# 区块链赋能物联网中联合资源分配与 控制的智能计算迁移研究

陈思光1) 王 倩1) 张海君2) 王 堃3

<sup>1)</sup>(南京邮电大学江苏省宽带无线通信和物联网重点实验室 南京 210003) <sup>2)</sup>(北京科技大学通信工程系 北京 100083)

3)(加州大学洛杉矶分校电子与计算机工程系 洛杉矶 CA90095 美国)

摘 要 大数据场景下,远程云服务器通常被部署用于数据处理与价值挖掘,但在面对时延敏感型或需要动态频繁交互的业务时,该种处理模式显得力不从心.作为对云计算模式的补充,雾计算因其可有效降低任务处理时延、能耗与带宽消耗而备受关注;同时,面向雾计算的计算迁移机制因其能有效缓解节点的处理负担并改善用户体验而成为领域研究焦点.在雾计算模式下,为了更好地满足计算密集型任务对时延与能耗的要求,基于区块链赋能物联网场景,本文提出了一种联合资源分配与控制的智能计算迁移方案.具体地,规划了一个在时延、能耗与资源约束下的最小化所有任务完成总成本的优化问题,其总成本构成综合考量了时延、能耗和挖掘成本,通过对通信、计算资源与迁移决策的联合优化,实现总成本的最小化.为完成任务迁移,终端以矿工的身份向雾节点挖掘(租借)计算资源,所提出的基于区块链技术的激励机制可充分调动终端和雾节点参与计算迁移的积极性并保障交易过程的安全性,设计的奖励分配规则可保证成功挖掘资源终端收获奖励的公平性.为解决上述规划的优化问题(即混合整数非线性规划问题),提出了一个联合通信、计算与控制的智能计算迁移算法,该算法融合深度确定性策略梯度算法思想,设计了基于反梯度更新的双"行动者-评论家"神经网络结构,使训练过程更加稳定并易于收敛;同时,通过对连读动作输出进行概率离散化运算,使其更加适用于混合整数非线性规划问题的求解.最后,仿真结果表明本文方案能以较快的速度收敛,且与其他三种基准方案相比,本文方案的总成本最低,例如,与其中性能最好的基于深度 Q 学习网络的计算迁移方案相比,总成本平均可降低 15.2%.

关键词 计算迁移;雾计算;区块链;深度强化学习;资源分配中图法分类号 TP393 **DOI**号 10.11897/SP.J.1016.2022.00472

# Resource Allocation and Control Co-aware Smart Computation Offloading for Blockchain-Enabled IoT

CHEN Si-Guang<sup>1)</sup> WANG Qian<sup>1)</sup> ZHANG Hai-Jun<sup>2)</sup> WANG Kun<sup>3)</sup>

<sup>1)</sup>(Jiangsu Key Lab of Broadband Wireless Communication and Internet of Things, Nanjing University of

Posts and Telecommunications, Nanjing 210003)

<sup>2)</sup>(Department of Communication Engineering, University of Science and Technology Beijing, Beijing 100083)

<sup>3)</sup>(Department of Electrical and Computer Engineering, University of California Los Angeles, Los Angeles CA90095 USA)

**Abstract** Under the scenario of big data, the remote cloud server is usually deployed for data processing and value mining, but in the face of delay-sensitive or dynamic and frequent update applications, this

收稿日期:2020-08-13;在线发布日期:2021-07-15.本课题得到国家自然科学基金(61971235,61771258)、江苏省"六大人才高峰"高层次人才项目(XYDXXJS-044)、江苏省"333高层次人才培养工程"、南京邮电大学'1311'人才计划、中国博士后科学基金(面上一等)(2018M630590)、江苏省博士后科研资助计划(2021K501C)和赛尔网络下一代互联网技术创新项目(NGII20190702)资助.陈思光(通信作者),博士,教授,中国计算机学会(CCF)会员(D6662M),主要研究领域为雾/边缘计算、智能物联网.E-mail:sgchen@njupt.edu.cn.王倩,博士研究生,中国计算机学会(CCF)学生会员(C6086G),主要研究领域为雾/边缘计算、智能物联网.张海君,博士,教授,主要研究领域为6G移动通信、人工智能和大数据.王堃,博士,教授,主要研究领域为边缘计算、区块链和分布式机器学习.

processing paradigm appears to be inadequate from reality. As a complement to cloud computing paradigm, fog computing has attracted great attention for which can effectively reduce task processing delay, energy and bandwidth consumptions. At the same time, fog computing based computation offloading mechanism has become a research focus, because it can effectively alleviate the processing burden of nodes and improve the experience of users. Under the fog computing paradigm, in order to better meet the requirements of delay and energy consumption for computation-intensive tasks, based on the blockchainenabled Internet of Things (IoT) scenario, this paper proposes a resource allocation and control co-aware smart computation offloading scheme. Specifically, an optimization problem is formulated to minimize the total cost of all tasks under the constraints of delay, energy consumption, and communication and computation resources. Based on the comprehensive consideration, the component of the total cost includes the delay, energy consumption and resource mining costs; it can achieve the minimization of the total cost by jointly optimizing the communication resource, computation resource and offloading decision. In order to inspire the active participation of terminals and the fog node in the computation offloading process, and more close to the needs of the real scenario, an incentive mechanism is designed in this paper. That is, to complete the offloading of tasks, the terminal mines (rents) computation resource from the fog node as a miner and the fog node charges a certain fee according to the needed resource of terminal. For the terminals that successfully obtain the resources to complete the tasks efficiently, the system will allocate the corresponding rewards according to the occupation ratios of gained computation resource, which ensures the fairness allocation of the rewards for the successful miners. This mechanism enables that the fog node and terminals both can win benefits in the computation offloading process, which promotes their collaboration. Meanwhile, this blockchain-based incentive mechanism guarantees the security of the transaction process. For the sake of solving the above formulated optimization problem (i.e., a mixed integer nonlinear programming problem), we propose a communication, computation and control co-aware smart computation offloading algorithm (3CC-SCO). By integrating the concept of deep deterministic policy gradient (DDPG) algorithm, our algorithm designs an inverting gradient update based double actor-critic neural networks structure to improve the stable and convergence rate in the training process. At the same time, it is more suitable for solving the mixed integer optimization problem by adopting probabilistic discrete operation of continuous action output. Finally, the simulation results demonstrate that the proposed scheme can converge to the optimal value quickly, and the total cost of the proposed scheme is the lowest compared with other three benchmark schemes, for example, as compared with the bestperforming scheme, i.e., deep Q-learning network (DQN) based computation offloading scheme, the total cost can be reduced by an average of 15.2%.

**Keywords** computation offloading; fog computing; blockchain; deep reinforcement learning; resource allocation

# 1 引 言

融合人工智能(Artificial Intelligence, AI)与5G技术的物联网(Internet of Things, IoT)使得万物智能互联成为可能,作为智能物联网(Artificial Intelligence Internet of Things, AIoT)的三大典型应用:工业物联网(Industrial Internet of Things, IIoT)、智慧医疗和自动驾驶,因其对任务处理时延的苛刻

要求所面临的挑战成为学者们关注的焦点.例如,在智能制造领域,各种 IIoT 设备根据各自的制造流程和控制需求生成不同的计算任务,而往往本地设备的计算能力有限,在承担大规模密集型任务计算时通常显得力不从心<sup>[1]</sup>.特别地,制造流水线对各个工序衔接的时延要求极为苛刻,一旦超过容忍延迟有可能造成极大的损失甚至是灾难<sup>[2]</sup>.若将任务迁移至远程云中心处理,由于长距离通信及各种可能

发生的核心网络拥塞都将对服务时延产生极大影响,难以满足 IIoT 对低时延的要求. 在当前大数据背景下,高效地处理延迟敏感型任务,不是通过对IoT、5G、AI 等技术简单的组合运用就能解决的,需要从网络体系架构层面做全新的设计. 传统的云计算模式体系由于离用户距离较远,无法满足用户对时延以及能耗的苛刻要求,该类问题的存在催生了雾/边缘计算模式<sup>[3]</sup>. 雾计算作为一种新的计算模式,将计算、存储资源下沉到靠近终端设备的网络边缘,可减少端到端的业务服务时延,缓解核心网流量压力和降低通信开销,即可通过借助近距离的边缘节点资源有效提升用户的服务体验<sup>[4]</sup>.

