, China)

(Jiangsu Collaborative Innovation Center of Atmospheric Environment and Equipment Technology (CICAEET),
Nanjing University of Information Science & Technology, Nanjing 210044, China)

Abstract: In order to meet the high computing demands caused by emerging compute-intensive applications in Mobile Edge Computing (MEC), this paper proposes a Digital Twin (DT)-empowered task offloading scheme where Reconfigurable Intelligent Surface (RIS) is used to enhance the communication links and extend the coverage. Firstly, the joint optimization of user offloading strategy, RIS phase-shift vector, beamforming vector, transmit power of users and computation capacity allocation are investigated with the aim of minimizing the total energy consumption of users and resource devices under the constraints of communication and computing resources. Then, the formulated non-convex combinational optimization problem is decomposed into three sub-problems, including RIS phase-shift design, binary optimization of transmit power, and computing resource allocation. In addition, the Double Deep Q Network (DDQN) approach is invoked to determine the offloading decisions and an alternating iteration optimization algorithm is designed to achieve the optimal solution. Simulation results show that the DDQN-based algorithm is able to train quickly and reduce effectively the total energy consumption of the system.

Key words: Reconfigurable Intelligent Surface (RIS); Digital Twin (DT); Mobile Edge Computing (MEC); Deep reinforcement learning

收稿日期: 2022-02-25; 改回日期: 2022-05-24; 网络出版: 2022-05-30

*通信作者: 李斌 bin.li@nuist.edu.cn

基金项目: 国家自然科学基金(62101277), 江苏省自然科学基金(BK20200822)

Foundation Items: The National Natural Science Foundation of China (62101277), The National Natural Science Foundation of Jiangsu Province (BK20200822)

1 引言

随着人工智能的快速发展,伴随着各种新兴的计算密集型应用,如图像处理、虚拟现实等,对移动设备的计算和处理需求正在空前增长,但由于移动设备自身的局限性,如无线通信状态不佳、计算资源不足等,用户的服务质量得不到保证。为了提高用户的满意度,需要增强其通信和计算能力。因此,智能超表面(Reconfigurable Intelligent Surface, RIS)^[1,2]和移动边缘计算(Mobile Edge Computing, MEC)^[3-5]作为两种新范式应运而生。

RIS具有低成本、易部署、可重构无线传播环境、无源波束成形等优点,引起了人们的高度关注^[6-10]。MEC技术允许移动用户通过无线链路将任务卸载到基站(Base Station, BS)的MEC服务器,将资源有限的用户从繁重的计算中解放出来。因此,将RIS应用到MEC系统中不仅可以增强用户通信能力,而且还是一种经济高效且环保的方式。目前,关于RIS辅助MEC的研究已取得许多有价值的研究成果^[11,12]。譬如,为了解决任务卸载速率低的问题,文献^[13]将RIS引入MEC系统中,通过联合优化任务卸载量,MEC服务器的计算资源以及RIS相移实现系统总时延最小化。为了探索RIS对MEC系统计算性能的影响,文献^[14]在能耗和时延的约束下,研究了用户最大卸载量的问题。为了最大化系统能效,文献^[15]通过联合优化发射功率,本地计算频率以及有源和无源波束成形,使得能量效率最大化。为了对比多址技术对RIS辅助MEC的影响,文献^[16]分别对比了非正交多址和时分多址协议之间的性能,文中数值实验表明非正交多址协议可以提升系统的性能。在边缘智能网络中,文献^[17]通过联合优化用户的发射功率、基站的接收波束形成向量和RIS的相移矩阵,将所有参与用户的最大学习误差降至最低。对于RIS辅助的MEC系统,公式化的问题是非凸且多优化变量紧密耦合的。因此,解决耦合的多变量问题是必要的。迭代算法能够提供接近最优的解决方案,但计算复杂度较高,这会阻碍它在实际网络中的使用。

近年来,数字孪生(Digital Twin, DT)技术引起了国内外学术界和工业界广泛关注,它是物理实体的数字副本,将真实的物理实体和环境实时地映射到虚拟空间。在虚拟空间中运用智能学习算法和DT提供的实时数据可以帮助物理实体做出更准确、更及时的卸载决策,降低用户决策的资源消耗^[18]。为此,DT为求解无线网络中高复杂度问题提供了一种新思路。鉴于DT的优势,已有相关工作将DT和MEC相结合,构建了数字孪生边缘网络(Digital

Twin Edge Network, DTEN)。为了研究DT在多任务MEC系统中的作用,文献^[19]提出了一种基于MEC在线卸载方案,并利用李雅普诺夫优化求解其能效最大化问题。为了研究DT在联邦学习中的作用,文献^[20]将DT引入区块链辅助的联邦学习系统中,并提出了一种低复杂度的训练方案。为了探究DT针对用户移动性系统的认知,文献^[21]将DT系统引入移动性系统中,有效地降低了用户的卸载时延。为了任务能够智能卸载到MEC服务器上,文献^[22]将DT引入进多小区网络中,进而优化了系统的性能。

上述研究展示出DT和RIS分别在提升MEC网络性能的优势,然而在DT结合MEC研究中物理实体的无线通信大多仅考虑视距链路,现实中物理实体的通信链路通常为非视距链路,同时DT需要物理实体实时地更新感知数据,因此对无线信道质量要求较高。本文提出一种RIS赋能的DTEN,通过RIS改善无线信道环境,DT实时地监控网络状态信息,以集中的视角为用户提供卸载决策,如何有效地求解RIS辅助DTEN中用户卸载决策是一个新的挑战。

本文的主要工作如下:(1)将RIS引入到DTEN中,从能耗和时延的角度联合优化RIS相移、波束成形矢量、计算资源分配和用户传输功率,构建一个非凸、多变量、耦合的能耗最小化问题。(2)为了求解该非凸优化问题,首先将用户卸载过程建模为马尔可夫决策过程(Markov Decision Process, MDP),并探索DT辅助智能任务卸载的方式,采用深度双Q网络(Double Deep Q Network, DDQN)算法求解卸载决策,并使用网络中物理设备的映射数据来训练神经网络模型。然后将交替优化算法求解RIS相移、波束成形矢量、传输功率和计算资源分配的部分嵌入到强化学习训练中,提出一种能耗最小化的资源分配算法。

