

## 能量收集辅助的矿山物联网智能计算卸载方法

闵明慧<sup>①②</sup> 张鹏<sup>①</sup> 朱浩鹏<sup>①</sup> 程志鹏<sup>③</sup> 马帅<sup>①</sup>  
李世银<sup>\*①</sup> 肖亮<sup>③</sup> 彭国军<sup>②</sup>

<sup>①</sup>(中国矿业大学信息与控制工程学院 徐州 221116)

<sup>②</sup>(武汉大学空天信息安全与可信计算教育部重点实验室 武汉 430072)

<sup>③</sup>(厦门大学信息学院 厦门 361005)

**摘要:** 针对计算、能量和内存资源受限的矿山物联网设备和大量时延敏感型计算任务需求的智慧矿山场景, 该文提出一种能量收集(EH)辅助的矿山物联网智能计算卸载方法。通过采用移动边缘计算(MEC)技术协助矿山物联网设备任务计算, 同时利用能量收集技术为能量受限的矿山物联网设备供电。基于Q-learning的智能计算卸载机制实现在不可精确获取矿井系统模型的情况下动态探索最优计算卸载策略。此外, 为处理复杂矿井环境下的维度灾难问题并减小策略离散化导致的离散化误差, 该文还提出一种基于深度确定性策略梯度(DDPG)的计算卸载机制来进一步提高井下任务计算卸载性能。理论分析与仿真结果表明所提机制可降低系统的能量损耗、计算时延和任务处理失败率, 有助于保障矿山物联网的安全和高效生产。

**关键词:** 矿山物联网; 移动边缘计算; 能量收集; Q-learning; 深度确定性策略梯度(DDPG)

中图分类号: TN929.5

文献标识码: A

文章编号: 1009-5896(2023)10-3547-11

DOI: [10.11999/JEIT220973](https://doi.org/10.11999/JEIT220973)

## Energy Harvesting Assisted Intelligent Computation Offloading Method for the IoT in Mining

MIN Minghui<sup>①②</sup> ZHANG Peng<sup>①</sup> ZHU Haopeng<sup>①</sup> CHENG Zhipeng<sup>③</sup>  
MA Shuai<sup>①</sup> LI Shiyin<sup>①</sup> XIAO Liang<sup>③</sup> PENG Guojun<sup>②</sup>

<sup>①</sup>(School of Information and Control Engineering, China University of Mining and Technology, Xuzhou 221116, China)

<sup>②</sup>(Key Laboratory of Aerospace Information Security and Trusted Computing, Ministry of Education, Wuhan University, Wuhan 430072, China)

<sup>③</sup>(School of Informatics, Xiamen University, Xiamen 361005, China)

**Abstract:** This paper proposes an Energy Harvesting (EH)-assisted mining IoT intelligent computation offloading method for the mine IoT device with limited computing, energy, and memory resources and smart mining scenario with a large number of latency-sensitive computation tasks. Mobile Edge Computing (MEC) technology is used to assist task computing of mine IoT devices, and EH technology is investigated to power energy-constrained mine IoT devices. The intelligent computation offloading mechanism based on Q-learning can dynamically explore and optimize computation offloading policy under the condition of an unknown precise mine system model. In addition, a computation offloading mechanism based on Deep Deterministic Policy Gradient (DDPG) is proposed. The curse of dimensionality in complex mining scenarios is resolved, the discretization error caused by policy discretization is reduced, and the computation offloading performance is further improved. Theoretical analysis and simulation results verify that the proposed mechanism can reduce energy consumption, computing delays, and task failure rate. This helps ensure safety and improve the

收稿日期: 2022-07-21; 改回日期: 2023-03-31; 网络出版: 2023-04-04

\*通信作者: 李世银 lishiyin@cumt.edu.cn

基金项目: 国家自然科学基金(62101557, 61971366), 中国博士后科学基金(2022M713378), 中央高校基本科研业务费专项资金(2042022kf0021)

Foundation Items: The National Natural Science Foundation of China(62101557, 61971366), China Postdoctoral Science Foundation(2022M713378), Fundamental Research Foundations for the Central Universities(2042022kf0021)

production efficiency of IoT in mining.

**Key words:** IoT in mining; Mobile edge computing; Energy harvesting; Q-learning; Deep Deterministic Policy Gradient (DDPG)

## 1 引言

为推进智慧矿山建设,实现矿山生产自动化、信息化和智能化,需要对矿山物联网中数据密集且时延敏感的计算任务进行及时处理和分析<sup>[1]</sup>。然而,一般情况下,矿山物联网设备计算、能量和内存资源受限<sup>[2]</sup>。因此,针对智慧矿山建设需求的高效任务计算卸载方法亟待研究。

移动边缘计算(Mobile Edge Computing, MEC)技术利用边缘服务器的计算、缓存和能量资源将矿山物联网设备产生的计算任务卸载至边缘端协助处理以降低计算时延和能耗,甚至提高智慧矿山的安全性<sup>[3,4]</sup>。Porambage等人<sup>[5]</sup>展开一项全面调查,介绍了MEC技术在不同物联网系统中的应用,并阐明了该技术可改善物联网性能。在工业物联网等场景,可以利用边缘计算技术有效地解决设备终端计算能力不足的问题,在靠近用户的边缘上提供满意的服务质量和灵活的计算资源<sup>[6]</sup>。采用智能计算卸载框架将单个物联网设备上的计算任务卸载到多个边缘服务器上,并对任务分配以及CPU频率进行联合优化,可最大限度地减少执行延迟和能耗<sup>[7]</sup>。本工作针对矿山物联网中丰富特殊的计算任务需求<sup>[8]</sup>,如实时检测甲烷浓度的时延敏感型计算任务、采煤工作面高效运行大量感知参数处理的数据密集型任务以及火灾/爆炸警报的生命保障型任务,运用MEC技术实现矿山物联网的高效任务处理和性能优化。

传统物联网系统利用电网供电必然导致大量的碳排放,这不符合节能减排的需求。在智慧矿山生产过程中,挖掘使用能量收集(Energy Harvesting, EH)技术,将井下可再生能源(包括风能<sup>[9]</sup>、无线射频信号能量<sup>[10]</sup>和机械能<sup>[11]</sup>等)作为绿色供电来源,这是响应国家倡导的“碳中和、碳达峰”目标的有效举措,可解决矿用电池不易更换的难题<sup>[9]</sup>,也可极大地提高井下工作生产的安全性。此外,将EH技术运用到边缘计算系统中,有助于实现可持续的任务计算效益。Sun等人<sup>[12]</sup>提出一种基于差分进化算法的能量收集和任务计算优化方案,该方案不仅优化效率高和能耗低,而且能有效缓解微型设备的能源短缺问题,延长设备的使用寿命。因此,本文采用EH技术辅助矿山物联网设备供电,以提高任务计算效益并延长设备使用寿命。

矿山物联网中井下通信环境相比于地面更加复

杂多变,严重的非视距通信场景导致边缘覆盖困难;此外,井下还存在无线信号衰落严重、传输距离受限等突出问题<sup>[13]</sup>。因此,本文考虑在矿山物联网设备近端布置多个边缘服务器提供视距通信且增强覆盖范围,并设计一对多的矿井MEC网络架构,即由一个矿山物联网设备和 $M$ 个井下边缘服务器组成的矿井MEC网络。此外,由于井下导致信号衰落的因素复杂多变、移动设备遮挡信号视距传输、多径衰落效应十分严重,目前尚欠缺针对复杂环境的无线通信理论来精确刻画井下通信模型。因此,传统的优化理论无法解决矿井下的计算卸载策略选择问题。本文将矿山物联网设备的计算卸载过程建模为马尔可夫决策过程(Markov Decision Process, MDP)<sup>[14]</sup>,利用强化学习(Reinforcement Learning, RL)算法可在不了解系统模型的情况下,根据上一个时隙中每条边缘链路的无线信道带宽、预测的EH再生能量和当前物联网设备的电池电量选择计算卸载策略(即选择边缘服务器和计算任务的卸载率)。通过与环境不断交互,矿山物联网设备可自适应复杂环境在动态试错中探索最优的计算卸载策略并获得最优长期效益。