当前,针对雾/边缘计算的研究已引起学术及工 业界的深度关注,特别地,针对雾/边缘计算场景下 的计算迁移问题成为研究焦点. 例如, 文献[5-7]从 时间与能耗的角度,提出了一个最小化时间及能耗 成本的计算迁移方案,基于传统数学规划算法,通 过优化迁移决策和资源分配,实现了任务完成时间 及能耗成本的最小化. 为克服上述传统数学规划求 解方法难以适应动态复杂网络环境问题, 文献[8]和 [9]将机器学习的深度学习理论应用到雾/边缘计算 迁移方案中,模拟人脑的多层神经网络使其在动态 网络环境中能够自适应地做出迁移决策, 有效解决 非凸优化问题,实现任务完成成本的最小化.上述 基于深度学习方案虽然可以适应动态的网络环境, 但其求解结果的好坏主要依赖于人工标记的大量训 练样本,其过程效率低下且成本过高.考虑到未来 环境的变化,深度学习方法需要持续标记新的数据, 很难适应未来可能发生的变化, 而具有不同需求任 务的迁移决策系统应具有长期规划和不断学习的能 力[10]. 强化学习可有效满足这一需求,无需"数据 喂养", 而是通过自己试错的方式学习[11], 自我探索 动态环境,根据状态给出惩罚或奖励,但无法适应 高维状态空间. 深度强化学习方法将深度学习与强 化学习相结合,深度学习对高维状态空间的感知能 力有效提高了强化学习的决策能力[12],更加适合具 有时变无线信道和动态变化服务请求的网络环境, 智能地与环境交互. 文献[13-15]基于深度强化学习 中的深度 Q 学习网络 (Deep Q-learning Network, DQN)解决计算迁移中迁移决策与资源优化分配问 题,进一步提高智能体在动态复杂的网络环境中感 知决策能力,实现高效智能计算迁移,提高用户体 验质量.

从上述分析可以发现,当前针对雾/边缘计算场景下的计算迁移研究面临两大挑战.(1)基于深度

学习方法的求解结果好坏过于依赖高成本的大量人工标记数据,虽然基于 DQN 的方案在一定程度上弥补了该缺点,但只适应于低维离散输出,难以实现高维或连续动作输出. (2)当前大多数研究方案均建立在理想的网络场景下,即雾节点会自愿向外部用户免费提供资源,这有悖于真实的网络场景,所以当前方案缺乏相关激励及公平性机制的考量,使得终端和雾节点真正参与迁移的积极性偏弱. 此外,虽然存在少数智能计算迁移研究方案融合考虑了基于区块链的激励机制,但缺乏对迁移决策、计算、通信等资源的综合优化考量.

针对上述挑战,在大数据物联网场景下本文提出了一种区块链赋能的联合资源分配与控制的智能 计算迁移方案,主要贡献总结如下:

- 研究了一种区块链赋能的计算迁移模型,基于此构建了一个最小化所有任务完成总成本的优化问题,其中,该总成本由完成时间、能耗及挖掘成本组成,通过对通信、计算资源与迁移决策的联合优化,在一定时延、带宽和能耗约束下,实现总成本的最小化.
- 为提升终端和雾节点参与迁移的积极性,更加贴近实际场景需求,本文设计了一种激励机制,即为完成任务迁移,终端以矿工的身份向雾节点挖掘(租借)计算资源,雾节点根据其所需资源收取一定费用,成功得到资源的终端因高效完成任务,系统按照其所得算力占比来分配奖励,保证了对挖掘成功的矿工收获奖励的公平性. 此机制使得雾节点和终端在迁移过程中都能得到收益,促进了二者的合作,同时基于区块链的激励机制保障了交易过程的安全性.
- 针对上述构建的混合整数非线性规划问题,本文提出了一个联合通信、计算与控制的智能计算迁移(Communication, Computation and Control Co-aware Smart Computation Offloading, 3CC-SCO)算法,该算法融合深度确定性策略梯度(Deep Deterministic Policy Gradient, DDPG)算法<sup>[16]</sup>思想,设计了基于反梯度更新的双"行动者-评论家"神经网络结构,使其训练过程更加稳定并易于收敛。同时,通过对连续迁移决策输出进行概率离散化处理,使其更加适用于混合整数非线性规划问题的求解。

最后,对本文所提出的方案进行了仿真评估, 并研究了各种系统参数的影响,仿真结果表明本文 方案能以较快的速度收敛,且与其他三种基准方案 相比,本文方案的总成本最低.

本文其余部分组织如下:第2节回顾了相关工作;第3节介绍了本文的系统模型;第4节是对优化问题的描述;第5节给出了本文所提3CC-SCO算法的具体介绍;第6节是仿真结果与分析;最后,第7节总结全文.

## 2 相关工作

新型物联网技术的迅速发展给人们带来极大便 利的同时,也使得用户提升了对高效计算服务的需 求,以更低的代价获得更高的服务质量成为普遍要 求,而基于雾/边缘计算的计算迁移机制成为实现这 一目标的有效途径,将计算任务迁移至离终端较近 的雾节点处理能更好地满足延迟或能耗敏感型任务 的需求,提升用户的体验.目前,关于计算迁移的 研究大致可分为以下三类:

第一类主要利用传统的数学规划方法优化计算 迁移中的资源分配和迁移决策,例如文献[17-22]. 在最小化任务完成能耗方面, 文献[17]提出了一种 雾计算迁移加速梯度算法, 在延迟约束下, 通过联 合优化迁移比和传输时间来最小化雾节点的能耗, 并且通过动态调整雾节点的计算速度,设计了一个 能耗交替最小化算法以适应动态变化网络场景的需 要,达到进一步提高系统性能的目的.在最小化任 务完成延迟方面, 文献[18-20]取得了显著效果. 例 如: 文献[18] 提出了一种递归算法来逼近最优解, 通过联合优化雾无线网络中任务的迁移和调度策 略,使任务的平均完成时间最小化. Mukherjee 等人 在文献[19]中构造了一种雾云混合的三层网络模 型,将任务完成时间最小化问题转换为二次约束二 次规划问题,并提出了一个近似算法优化雾节点的 计算和通信资源,实现延迟最小化的目标.上述文 献[19]的作者在文献[20]中进一步地考虑了雾节点 中任务排队对延迟的影响,并提出一种新的求解方 法——基于半定松弛的延迟最小化算法,可获得最 小延迟的近似解. 在时延和能耗的联合优化方面, 文献[21]和[22]给出了有效的解决方案. 具体地, 文 献[21]设计了一种物-雾-云架构,通过优化迁移决策 和传输功率来实现时延和能量加权和成本最小化, 针对此混合整数非线性规划问题, 先进行松弛处理 分解为两个子问题,并提出了一种基于牛顿迭代算 法的能耗和时延高效计算迁移资源分配算法,可有 效地降低总成本. 文献[22]研究了物联网中任务和 能量的联合迁移机制,允许任务节点以时分多址方 式将能量和任务分配给多个相邻的辅助雾节点,充 分利用能量收集的优势,基于李雅普诺夫优化理论设计了一种在线迁移策略,使任务节点的执行延迟和能耗最小化. 综上可知,以上研究方案取得了良好的性能提升效果,为传统数学规划方法求解该类问题奠定了坚实的基础,但该类求解方法存在一个共同的缺点,即采用传统梯度优化方法具有收敛速度慢的特点,在处理大规模数据时算法耗时长,难以适应大数据网络数据处理时效性需求. 同时,这些方法均建立在理想的网络环境下,即需提前获知网络状态参数,而这在实际情况中是比较困难的.