2 系统模型及问题描述

2.1 系统模型

如图1所示,本文考虑两层DTEN模型,分别为物理实体层和数字孪生层。其中, K 个用户设备、 Q 个资源设备(Resource Device, RD)、1面具有 N 个反射单元的RIS和1个装有 M 根天线并搭载MEC服务器的基站共同构成了物理实体层,所有物理实体的数字孪生和无线通信环境共同构成了数字孪生层。物理实体层中的设备单元都通过实时通道将当前运行状态发送到DT,DT系统根据物理实体层的数据构建了真实世界的虚拟模型,在该模型

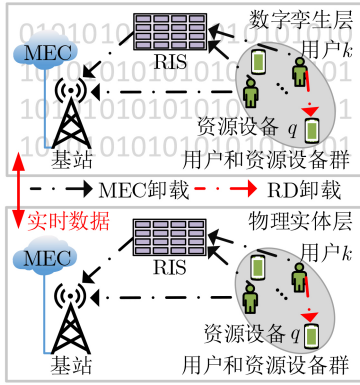


图1 智能超表面辅助DTEN模型图

中能够对用户卸载策略进行模拟，并直观地评估系统总能耗。假设用户和基站之间直连链路为非视距信道其余链路均为视距信道，因此部署RIS以增强用户和基站之间的网络的性能。此外，基站和资源设备的位置固定，用户的位置随时间变化，用户和资源设备只考虑装配单根天线。为了便于表达和分析，定义用户、资源设备和RIS反射单元的集合分别为 $\forall k \in \mathcal{K} \triangleq \{1, 2, \dots, K\}$ ， $\forall q \in \mathcal{Q} \triangleq \{1, 2, \dots, Q\}$ 和 $\forall n \in \mathcal{N} \triangleq \{1, 2, \dots, N\}$ 。

(1) DT模型。本文考虑了3种类型的数字孪生：基站、用户和资源设备。由于数字孪生重现了物理实体的运行，因此维护海量设备的数字孪生将会消耗大量资源。为了使数字孪生网络能够平稳运行，所有数字孪生构建在基站MEC中。在 t 时刻，计算实体关于其计算频率的数字孪生可表示为

$$DT^t = \theta(f^t, \tilde{f}^t) \quad (1)$$

其中， f^t 是DT对真实计算频率的估计值； \tilde{f}^t 是真实计算频率与DT估计值之间的误差。由于DT和真实设备之间数据交互存在时延误差，DT有时不能准确反映设备的真实状态，但是真实计算时延和DT估计时延之间的误差可以提前获得^[21]。假设 t 时刻任务数据量大小为 D^t ，计算单位比特数据所需计算周期为 C^t ，真实计算时延和DT估计时延的误差可表示为

$$\begin{aligned} \Delta T &= \frac{D^t C^t}{f^t + \tilde{f}^t} - \frac{D^t C^t}{f^t} \\ &= \frac{-D^t C^t \tilde{f}^t}{(f^t + \tilde{f}^t) f^t} \end{aligned} \quad (2)$$

(2) 卸载计算模型。在 t 时刻，每个用户随机生成不可分割的计算任务记为 $M_k^t = \{D_k^t, C_k^t, T_k^t\}$ ，其中， D_k^t 为任务数据量大小， C_k^t 为计算单位比特数据所需的计算周期， T_k^t 为任务时延限制。本文采用时分多址技术避免任务卸载时不同用户之间的干

扰。将任务卸载周期 T 等分为 K 个时隙，用户 k 在所分得时隙 $\tau_k = TK^{-1}$ 内进行任务卸载。定义卸载决策集合为 $\forall j \in \mathcal{J} \triangleq \{0, 1, \dots, Q+1\}$ ，用户 k 根据DT的卸载因子 $b_{k,j} \in \{0, 1\}$ 来决定任务的卸载决策， $b_{k,0} = 1$ 表示任务在本地计算， $b_{k,q} = 1$ 表示任务在资源设备计算， $b_{k,Q+1} = 1$ 表示任务在边缘服务器计算。由于用户任务不可分割性，采用完全卸载方式，卸载因子需满足 $\sum_{i=0}^{Q+1} b_{k,i} = 1$ 。

本地设备处理任务时，用户 k 所需计算资源记为 f_k^l 。数字孪生层中用户 k 的估计时延可计算为 $\tilde{T}_k^l = D_k^t C_k^t (f_k^l)^{-1}$ 。根据DT模型，由式(2)可得本地计算时延误差 ΔT_k^l 。因此，用户 k 在本地设备处理任务的总时延为 $T_k^{\text{all}} = \tilde{T}_k^l + \Delta T_k^l$ 。

当用户选择将任务卸载至基站时，设时隙 τ_k 内用户 k 传输功率为 p_k ，用户 k 发送信号为 s_k 且满足 $E|s_k| = 1$ ，则基站在时隙 τ_k 内接收到信号为 $y_k = (\mathbf{G}\Phi_k \mathbf{g}_{r,k} + \mathbf{g}_{d,k}) \sqrt{p_k} s_k + \mathbf{n}_k$ ，其中， $\Phi_k \triangleq \text{diag}(e^{j\theta_{k,1}}, e^{j\theta_{k,2}}, \dots, e^{j\theta_{k,N}})$ 为RIS的相移矩阵， $\theta_{k,n} \in [0, 2\pi)$ 代表RIS上第 n 个反射单元的相移； $\mathbf{g}_{r,k} \in \mathbb{C}^{N \times 1}$ ， $\mathbf{G} \in \mathbb{C}^{M \times N}$ 和 $\mathbf{g}_{d,k} \in \mathbb{C}^{M \times 1}$ 分别为用户 k 到RIS，RIS到基站，用户 k 到基站的信道增益； $\mathbf{n}_k \sim \mathcal{CN}(0, \sigma_k^2 \mathbf{I})$ 为加性高斯白噪声。在接收信号 y_k 时，基站采用线性波束成形技术，用户 k 到基站的接收波束成形矢量为 $\mathbf{w}_k \in \mathbb{C}^{M \times 1}$ ，则基站在时隙 τ_k 内信号估计为 $\tilde{y}_k = \mathbf{w}_k^H y_k$ 。根据香农定理，时隙 τ_k 内用户 k 至基站的传输速率为

$$r_k = B \log_2 \left(1 + \frac{p_k |\mathbf{w}_k^H (\mathbf{G}\Phi_k \mathbf{g}_{r,k} + \mathbf{g}_{d,k})|^2}{\sigma^2 \mathbf{w}_k^H \mathbf{w}_k} \right) \quad (3)$$