综上所述,本文提出了融合能量收集和强化学习的智能计算卸载机制,建立井下一对多MEC网络模型,对能耗、时延和任务失败等问题进行具体分析。本文首先提出了一种基于RL—Q-learning算法的智能计算卸载机制,矿山物联网设备作为智能体在不能精确获取井下系统模型的情况下动态探索最优计算卸载策略。但是,传统的强化学习技术面临维度灾难问题,当动作状态空间很大时,Q-learning算法的计算效率会大大降低。并且计算卸载率是连续的策略空间,Q-learning算法需先将该策略离散化才可进行学习。因此,本文还提出了一种基于深度确定性策略梯度算法(Deep Deterministic Policy Gradient, DDPG)的计算卸载机制,可实现在连续空间内选择计算卸载策略,减小离散化误差<sup>[15]</sup>。再者,利用卷积神经网络解决了维度灾难问题,能够更好地捕获井下环境特征,进一步优化计算卸载策略选择,提升系统性能。仿真结果表明所提机制可减少能量损耗、降低计算时延和任务计算失败率,提高计算卸载效益。

## 2 系统模型

图1给出了由一个矿山物联网设备和 $M$ 个井下

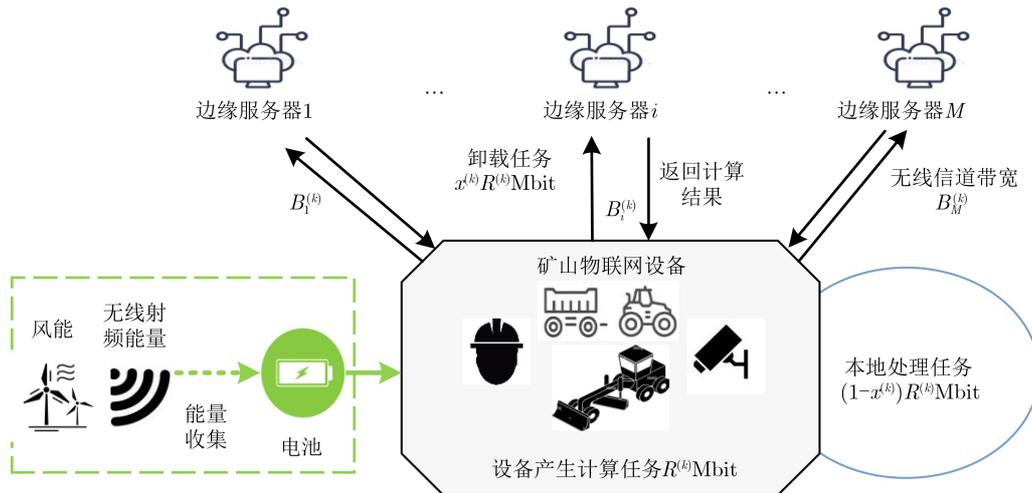


图1 矿山物联网中的MEC架构

边缘服务器组成的矿井MEC网络架构。矿山物联网设备为井下多种传感设备，如巷道巡检机器人、煤炭采掘机和采煤工人安全头盔等。这些设备将产生的计算任务卸载至边缘服务器以协助任务处理。矿山物联网设备配备了射频能量采集器、风力涡轮机等储能元件和EH组件，利用能量收集技术产生电能。物联网设备既可以在本地处理计算任务，也可以将部分或全部任务卸载至井下边缘服务器。另外，任务处理的最小时间单位为时隙，假设物联网设备在第 $k$ 个时隙生成 $R^{(k)}$ Mbit的计算任务数据量。

## 2.1 移动边缘计算卸载模型

$M$ 个边缘服务器部署在矿山物联网设备可视距通信的范围内。物联网设备至边缘服务器的无线信道链路带宽为 $B_i^{(k)}$ ，其中 $1 \leq i \leq M$ 。物联网设备在时隙 $k$ 内选择信道较好的井下边缘服务器 $i$ 并将 $R^{(k)}x^{(k)}$ Mbit计算任务量卸载至该服务器，其中 $x^{(k)}$  ( $0 \leq x^{(k)} \leq 1$ )为卸载率。具体来说，当 $x^{(k)} = 0$ 时，物联网设备本地处理全部计算任务；当 $x^{(k)} = 1$ 时，设备将所有计算任务卸载到边缘服务器；当 $0 \leq x^{(k)} \leq 1$ 时，该设备将 $R^{(k)}x^{(k)}$ Mbit数据卸载至井下边缘服务器 $i$ ，剩余的 $(1-x^{(k)})R^{(k)}$ Mbit数据在本地处理。于是，矿山物联网设备在时隙 $k$ 内选择卸载策略 $\mathbf{a}^{(k)} = [i^{(k)}, x^{(k)}] \in \mathbf{A}$ ，其中 $\mathbf{A}$ 表示所有可能的计算卸载策略集合。

矿山物联网设备的CPU用作本地计算，处理1 Mbit数据需要 $N$ 个CPU周期，则本地计算需要 $(1-x^{(k)})R^{(k)}N$ 个CPU周期，CPU周期频率为 $f$ ，用 $S_1 = N/f$ 表示本地计算1 Mbit数据所需的时间。 $e_0$ 表示物联网设备本地处理1 bit数据消耗的能量。 $T_0^{(k)}$ 和 $E_0^{(k)}$ 分别表示本地计算的时延和能耗，表达式为

$$T_0^{(k)} = (1-x^{(k)})R^{(k)}S_1 \quad (1)$$

$$E_0^{(k)} = (1-x^{(k)})R^{(k)}e_0 \quad (2)$$

由于动态复杂矿井环境下的无线信道状态信息难以精确获取，本文将信道带宽 $B$ 建模成马尔可夫链<sup>[16]</sup>，并用来简化和表征信号传输速率。矿山物联网设备通过带宽为 $B_i^{(k)}$ 的无线信道，在时隙 $k$ 内卸载 $R^{(k)}x^{(k)}$ bit计算任务至边缘服务器 $i$ 。 $S_2$ 表示井下边缘服务器计算1 Mbit数据所需的时间，且有 $S_1 \gg S_2$ ，表明边缘服务器具有更强的计算处理能力。 $P$ 为物联网设备的发射功率。 $T_1^{(k)}$ 和 $E_1^{(k)}$ 分别表示任务卸载产生的时延和能耗，表达式为

$$T_1^{(k)} = \frac{R^{(k)}x^{(k)}}{B_i^{(k)}} + R^{(k)}x^{(k)}S_2 \quad (3)$$

$$E_1^{(k)} = \frac{R^{(k)}x^{(k)}P}{B_i^{(k)}} \quad (4)$$

## 2.2 能量收集与电池模型

矿山物联网设备使用EH技术，捕获井下可再生能源(包括风能、无线环境中的射频信号、机械能等)作为绿色供电能源。EH技术配合电池充电，来补给物联网设备计算任务的电量需求。

令时隙 $k$ 内收集的能量为 $g^{(k)}$ ，且本地计算和计算卸载的总能量损耗为 $E^{(k)} = (E_0^{(k)} + E_1^{(k)})$ 。电池电量表示为 $b^{(k)}$ ， $b^{(k+1)}$ 则表示获取EH产能后下一时刻的电池电量，其表达式为

$$b^{(k+1)} = \max \{ b^{(k)} - E^{(k)} + g^{(k)}, 0 \} \quad (5)$$

本文以无线信道中的射频信号能量为例对EH过程建模<sup>[17]</sup>，在时隙 $k$ 内EH产能为

$$g^{(k)} = \nu\eta^{(k)}(d^{(k)})^{-\tau}G \quad (6)$$

其中,  $G$ 表示矿山物联网设备上的天线部分和射频能量发射器天线的联合增益;  $\nu \in (0, 1)$ 表示井下再生能量的转换效率, 由于井下无线信道条件较差, 能源转换效率相对地面场景较低;  $\eta^{(k)}$ 表示矿山物联网设备上EH模块的传输功率;  $d^{(k)}$ 代表在时隙  $k$ 内EH模块和对应射频能量发射器之间的距离;  $\tau \geq 2$ 表示再生能量传输过程中的路径损耗因子。此外, 在动态复杂矿井环境下, 根据文献[18]中模型评估实际产能 $g^{(k)}$ 并用马尔可夫链模型表示其产能随时间的变化。

