

面向绿色计算的车辆协同任务卸载方法

张红霞* 吕智豪 席诗语 刘佳敏 郭加树 张培颖

(中国石油大学(华东)青岛软件学院 青岛 266000)
(中国石油大学(华东)计算机科学与技术学院 青岛 266000)

摘要: 车辆边缘计算(VEC)为处理计算密集、延迟敏感型任务提供了新的范式,然而边缘服务器在整合可再生能源方面的能力较差。因此,为了提高边缘服务器的能效,该文设计了一种面向绿色计算的车辆协同任务卸载框架。在该框架中,车辆配备能源收集(EH)设备,通过彼此间共享绿色能源和计算资源协作执行任务。为有效促进车辆的参与积极性,该文通过动态定价激励车辆,并综合考虑了车辆的移动性、任务优先级等。为了使卸载决策适应动态环境的变化,该文提出了一种基于双延迟深度确定性策略梯度(TD3)的任务卸载方法,以在最大化所有车辆平均任务完成效用的同时减少边缘端电网电力的使用。最后,仿真结果验证了该方法的有效性,相比基于深度确定性策略梯度(DDPG)和基于贪心原则(GPE)的方法在性能上分别提升了7.34%和37.47%。

关键词: 车辆边缘计算; 任务卸载; 能源收集; 车辆协同; 动态定价

中图分类号: TN919

文献标识码: A

文章编号: 1009-5896(2024)01-0175-09

DOI: 10.11999/JEIT230051

A Method for Offloading Vehicle Collaborative Tasks for Green Computing

ZHANG Hongxia LÜ Zhihao XI Shiyu LIU Jiamin

GUO Jiashu ZHANG Peiying

(Qingdao Institute of Software, China University of Petroleum (East China), Qingdao 266000, China)

(College of Computer Science and Technology, China University of Petroleum (East China), Qingdao 266000, China)

Abstract: Vehicular Edge Computing (VEC) has become a promising and prospective paradigm for computation-intensive and delay-sensitive tasks. However, edge servers are less capable of integrating renewable energy. Therefore, in order to improve the energy efficiency of edge servers, a green computing oriented vehicle collaborative task offloading framework is proposed. In this framework, vehicles equipped with Energy Harvest (EH) devices cooperate to perform tasks by sharing green energy and computing resources with each other. To effectively enhance the participation enthusiasm of vehicles, dynamic pricing is adopted to motivate vehicles, and the mobility and task priority are also considered comprehensively. In order to adapt the offloading decisions to the dynamic environment, a Twin Delayed Deep Deterministic policy gradient (TD3) based task offloading method is proposed to maximize the average task completion utility of all vehicles while reducing the use of grid power. Finally, simulation results verify the effectiveness of the proposed method, and the performance achieves 7.34% and 37.47% improvement respectively compared with Deep Deterministic Policy Gradient (DDPG) based method and Greedy Principle Execution (GPE) method.

Key words: Vehicular Edge Computing (VEC); Task offloading; Energy Harvest(EH); Vehicle collaboration; Dynamic pricing

1 引言

随着智能交通的深入研究和快速发展,增强现

实、自动辅助驾驶等新兴应用不断被提出。为了支持这些应用,通常需要在车辆端快速处理复杂任务。随着技术的不断成熟,新型车辆将拥有强大的计算能力,同时新能源汽车将通过配备能源收集(Energy Harvest, EH)设备收集环境中的太阳能等绿色能源,促进车载设备的自我可持续性和不间断运行,实现绿色节能的目的^[1]。然而绿色能源在时间和地域上的不确定性使得车辆无法完全依赖,结

收稿日期: 2023-02-14; 改回日期: 2023-05-12; 网络出版: 2023-05-22

*通信作者: 张红霞 zhanghx@upc.edu.cn

基金项目: 山东省自然科学基金(ZR2020MF006, ZR2022LZH015)

Foundation Items: The Natural Science Foundation of Shandong Province (ZR2020MF006, ZR2022LZH015)

合车联网和移动边缘计算优点的车辆边缘计算(Vehicular Edge Computing, VEC)为该问题的解决提供了新的范式,通过将计算资源下沉到车辆附近,在不显著增加网络传输时延的同时实现车辆能耗的有效降低^[2]。