如今利用机器学习框架来实现自动和有效的网 络资源分配管理成为了一种新的趋势[23],为了更好 地适应复杂动态的网络环境并克服上述研究方案的 缺陷,第二类研究方案主要基于深度学习方法求解 计算迁移问题,例如文献[24-28]. 文献[24]研究了建 立在集中式单元-分布式单元架构之上支持深度学 习的边缘服务迁移架构,提出了一种最小化总传输 时延的启发式迁移方法,通过初始-目标边缘服务器 距离的估计和启发式搜索,获得合适的迁移策略, 从而有效地缩短 5G 网络中深度学习边缘服务的时 延. Yang 等人[25]针对工业物联网环境,提出了一种 基于移动边缘计算(Mobile Edge Computing, MEC) 的分层机器学习任务分配新框架,并在设备和边缘 节点中部署深度神经网络模型,经过预训练的深度 神经网络使其能够动态地确定迁移策略,实现对任 务完成总延迟的最小化. 文献[26]提出了一种能量 与延迟协同感知的雾计算迁移机制,即为实现高效 的任务计算,构建了一个任务完成时间和能量消耗 权重和最小化问题,并提出了一种基于深度学习的 联合迁移决策和资源分配算法,通过联合优化迁移 决策、本地 CPU、带宽和外部 CPU 占用率来解决该 问题,提高了网络效率.文献[27]和[28]提出了一种 MEC 中基于深度学习的计算迁移方案,综合考虑了 用户设备的剩余能量、本地和服务器执行的成本、 迁移决策和变化的网络条件对任务执行总成本的影 响,并通过贪婪法得到的标签数据对边缘节点中深 度神经网络模型进行训练,经过不断学习最终给出 最优的迁移决策, 使每个任务的总执行成本降至最 低. 上述研究方案均采用单一的深度学习算法来求 解最优的迁移策略,能较好地适应动态的网络环境, 但深度神经网络需要大量的样本数据对其进行训 练,而样本数据的生成需要人工获得,且其规模和 质量直接影响到深度学习的性能.

第三类研究方案将深度学习的特征感知能力与 强化学习的决策能力相融合,使其能够有效解决复

杂网络系统中感知与决策问题,例如:在大数据场 景下, 文献[29]为满足用户对高效计算服务需求, 通过对迁移决策、计算资源和带宽的综合性考量, 构建了一个最小化所有用户任务执行延迟与能耗权 重和的优化问题,提出了一个异步云边协同的深度 强化学习算法,该算法通过自适应地调整迁移策略, 实现了系统总成本的最小化. 为了解决物联网中缓 存、计算和无线电资源的分配问题, 文献[30]采用 深度强化学习中的"行动者-评论家"算法,通过不 断地学习,得出最优的迁移决策以及资源调度策略 来最大程度地减少服务延迟. 针对 MEC 超密集网 络, Chen 等[31]在该网络中选择多个边缘节点中的一 个来进行计算迁移,并将最优迁移策略问题建模为 马尔可夫决策过程,目标是最小化长期成本,并且 基于移动用户和边缘节点之间的信道质量、能量队 列,提出一种基于 DQN 的迁移策略算法,以获得最 优迁移决策. 为了使任务迁移决策和无线电资源分 配最佳地适应时变无线信道, Huang 等[32]提出了一 种基于深度强化学习的在线分流算法,大大降低了 计算复杂性,实现了自适应地做出最优迁移决策.

此外,为保证智能计算迁移过程中的安全问题, 文献[33]构建了一个基于区块链的 MEC 系统协作计 算迁移和资源分配框架,提出了一种基于异步优势 "行动者-批评家"的协作计算迁移和资源分配算 法,在确保 MEC 系统中数据可靠和不可逆的基础 上,有效提高了 MEC 系统的计算速度和区块链系统 的吞吐量. 类似地, 文献[34]研究了基于区块链和移 动边缘计算的未来无线网络自适应资源分配方案, 利用 Double-Dueling DQN 算法求解最优的资源分 配以联合优化 MEC 系统的计算速度和区块链系统 的吞吐量. 为了激励参与计算迁移的积极性, 文献 [35-37]引入了基于区块链的激励机制. Zhang 等<sup>[35]</sup> 设计了一个基于区块链的设备到设备边缘缓存管理 系统,该系统包含一个基于区块链的内容缓存和交 付市场,将其作为缓存共享激励机制,并提出了基 于 DRL 的动态缓存管理优化算法以实现任务延迟 最小化. 文献[36]在区块链赋能的 MEC 系统中提出 了一种新的基于无模型 DRL 的在线计算迁移方法, 通过优化迁移决策实现迁移成本的最小化. 文献 [37] 将任务迁移、用户隐私保护和挖掘利润作为一 个联合优化问题,为了获得最优迁移决策,提出了 一个基于 DQN 的决策算法,在最小化迁移延迟和能 耗的同时,获得最佳的隐私性能和挖掘利润.

上述计算迁移方案在资源优化和成本节约等方面取得了优越的效果,较适应复杂动态的大数据网

络环境,尤其将区块链技术结合到计算迁移中,可有效提升系统的安全性.然而,上述大部分研究方案建立在理想的环境基础之上,即雾节点自愿向终端免费提供资源,雾节点和终端参与整个迁移过程的积极性偏弱.尽管存在少数研究方案将基于区块链的激励机制融入到智能计算迁移中,提升了终端和雾节点参与计算迁移的积极性,使其更加贴近实际网络场景,但是该类方案对完成任务计算的网络资源优化考量较为单一,缺乏通信、计算与控制的协同综合考量.

## 3 系统模型

## 3.1 网络模型

本文构建了一个区块链赋能的雾计算迁移模型,自下而上包含两部分——物联网设备层和雾层,如图 1 所示.

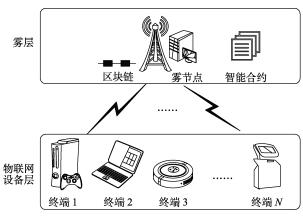


图 1 网络模型

假设物联网设备层有 N 个智能终端,终端  $i \in \{1,2,3,...,N\}$  中存在计算密集型任务,但可能由于 其自身计算资源受限,在本地处理并不能满足其需求,为提高服务质量与用户体验,该任务可通过无线接入点迁移到雾节点处理,假设该 N 个终端均在 此雾节点的通信覆盖范围内。同时,为促进终端和雾节点参与计算迁移的积极性,本文考虑区块链赋能的迁移模型,当雾节点接收到来自物联网设备层的资源挖掘(租借)请求时,雾节点根据智能合约中的规定,在保证接入终端身份合法的情况下,做出智能迁移决策,对于需要迁移的任务,终端以矿工身份至雾节点向其租借计算资源以完成任务,并支付给雾节点相应的租赁费用,雾节点通过优化计算资源和通信资源的分配,使得任务完成的总成本最低。当矿工利用挖掘(租借)到的算力完成任务时,

将交易过程记入任务区块,并广播到整个区块链网络与其他矿工达成共识,最后添加到区块链中,并因其为整个网络性能提升所做的贡献得到一定的奖励.