### 2.3 问题描述

针对矿山物联网中大量且特殊的计算任务需求, 并充分考虑矿山特殊通信环境对物联网智能计算卸载的影响, 本文构建EH辅助的一对多矿井MEC网络框架, 实现矿山物联网的智能计算卸载。此外, 为了充分满足矿山物联网应用中特殊任务处理的服务质量需求, 本文设计了由任务共享收益、能量损耗、计算时延和任务失败损失共同组成的目标效益函数。再者, 考虑到动态复杂矿井环境中的无线通信模型无法精确获取, 传统的优化理论无法解决矿井下的计算卸载策略选择问题, 研究基于强化学习的动态计算卸载机制, 最大化系统的长期折扣累计效益, 提升矿山物联网智能计算卸载性能。

## 3 基于Q-learning的矿山物联网计算卸载机制

矿山物联网设备的下一时刻系统状态仅依赖于当前时刻的状态和计算卸载策略, 而与过去的历史情况无关; 因此, 动态计算卸载过程可以看作一个MDP过程。于是, 本文首先采用RL技术Q-learning, 提出了基于Q-learning的矿山物联网智能计算卸载机制, 记为RLOM。矿山物联网设备作为学习智能体可在不精确获取MEC模型、能耗模型和计算时延模型的情况下根据系统状态动态探索计算卸载策略 $\mathbf{a}^{(k)} = [i^{(k)}, x^{(k)}]$ 。其中, 系统状态 $\mathbf{s}^{(k)}$ , 即 $\mathbf{s}^{(k)} = [B_1^{(k-1)}, \dots, B_M^{(k-1)}, g^{(k)}, b^{(k)}]$ , 由物联网设备到边缘服务器的无线信道带宽、EH产能和电池电量组成。

矿山物联网设备通过带宽为 $B_i^{(k)}$ 的无线信道将 $R^{(k)}x^{(k)}$ bit计算任务卸载至边缘服务器 $i$ , 产生的传输和任务计算总时延 $T_1^{(k)}$ 如式(3)所示。由于计算任务被卸载至边缘服务器处理, 物联网设备不仅可得到任务共享收益, 而且可降低任务传输至云端的损耗以及本地无法处理的任务失败率。此外, 设备本地处理部分任务时产生的任务计算时延 $T_0^{(k)}$ 如式(1)所示。物联网设备评估效益时, 对任务共享收益、时延、能耗与任务失败率进行综合衡量。其中, 由

于本地和卸载处理计算任务同时进行, 那么执行部分卸载方案时的总时延为本地处理时延与卸载时延的较大值, 即MEC卸载模型总的计算时延定义为 $T^{(k)} = \max(T_1^{(k)}, T_0^{(k)})$ ; 根据式(2)和式(4), 总的能量损耗为 $E^{(k)} = (E_0^{(k)} + E_1^{(k)})$ 。指示函数 $I(b^{(k+1)} = 0)$ 表示任务能否成功处理: 若 $I(b^{(k+1)} = 0) = 0$ 表示任务能否成功处理: 若 $I(b^{(k+1)} = 0) = 0$ 中函数体不成立时函数值为0, 表示电量充足任务可成功处理; 反之, 则表示任务因电量不足而处理失败。因此, 效益函数可用如下分段函数表示(其中 $\theta_x$ 作为各项指标的权重值):

$$U^{(k)} = \begin{cases} \theta_0 x^{(k)} R^{(k)} - \theta_1 E^{(k)} - \theta_2 T^{(k)}, & I(b^{(k+1)} = 0) = 0 \\ -\theta_1 E^{(k)} - \theta_2 T^{(k)} - \theta_3, & I(b^{(k+1)} = 0) = 1 \end{cases} \quad (7)$$

由式(7)可见, 当 $I(b^{(k+1)} = 0) = 0$ 时, 矿山物联网设备将计算任务卸载至边缘服务器, 可降低资源有限的矿山物联网设备的任务处理负担, 得到任务共享收益 $\theta_0 x^{(k)} R^{(k)}$ 。当 $I(b^{(k+1)} = 0) = 1$ 时, 计算任务因电量不足而处理失败, 无任务共享收益。但是, 在这个过程中产生了相应的能量损耗 $E^{(k)}$ 和计算时延 $T^{(k)}$ , 并且效益函数也由于计算任务处理失败而降低。

矿山物联网设备在状态 $\mathbf{s}^{(k)}$ 下基于 $\epsilon$ -greedy选择计算卸载策略 $\mathbf{a}^{(k)}$ , 将计算任务卸载至边缘服务器 $i$ , 得到效益 $U^{(k)}$ , 并进入下一个状态 $\mathbf{s}^{(k+1)}$ 。Q值 $Q(\mathbf{s}^{(k)}, \mathbf{a}^{(k)})$ 表示井下MEC系统的每个计算卸载动作-状态对的长期期望折扣效益值, Q值在每个时隙内通过贝尔曼方程完成迭代更新, 具体为

$$Q(\mathbf{s}^{(k)}, \mathbf{a}^{(k)}) \leftarrow (1 - \alpha)Q(\mathbf{s}^{(k)}, \mathbf{a}^{(k)}) + \alpha(U^{(k)} + \gamma \max_{\mathbf{a}' \in \mathbf{A}} Q(\mathbf{a}', \mathbf{s}^{(k+1)})) \quad (8)$$

其中, 学习率 $\alpha \in (0, 1]$ 表示当前Q值所占权重; 折扣因子 $\gamma \in [0, 1]$ 表示未来效益的衰减。 $\gamma$ 越小表示物联网设备越重视当前的效益, 而忽略未来长期效益。基于Q-learning的计算卸载机制的核心步骤如算法1所示。

## 4 基于DDPG的矿山物联网计算卸载机制

在智慧矿山建设的大背景下, 矿山物联网环境更加复杂, 使用传统的强化学习如Q-learning算法存在维度灾难。再者, 实际的计算卸载策略空间为高维且连续的向量。第3节提出的RLOM机制需要将计算卸载策略离散化, 这将不可避免地导致离散化误差, 使得最终的学习结果并非真实最优。为了能够解决复杂矿山物联网环境下的维度灾难问题并更好地处理连续计算卸载策略空间的策略选择问题, 满足井下计算密集型和时延敏感型任务的快速

计算需求, 本文提出了一种基于DDPG的计算卸载机制, 记为DDRLOM。

DDRLOM机制结合了深度强化学习DQN和Actor-Critic算法来处理矿山物联网设备的高维且连续的计算卸载策略选择问题。该机制采用深度神经网络(Deep Neural Network, DNN)压缩系统的状态空间, 解决维度灾难问题<sup>[19,20]</sup>, 并且可更好地捕获井下环境特征, 进一步优化计算卸载策略选择, 提升矿山物联网的任务计算卸载性能。此外, DNN作为非线性拟合器, 拟合高维且连续的计算卸载策略与Q值以加快学习速度; 然后, 利用一个参数化的动作价值函数 $\mathbf{a}^{(k)} = \mu(\mathbf{s}^{(k)}; \xi)$ 进行计算卸载策略选择, 提高算法对连续动作空间的处理能力, 该动作函数实现给定输入系统状态 $\mathbf{s}^{(k)}$ , 输出一个最优的计算卸载策略 $\mathbf{a}^{(k)}$ 。

如图2所示, DDRLOM机制采用Actor-Critic架构体系, 包含4个神经网络, 即Actor网络和Critic网络; Target Actor网络和Target Critic网络。其中, Actor网络对应网络参数为 $\xi_2$ 将当前系统状态作为输入, 输出计算卸载策略, 即选择协助计算的边缘服务器和矿山物联网设备的卸载率。为了使模型更有效地更新网络参数和实现优化的收敛效果, 增加算法的探索能力, 在计算卸载策略选择时增加了OU(Ornstein-Hollenbeck)过程<sup>[15]</sup>即噪声函数 $\mathcal{N}$ , 加入扰动噪声后的计算策略生成如式(9)

$$\mathbf{a}^{(k)} = \mu(\mathbf{s}^{(k)} | \xi_2^{(k)}) + \mathcal{N} \quad (9)$$

Critic网络对应网络参数为 $\xi_1$ , 作为价值网络给出逼近状态动作对的值函数 $Q(\mathbf{s}^{(k)}, \mathbf{a}^{(k)})$ , 从而提供优化的梯度信息, 帮助Actor网络采取更好的策略。