其中， B 是通信带宽。用户 k 卸载任务至基站的传输时延为 $t_k^c = D_k^t (r_k)^{-1}$ ，基站收到用户 k 的数据后，MEC会为用户 k 分配计算资源 f_k^c ，此时数字孪生层中MEC的估计时延可计算为 $\tilde{T}_k^c = D_k^t C_k^t (f_k^c)^{-1}$ 。根据DT模型，由式(2)可得MEC计算时延误差 ΔT_k^c 。因此，MEC的真实计算时延为 $T_k^c = \tilde{T}_k^c + \Delta T_k^c$ ，则用户 k 卸载任务至基站总时延为 $T_k^{\text{all}} = t_k^c + T_k^c$ 。

用户 k 的任务除了卸载至基站或者本地， Q 个资源设备也可以通过D2D(Device-to-Device)通信为用户提供计算服务。在时隙 τ_k 内，用户 k 在 Q 个资源设备中选择一处卸载，其余用户、 $Q-1$ 个资源设备和基站处于静默状态。假设时隙 τ_k 内用户 k 至资源设备 q 的信道增益为 $h_{k,q} = \beta_0 (d_{k,q})^{-2}$ ，其中， $d_{k,q}$ 为时隙 τ_k 内用户 k 至资源设备 q 的距离， β_0 为每单位米的路径损耗因子。根据香农定理，时隙 τ_k 内用户 k 至资源设备 q 的数据传输速率为

$$r_{k,q} = B \log_2 \left(1 + \frac{p_k h_{k,q}}{\sigma^2} \right) \quad (4)$$

用户 k 卸载任务至资源设备 q 的传输时延为 $t_{k,q}^t = D_k^t (r_{k,q})^{-1}$ 。资源设备 q 收到用户 k 的数据后，会为用户 k 分配计算资源 $f_{k,q}^c$ 。数字孪生层中资源设备 q 的估计时延可计算为 $\tilde{T}_{k,q}^r = D_k^t C_k^t (f_{k,q}^c)^{-1}$ 。根据DT模型，由式(2)可得资源设备计算时延误差 $\Delta T_{k,q}^r$ 。因此，资源设备 q 的真实计算时延为 $T_{k,q}^r = \tilde{T}_{k,q}^r + \Delta T_{k,q}^r$ ，则用户 k 卸载任务至资源设备总时延为 $T_{k,q}^{\text{rall}} = t_{k,q}^t + T_{k,q}^r$ 。

(3) 能耗模型。通常，BS有充足的电力资源，BS的能耗对系统影响较小，本文只考虑用户和资

源设备的能耗。当任务在本地设备处理时，用户 k 的计算能耗为 $E_k^l = k_1 (f_k^l)^2 D_k^t C_k^t$ ；当任务卸载至基站时，用户 k 的卸载能耗为 $E_k^e = t_{k,q}^t p_k$ ；当任务卸载至资源设备时，能耗分为两部分：用户 k 的卸载能耗为 $E_{k,q}^{\text{off}} = t_{k,q}^t p_k$ ，资源设备 q 的计算能耗为 $E_{k,q}^{\text{com}} = k_q (f_{k,q}^c)^2 D_k^t C_k^t$ ，其中， k_1 和 k_q 为功率消耗因子。

2.2 问题描述

本文目标是在时延、功率、计算资源和能量的限制条件下，通过联合优化用户卸载决策、智能超表面相移、基站端接收波束成矢量、用户传输功率和计算资源分配以最小化任务周期 T 内用户和资源设备处理任务的总能耗，因此该优化问题可以表述为

$$\left. \begin{aligned} & \min_{b_{k,j}, \theta_{k,n}, \mathbf{w}_k, p_k, f_k^l, f_{k,q}^c, f_k^e} \sum_{k \in \mathcal{K}} \left(b_{k,0} E_k^l + \sum_{q \in \mathcal{Q}} b_{k,q} (E_{k,q}^{\text{off}} + E_{k,q}^{\text{com}}) + b_{k,Q+1} E_k^e \right) \\ & \text{s.t. C1: } \sum_{i=0}^{Q+1} b_{k,i} = 1, b_{k,j} \in \{0, 1\}, \quad \forall k \in \mathcal{K}, \forall j \in \mathcal{J} \\ & \text{C2: } |\exp(j\theta_{k,n})| = 1, \quad \forall k \in \mathcal{K}, \forall n \in \mathcal{N} \\ & \text{C3: } 0 \leq p_k \leq p^{\max}, \quad \forall k \in \mathcal{K} \\ & \text{C4: } 0 \leq f_k^l \leq f_1^{\max}, \quad \forall k \in \mathcal{K} \\ & \text{C5: } 0 \leq f_k^e \leq f_e^{\max}, \quad \forall k \in \mathcal{K} \\ & \text{C6: } 0 \leq f_{k,q}^c \leq f_r^{\max}, \quad \forall k \in \mathcal{K}, \forall q \in \mathcal{Q} \\ & \text{C7: } 0 \leq b_{k,0} T_k^{\text{lall}} + b_{k,q} T_k^{\text{rall}} + b_{k,Q+1} T_k^{\text{eall}} \leq T_k^{\text{min}}, \quad \forall k \in \mathcal{K}, \forall q \in \mathcal{Q} \\ & \text{C8: } 0 \leq b_{k,0} E_k^l + b_{k,q} E_{k,q}^{\text{off}} + b_{k,Q+1} E_k^e \leq E_1^{\max}, \quad \forall k \in \mathcal{K}, \forall q \in \mathcal{Q} \\ & \text{C9: } 0 \leq b_{k,Q+1} E_{k,q}^{\text{com}} \leq E_r^{\max}, \quad \forall k \in \mathcal{K}, \forall q \in \mathcal{Q} \end{aligned} \right\} \quad (5)$$