## 3.2 计算模型

定义每个终端向雾节点发送的请求信息为  $(l_i,d_i,T_i^{\max},E_i^{\max})$ ,其中  $l_i$  表示终端 i 的任务大小 (bits), $d_i$  为完成该任务所需要的 CPU 周期数,  $T_i^{\max}$  和  $E_i^{\max}$  分别表示完成该任务的最大容忍延迟 和容忍能耗. 雾节点根据请求信息以及计算资源和 通信资源可用情况,并依照智能合约中合法访问及 收费等规则给出智能迁移决策  $x_i$ . 当  $x_i$  = 0 时,表示任务在本地终端 i 处理,当  $x_i$  = 1 时,表示终端 i 以 矿工身份,将任务迁移至雾节点处理.

## (1) 本地计算模型

当  $x_i = 0$  时,任务的完成时间和能耗只与本地的计算能力  $f_i^{local}$  (cycles/s)有关,因此终端 i 的任务在本地处理时间和能耗成本分别表示为

$$T_i^{\text{local}} = \frac{d_i}{f_i^{\text{local}}} \tag{1}$$

$$E_i^{\text{local}} = k^{\text{local}} (f_i^{\text{local}})^2 d_i$$
 (2)

其中,公式(2)中的  $k^{local}$  为本地终端的能耗系数. 此时,处理该任务所需的总成本即时间和能耗 成本的加权和为

$$C_i^{\text{local}} = \lambda_i^t T_i^{\text{local}} + \lambda_i^e E_i^{\text{local}}$$
 (3)

其中  $\lambda_i^t$  为时间成本的单位价格(tokens/s ),  $\lambda_i^e$  为能 耗成本的单位价格(tokens/J ) .

#### (2) 雾计算模型处理

当  $x_i$  = 1时,终端 i 的任务被迁移至雾节点处理,此时的总成本将会由三部分组成,即时间成本、能耗成本以及挖掘成本.

任务的完成时间由两部分组成,即传输时间和在雾节点的计算时间,由于挖掘请求和计算结果的数据大小相对于任务的数据大小来说小很多,因此挖掘请求时间和计算结果返回时间可忽略不计.同时,本文考虑计算结果的验证可在极短的时间内完成并达成共识,所以共识验证时间不纳入总时间成本,则任务的完成时间可表示为

$$T_i^{\text{fog}} = \frac{l_i}{\psi_i} + \frac{d_i}{f_i^{\text{fog}}} \tag{4}$$

其中,  $f_i^{\text{fog}}$  (cycles/s ) 为分配给终端 i 任务的计算资源,  $\psi_i$  表示终端 i 与雾节点之间的传输速度,即

$$\psi_i = b_i \log_2 \left( 1 + \frac{p_i h_i^2}{N_0 b_i} \right)$$
 (5)

其中 $b_i$ 是分配给终端i的带宽, $p_i$ 表示终端i上传

数据的传输功率, $h_i$  表示终端 i 在无线信道中的信道增益, $N_0$ 表示信道噪声功率.

相应地,完成该任务的能耗可表示为

$$E_i^{\text{fog}} = p_i \frac{l_i}{\psi_i} + k^{\text{fog}} (f_i^{\text{fog}})^2 d_i$$
 (6)

其中 $k^{\text{fog}}$ 表示雾节点的计算能耗系数.

作为外部的雾计算资源并不是可以随意使用的,而需要支付一定租赁费用. 在区块链网络中,定义终端 i 以矿工身份到雾节点挖掘资源,租借到的相应算力占比 $\varphi_i$ 为

$$\varphi_i = \frac{f_i^{\text{fog}}}{\sum_{i=1}^{N} f_i^{\text{fog}}}$$
 (7)

矿工利用所得算力执行任务,首先完成任务的矿工 将其打包的区块广播到整个区块链网路中,与其他 矿工达成共识者会获得一定奖励 R,若未达成共识成 为孤立块从而被遗弃,即任务区块无法添加到区块链 中,那么矿工也就无法得到奖励,其发生概率<sup>[38]</sup>为

$$P_i = 1 - e^{-\mu g(v_i)}, (8)$$

其中 $\mu$ 为常数,表示区块成功添加到当前区块链的概率遵循泊松过程<sup>[39]</sup>, $g(v_i)$ 是关于区块大小 $v_i$ 的函数,表示区块形成之后广播到区块链中所需的时间,可记为 $\delta*v_i$ ,其中 $\delta$ 为常数,表示延迟系数. 因此成功挖掘并加入到区块链中的概率即为 $(1-P_i)$ . 矿工完成整个挖掘过程获得奖励,该奖励可用于支持后续租借雾节点计算资源的租赁费.

若终端 i 成功完成挖矿并得到奖励,为保证每个成功挖矿的矿工得到奖励的公平性,其奖励按照算力占比来分配,具体定义为  $R \times \varphi_i \times (1-P_i)$ ,则完成此任务的总挖掘成本为

$$C_i^{\text{mining}} = c \times d_i - R \times \varphi_i \times (1 - P_i)$$
 (9)

其中 c 为雾节点向终端 i 收取的单位 CPU 周期价格,此处  $C_i^{\text{mining}} \ge 0$ .

因此,完成该迁移任务的总成本可以表示为

$$C_i^{\text{fog}} = \lambda_i^t T_i^{\text{fog}} + \lambda_i^e E_i^{\text{fog}} + C_i^{\text{mining}}$$
 (10)

## 4 优化问题描述

通常,时延、能耗和挖掘成本是衡量区块链网络性能的重要指标.本文的优化目标为最小化物联网设备层所有任务完成的总成本,由时延、能耗和挖掘成本构成,即通过联合优化迁移决策  $x_i$  、信道带宽分配  $b_i$  以及雾节点计算资源分配  $f_i^{fog}$  来实现.该优化问题具体构建如下:

$$\min_{x_i, b_i, f_i^{\text{fog}}} C = \min_{x_i, b_i, f_i^{\text{fog}}} \sum_{i=1}^{N} [x_i C_i^{\text{local}} + (1 - x_i) C_i^{\text{fog}}]$$
 (11)

s.t. 
$$(1-x_i)T_i^{\text{local}} + x_i T_i^{\text{fog}} \le T_i^{\text{max}}$$
 (11a)

$$(1-x_i)E_i^{\text{local}} + x_i E_i^{\text{fog}} \le E_i^{\text{max}}$$
 (11b)

$$\sum_{i=1}^{N} x_i b_i \leqslant B \tag{11c}$$

$$\sum_{i=1}^{N} x_i f_i^{\text{fog}} \leqslant F \tag{11d}$$

$$x_i \in \{0,1\}$$
 (11e)

目标函数(11)即为最小化所有终端任务的完成总成本. 约束(11a)表示终端 i 任务完成的时间不得超过其最大容忍延迟  $T_i^{\max}$ . 约束(11b)表示终端 i 任务完成的能耗不得超过最大容忍能耗  $E_i^{\max}$ . 约束(11c)表示通信资源分配的约束,即分配给所有需要迁移任务的带宽之和不能超过信道的总带宽 B. 约束(11d)表示计算资源分配的约束,即分配给所有迁移任务的计算资源之和不得超过雾节点的总计算资源 F. 约束(11e)表示迁移决策的取值为 0 或 1,即任务在本地处理或迁移到雾节点处理.

## 5 3CC-SCO 算法

针对上述优化问题,由于网络环境的复杂动态 变化特性,通常情况下,难以提前准确获知相关网 络状态参数, 使得传统的求解方法往往显得力不从 心[40], 而作为机器学习方法的典型代表——深度强 化学习算法,可有效解决复杂动态网络场景中的智 能感知与迁移决策问题. 本文基于深度强化学习理 论解决上述混合整数非线性规划问题,即融合 DDPG 算法思想,提出了 3CC-SCO 算法. 为使训练 过程中更加稳定且易于收敛, 该算法设计了基于反 梯度更新的双"行动者-评论家"神经网络结构. 同 时,为有效解决连续-离散混合动作空间的优化问 题,对连续动作输出进行概率离散运算,克服传统 DDPG 算法的只适应于优化纯连续动作的局限性. 此外,为提升对连续动作搜索的效率并增大学习的 覆盖率,在原本连续动作的基础上加入 Ornstein-Uhlenbeck (OU)噪声,使得智能体对连续动作空 间的搜索有较好的效果.