边缘服务器的能效是实现VEC可持续计算的瓶颈,边缘服务器通常位于难以吸纳可再生能源的城市与人群中心,使用化石燃料燃烧产生的电网电力^[3],仅依赖边缘服务器无法充分挖掘绿色能源的潜力。尽管车辆协同任务卸载对节约能源成本有帮助,但从经济角度看车辆没有义务在无任何补偿的情况下作为服务提供者执行其他车辆的任务。考虑到自身资源消耗和出于安全的目的,可以预期车辆不愿意在没有任何激励的情况下贡献其空闲资源^[4]。

不同车辆会根据其乘客偏好生成特定类型的任务请求。例如,对于乘客偏好路线规划的车辆会频繁地请求交通状况信息,对于乘客偏好舒适交通体验的车辆会频繁地请求交互式信息^[5]。由于任务类型不同,其执行的必要程度也不同。然而,现有工作很少考虑到不同类型的任务给车辆带来的体验差异,所有任务都有相同的概率得到所需资源。当车辆资源不足时,可能导致一些关键任务的卸载失败,造成严重后果。此外考虑到任务请求在时空上的动态性,在追求即时性能时不能忽略长期性能。然而车辆的高速移动性、时变网络环境下的复杂资源分配对确保系统的长期性能提出了挑战。

为了解决上述问题,促进VEC系统的可持续发展,本文提出了面向绿色计算的车辆协同任务卸载方法,主要贡献包括3个方面:

(1)本文设计一种“绿色能源-电网”混合能源供应模式下的车辆协同任务卸载框架。在该框架下,车辆之间使用车对车(Vehicle-to-Vehicle, V2V)技术共享资源,节省边缘服务器能源成本,并根据任务类型分别采用不同的效用函数,保证了高优先级任务的优先执行。

(2)本文引入一种动态定价方案,动态调整价格促进用户和服务提供者之间的合作,缓解了任务车辆和服务车辆之间存在的利益冲突,激励车辆按需共享其空闲资源,提高资源利用效率。

(3)本文提出一种基于双延迟深度确定性策略梯度(Twin Delayed Deep Deterministic policy gradient, TD3)^[6]的在线任务卸载方法。实验结果表明,本文所提方法在性能上相较于深度确定性策略梯度(Deep Deterministic Policy Gradient, DDPG)和基于贪心原则(Greedy Principle Execution, GPE)的方法分别提升了7.34%和37.47%。

2 相关工作

2.1 基于VEC的任务卸载方法研究

考虑到车辆拥有一定的闲置计算资源时本身可以被看作边缘计算节点,针对车辆可能在收到任务处理结果前驶出路侧单元覆盖范围的问题,文献^[7]构造了一个基于车辆卸载决策的博弈以最小化计算开销。文献^[8]制定了一个具有通信、计算、缓存和协作功能的统一框架,车辆可以通过V2V通信进行数据中继和计算的相互协作,并开发了一种调度方案以最小化系统范围内的数据处理成本。但是能效限制仍然是制约VEC的一个关键问题,上述工作并没有充分考虑边缘服务器的能效。

2.2 面向绿色计算的任务卸载方法研究

大多数设备的能源有限,因此提供自主的能源至关重要,研究人员通过实施EH技术作为电池的可行且经济实用的替代方案。文献^[9]针对具有EH模块的物联网设备提出了一种基于强化学习的卸载方法,根据当前电池电量、先前对每个边缘设备的无线电传输速率以及预测捕获的能量量来选择边缘设备和卸载速率。文献^[10]研究了多层次边缘计算系统中的联合任务卸载和能源调度问题,通过调用李雅普诺夫技术将长期优化问题分解为一系列仅使用当前系统信息的单时隙优化问题。文献^[11]提出了一种近邻感知的分布式任务卸载方法,其中物联网设备兼顾考虑其能源状态和近邻设备的决策,解决了大量移动设备同时将任务卸载到边缘云上时任务无法在预期时间内完成的问题。然而现有的工作大多是在物联网场景下,不适用于本文提出的动态异构VEC环境下车辆间通过共享绿色能源和计算资源协作执行任务的情况。