其中， $T_k^{\text{min}} = \min \{\tau_k, T_k^t\}$ 。C1约束为用户卸载点限制，C2约束为RIS相移限制，C3约束为用户传输功率限制，C4~C6约束为用户、MEC和资源设备计算资源限制，C7约束为任务计算时延限制，C8~C9约束为用户和资源设备能耗限制。由于目标函数的非线性和约束条件的非凸性，使得问题式(5)是一个多变量耦合的非凸问题，难以直接求解。

3 优化问题求解

为了有效求解上述优化问题，将问题式(5)分

$$\left. \begin{aligned} & \min_{b_{k,j}} \sum_{k \in \mathcal{K}} \left(b_{k,0} E_k^l + \sum_{q \in \mathcal{Q}} b_{k,q} (E_{k,q}^{\text{off}} + E_{k,q}^{\text{com}}) + b_{k,Q+1} E_k^e \right) \\ & \text{s.t. C1: } \sum_{i=0}^{Q+1} b_{k,i} = 1, b_{k,j} \in \{0, 1\}, \quad \forall k \in \mathcal{K}, \forall j \in \mathcal{J} \end{aligned} \right\} \quad (6)$$

由于问题式(6)是一个整数优化问题，难以求解用户最优卸载策略。本节提出基于DDQN算法的用户卸载策略优化方案。首先介绍了强化学习中MDP的基本要素，然后给出了基于DDQN算法优化用户决策的算法流程。

解为3个子问题。首先，使用DDQN算法输出用户的卸载决策；然后，在DRL训练环境中，使用交替迭代的方法，在给定用户决策、传输功率和计算频率下优化RIS相移和波束成矢量；最后，在给定用户决策、RIS相移和波束成矢量下优化传输功率和计算频率。为方便讨论，本文首先阐述了基于DDQN的卸载决策算法流程，然后提出了在DRL环境中交替迭代的整体算法。

3.1 用户卸载决策优化

在给定RIS相移、波束成形矢量、传输功率和计算频率时，问题式(5)可以描述为

3.1.1 MDP基本要素定义

MDP是一种对智能体和动态环境交互过程进行建模的方法。本文考虑的模型中，基站和资源设备不需要任何环境的先验信息，只能从环境状态中获取因果关系，所以模型中转移概率是未知的。因

此可以建模为无模型、无转移概率的马尔可夫决策过程。下面依次对MDP的3要素,即状态空间、动作空间和奖励函数分别进行了定义。

(1) 状态空间: $S[t] = \{s[t] | s[t] = \{M_k^t, d_k^t\}\}$ 。其中, M_k^t 为 t 时刻用户 k 的任务信息, $d_k^t = \{d_{k,0}^t, d_{k,1}^t, \dots, d_{k,Q+1}^t\}$ 表示 t 时刻用户 k 与基站、 Q 个资源设备以及RIS之间的距离向量。

(2) 动作空间: $A[t] = \{a[t] | a[t] = b_{k,j}, b_{k,j} \in \{0,1\}, \sum_{j=0}^{Q+1} b_{k,j} = 1\}$ 。其中, $a[t]$ 是用户 k 在 t 时刻卸载方式。

(3) 奖励函数: $r(t) = -E_k^t - \vartheta - C$ 。其中, $E_k^t = b_{k,0}E_k^1 + \sum_{q \in \mathcal{Q}} b_{k,q}(E_{k,q}^{\text{off}} + E_{k,q}^{\text{com}}) + b_{k,Q+1}E_k^c$ 是处理用户 k 任务所消耗的能量; ϑ 是任务卸载至本地,资源设备和基站的开销向量; C 是任务处理失败的惩罚值。

3.1.2 基于DDQN的能耗最小化算法

本文所提基于DDQN的能耗最小化算法与深度Q网络(Deep Q Network, DQN)学习算法一样,构建了两个结构相同但参数不同的神经网络:主网络和目标网络。其中主网络的参数实时更新,而目标网络的参数每隔 Z 步进行更新,这样可以缓解DQN算法带来的 Q 值高估。同时加入回放经验数组来解决数据之间的关联性,把每个时间段中用户与环境交互得到的记忆存储到回放经验数组,训练时随机选出一部分样本来训练。定义主网络损失函数为

$$L(\omega) = \mathbb{E} \left[\left(Y_t^{\text{DDQN}} - Q(s[t], a[t]; \omega) \right)^2 \right] \quad (7)$$

其中, ω 为主网络的参数; Y_t^{DDQN} 为目标网络的输出,用来计算采取所有可能动作后的目标 Q 值; $Q(s[t], a[t]; \omega)$ 为主网络的输出,用来计算当前状态 $s[t]$ 下执行动作 $a[t]$ 的 Q 值。

基于DQN的算法中,目标网络的输出为 $Y_t^{\text{DQN}} = r[t] + \gamma \max_a Q(s[t+1], a; \omega^-)$,其中, ω^- 为目标网络参数, γ 为折扣因子。选择最大 Q 值的动作和计算此动作对应 Q 值均使用目标网络,这会导致 Q 值高估问题。为了解决该问题,在DDQN算法中,目标网络输出修改为 $Y_t^{\text{DDQN}} = r[t] + \gamma Q(s[t+1], a_{\max}; \omega^-)$ 。首先利用主网络求出 Q 值最大的动作 $a_{\max} = \arg \max_a Q(s[t+1], a; \omega)$;然后使用目标网络计算动作 a_{\max} 的 Q 值;最后根据式(7)计算损失函数,使用梯度下降反向更新主网络参数 ω 。此外,基于DDQN的算法中, Q 值的更新满足贝尔曼方程 $Q(s[t], a[t]) = Q(s[t], a[t]) + \beta [r[t+1] + \gamma Q(s[t+1], a_{\max}) - Q(s[t], a[t])]$,其中, β 为学习率。