#### 5.1 3CC-SCO算法要素定义

在 3CC-SCO 算法中,雾节点作为智能体收集来自所有终端在时隙 t 的状态信息  $s_t$  并执行动作  $a_t$  ,与环境进行交互后转换到下一可能状态  $s_{t+1}$  ,且获得即时奖励  $r_t$  .智能体以最大化累积奖励为目标,不断地学习并调整迁移策略,最终获得最优迁移策

略. 下面将具体地给出状态、动作以及奖励三要素的定义.

状态空间  $S = \{s_0, s_1, ..., s_t, s_{t+1}, ..., s_T\}$ :  $s_t$  表示时隙 t 的状态, $s_{t+1}$  表示下一状态,假设共有 T 个时隙。  $s_t = \{C_1(t), C_2(t), ..., C_N(t)\}$ ,其中, $C_i(t)$  表示终端 i 的任务完成成本.

动作空间  $A = \{a_0, a_1, ..., a_t, a_{t+1}, ..., a_T\}$ :  $a_t$  表示在时隙 t 所采取的动作, $a_{t+1}$  表示下一状态动作。  $a_t = (a_t^1, a_t^2)$ ,其中离散动作  $a_t^1 = (x(t))$ , x(t) 表示终端任务的迁移决策集,连续动作  $a_t^2 = (b(t), f^{fog}(t))$ ,b(t) 表示任务被分配到的信道带宽集合, $f^{fog}(t)$  表示任务被分配到的雾节点计算资源集合。

$$x(t) = \{x_1(t), x_2(t), ..., x_N(t)\}$$
 (12)

$$b(t) = \{b_1(t), b_2(t), ..., b_N(t)\}$$
 (13)

$$f^{\text{fog}}(t) = \{ f_1^{\text{fog}}(t), f_2^{\text{fog}}(t), ..., f_N^{\text{fog}}(t) \}$$
 (14)

奖励  $\Re$ : 在时隙 t 执行动作  $a_t$  得到的即时奖励为  $r_t$ , 一般情况下,可以将目标函数作为即时奖励函数,但是,智能体的目标是最大化奖励,而本文优化目标是最小化所有终端任务完成的总成本,因此,本文将即时奖励函数定义为目标函数取负值,即

其中 C(t) 表示在时隙 t 所有终端任务完成的总成本,  $\kappa(\kappa < 0)$  是一个常数,表示在不满足(11a)-(11d)的约束条件时,具体环境给出的奖励值.

## 5.2 3CC-SCO算法

3CC-SCO 算法的具体架构如图 2 所示,它有类似于 Double DQN (DDQN)的双神经网络结构,亦有行动者(Actor)和评论家(Critic)网络,因此其结构由 Actor 当前网络, Actor 目标网络, Critic 当前网络和 Critic 目标网络四个网络组成. Actor 网络输出的是动作策略,而 Critic 网络输出动作-状态值函数评估动作的好坏. DDQN 双网络参数定期完全复制,而本算法属于离线策略,目标网络参数是当前网络参数的"软更新",使得学习过程更加稳定.

在时隙 t,雾节点收集来自所有终端用户的状态信息  $s_t$ ,基于该状态,Actor 当前网络给出相应的迁移策略  $\pi_{\theta}(s_t)$ ,由于本文构建的优化问题为混合整数非线性规划问题,即动作空间应为混合连续-离散动作空间,而 DDPG 算法受限于纯连续动作输出.为此,本文定义一个概率  $\eta$  表示当前行动者网络的连续迁移决策输出,并对该连续动作输出执行概率离散化运算.据此,可通过如下概率离散运算得到

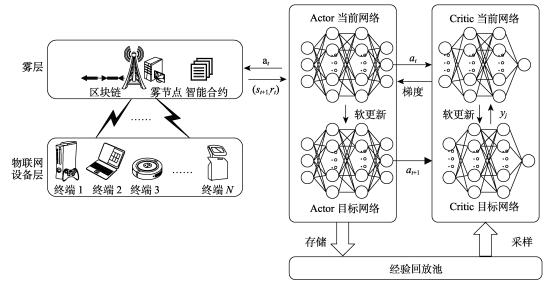


图 2 3CC-SCO 算法架构

离散动作  $a_t^1$ :

$$a_t^1 = \begin{cases} 0, \, \eta < 0.5, \\ 1, \, \eta \ge 0.5 \end{cases} \tag{16}$$

同时,为了增加连续动作搜索的随机性,会在原动作基础上加 OU 噪声  $\aleph_t$ ,即,最终资源分配策略表示为  $a_t^2 = \pi_\theta(s_t) + \varsigma \aleph_t$ ,其中  $\varsigma$  是噪声的退火因子,其值会随着迭代步数逐渐减小,这是为了保证神经网络训练的收敛性能,即随着网络训练的越来越好,减小噪声的扰动,保证网络的稳定性.

雾节点作为智能体执行动作  $a_t$  并与环境交互转移到下一状态  $s_{t+1}$  以及获得奖励  $r_t$  ,将  $(s_t, a_t, s_{t+1}, r_t)$  存入经验回放池. 从经验回放池中随机采样 M 个样本,然后 Actor 目标网络将根据下一状态  $s_{t+1}$  预测相应的下一动作  $a_{t+1}$  ,这里的 Actor 目标网络和当前网络结构完全相同,且其网络参数  $\theta'$  定期从 Actor 当前网络参数  $\theta$  复制.

Critic 目标网络计算第  $j \in \{1,2,...,M\}$ 个样本的价值函数  $Q'(s_i,a_i,\omega')$ ,则目标 Q 值为

$$y_{j} = r_{j} + \gamma Q'(s_{j+1}, a_{j+1}, \omega')$$
 (17)

其中 $\gamma$ 为折扣系数,由于 Critic 的两个网络结构也 完全相同,则 Critic 目标网络的网络参数 $\omega$ '也会定 期复制 Critic 当前网络的 $\omega$ . Critic 当前网络主要负责计算当前动作的价值函数 $Q(s_j,a_j,\omega)$ ,定义均方 差损失函数如下:

$$Loss(\omega) = \frac{1}{M} \sum_{i=1}^{M} (y_j - Q(s_j, a_j, \omega))^2$$
 (18)

通过梯度下降法最小化  $Loss(\omega)$  来更新神经网络的网络参数  $\omega$ ,更新定义如下:

$$\omega \leftarrow \omega + \alpha \nabla_{\omega} Loss(\omega) \tag{19}$$

其中 $\alpha$ 为 $\omega$ 的学习率.

Actor 当前网络根据 Critic 当前网络对其给出策略的反馈, 计算策略梯度, 并更新其网络参数  $\theta$ , 使其得到更大的反馈 Q 值, 策略梯度定义为

$$\nabla J(\theta) = \frac{1}{M} \sum_{i=1}^{M} \left[ \nabla_a Q(s_j, a_j, \omega)_{\mid a = \pi_{\theta}(s_j)} \nabla_{\theta} \pi_{\theta}(s)_{\mid s = s_j} \right] (20)$$

如果 Critic 反馈的梯度  $\nabla_a Q(s_t, a_t, \theta)$  不断鼓励 Actor 增大其输出的动作值,那么该动作值将逐渐趋 向甚至超出其边界值.为防止此类情况发生,本算 法需要对原梯度进行转换. 当动作值接近其取值范 围的边界值时,梯度减小;当动作值超过动作范围 时,则梯度为原来的反梯度,具体转换定义如下:

$$\nabla_a Q(s_t, a_t, \theta) =$$

$$\nabla_{a}Q(s_{t}, a_{t}, \theta) * \begin{cases} (a_{\max} - a_{t}) / (a_{\max} - a_{\min}), \\ \nabla_{a}Q(s_{t}, a_{t}, \theta) \ge 0 \\ (a_{t} - a_{\min}) / (a_{\max} - a_{\min}), \\ \nabla_{a}Q(s_{t}, a_{t}, \theta) < 0 \end{cases}$$
(21)

其中 $\nabla_a Q(s_t, a_t, \theta)$ 表示 Critic 对动作 $a_t$ 的梯度, $a_{\max}$ 和 $a_{\min}$ 分别为动作的上界和下界.