算法1 基于Q-learning的计算卸载机制

```

初始化系统参数 $\alpha, \gamma, B_1^0, \dots, B_M^0, g^{(1)}, b^0$ ; 设置学习迭代次数
1: For  $k = 1, 2, 3, \dots$  do
2:   预测能量收集的产能 $g^{(k)}$ , 观察得到 $b^{(k)}$ 和 $B_1^{(k-1)}, \dots, B_M^{(k-1)}$ 
3:   得到状态 $\mathbf{s}^{(k)} = [B_1^{(k-1)}, \dots, B_M^{(k-1)}, g^{(k)}, b^{(k)}]$ 
4:   基于 $\epsilon$ -greedy选择计算卸载策略 $\mathbf{a}^{(k)} = [i^{(k)}, x^{(k)}]$ , 并卸载任务量 $R^{(k)}x^{(k)}$ 至边缘服务器 $i$ 
5:   与环境交互, 得到电量变化 $b^{(k+1)} = \max\{b^{(k)} - E^{(k)} + g^{(k)}, 0\}$ , 计算评估 $E^{(k)}, T^{(k)}$ 和 $I(b^{(k+1)} = 0)$ 
6:   通过式(7)获得效益 $U^{(k)}$ , 通过式(8)更新Q值 $Q(\mathbf{s}^{(k)}, \mathbf{a}^{(k)})$ 
7: End
    
```

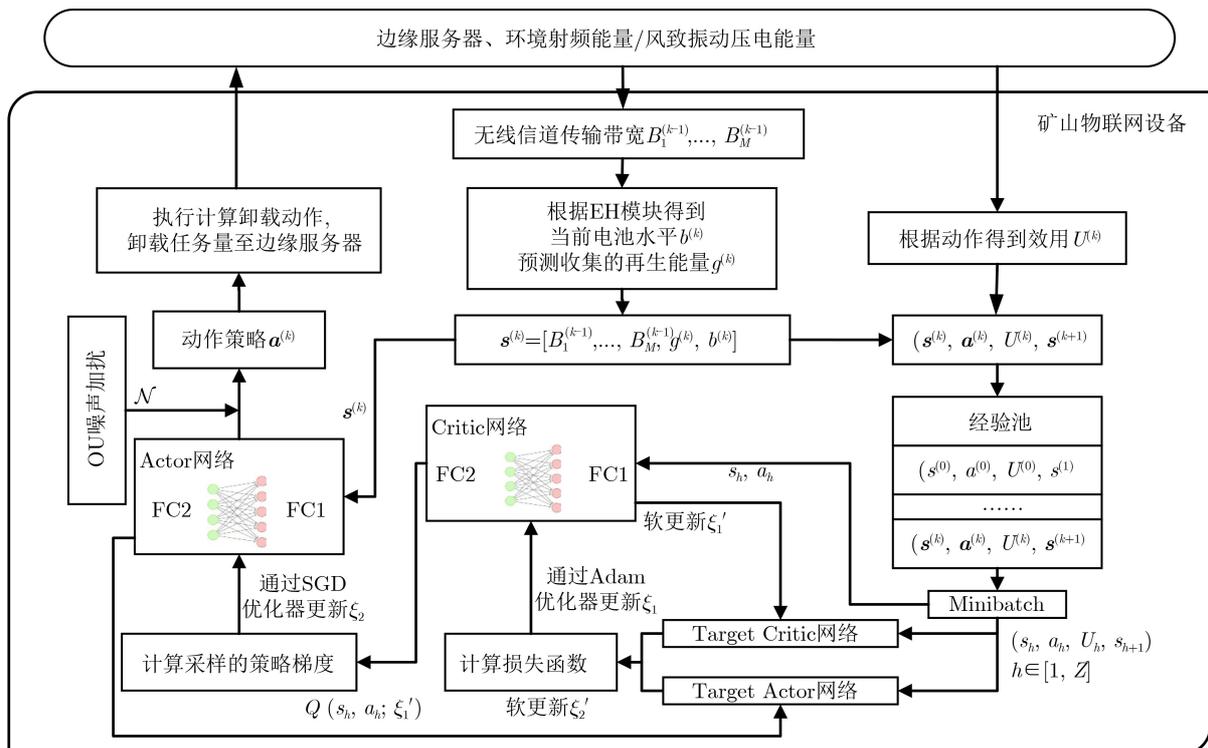


图2 基于DDPG的计算卸载机制(DDRLOM)框架

DDRLM机制的策略选择过程中, Actor网络基于当前输入的状态 $\mathbf{s}^{(k)}$ 输出确定性动作 $\mu(\mathbf{s}^{(k)}|\xi_2^{(k)})$ , 矿山物联网设备基于 $\mu(\mathbf{s}^{(k)}|\xi_2^{(k)})$ 和OU过程采样得到的噪声 $\mathcal{N}$ 选择计算卸载策略 $\mathbf{a}^{(k)}$ , 即选择对应的边缘服务器 $i^{(k)}$ 和卸载率 $x^{(k)}$ 。矿山物联网设备执行当前状态 $\mathbf{s}^{(k)}$ 下的卸载策略 $\mathbf{a}^{(k)}$ 后, 通过式(7)评估系统效益值 $U^{(k)}$ , 并获得下一时刻的系统状态 $\mathbf{s}^{(k+1)}$ , 上述元素组成计算卸载经验 $(\mathbf{s}^{(k)}, \mathbf{a}^{(k)}, U^{(k)}, \mathbf{s}^{(k+1)})$ 存储到经验池。

基于经验回放技术, 矿山物联网设备从经验池中随机采样 $Z$ 组 $(\mathbf{s}_h, \mathbf{a}_h, U_h, \mathbf{s}_{h+1}), h \in [1, Z]$ 计算卸载经验。Target Actor网络根据输入的 $\mathbf{s}_{h+1}$ 估计得出下一个时隙的卸载策略 $\mathbf{a}_{h+1}$ 。接着Target Critic网络由输入的 $\mathbf{s}_{h+1}$ 和 $\mathbf{a}_{h+1}$ 输出下一个系统状态下的 $Q(\mathbf{s}_{h+1}, \mathbf{a}_{h+1}; \xi_1')$ , 同时Critic网络输出当前系统状态下的 $Q(\mathbf{s}_h, \mathbf{a}_h; \xi_1)$ 。利用Adam优化器作为梯度下降算法, 使 $Q$ 估计值 $y_h$ 与 $Q$ 真实值 $Q(\mathbf{s}_h, \mathbf{a}_h|\xi_1)$ 之间的均方误差达到最小, 从而更新Critic网络参数 $\xi_1$ , 具体公式为

$$y_h = E_{\xi_1'} [U + \gamma \max_{\mathbf{a} \in \mathcal{A}} Q(\mathbf{s}_{h+1}, \mathbf{a}_{h+1}; \xi_1'), \mathbf{a}_h] \quad (10)$$

$$\xi_1 \leftarrow \xi_1 - \alpha_1 \frac{\partial \left( \frac{1}{Z} \sum_{h=1}^Z (y_h - Q(\mathbf{s}_h, \mathbf{a}_h|\xi_1))^2 \right)}{\partial \xi_1} \quad (11)$$

此外, 通过随机梯度下降(Stochastic Gradient Descent, SGD)<sup>[15]</sup>算法来更新Actor网络参数 $\xi_2$ , 具体为

$$\xi_2 \leftarrow \xi_2 - \frac{\alpha_2}{Z} \sum_{h=1}^Z \nabla_{\mathbf{a}} Q(\mathbf{s}_h, \mathbf{a}_h|\xi_1) \nabla_{\xi_2} \mu(\mathbf{s}_h|\xi_2) \quad (12)$$

另外, 对应的两个Target网络可以提高算法的稳定性<sup>[21]</sup>, 其参数 $\xi_1', \xi_2'$ 则以学习率 $\kappa (\kappa \ll 1)$ 进行软更新来放慢对两个Online网络的追踪速度, 具体软更新过程如下:

$$\xi_1' = \kappa \xi_1 + (1 - \kappa) \xi_1' \text{ and } \xi_2' = \kappa \xi_2 + (1 - \kappa) \xi_2' \quad (13)$$