2.3 VEC中车辆协同激励机制的设计

目前车辆协同激励机制常见的设计思路大都受经济学中契约理论、拍卖理论等的启发。文献^[12]利用契约理论使得路侧单元能够根据资源共享车辆的贡献和独特特性为其提供量身定制的合同,从而获得最适宜的奖励。文献^[13]为鼓励车辆共享资源制定了一个反向拍卖机制,并开发了一种基于单边匹配的方法,利用整数线性规划提供了具有个体理性以及匹配稳定性的解。然而,这些工作大都基于一个假设,即所有参与者都会公开自己的私人信息,然后依据相关指标对车辆类型进行划分。计算资源分配策略应该实时响应环境变化,上述工作因很难及时获得完整的系统模型和环境动态并做出反应而不太适用。而动态定价基于平衡计算需求和资源的关系,能使价格随着资源的供需关系动态调整,展现出强大的激励潜力。文献^[14]通过动态定价,由

运营商租赁服务车辆实现了运营商收益最大化。文献[15]研究了在多个服务提供者存在竞争的情况下如何动态定价。

3 系统模型

3.1 系统概览

如图1所示，本文提出一个面向绿色计算的车辆协同任务卸载框架(Green Computing Oriented Vehicle Collaborative task offloading framework, GCOVC)。在所提框架中，考虑一个城市双向直行车道上，1个配备边缘服务器的基站和多个车辆组成的网络。每辆车都配有1个太阳能电池板，从环境中收集太阳能储存在电池中作为提供持续能源的唯一来源。绿色能源可用于任务的本地执行，帮助其他车辆，以及任务卸载产生的通信开销，且不计入系统内的能耗成本。基站的通信距离远大于车载通信距离，鉴于基站具有全局信息，其被信任为所有车辆做出全面的卸载决策。如果车辆进入基站的覆盖区域后愿意参与协作，它将持续向基站发送包含其位置、速度、可用绿色能源量、计算能力等消息，基站可由此获知其覆盖范围内的道路交通情况。系统内存在两类角色：一类是任务车辆，即产生任务请求的车辆；另一类是服务车辆，即拥有大量绿色能源，可以对外提供服务以获取报酬的车辆。本文将时间离散为多个时隙，在每个时隙，任务车辆首先向基站发送请求，然后基站决定任务分配，并将分配消息发送回任务车辆和所选服务车辆。假设在某一时段期间，基站的通信范围内有 M 辆车，集合表征为 M ，车辆的计算能力表示为 $F_m, m \in M$ 。每一时隙每个车辆都会产生一个任务请求，任务车辆 m 在时隙 l 产生的任务 k_m^l 可以用4元

组 $\{D_m^l, C_m^l, \alpha_m^l, \beta_m^l\}$ 表示，其中 D_m^l 表示任务大小， C_m^l 表示任务所需的CPU周期， α_m^l 表示任务延迟容忍， β_m^l 表示任务类型。使用 $a_{k_m^l, d}, d \in \{0, 1, \dots, M\}$ 表示任务车辆 m 的3类执行方式，其中 $a_{k_m^l, d} = 1, d = 0$ 表示任务被卸载到基站， $a_{k_m^l, d} = 1, d = m$ 表示任务在本地执行， $a_{k_m^l, d} = 1, d = n, n \neq m$ 表示任务被卸载到服务车辆 n 。

3.2 能源队列模型

由于自然界中绿色能源的可捕获量是随机和突发的，为了体现绿色能源收集过程的不确定性，假定车辆 m 收集到的绿色能源服从最大值为 R_m 的均匀分布，并且每个时隙是独立同分布的。此外，EH模块的充电和放电过程是可以同时进行的，所收集的能源被缓存在一个存储队列中，定义 Q_m^l 为时隙 l 时车辆 m 的能源存储队列状态， r_m^l 表示成功收集到的绿色能源， E_m^l 表示时隙内所消耗的绿色能源。基于上述定义和假设，车辆 m 的EH模块的动态变化为

$$Q_m^{l+1} = \max \{Q_m^l + r_m^l - E_m^l, 0\} \quad (1)$$

3.3 通信模型

车辆通信方式可分为两种：当整个系统内绿色能源充足时，为了节约能源成本，绿色能源供应不足的任务车辆可优先通过V2V方式将任务卸载到供应充足的相邻车辆上执行；但当由于天气原因使环境中可收集的绿色能源减少，整个系统内可用绿色能源无法满足所有任务要求时，车辆可采用车对基础设施(Vehicle-to-Infrastructure, V2I)方式通过蜂窝链接将任务卸载到基站的边缘服务器上执行。