为了更好地理解,可将上式中的 $J(\theta)$ 看作Actor的损失,具体定义如下:

$$J(\theta) = -\frac{1}{M} \sum_{j=1}^{M} Q(s_j, a_j, \theta)$$
 (22)

由此可以看出,得到的反馈 Q 值越大,损失越小.利用梯度下降法更新参数  $\theta$ ,如下所示:

$$\theta \leftarrow \theta + \beta \, \nabla_{\theta} J(\theta) \tag{23}$$

其中 $\beta$ 为 $\theta$ 的学习率.

最后分别对 Actor 和 Critic 目标网络的网络参数进行更新,本算法采取"软更新"的方式,即

$$\theta' \leftarrow \sigma\theta + (1 - \sigma)\theta' \tag{24}$$

$$\omega' \leftarrow \sigma\omega + (1 - \sigma)\omega' \tag{25}$$

其中 $\sigma$ 是参数的更新速度,一般取非常小的值,表示网络参数缓慢地更新,可使训练过程更加稳定,

"软更新"便体现在这里.通过不断地学习,最终训练到合适的参数,雾节点能获得最优迁移策略.

为了更好地理解本文所提求解方法的思想与内涵,算法1将上述求解过程高度总结凝练如下: **算法1** 联合通信 计算与控制的智能计算迁移算法

**算法 1**. 联合通信、计算与控制的智能计算迁移算法输入: 任务大小 $l_i$ , $i \in \{1,2,...,N\}$ ;

所需要的计算资源  $d_i$ ,  $i \in \{1, 2, ..., N\}$ ;

时间成本单价  $\lambda_i^T$ ,  $i \in \{1, 2, ..., N\}$ ;

能耗成本单价  $\lambda_i^E, i \in \{1, 2, ..., N\}$ .

输出:最优的迁移策略  $(x_i^*, b_i^*, f_i^{\text{fog}^*})$  与最小总成本  $C^*$ .

- 1. BEGIN
- 2. 初始化网络参数 $\theta'=\theta$ ,  $\omega'=\omega$ , 清空经验回放池;
- 3. FOR i=0 to maximum episode Z DO
- 4. 初始化 OU 噪声和初始状态  $s_t$ ;
- 5. FOR t = 0 TO T DO
- 6. Actor 当前网络基于状态  $s_t$ 选择动作  $a_t$ ;
- 7. 执行  $a_t$ , 与环境交互得到  $s_{t+1}$ 和  $r_t$ ;
- 8. 将  $(s_t, a_t, s_{t+1}, r_t)$  存入经验回放池;
- 9. 状态更新  $s_t = s_{t+1}$ ;
- 10. 从经验回放池中取M个样本,基于公式(17)计算目标O值;
- 11. 基于公式(18)和(19)更新 Critic 网络参数ω;
- 12. 基于公式(20)和(23)更新 Actor 网络参数θ;
- 13. 基于公式(24)和(25)分别更新目标网络参数 $\theta$ 和 $\omega$ ;
- 14. END FOR
- 15. END FOR
- 16. 获得最优迁移策略  $(x_i^*, b_i^*, f_i^{\text{fog}^*})$  及最小总成本  $C^*$ .
- 17. END

# 6 仿真结果与分析

本节通过仿真实验评估 3CC-SCO 算法的有效性,并与几种典型的相关方案进行对比,验证本文计算迁移方案的性能优势.

在本仿真环境中,设定信道总带宽 B 为 10 Mb/s,终端 i 上传数据的传输功率为 0.1 W;本地终端计算能力  $f_i^{\text{local}}$  的取值在 10 到 80 Mcycles/s之间随机生成,本地终端的能耗系数  $k^{\text{local}}$  为  $10^{-29}$ ;雾节点的计算能力 F 为 5 Gcycles/s,雾节点计算的能耗系数  $k^{\text{fog}}$  为  $10^{-26}$ ,雾节点中 CPU 的收费单价 c

为 5 tokens/Gcycles. 假定物联网设备层有 4 个终端,终端 i 任务的大小  $l_i$  在 100 Kb 到 500 Kb 之间随机生成. 为了便于计算,将每个任务的最大容忍能耗和最大容忍延迟均分别设置为 0.002 J 和 0.2 s.

图 3(a)描绘的是 Critic 网络不同学习率对损失函数收敛性能的影响. 从该曲线图可以看出,学习率 $\alpha$ 越低,损失函数收敛的越慢,其中, $\alpha$  = 0.001和 $\alpha$  = 0.0001的收敛速度十分接近. 在迭代初期,智能体学习到的策略与理想值相差较大,则损失函数值也会比较大,随着迭代次数的增加,智能体逐步学习到最优策略并达到较稳定的状态,因此损失函数值就会变小,最终收敛.

图 3(b)呈现的是 Critic 网络不同学习率对奖励函数收敛性能的影响. 图中三种不同学习率的奖励函数值均在 200 次迭代内快速上升,并在之后趋于收敛. 当  $\alpha=0.001$  时,奖励函数的波动比较大,且很难收敛到最优值,而学习率越小奖励函数波动越小. 因此,Critic 网络的学习率过快,学习的波动性较大,可能导致奖励函数收敛不到最大值.

结合图 3(a)和 3(b)分析可知,虽然 Critic 网络的学习率越大,损失函数收敛的越快,但会导致奖励

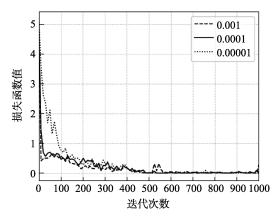


图3(a) Critic网络不同学习率的损失函数收敛情况

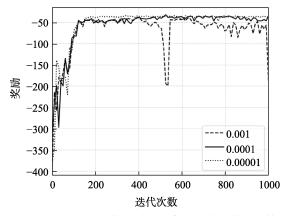


图3(b) Critic网络不同学习率的奖励函数收敛情况

函数波动较大,可能发生收敛不到最优值;同时, 学习率太小,亦会导致收敛速度达不到理想效果.综 上所述,在后续仿真实验中设置 Critic 网络的学习 率为 0.0001.

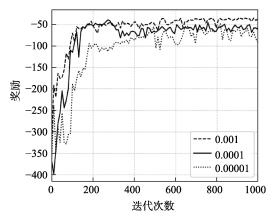


图4(a) Actor网络不同学习率的奖励函数收敛情况

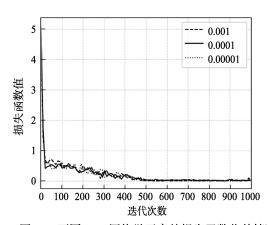


图4(b) 不同Actor网络学习率的损失函数收敛情况

接下来的仿真实验是通过将本文所提方案与其他三种相关方案进行对比分析,验证本文方案的性能优势. "Full-offloading"代表任务全部迁移到雾节点处理;"Full-local"代表所有任务均在本地终端处理;"DQN"代表文献[41]所提出的基于 DQN 的计算迁移方案;"3CC-SCO"代表本文所提方案. 如下仿真结果均为经过 50 次仿真实验取的平均值.