基于DDPG的计算卸载机制的核心步骤如算法2所示。

## 5 理论分析

本节分析3种典型场景下的静态计算卸载策略, 从理论上推导并证明了不同条件下的最优卸载策略选择及其效益函数的理论界, 以指导基于RL的计算卸载机制的学习参数设置和环境配置。下述分析中假设计算卸载策略选择和状态更新均在一个时隙内完成, 故省略时隙 $k$ 。

### 5.1 计算任务部分卸载到边缘服务器处理

**定理1** 当矿山物联网设备至无线信道状态最好的边缘服务器 $i^*$ 的信道状态一般, 即 $B_{\min}' \leq B_{i^*} \leq B_{\max}'$ 时, 其中:

$$B_{\max}' = \frac{\theta_1 P + \theta_2}{\theta_0 + \theta_1 e_0 - \theta_2 S_2} \quad (14)$$

$$B_{\min}' = \max \left\{ \frac{\theta_1 P}{\theta_0 + \theta_1 e_0 + \theta_2 S_1}, \frac{R(e_0 + S_1 P) - (b + g)}{S_1(b + g)} \right\} \quad (15)$$

矿山物联网设备将生成的计算任务量部分卸载至无线信道带宽最大的边缘服务器 $i^*$ 处理,  $i^* = \arg \max_{1 \leq i \leq M} B_i$ , 其余在本地处理。此时, 最优计算卸载策略为 $a^* = [i^*, S_1 B_{i^*} / (S_1 B_{i^*} + S_2 B_{i^*} + 1)] g^{(k)}$ , 矿山物联网设备的效益、能量损耗和计算时延分别为

$$U = \theta_0 R \frac{S_1 B_{i^*}}{(S_1 B_{i^*} + S_2 B_{i^*} + 1)} - \theta_1 R \frac{e_0(1 + S_2 B_{i^*}) + S_1 P}{(S_1 B_{i^*} + S_2 B_{i^*} + 1)} - \theta_2 R \frac{S_1 + S_1 S_2 B_{i^*}}{(S_1 B_{i^*} + S_2 B_{i^*} + 1)} \quad (16)$$

$$E = R \frac{e_0(1 + S_2 B_{i^*}) + S_1 P}{(S_1 B_{i^*} + S_2 B_{i^*} + 1)} \quad (17)$$

### 算法2 基于DDPG的计算卸载机制

初始化学习参数 $\alpha, \lambda, \xi_1, \xi_2, Z, \kappa$ ; OU噪声 $\mathcal{N}$ 的参数; 设置学习迭代次数

1: 观测得到 $g^{(k)}, b^{(k)}, B_1^{(k-1)}, \dots, B_M^{(k-1)}$ , 组成状态 $\mathbf{s}^{(k)} = [B_1^{(k-1)}, \dots, B_M^{(k-1)}, g^{(k)}, b^{(k)}]$

2: For  $k = 1, 2, 3, \dots$  do

3: 基于状态 $\mathbf{s}^{(k)}$ 输出确定性动作, 加入噪声 $\mathcal{N}$ 后, 产生卸载策略 $\mathbf{a}^{(k)} = \mu(\mathbf{s}^{(k)}|\xi_2^{(k)}) + \mathcal{N}$

4: 卸载任务量 $R^{(k)} x^{(k)}$ 至边缘服务器 $i$ , 评估得到 $\mathbf{s}^{(k+1)}$ 和 $U^{(k)}$

5: 将 $(\mathbf{s}^{(k)}, \mathbf{a}^{(k)}, U^{(k)}, \mathbf{s}^{(k+1)})$ 存入经验池

6: 从经验池中抽取 $Z$ 组学习经验 $(\mathbf{s}_h, \mathbf{a}_h, U_h, \mathbf{s}_{h+1}), h \in [1, Z]$ , 更新网络参数

7: 由式(10)和式(11)更新Critic网络参数 $\xi_1$ , 由式(12)更新Actor网络参数 $\xi_2$

8: 由式(13)对Target网络参数 $\xi_1', \xi_2'$ 进行软更新

9:  $\mathbf{s}^{(k)} = \mathbf{s}^{(k+1)}$

10: End for

$$T = R \frac{S_1 + S_1 S_2 B_{i^*}}{(S_1 B_{i^*} + S_2 B_{i^*} + 1)} \quad (18)$$

**证明** 由式(7)可知, 计算卸载策略的效益值因任务成功处理而获得计算任务共享收益, 即 $\theta_0 x^{(k)} R^{(k)} - \theta_1 E^{(k)} - \theta_2 T^{(k)} > -\theta_1 E^{(k)} - \theta_2 T^{(k)} - \theta_3$  成立, 那么, 若 $I(b^{(k+1)} = 0) = 0$ , 即 $b^{(k)} - E^{(k)} + g^{(k)} \geq 0$ 。代入运算得到 $B$ 的范围为

$$B \geq \frac{R(e_0 + S_1 P) - (b + g)}{S_1(b + g)} \quad (19)$$

在该条件下得到的效益函数 $U$ 较大, 表达式为

$$U(x) = xR\theta_0 - R\theta_1 \left[ (1-x)e_0 + \frac{xP}{B} \right] - R\theta_2 \max \left\{ (1-x)S_1, \frac{x}{B} + xS_2 \right\} \quad (20)$$

其单调性分析如式(21)

$$\frac{\partial U}{\partial x} = \begin{cases} R\theta_0 - R\theta_1 \left( -e_0 + \frac{P}{B} \right) + R\theta_2 S_1, & x < \frac{BS_1}{1 + BS_1 + BS_2} \\ R\theta_0 - R\theta_1 \left( -e_0 + \frac{P}{B} \right) - R\theta_2 \left( \frac{1}{B} + S_2 \right), & \text{other} \end{cases} \quad (21)$$

当 $x < BS_1/(1 + BS_1 + BS_2)$ 时, 若 $\partial U/\partial x$ 最小值大于等于0, 那么 $\partial U/\partial x \geq 0$ 在该范围内恒成立, 此时 $U$ 关于 $x$ 单调递增。当 $x \geq BS_1/(1 + BS_1 + BS_2)$ 时, 若 $\partial U/\partial x$ 最大值小于0, 那么 $\partial U/\partial x < 0$ 在该范围内恒成立, 此时 $U$ 关于 $x$ 单调递减。由于 $\partial U/\partial x$ 关于 $B$ 单调递增, 根据式(21)有:

由 $\partial U/\partial x \geq 0$ ,  $x < BS_1/(1 + BS_1 + BS_2)$ , 得 $B_{\min} = \theta_1 P / (\theta_0 + \theta_1 e_0 + \theta_2 S_1)$ ;

由 $\partial U/\partial x < 0$ ,  $x \geq BS_1/(1 + BS_1 + BS_2)$ , 得 $B_{\max} = (\theta_1 P + \theta_2) / (\theta_0 + \theta_1 e_0 - \theta_2 S_2)$ 。因此可得下式:

$$\frac{\theta_1 P}{\theta_0 + \theta_1 e_0 + \theta_2 S_1} \leq B < \frac{\theta_1 P + \theta_2}{\theta_0 + \theta_1 e_0 - \theta_2 S_2} \quad (22)$$

综合式(20)、式(21)和式(22)得到 $U$ 在 $x^*$ 处取最大值, 且

$$x^* = \frac{B_{i^*} S_1}{1 + B_{i^*} S_1 + B_{i^*} S_2}, x^* \in (0, 1) \quad (23)$$

由式(23)得 $x^*$ 的取值关于 $B_{i^*}$ 单调递增, 当带宽值最大时, 取得的 $x$ 值对应最大的 $U$ 值, 即 $i^* = \arg \max_{1 \leq i \leq M} B_i$ 得证。接着对带宽的限定条件取交集可得式(15)。

式(23)表明在 $B_{\min}' \leq B \leq B_{\max}$ 范围内, 矿山物联网设备通过强化学习算法不断试错, 当 $i^* = \arg \max_{1 \leq i \leq M} B_i$ 时矿山物联网的效益值 $U$ 在采取最佳卸载策略 $a^* = [i^*, S_1 B_{i^*} / (S_1 B_{i^*} + S_2 B_{i^*} + 1)]$ 时达