基于DDQN卸载决策算法与DRL训练框架如图2所示。数字孪生层主要由智能体、回放经验数组、主网络、目标网络和损失函数五大模块构成,物理实体层由用户和卸载方式模块组成。在 t 时刻,用户将任务信息和坐标数据发送给DT,DT根据用户的数据和对实体环境监控的信息使用DDQN算法训练用户的卸载策略,训练完成后智能体将卸载策略发送给用户,用户根据卸载策略来判断任务的执行方式。

3.2 RIS相移和波束赋形优化

本节固定传输功率、计算频率和卸载决策,优化RIS相移和波束成形矢量。去除目标函数与约束条件中与RIS相移和波束成形矢量无关的部分,优化RIS相移和波束成形矢量子问题可以表示为

$$\left. \begin{aligned} \max_{\theta_{k,n}, \mathbf{w}_k} & \left| \mathbf{w}_k^H (\mathbf{G} \Phi_k \mathbf{g}_{r,k} + \mathbf{g}_{d,k}) \right|^2 \\ \text{s.t. } & C2: |\exp(j\theta_{k,n})| = 1, \quad \forall k \in \mathcal{K}, \forall n \in \mathcal{N} \end{aligned} \right\} \quad (8)$$

问题式(8)是一个非凸问题,难以直接求解。通过固定波束成形矢量,将目标函数转换为 $|\mathbf{w}_k^H (\mathbf{G} \Phi_k \mathbf{g}_{r,k} + \mathbf{g}_{d,k})|^2 = |\mathbf{w}_k^H \mathbf{G} \text{diag}(\mathbf{g}_{r,k}) e^{j\theta_k} + \mathbf{w}_k^H \mathbf{g}_{d,k}|^2$,其中, $\theta_k = [\theta_{k,1}, \theta_{k,2}, \dots, \theta_{k,N}]$ 。对转换后的目标函数使用三角不等式 $|\mathbf{w}_k^H \mathbf{G} \text{diag}(\mathbf{g}_{r,k}) e^{j\theta_k} + \mathbf{w}_k^H \mathbf{g}_{d,k}| \leq \sum_n |\mathbf{w}_k^H \mathbf{g}_{d,k}| + |(\mathbf{w}_k^H \mathbf{G} \text{diag}(\mathbf{g}_{r,k})) [n]|$ 可得目标函数上界。当等式成立时,满足 $|\exp(j\theta_{k,n})| = 1$,此时最优相移为 $(\theta_{k,n})^* = \arg(\mathbf{w}_k^H \mathbf{g}_{d,k}) - \arg((\mathbf{w}_k^H \mathbf{G} \text{diag}(\mathbf{g}_{r,k})) [n])$ 。

求出最优RIS相移后,采用合并最大比准则^[23],最佳接收波束成形的表述形式为

$$(\mathbf{w}_k)^* = \frac{\mathbf{G} \Phi_k \mathbf{g}_{r,k} + \mathbf{g}_{d,k}}{\|\mathbf{G} \Phi_k \mathbf{g}_{r,k} + \mathbf{g}_{d,k}\|} \quad (9)$$

3.3 传输功率和计算频率优化

在给定卸载决策、RIS最优相移和最佳波束成形矢量时,优化传输功率和计算频率的子问题可以描述为

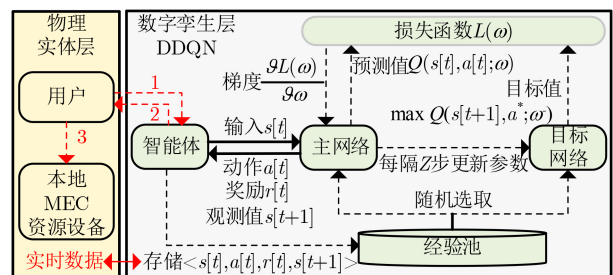


图2 基于DDQN的DRL训练框架图

$$\min_{p_k, f_k^l, f_{k,q}^r, f_k^e} \left\{ \sum_{k \in \mathcal{K}} \left(b_{k,0} E_k^l + \sum_{q \in \mathcal{Q}} b_{k,q} (E_{k,q}^{\text{off}} + E_{k,q}^{\text{com}}) + b_{k,Q+1} E_k^e \right) \right\} \quad (10)$$

s.t. C3 ~ C9

上述问题中目标函数关于变量 $p_k, f_k^l, f_{k,q}^r, f_k^e$ 递减，求其解分为3种情况。当本地计算时， $(p_k)^* = (f_{k,q}^r)^* = (f_k^e)^* = 0$ ，约束条件为线性约束，由时延约束C7可得最优计算频率 $(f_k^l)^* = D_k^t C_k^t (T_k^{\min})^{-1}$ ；当卸载至基站时， $(f_k^l)^* = (f_{k,q}^r)^* = 0$ ，目标函数与变量 f_k^e 无关，MEC以最大计算频率 $(f_k^e)^* = f_e^{\max}$ 处理任务可使计算时延最小化，将其代入时延约束C7可得用户最优传输功率 $(p_k)^* = (2^{c_k} - 1) \sigma^2 \mathbf{w}_k^H \mathbf{w}_k |h_k|^{-2}$ (其中， $c_k = D_k^t (B(T_k^{\min} - D_k^t C_k^t (f_e^{\max})^{-1}))^{-1}$)；当卸载至资源设备时， $(f_k^l)^* = (f_k^e)^* = 0$ ，时延约束C7耦合了传输功率和计算频率，当能耗最小时，满足 $t_{k,q}^r + T_{k,q}^r = T_k^{\min}$ ，使用二分法求解变量 p_k 和 $f_{k,q}^r$ 次优解。

基于前文分析，在DRL环境中交替迭代算法的具体过程如表1所示。

算法1中深度强化学习的时间复杂度通常以迭代中乘法的计算次数来衡量。因此，算法1中深度强化学习的复杂度为 $O\left(\sum_l^L n_l n_{l-1}\right)$ ，其中 n_l 是第 l 隐藏层的神经单元数量。本文算法中隐藏层的大小相近，因此深度强化学习的复杂度可以近似为

$O((l-1)n_l^2) \sim O(n_l^2)$ 。此外，精度 ε 下二分法的复杂度为 $O(\log_2 \varepsilon^{-1})$ ，交替迭代算法执行 P 次，则总的算法时间复杂度为 $O\left(\sum_l^L n_l n_{l-1} + P \log_2 \varepsilon^{-1}\right)$ 。