图 5 展示了四种方案的总成本与终端数量的关

系. 由图可知,四种方案的成本均随着任务数量的增多而上升,其中本文所提方案的总成本最低,"Full-offloading"的总成本最高. 在任务数量很少时(例如 $N \leq 2$ ),除了"Full-offloading",其它三种方案的总成本基本一致,因为任务都在本地处理的总成本最低,但随着任务数量增多,将部分任务迁移到雾节点处理会使总成本降低. 虽然任务迁移可以降低系统总成本,但持续的迁移也会导致通信成本、挖掘费用的持续上升. 因此本文通过联合优化计算与通信资源及迁移决策,使得总成本最低,而"DQN"虽然优化了迁移决策却没有考量计算与通信资源优化分配问题,其成本要高于本文方案.

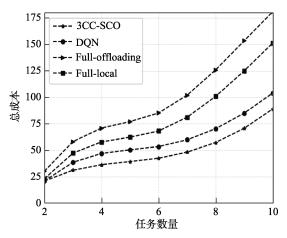


图 5 不同方案的总成本随任务数量变化的对比

图 6 描述了不同方案的总成本随雾节点计算能力变化的对比情况. 从图中可发现,除了"Full-local"之外,其他三种方案的总成本均随着雾节点计算能力的增大逐渐降低,这是因为分配给任务的计算资源越来越多,其时间成本会减小.同时,很明显本文方案的总成本比其他方案都要低,因为本文方案对雾节点计算资源进行了优化分配,使其总成本最低.再者,从图中曲线走势可以看出,随着

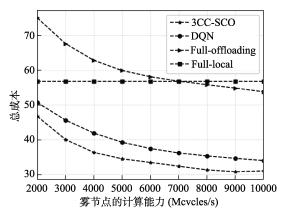


图 6 不同方案的总成本随雾节点计算能力变化的对比

雾节点计算能力的增大,总成本降低幅度逐渐趋于平缓,这是因为雾节点计算能力增大所带来的计算时间降低会逐渐被增加的通信开销和挖掘成本所抵消,导致总成本降低越趋于平缓.此外,雾节点计算能力的变化并不会对"Full-local"的总成本产生影响,因为该方案的总成本只与终端的计算能力有关.

任务数量的变化会引起雾节点计算资源分配产生变化,而矿工得到的算力占比也会随之改变,从而对挖掘成本产生影响,图 7 描绘了任务数量的变化对不同方案挖掘成本的影响情况. 当任务数量较少时,任务全部在本地处理,因此无需支付挖掘费用. 随着任务数量增多,需要迁移的任务也会增多,挖掘成本就会增加,尤其是将任务全部迁移到雾节点处理,挖掘成本最高. 同时,迁移任务数量增多也会使每个任务分得的奖励减少,因此总成本会增加的越来越快. 本文方案通过优化计算资源的分配,使矿工所得算力也得到优化,同时还有对通信资源和决策的优化控制,使其在三种方案中挖掘成本最低,验证了本文方案在节约挖掘成本方面的优越性.

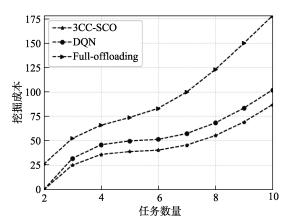


图 7 不同方案的挖掘成本随任务数量变化的对比

## 7 总 结

为进一步满足雾计算场景中计算密集型任务对时延与能耗的要求,本文提出了一种区块链赋能的联合资源分配与控制的智能计算迁移方案.通过对总成本中时延、能耗和挖掘成本的综合考量,构建了一个最小化所有任务完成总成本的优化问题.为促使终端和雾节点积极参与整个计算迁移过程,研究了基于区块链技术的激励机制,特别地,挖掘成本中奖励的分配规则保证了成功挖掘资源终端收获奖励的公平性.基于上述最小化总成本的优化目标,提出了3CC-SCO算法,为提升收敛效果并更加

适用于该类混合整数非线性规划问题的求解,设计了基于反梯度更新的双"行动者-评论家"神经网络结构和连续动作输出概率离散化运算.最后,经过大量仿真实验与分析,其结果表明,与其他三种基准方案相比,本文方案能以较快的速度收敛到最低总成本.在未来工作中,计划将本文所研究的智能计算迁移机制应用于智慧医疗的智能预诊网络系统上,以提升该网络系统的服务性能.

## 参考文献

- Hong Z, Chen W, Huang H, et al. Multi-hop cooperative computation offloading for industrial IoT-edge-cloud computing environments. IEEE Transactions on Parallel and Distributed Systems, 2019, 30(12): 2759-2774
- [2] Aazam M, Zeadally S, Harras K. Deploying fog computing in industrial Internet of Things and industry 4.0. IEEE Transactions on Industrial Informatics, 2018, 14(10): 4674-4682
- [3] Pan J and McElhannon J. Future edge cloud and edge computing for Internet of Things applications. IEEE Internet of Things Journal, 2018, 5(1): 439-449
- [4] Ma X, Wang S, Zhang S, et al. Cost-efficient resource provisioning for dynamic requests in cloud assisted mobile edge computing. IEEE Transactions on Cloud Computing, 2021, 9(3): 968-980
- [5] Chen S, Zhu X, Zhang H, et al. Efficient privacy preserving data collection and computation offloading for fog-assisted IoT. IEEE Transactions on Sustainable Computing, 2020, 5(4): 526-540
- [6] Li S, Tao Y, Qin X, et al. Energy-aware mobile edge computation offloading for IoT over heterogeneous networks. IEEE Access, 2019, 7: 13092-13105
- [7] Wang J, Feng D, Zhang S, et al. Computation offloading for mobile edge computing enabled vehicular networks. IEEE Access, 2019, 7: 62624-62632
- [8] Li L, Ota K, Dong M, et al. Deep learning for smart industry: efficient manufacture inspection system with fog computing. IEEE Transactions on Industrial Informatics, 2018, 14(10): 4665-4673
- [9] Yu S, Wang X, Langar R. Computation offloading for mobile edge computing: a deep learning approach//Proceedings of the IEEE International Symposium on Personal, Indoor and Mobile Radio Communications. Montreal, Canada, 2017: 1-6
- [10] Qi Q, Wang J, Ma Z, et al. Knowledge-driven service offloading decision for vehicular edge computing: a deep reinforcement learning approach. IEEE Transactions on Vehicular Technology, 2019, 68(5): 4192-4203
- [11] Wang S, Guo Y, Zhang N, et al. Delay-aware microservice coordination in mobile edge computing: a reinforcement learning approach. IEEE Transactions on Mobile Computing, 2021, 20(3): 939-951
- [12] Shen S, Han Y, Wang X, et al. Computation offloading with multiple agents in edge-computing–supported IoT. ACM Transactions on Sensor Networks, 2019, 16(1):1-27