到最大值, 同时获得系统最优性能如式(16)—式(18)所示。证毕

当物联网设备到边缘服务器的无线信道状态一般时, 物联网设备本地处理部分计算任务, 并将剩余的计算任务卸载给边缘服务器, 以权衡能耗和时延, 从而使其效益最大化如式(16)所示。此时, 能量损耗和计算时延与生成的任务大小 $R$ 正相关, 如式(17)和式(18)所示。

## 5.2 计算任务全部卸载至边缘服务器处理

**定理2** 当矿山物联网设备至无线信道状态最好的边缘服务器 $i^*$ 的信道状态良好, 即 $B_{i^*} \geq B'_{\max}$ , 其中

$$B'_{\max} = \max \left\{ \frac{\theta_1 P + \theta_2}{\theta_0 + \theta_1 e_0 - \theta_2 S_2}, \frac{R(e_0 + S_1 P) - (b + g)}{S_1(b + g)} \right\} \quad (24)$$

时, 矿山物联网设备将生成的任务全部卸载至无线信道状态最好的边缘服务器 $i^*$ 处理,  $i^* = \arg \max_{1 \leq i \leq M} B_i$ , 矿山物联网设备的最佳计算卸载策略为 $a^* = [i^*, 1]$ 。此时, 物联网设备的效益、能量损耗和计算时延分别为

$$U = \theta_0 R - R \frac{\theta_1 P}{B_{i^*}} - R \left( \frac{\theta_2}{B_{i^*}} + S_2 \right) \quad (25)$$

$$E = R \frac{P}{B_{i^*}} \quad (26)$$

$$T = R \left( \frac{1}{B_{i^*}} + S_2 \right) \quad (27)$$

**证明** 当 $B \geq B'_{\max}$ 时, 证明过程类似定理1, 在式(24)成立的条件下, 由式(22)可得 $\partial U/\partial x > 0$ 恒成立, 即 $B \geq B_{\max}'$ 时 $U$ 关于 $x$ 单调递增, 所以

$$U(1) = \theta_0 R - R \frac{\theta_1 P}{B_{i^*}} - R \left( \frac{\theta_2}{B_{i^*}} + S_2 \right) \geq U(x) \quad (28)$$

式(28)表明矿山物联网设备的效益值随着 $B \geq B'_{\max}$ 递增, 且于 $i^* = \arg \max_{1 \leq i \leq M} B_i$ 达到最大值。由式(28)可知,  $a^* = [i^*, 1]$ 是该场景下的最佳计算卸载策略, 且矿山物联网设备获得最大效益值 $U(1)$ , 同时获得系统最优性能如式(25)—式(27)所示。证毕

若矿山物联网设备与至少一个边缘服务器之间的无线信道状态良好, 那么物联网设备将所有计算任务卸载至边缘服务器处理。此时, 最佳计算卸载策略为 $a^* = [i^*, 1]$ , 可取得较高的计算任务共享收益; 此外, 无线信道状态良好时, 计算卸载过程中的时延和能耗都较小, 如式(26)和式(27)所示。

## 5.3 计算任务全部本地处理

**定理3** 当矿山物联网设备至无线信道状态最

好的边缘服务器  $i^*$  的信道状态较差, 即  $B_{i^*} \leq B''_{\min}$  时,

$$B_{\min}'' = \min \left\{ \frac{\theta_1 P}{\theta_0 + \theta_1 e_0 + \theta_2 S_1}, \frac{R(e_0 + S_1 P) - (b + g)}{S_1(b + g)} \right\} \quad (29)$$

矿山物联网设备将生成的计算任务全部本地处理, 即最佳计算卸载策略为  $a^* = [i^*, 0]$ 。此时, 物联网设备的效益、能量损耗和计算时延分别为

$$U = -R\theta_1 e_0 - R\theta_2 S_1 \quad (30)$$

$$E = R e_0 \quad (31)$$

$$T = R S_1 \quad (32)$$

**证明** 由于  $B_{i^*} \leq B_{\min}''$ , 最佳的无线信道带宽小于有效带宽最小范围值。在该条件下, 由式(7)和式(22)可得:  $\partial U / \partial x < 0$  恒成立, 即  $B_{i^*} \leq B''_{\min}$  时  $U$  关于  $x$  单调递减, 所以

$$U(0) = -R\theta_1 e_0 - R\theta_2 S_1 \geq U(x) \quad (33)$$

式(33)表明矿山物联网设备的效益值在卸载率为0时达到最大值。因此,  $a^* = [i^*, 0]$  是矿山物联网设备在该场景下的最佳卸载策略, 同时获得系统最优性能如式(30)—式(32)所示。证毕

若无线信道状态较差, 即式(3)和式(4)中无线信道带宽  $B$  较小 ( $B \leq B''_{\min}$ ) 时, 传输过程中的时延和能耗较大。如式(33)所示, 物联网设备将在本地处理所有的计算任务, 以避免任务卸载导致过大的任务处理时延和能量消耗。

## 6 仿真结果与分析

本节通过仿真验证所提方法的性能优势。仿真中考虑矿井MEC系统包含3个边缘服务器和1个矿山物联网设备。其中矿山物联网设备的EH产能由无线射频能量转换而来, 由于井下无线通信环境复杂且信号衰落与地面相比更加严重, 井下EH的平均能量转换效率设置为0.3<sup>[22]</sup>。物联网设备上的电池最大电量为1000 mAh。边缘服务器的计算能力更强<sup>[23]</sup>, 量化体现如下, 边缘服务器处理1 Mbit数据所需要的时间为0.05 ms, 而本地设备处理1 Mbit数据所需要的时间为0.4 ms。根据防爆标准GB/T 3836.1—2021中无线射频信号的规定<sup>[24]</sup>, 设置物联网设备的发射功率为5.5W。

在学习算法中, 设置学习率  $\alpha = 0.001$ 、折扣因子  $\gamma = 0.9$ 。表1列出了DDRLOM机制的超参数设置。我们将能量损耗、计算时延和任务失败率作为边缘计算卸载的各项性能指标, 其中任务失败率定义为失败的计算任务占总处理任务的比例。根据

式(7)设定权重系数  $\theta_0 = 15, \theta_1 = 4, \theta_2 = 5, \theta_3 = 7$ 。仿真评估选取了DRLO<sup>[18]</sup>和全本地处理方案Full Local作为对比机制。

### 6.1 不同计算卸载机制的收敛性能评估

图3给出了给定训练回合数下矿山物联网设备采用不同计算卸载机制的收敛性能结果, 其中最大回合数为200, 每回合最大时隙数为1000。将不同的智能计算卸载机制的收敛性能进行对比, 可以看出DDRLOM机制的性能优于RLOM机制。首先, DDRLOM机制相比于RLOM机制具有更快的收敛速度。例如, 在相同的学习回合数下, DDRLOM机制基本已经实现收敛, 而RLOM机制的效益值仍然呈现逐步上升趋势。这是因为DDRLOM机制通过DNN压缩状态空间来加速学习过程, 提高计算卸载性能。此外, DDRLOM机制收敛后的计算卸载效益也明显优于DRLO机制。例如, 在第40回合数处, DDRLOM机制的平均效益值比DRLO机制约高24.6%。因为DDRLOM结合了DQN和Actor-Critic架构来处理矿山物联网设备的连续高维计算卸载策略, 减小了离散误差, 优化了计算卸载策略选择。

### 6.2 能量收集产能对计算卸载性能的影响

图4给出了矿井MEC系统中EH产能对计算任务失败率的影响。可以明显看出, 随着EH平均产能的增加, 矿山物联网设备的任务失败率显著降低。例如, 当平均收集产能从2.00 J增加到3.00 J时, RLOM机制的任务失败率降低了52.5%。当平

表1 DDRLOM机制的超参数设置

参数	数值
Actor/Critic学习率	0.0001/0.002
折扣因子	0.9
OU噪声均值/标准差/均值回归速度	5/0.015/40
batch-size	64
激活函数	Leaky ReLU
卷积网络隐藏层数/隐藏的单元数	2层/32,16