4 仿真结果

本节将通过Pytorch框架仿真并分析所提方案的有效性。以(0,0,0)建立3维坐标系，8个用户和4个资源设备分别随机分布在半径10 m和20 m的圆内，圆心水平坐标为(50,20) m，RIS的坐标为(200,0,10) m，配有MEC的基站在(0,400,5) m，任务的数据量 $D_k^t \in [0, 5 \times 10^5]$ bit，单位比特平均计算次数 $C_k^t \in [150, 200]$ cycles/bit，任务最大时延 $T_k \in [0.02, 0.04]$ s，DT的误差率设为0.9。时间 T 为1 s，带宽 B 为 10^7 Hz，最大传输功率 p^{\max} 为23 dBm，噪声功率 σ^2 为-60 dBm，信道增益 β_0 为-30 dB。假设所有用户的能耗预算 E_1^{\max} 和最大计算频率 f_1^{\max} 分别为10 J和0.6 GHz，资源设备的能耗预算 E_r^{\max} 和最大计算频率 f_r^{\max} 分别为20 J和2.4 GHz，MEC最大计算频率 f_e^{\max} 为5 GHz，CPU电容系数 k_1 和 k_q 均为 10^{-26} 。用户和基站之间的通信采用文献[24]的Rician信道模型。DDQN训练参数如表2所示。

表 1 基于DDQN能耗最小化算法(算法1)

输入：最大回合数 E ，学习率 β ，折扣回报 γ ，用户数 K ，二分法精度 ε 。	
步骤1	初始化主网络参数 ω ，目标网络参数 ω^- ，经验数组，用户、基站和资源设备位置，用户任务信息 M_k^t
步骤2	for $t = 1 : E$
	更新 t 时刻用户位置和任务信息；
	for step = 1 : K
	根据当前状态 $s[t]$ 和策略 π_ω 选择动作 $a[t]$ ；
	根据3.2.3.3节讨论和动作 $a[t]$ ，使用交替迭代法和二分法求解 $\theta_{k,n}, \mathbf{w}_k, p_k, f_k^l, f_{k,q}^r, f_k^e$ ；
	根据奖励函数计算 $r(t)$ ，观察下一个状态 $s[t+1]$ ，并存储 $(s[t], a[t], r[t], s[t+1])$ 到经验数组；
	从经验回放数组随机取出一组经验，根据式(7)计算损失函数，并更新当前主网络参数 ω ；
	每隔 Z 步更新目标网络参数 ω^- ；
步骤3	输出任务卸载策略网络参数 ω 。

表 2 DDQN训练参数

参数	值	参数	值
隐藏层数量 L	3	惩罚值 C	100
折扣回报 γ	0.9	开销值 ϑ	(0,1,5)
最大回合数 E	100 k	二分法精度 ε	10^{-6}
学习率 β	5×10^{-3}	经验回放数组大小	2^{15}
目标网络更新频率 Z	320	贪婪策略比例	0.1

图3验证了DDQN算法的收敛性。从中观察到,随着训练次数的增加,智能体对用户的任务卸载决策逐渐变好,曲线振荡有明显的下降趋势,最终奖励值稳定在-75左右,说明此时智能体已经能为用户提供比较好的卸载决策。为验证学习率对算法收敛性的影响,本文还比较了不同学习率下的曲线收敛情况。由图可知,当学习率为0.001时,在经过75 k步后趋近于收敛;当学习率为0.005时,在经过10 k步后趋近于收敛。两者曲线收敛后奖励相差不大且稳定,这说明学习率对算法收敛速度有影响,但对性能影响不大。

图4描述了不同方案的收敛性能。由图可知,随着迭代次数的增加,4种方案下奖励值逐渐趋于稳定。算法1在经过10 k步后趋近于收敛,而基于DQN方法的算法1在经过12 k步后趋近于收敛,DDQN算法收敛速度快于DQN算法,这与3.1.1节中分析吻合,进一步证明了所提算法能够有效缓解DQN算法的高估问题。无RIS方案收敛后奖励值最低。这是因为用户和基站之间距离较远,无线信道质量较差,导致用户和基站之间无线传输速率较慢。当卸载至基站时,任务处理超时,智能体受惩罚,所以无RIS方案稳定后的奖励值低于其他3种方案。RIS相移随机方案在经过17 k步后趋近于收敛,收敛后的奖励值略低于算法1但高于无RIS方案。这是因为随机相移方案下用户到基站的无线速

率始终低于算法1,导致此方案下用户卸载任务的开销增大,奖励值偏低。仿真结果充分说明了将RIS引入到DTEN中有较好的性能提升。

图5描述了RIS反射单元数目与用户数量对系统能耗的影响。设置用户任务量为500 kbit,单位比特数据所需计算周期为200 cycles。由图可知,在用户数目固定时,随着RIS反射单元数目的增加,信道增益逐渐增大,无线信道质量明显提升,用户总能耗逐渐下降。当RIS反射单元数目和任务卸载周期 T 不变时,随着用户数量的增大,每位用户分得的时隙变小,因此用户需在时隙内以较大的功率进行传输任务,用户的总能耗增加。此外,当RIS反射单元个数为36时,用户总能耗始终保持最低。

为进一步探究DT对系统的提升作用,图6描述了不同方案下DT对系统能耗的影响。设置用户数 K 为6。所提算法1和采用贪心算法方案分别在DT和无DT下进行了对比。从中观察到,随着任务量的增加,系统能耗也逐渐增加,无DT方案和无DT方案系统能耗差距越来越明显。当任务量为280 kbit时,算法1的DT方案比无DT方案能耗降低约18.9%。这是由于DT根据用户传过来的信息辅助其进行决策,用户只需要执行DT发送过来的指令,节省了用户寻找最佳卸载方式的能耗。当使用贪心算法进行用户决策时,DT和无DT的情况下能耗均比算法