- [13] Liu X, Yu J, Wang J, et al. Resource allocation with edge computing in IoT networks via machine learning. IEEE Internet of Things Journal, 2020, 7(4): 3415-3426
- [14] Chen X, Zhang H, Wu C, et al. Optimized computation offloading performance in virtual edge computing systems via deep reinforcement learning. IEEE Internet of Things Journal, 2019, 6(3): 4005-4018
- [15] Li J, Gao H, Lv T, et al. Deep reinforcement learning based computation offloading and resource allocation for MEC// Proceedings of the IEEE Wireless Communications and Networking Conference. Barcelona, Spain, 2018:1-6
- [16] Li Y, Qi F, Wang Z, et al. Distributed edge computing offloading algorithm based on deep reinforcement learning. IEEE Access, 2020, 8: 85204-85215
- [17] Chen S, Zheng Y, Lu W, et al. Energy-optimal dynamic computation offloading for industrial IoT in fog computing. IEEE Transactions on Green Communications and Networking, 2020, 4(2): 566-576
- [18] Guo K, Sheng M, Quek T, et al. Task offloading and scheduling in fog RAN: a parallel communication and computation perspective. IEEE Wireless Communications Letters, 2020, 9(2): 215-218
- [19] Mukherjee M, Kumar S, Shojafar M, et al. Joint task offloading and resource allocation for delay-sensitive fog networks// Proceedings of the IEEE International Conference on Communications. Shanghai, China, 2019: 1-7
- [20] Mukherjee M, Kumar S, Mavromoustakis C, et al. Latency-driven parallel task data offloading in fog computing networks for industrial applications. IEEE Transactions on Industrial Informatics, 2020, 16(9): 6050-6058
- [21] Sun H, Yu H, Fan G, et al. Energy and time efficient task offloading and resource allocation on the generic IoT-fog-cloud architecture. Peer-to-Peer Networking and Applications, 2020, 13(2): 548-563
- [22] Cai P, Yang F, Wang J, et al, JOTE: joint offloading of tasks and energy in fog-enabled IoT network. IEEE Internet of Things Journal, 2020, 7(4): 3067-3082
- [23] Min M, Xiao L, Chen Y, et al. Learning-based computation offloading for IoT devices with energy harvesting. IEEE Transactions on Vehicular Technology, 2019, 68(2): 1930-1941
- [24] Xu X, Li D, Dai Z, et al. A Heuristic offloading method for deep learning edge services in 5G networks. IEEE Access, 2019, 7: 67734-67744
- [25] Yang B, Cao X, Li X, Zhang, et al. Mobile-edge-computing-based hierarchical machine learning tasks distribution for IIoT. IEEE Internet of Things Journal, 2020, 7(3): 2169-2180
- [26] Zhu X, Chen S, Chen S, et al. Energy and delay co-aware computation offloading with deep learning in fog computing networks//Proceedings of the International Performance Computing and Communications Conference. London, UK, 2019: 1-6
- [27] Ali Z, Jiao L, Baker T, et al. A deep learning approach for energy efficient computational offloading in mobile edge computing. IEEE Access, 2019, 7: 149623-149633

- [28] Gong Y, Lv C, Cao S, et al. Deep learning-based computation offloading with energy and performance optimization. EURASIP Journal on Wireless Communications and Networking, 2020, 2020(1): 1-8
- [29] Chen S, Chen J, Zhao C. Deep reinforcement learning based cloud-edge collaborative computation offloading mechanism. Acta Electronica Sinica, 2021, 49 (1): 157-166 (in Chinese) (陈思光,陈佳民,赵传信.基于深度强化学习的云边协同计算迁移研究.电子学报, 2021, 49 (1): 157-166)
- [30] Wei Y, Yu F, Song M, et al. Joint optimization of caching, computing, and radio resources for fog-enabled IoT using natural actor–critic deep reinforcement learning. IEEE Internet of Things Journal, 2019, 6(2): 2061-2073
- [31] Chen X, Zhang H, Wu C, et al. Performance optimization in mobile-edge computing via deep reinforcement learning// Proceedings of the IEEE Vehicular Technology Conference. Chicago, USA, 2018: 27-30
- [32] Huang L, Bi S, Zhang Y. Deep reinforcement learning for online computation offloading in wireless powered mobile-edge computing networks. IEEE Transactions on Mobile Computing, DOI: 10.1109/TMC.2019.2928811, 2019
- [33] Feng J, Yu F, Pei Q, et al. Cooperative computation offloading and resource allocation for blockchain-enabled mobile-edge computing: a deep reinforcement learning approach. IEEE Internet of Things Journal, 2020, 7(7): 6214 6228
- [34] Guo F, Yu F, Zhang H, et al. Adaptive resource allocation in future wireless networks with blockchain and mobile edge computing. IEEE Transactions on Wireless Communications, 2020, 19(3): 1689-1703
- [35] Zhang R, Yu F, Liu J, et al. Deep reinforcement learning (DRL)-based device-to-device (D2D) caching with blockchain and mobile edge computing. IEEE Transactions on Wireless Communications, 2020, 19(10): 6469-6485
- [36] Qiu X, Liu L, Chen W, et al. Online deep reinforcement learning for computation offloading in blockchain-empowered mobile edge computing. IEEE Transactions on Vehicular Technology, 2019, 68(8): 8050-8062
- [37] Dinh C, Pubudu N, Ming D, et al. Privacy-preserved task offloading in mobile blockchain with deep reinforcement learning. IEEE Transactions on Network and Service Management, 2020, 17(4): 2536-2549
- [38] Li Z, Yang Z, Xie S, et al. Credit-based payments for fast computing resource trading in edge-assisted Internet of Things. IEEE Internet of Things Journal, 2019, 6(4): 6606-6617
- [39] Houy N. The bitcoin mining game. Ledger, 2016, 1: 53-68
- [40] Zhan W, Luo C, Wang J, et al. Deep-reinforcement-learning-based offloading scheduling for vehicular edge computing. IEEE Internet of Things Journal, 2020, 7(6): 5449-5465
- [41] Yang L, Li M, Si P, et al. Energy-efficient resource allocation for blockchain-enabled industrial internet of things with deep reinforcement learning. IEEE Internet of Things Journal, 2021, 8(4):2318-2329



**CHEN Si-Guang**, Ph.D., associate professor. His main research areas include fog/edge computing and AIoT.

**WANG Qian**, Ph.D. candidate. Her main research areas include fog/edge computing and AIoT. **ZHANG Hai-Jun**, Ph.D., professor. His main research areas include 6G mobile communication, artificial intelligence and big data.

**WANG Kun**, Ph.D., professor. His main research areas include edge computing, blockchain and distributed machine learning.

#### **Background**

The research topic of this paper focuses on the computation offloading problem in fog computation scenario. This topic becomes a research hotspot for which can effectively improve users' service experience and system overhead with the help of the edge node resources. Currently, many related schemes were proposed and made outstanding contributions to reduce delay and energy consumption by optimizing offloading decision and resource allocation. However, the developed schemes are difficult to adapt the solution of complex mixed integer programming problem, and all of them are based on an ideal network scenario, i.e., the fog node volunteered to provide resources to external terminal free of charge.

According to above challenges, in this paper, we propose a resource allocation and control co-aware smart computation offloading scheme. First, we formulate a minimization problem of the total cost of all tasks, in which a blockchain-based mining cost is integrated into motivate the participation enthusiasm of terminals and the fog in computation offloading. Furthermore, based on the joint design of inverting gradient update based double actor-critic neural network structure and probabilistic discrete of continuous action output, we propose a 3CC-SCO algorithm

to address above minimization problem and achieve better convergence effect and adaptability.

This work is partially supported by the National Natural Science Foundation of China (Nos. 61971235, 61771258), the Six Talented Eminence Foundation of Jiangsu Province (No. XYDXXJS-044), the 333 High-level Talents Training Project of Jiangsu Province, the 1311 Talents Plan of NUPT, the China Postdoctoral Science Foundation (No. 2018M630590), the Jiangsu Planned Projects for Postdoctoral Research Funds (No. 2021K501C), and the CERNET Innovation Project (No. NGII20190702). These projects aim to alleviate the mismatching pressure on increasing rate between data scale and storage capacity, provide secure, high quality and low latency data services with minimum energy consumption, and achieve smart computation offloading decision for improving the adaptability of the whole big data network. Several papers have been published on respectable journals, such as IEEE Internet of Things Journal, Engineering, IEEE Transactions on Green Communications and Networking, IEEE Transactions on Sustainable Computing, Applied Mathematical Modelling, Journal of Network and Computer Applications, Computer Networks and Acta Electronica Sinica.