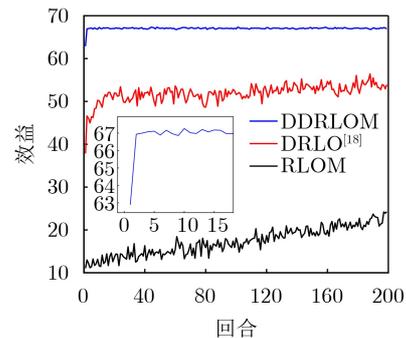


图3 不同智能计算卸载机制的收敛性能

均收集产能为2.5 J时，DDRLOM机制比DRLO机制的任务失败率约少99.9%。此外，在平均收集产能大于2.75 J时DDRLOM机制的任务失败率几乎达到0。

### 6.3 边缘服务器计算能力对计算卸载性能的影响

图5给出了矿井MEC系统中边缘服务器计算能力对计算时延的影响。边缘服务器的处理能力用处理1 Mbit所需要的时间 $S_2$ 来衡量， $S_2 \in [0.001, 0.025, 0.050, 0.075, 0.100]$ 。如图5所示，由于Full Local方案将计算任务全部本地处理不受 $T_1$ 的变化而变化，其他机制下矿山物联网设备的计算时延随着 $T_1$ 的降低而减小，且远远小于Full Local方案的时延。例如，当 $T_1$ 从0.1降低到0.001时，DDRLOM机制的计算时延降低57.7%。

### 6.4 任务量大小对计算卸载性能的影响

本工作对动态计算卸载中前3000个时隙的平均性能进行了评估，分析了计算卸载性能和总计算任务量的关系。如图6所示，可以明显看出，随着矿山MEC系统中计算任务量的增加，矿山物联网设备的能量损耗、计算时延和任务失败率也在增加。例如，当总计算任务量从4.2 Mbit增加到5.8 Mbit时，采用Full Local的计算卸载方案平均性能指标增加明显。当计算任务量为5 Mbit时，DDRLOM

机制相对于DRLO机制，能量损耗约降低14.2%，计算时延约降低7.7%，任务失败率约降低30.9%。可见，针对计算任务需求更强的矿山物联网环境，采用DDRLOM机制可以在保证能耗、时延和任务失败率都增加不多的前提下取得更好的计算卸载性能。

## 7 结束语

该文提出一种融合能量收集和强化学习技术的矿山物联网智能计算卸载方法，以应对井下通信环境复杂和设备资源受限的挑战，并提高计算密集和时延敏感任务的计算效率。在不可精确获取井下系统模型的情况下，针对计算能力较弱的矿山物联网设备提出基于复杂度较低的Q-learning算法的RLOM机制，实现动态探索最优计算卸载策略；此外，对于复杂的矿山物联网场景下高维且连续的策略空间问题，该文设计了基于DDPG的智能计算卸载机制DDRLOM，减小了计算卸载策略选择过程中的离散化误差，进一步优化计算卸载策略选择。针对3种典型场景的静态计算卸载策略进行了理论分析，可助于指导基于RL的智能计算卸载机制的学习参数设置和环境配置。仿真结果表明，所提方法可提高任务计算卸载性能，提升矿山物联网的安全性和生

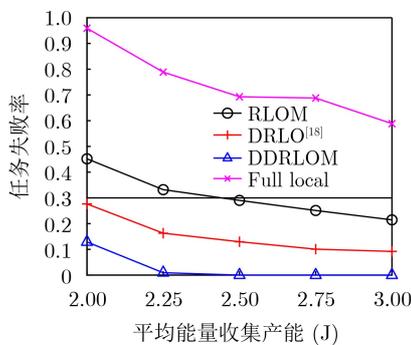


图 4 能量收集产能对计算任务失败率的影响

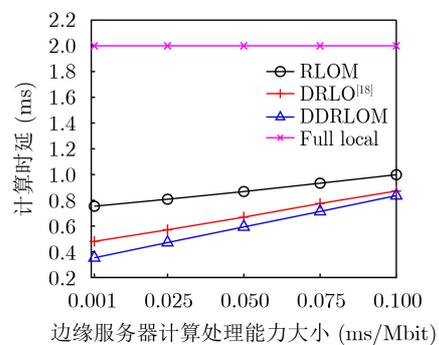


图 5 边缘服务器计算能力对计算时延的影响

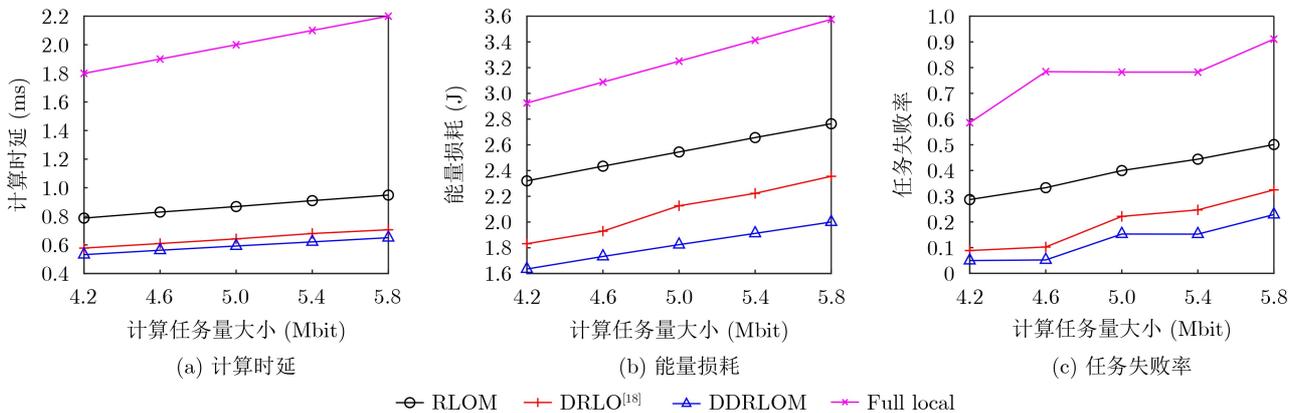


图 6 计算卸载的平均性能与计算任务量之间的关系

产效率。针对复杂受限空间的矿井环境,引入智能超表面辅助井下通信,完善井下无线通信理论并优化智能计算卸载方法是下一步研究工作的重点,为5G助力智慧矿山数字化转型提供有力支撑。