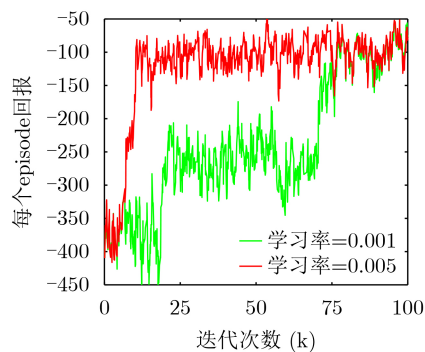


图3 DDQN算法收敛性图

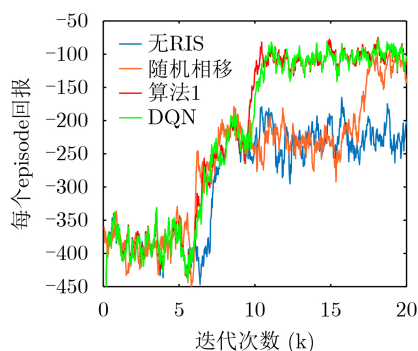


图4 不同方案性能对比图

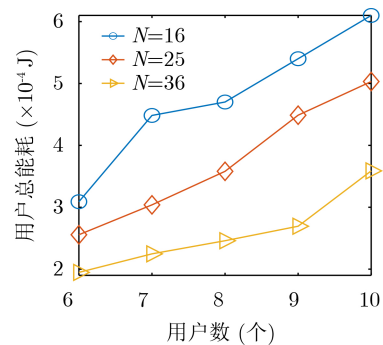


图5 用户数量与用户总能耗间的关系

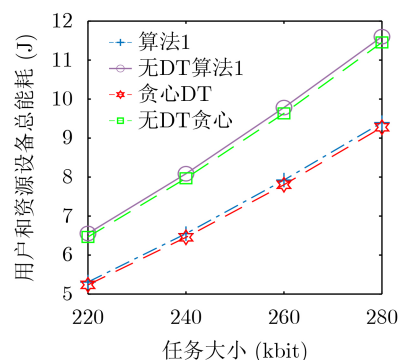


图6 不同方案下系统能耗与任务量关系图

1降低1.2%左右。这是因为使用贪心算法可获得最优解,但是使用贪心算法求解用户卸载决策的时间复杂度较高,用户数量较大时可能会影响用户体验。算法1虽然不能达到最优解,但随着DT中模型不断的训练,DT能够在较短的时间内为用户提供一个比较好的卸载决策,节省用户寻找最佳卸载点的能耗。仿真结果进一步验证了DT可以对系统带来较好的性能。

5 结束语

本文研究了RIS辅助DTEN中计算任务卸载方案。以用户和资源设备的总能耗最小化为目标,通过联合优化用户卸载策略、RIS相移、波束成形矢量、传输功率和计算资源分配,实现用户任务的智能卸载,并提出了一种基于DDQN和交替迭代的任务卸载算法。仿真结果表明,在强化学习中嵌入优化方法能够快速得到优化解。在未来工作中,考虑多用户任务卸载过程中信道时变的DTEN方案。

参 考 文 献

- [1] ZHANG Jiayi, LIU Heng, WU Qingqing, *et al.* RIS-aided next-generation high-speed train communications: Challenges, solutions, and future directions[J]. *IEEE Wireless Communications*, 2021, 28(6): 145–151. doi: [10.1109/MWC.001.2100170](https://doi.org/10.1109/MWC.001.2100170).
- [2] 徐勇军, 高正念, 王笛竹, 等. 基于智能反射面辅助的无线供电通信网络鲁棒能效最大化算法[J]. 电子与信息学报, 待发表. doi: [10.11999/JEIT210714](https://doi.org/10.11999/JEIT210714).
XU Yongjun, GAO Zhengnian, WANG Qianzhu, *et al.* Robust energy efficiency maximization algorithm for intelligent reflecting surface-aided wireless powered-communication networks[J]. *Journal of Electronics & Information Technology*, To be published. doi: [10.11999/JEIT210714](https://doi.org/10.11999/JEIT210714).
- [3] MAO Yuyi, YOU Changsheng, ZHANG Jun, *et al.* A survey on mobile edge computing: The communication perspective[J]. *IEEE Communications Surveys & Tutorials*, 2017, 19(4): 2322–2358. doi: [10.1109/COMST.2017.2745201](https://doi.org/10.1109/COMST.2017.2745201).
- [4] XU Yongjun, GUI Guan, GACANIN H, *et al.* A survey on resource allocation for 5G heterogeneous networks: Current research, future trends, and challenges[J]. *IEEE Communications Surveys & Tutorials*, 2021, 23(2): 668–695. doi: [10.1109/COMST.2021.3059896](https://doi.org/10.1109/COMST.2021.3059896).
- [5] XU Yongjun, GU Bowen, HU R Q, *et al.* Joint computation offloading and radio resource allocation in MEC-based wireless-powered backscatter communication networks[J]. *IEEE Transactions on Vehicular Technology*, 2021, 70(6): 6200–6205. doi: [10.1109/TVT.2021.3077094](https://doi.org/10.1109/TVT.2021.3077094).
- [6] CHEN Yuanbin, WANG Ying, ZHANG Jiayi, *et al.* QoS-driven spectrum sharing for reconfigurable intelligent surfaces (RISs) aided vehicular networks[J]. *IEEE Transactions on Wireless Communications*, 2021, 20(9): 5969–5985. doi: [10.1109/TWC.2021.3071332](https://doi.org/10.1109/TWC.2021.3071332).
- [7] SHI Enyu, ZHANG Jiayi, CHEN Shuaifei, *et al.* Wireless energy transfer in RIS-aided cell-free massive MIMO systems: Opportunities and challenges[J]. *IEEE Communications Magazine*, 2022, 60(3): 26–32. doi: [10.1109/MCOM.001.2100671](https://doi.org/10.1109/MCOM.001.2100671).
- [8] ZHANG Yan, ZHANG Jiayi, DI RENZO M, *et al.* Reconfigurable intelligent surfaces with outdated channel state information: Centralized vs. distributed deployments[J]. *IEEE Transactions on Communications*, 2022, 70(4): 2742–2756. doi: [10.1109/TCOMM.2022.3146344](https://doi.org/10.1109/TCOMM.2022.3146344).
- [9] JIN Yu, ZHANG Jiayi, HUANG Chongwen, *et al.* Multiple residual dense networks for reconfigurable intelligent surfaces cascaded channel estimation[J]. *IEEE Transactions on Vehicular Technology*, 2022, 71(2): 2134–2139. doi: [10.1109/TVT.2021.3132305](https://doi.org/10.1109/TVT.2021.3132305).
- [10] JIN Yu, ZHANG Jiayi, ZHANG Xiaodan, *et al.* Channel estimation for semi-passive reconfigurable intelligent surfaces with enhanced deep residual networks[J]. *IEEE Transactions on Vehicular Technology*, 2021, 70(10): 11083–11088. doi: [10.1109/TVT.2021.3109937](https://doi.org/10.1109/TVT.2021.3109937).
- [11] YANG Zhaohui, XU Wei, HUANG Chongwen, *et al.* Beamforming design for multiuser transmission through reconfigurable intelligent surface[J]. *IEEE Transactions on Communications*, 2021, 69(1): 589–601. doi: [10.1109/TCOMM.2020.3028309](https://doi.org/10.1109/TCOMM.2020.3028309).
- [12] HUANG Chongwen, ZAPPONE A, ALEXANDROPOULOS G C, *et al.* Reconfigurable intelligent surfaces for energy efficiency in wireless communication[J]. *IEEE Transactions on Wireless Communications*, 2019, 18(8): 4157–4170. doi: [10.1109/TWC.2019.2922609](https://doi.org/10.1109/TWC.2019.2922609).
- [13] BAI Tong, PAN Cunhua, DENG Yansha, *et al.* Latency minimization for intelligent reflecting surface aided mobile edge computing[J]. *IEEE Journal on Selected Areas in Communications*, 2020, 38(11): 2666–2682. doi: [10.1109/JSAC.2020.3007035](https://doi.org/10.1109/JSAC.2020.3007035).
- [14] CHU Zheng, XIAO Pei, SHOJAFAR M, *et al.* Intelligent reflecting surface assisted mobile edge computing for internet of things[J]. *IEEE Wireless Communications Letters*, 2021, 10(3): 619–623. doi: [10.1109/LWC.2020.3040607](https://doi.org/10.1109/LWC.2020.3040607).
- [15] WANG Qun, ZHOU Fuhui, HU Han, *et al.* Energy-efficient design for IRS-assisted MEC networks with NOMA[C]. 2021 13th International Conference on Wireless Communications and Signal Processing (WCSP), Changsha, China, 2021: 1–6. doi: [10.1109/WCSP52459.2021.9613291](https://doi.org/10.1109/WCSP52459.2021.9613291).