### 参考文献

- [1] 丁恩杰, 俞啸, 夏冰, 等. 矿山信息化发展及以数字孪生为核心的智慧矿山关键技术[J]. 煤炭学报, 2022, 47(1): 564–578. doi: [10.13225/j.cnki.jccs.yg21.1930](https://doi.org/10.13225/j.cnki.jccs.yg21.1930).  
DING Enjie, YU Xiao, XIA Bing, *et al.* Development of mine informatization and key technologies of intelligent mines[J]. *Journal of China Coal Society*, 2022, 47(1): 564–578. doi: [10.13225/j.cnki.jccs.yg21.1930](https://doi.org/10.13225/j.cnki.jccs.yg21.1930).
- [2] MENG Yifan and LI Jingzhao. Task offloading and resource allocation mechanism of moving edge computing in mining environment[J]. *IEEE Access*, 2021, 9: 155534–155542. doi: [10.1109/ACCESS.2021.3129464](https://doi.org/10.1109/ACCESS.2021.3129464).
- [3] 赵小虎, 张凯, 赵志凯, 等. 矿山物联网网络技术的发展趋势与关键技术[J]. 工矿自动化, 2018, 44(4): 1–7. doi: [10.13272/j.issn.1671-251x.17324](https://doi.org/10.13272/j.issn.1671-251x.17324).  
ZHAO Xiaohu, ZHANG Kai, ZHAO Zhikai, *et al.* Developing trend and key technologies of network technology of mine Internet of things[J]. *Industry and Mine Automation*, 2018, 44(4): 1–7. doi: [10.13272/j.issn.1671-251x.17324](https://doi.org/10.13272/j.issn.1671-251x.17324).
- [4] DENG Xiaoheng, YIN Jian, GUAN Peiyuan, *et al.* Intelligent delay-aware partial computing task offloading for multiuser industrial Internet of Things through edge computing[J]. *IEEE Internet of Things Journal*, 2023, 10(4): 2954–2966. doi: [10.1109/JIOT.2021.3123406](https://doi.org/10.1109/JIOT.2021.3123406).
- [5] PORAMBAGE P, OKWUIBE J, LIYANAGE M, *et al.* Survey on multi-access edge computing for internet of things realization[J]. *IEEE Communications Surveys & Tutorials*, 2018, 20(4): 2961–2991. doi: [10.1109/COMST.2018.2849509](https://doi.org/10.1109/COMST.2018.2849509).
- [6] SUN Lu, WANG Jie, and LIN Bin. Task allocation strategy for MEC-enabled IIoTs via Bayesian network based evolutionary computation[J]. *IEEE Transactions on Industrial Informatics*, 2021, 17(5): 3441–3449. doi: [10.1109/TII.2020.3019572](https://doi.org/10.1109/TII.2020.3019572).
- [7] DINH T Q, TANG Jianhua, LA Q D, *et al.* Offloading in mobile edge computing: Task allocation and computational frequency scaling[J]. *IEEE Transactions on Communications*, 2017, 65(8): 3571–3584. doi: [10.1109/TCOMM.2017.2699660](https://doi.org/10.1109/TCOMM.2017.2699660).
- [8] 袁亮, 俞啸, 丁恩杰, 等. 矿山物联网人-机-环状态感知关键技术研究[J]. 通信学报, 2020, 41(2): 1–12. doi: [10.11959/j.issn.1000-436x.2020036](https://doi.org/10.11959/j.issn.1000-436x.2020036).  
YUAN Liang, YU Xiao, DING Enjie, *et al.* Research on key technologies of human-machine-environment states perception in mine Internet of Things[J]. *Journal on Communications*, 2020, 41(2): 1–12. doi: [10.11959/j.issn.1000-436x.2020036](https://doi.org/10.11959/j.issn.1000-436x.2020036).
- [9] 周代勇. 井下风致振动压电能量收集技术[J]. 煤矿安全, 2021, 52(9): 153–156. doi: [10.13347/j.cnki.mkaq.2021.09.024](https://doi.org/10.13347/j.cnki.mkaq.2021.09.024).  
ZHOU Daiyong. Wind-induced vibration piezoelectric energy harvesting technology in underground mine[J]. *Safety in Coal Mines*, 2021, 52(9): 153–156. doi: [10.13347/j.cnki.mkaq.2021.09.024](https://doi.org/10.13347/j.cnki.mkaq.2021.09.024).
- [10] YE Junliang and GHARAVI H. Deep reinforcement learning-assisted energy harvesting wireless networks[J]. *IEEE Transactions on Green Communications and Networking*, 2021, 5(2): 990–1002. doi: [10.1109/TGCN.2020.3045075](https://doi.org/10.1109/TGCN.2020.3045075).
- [11] RODRIGUEZ J C, NICO V, and PUNCH J. Powering wireless sensor nodes for industrial IoT applications using vibration energy harvesting[C]. IEEE 5th World Forum on Internet of Things, Limerick, Ireland, 2019: 392–397. doi: [10.1109/WF-IoT.2019.8767352](https://doi.org/10.1109/WF-IoT.2019.8767352).
- [12] SUN Yingying, SONG Chunhe, YU Shimao, *et al.* Energy-efficient task offloading based on differential evolution in edge computing system with energy harvesting[J]. *IEEE Access*, 2021, 9: 16383–16391. doi: [10.1109/ACCESS.2021.3052901](https://doi.org/10.1109/ACCESS.2021.3052901).
- [13] RANJAN A, SAHU H B, and MISRA P. Wave propagation model for wireless communication in underground mines[C]. 2015 IEEE Bombay Section Symposium (IBSS), Mumbai, India, 2015: 1–5. doi: [10.1109/IBSS.2015.7456655](https://doi.org/10.1109/IBSS.2015.7456655).
- [14] LEI Lei, XU Huijuan, XIONG Xiong, *et al.* Multiuser resource control with deep reinforcement learning in IoT edge computing[J]. *IEEE Internet of Things Journal*, 2019, 6(6): 10119–10133. doi: [10.1109/JIOT.2019.2935543](https://doi.org/10.1109/JIOT.2019.2935543).
- [15] XIAO Liang, LU Xiaozhen, XU Tangwei, *et al.* Reinforcement learning-based mobile offloading for edge computing against jamming and interference[J]. *IEEE Transactions on Communications*, 2020, 68(10): 6114–6126. doi: [10.1109/TCOMM.2020.3007742](https://doi.org/10.1109/TCOMM.2020.3007742).
- [16] MIN Minghui, WAN Xiaoyue, XIAO Liang, *et al.* Learning-based privacy-aware offloading for healthcare IoT with energy harvesting[J]. *IEEE Internet of Things Journal*, 2019, 6(3): 4307–4316. doi: [10.1109/JIOT.2018.2875926](https://doi.org/10.1109/JIOT.2018.2875926).
- [17] WANG Hao and HUANG Jianwei. Incentivizing energy trading for interconnected microgrids[J]. *IEEE Transactions on Smart Grid*, 2018, 9(4): 2647–2657. doi: [10.1109/TSG.2016.2614988](https://doi.org/10.1109/TSG.2016.2614988).
- [18] MIN Minghui, XIAO Liang, CHEN Ye, *et al.* Learning-based computation offloading for IoT devices with energy harvesting[J]. *IEEE Transactions on Vehicular Technology*,

- 2019, 68(2): 1930–1941. doi: [10.1109/TVT.2018.2890685](https://doi.org/10.1109/TVT.2018.2890685).
- [19] WANG Jiadai, ZHAO Lei, LIU Jiajia, *et al.* Smart resource allocation for mobile edge computing: A deep reinforcement learning approach[J]. *IEEE Transactions on Emerging Topics in Computing*, 2021, 9(3): 1529–1541. doi: [10.1109/TETC.2019.2902661](https://doi.org/10.1109/TETC.2019.2902661).
- [20] QIU Chengrun, HU Yang, CHEN Yan, *et al.* Deep deterministic policy gradient (DDPG)-based energy harvesting wireless communications[J]. *IEEE Internet of Things Journal*, 2019, 6(5): 8577–8588. doi: [10.1109/JIOT.2019.2921159](https://doi.org/10.1109/JIOT.2019.2921159).
- [21] REN Jieying and XU Shaoyi. DDPG based computation offloading and resource allocation for MEC systems with energy harvesting[C]. 2021 IEEE 93rd Vehicular Technology Conference (VTC2021-Spring), Helsinki, Finland, 2021: 1–5. doi: [10.1109/VTC2021-Spring51267.2021.9448922](https://doi.org/10.1109/VTC2021-Spring51267.2021.9448922).
- [22] MAO Yuyi, ZHANG Jun, and LETAIEF K B. Dynamic computation offloading for mobile-edge computing with energy harvesting devices[J]. *IEEE Journal on Selected Areas in Communications*, 2016, 34(12): 3590–3605. doi: [10.1109/JSAC.2016.2611964](https://doi.org/10.1109/JSAC.2016.2611964).
- [23] QIN Meng, CHENG Nan, JING Zewei, *et al.* Service-oriented energy-latency tradeoff for IoT task partial offloading in MEC-enhanced multi-RAT networks[J]. *IEEE Internet of Things Journal*, 2021, 8(3): 1896–1907. doi: [10.1109/JIOT.2020.3015970](https://doi.org/10.1109/JIOT.2020.3015970).
- [24] 国家市场监督管理总局, 国家标准化管理委员会. GB/T 3836.1-2021 爆炸性环境 第1部分: 设备 通用要求[S]. 北京: 中国标准出版社, 2021.
- State Administration for Market Regulation and Standardization Administration of the People's Republic of China. GB/T 3836.1-2021 Explosive atmospheres—Part 1: Equipment—General requirements[S]. Beijing: Standards Press of China, 2021.
- 闵明慧: 女, 讲师, 研究方向为无线通信、网络安全、隐私保护等.  
张 鹏: 男, 博士生, 研究方向为无线通信、边缘计算.  
朱浩鹏: 男, 硕士生, 研究方向为无线通信、隐私保护.  
程志鹏: 男, 博士生, 研究方向为无线通信、边缘计算.  
马 帅: 男, 副教授, 研究方向为语义通信、可见光通信等.  
李世银: 男, 教授, 研究方向为煤矿信息化、移动目标定位等.  
肖 亮: 女, 教授, 研究方向为无线通信、网络安全、强化学习等.  
彭国军: 男, 教授, 研究方向为网络安全、恶意代码、可信软件等.

责任编辑：陈 倩