- [16] LI Zhiyang, CHEN Ming, YANG Zhaohui, *et al.* Energy efficient reconfigurable intelligent surface enabled mobile edge computing networks with NOMA[J]. *IEEE Transactions on Cognitive Communications and Networking*, 2021, 7(2): 427–440. doi: [10.1109/TCCN.2021.3068750](https://doi.org/10.1109/TCCN.2021.3068750).
- [17] HUANG Shanfeng, WANG Shuai, WANG Rui, *et al.* Reconfigurable intelligent surface assisted mobile edge computing with heterogeneous learning tasks[J]. *IEEE Transactions on Cognitive Communications and Networking*, 2021, 7(2): 369–382. doi: [10.1109/TCCN.2021.3056707](https://doi.org/10.1109/TCCN.2021.3056707).
- [18] WU Yiwen, ZHANG Ke, and ZHANG Yan. Digital twin networks: A survey[J]. *IEEE Internet of Things Journal*, 2021, 8(18): 13789–13804. doi: [10.1109/JIOT.2021.3079510](https://doi.org/10.1109/JIOT.2021.3079510).
- [19] DAI Yueyue, ZHANG Ke, MAHARJAN S, *et al.* Deep reinforcement learning for stochastic computation offloading in digital twin networks[J]. *IEEE Transactions on Industrial Informatics*, 2020, 17(7): 4968–4977. doi: [10.1109/TII.2020.3016320](https://doi.org/10.1109/TII.2020.3016320).
- [20] LU Yunlong, HUANG Xiaohong, ZHANG Ke, *et al.* Low-latency federated learning and blockchain for edge association in digital twin empowered 6G networks[J]. *IEEE Transactions on Industrial Informatics*, 2020, 17(7): 5098–5107. doi: [10.1109/TII.2020.3017668](https://doi.org/10.1109/TII.2020.3017668).
- [21] SUN Wen, ZHANG Haibin, WANG Rong, *et al.* Reducing offloading latency for digital twin edge networks in 6G[J]. *IEEE Transactions on Vehicular Technology*, 2020, 69(10): 12240–12251. doi: [10.1109/TVT.2020.3018817](https://doi.org/10.1109/TVT.2020.3018817).
- [22] LIU Tong, TANG Lun, WANG Weili, *et al.* Digital-twin-assisted task offloading based on edge collaboration in the digital twin edge network[J]. *IEEE Internet of Things Journal*, 2022, 9(2): 1427–1444. doi: [10.1109/JIOT.2021.3086961](https://doi.org/10.1109/JIOT.2021.3086961).
- [23] WU Qingqing and ZHANG Rui. Intelligent reflecting surface enhanced wireless network via joint active and passive beamforming[J]. *IEEE Transactions on Wireless Communications*, 2019, 18(11): 5394–5409. doi: [10.1109/TWC.2019.2936025](https://doi.org/10.1109/TWC.2019.2936025).
- [24] HU Xiaoyan, MASOUIROS C, WONG K K, *et al.* Reconfigurable intelligent surface aided mobile edge computing: From optimization-based to location-only learning-based solutions[J]. *IEEE Transactions on Communications*, 2021, 69(6): 3709–3725. doi: [10.1109/TCOMM.2021.3066495](https://doi.org/10.1109/TCOMM.2021.3066495).
- 苏 健: 男, 1986年生, 副教授, 硕士生导师, 研究方向为智能物联网.
- 钱 震: 男, 1999年生, 硕士生, 研究方向为智能反射面技术、移动边缘计算.
- 李 斌: 男, 1987年生, 副教授, 硕士生导师, 研究方向为智能反射面技术、移动边缘计算.

责任编辑: 余 蓉