3.3.1 V2I通信

以道路为 x 轴，道路到基站的垂直连线为 y 轴，

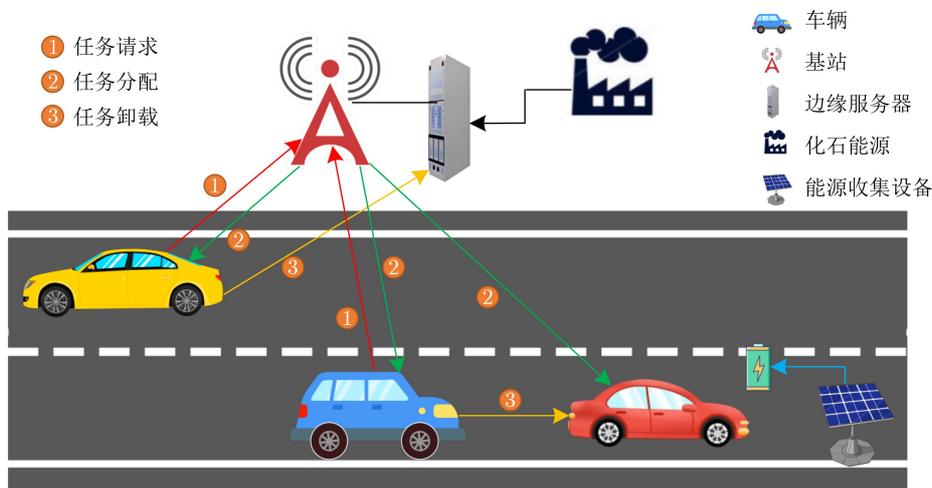


图1 GCOVC架构

建立2维坐标系。假设基站位于 $(0, h)$ 处, h 表示基站与道路的垂直距离, 车辆在道路上匀速行驶, 车辆 m 的移动模式可由一个2元组 $\{(x_m, 0), \mathbf{v}_m\}$ 表示, 其中 $(x_m, 0)$ 表示车辆 m 的初始位置, $|x_m|$ 表示车辆与基站的横向距离, 矢量 \mathbf{v}_m 表示车辆的速度。假设无线信道状态在任务的数据传输期间保持静态, 上行链路的数据传输速率为

$$V_{m,0} = \omega \log_2 (1 + P_m g_{m,0} d_{m,0}^{-\alpha} / N_0) \quad (2)$$

其中, ω 为分配的传输信道带宽, P_m 为任务车辆 m 的传输功率, $g_{m,0}$ 为参考距离处的信道功率增益, α 为路径损耗指数, N_0 为加性高斯白噪声, $d_{m,0}$ 为任务车辆 m 与基站之间的距离, 定义为

$$d_{m,0} = \sqrt{h^2 + (x_m + \mathbf{v}_m l)^2} \quad (3)$$

则任务车辆 m 与基站间的通信时延为

$$t_{m,0}^{l,\text{trans}} = D_m^l / V_{m,0} \quad (4)$$

3.3.2 V2V通信

假设V2V通信采用正交频率, 忽略由其他V2V传输链接引入的干扰, 从任务车辆 m 到服务车辆 n 的数据传输速率为

$$V_{m,n} = \omega \log_2 (1 + P_m g_{m,n} d_{m,n}^{-\alpha} / N_0) \quad (5)$$

其中, $d_{m,n}$ 为任务车辆 m 与服务车辆 n 之间的距离, 定义为

$$d_{m,n} = \sqrt{(x_m - x_n)^2 + [(\mathbf{v}_m - \mathbf{v}_n) l]^2} \quad (6)$$

由于车辆传输范围有限且移动性强导致传输链路不稳定, 定义链路失效时间为

$$\rho_{m,n} = [R - (x_m - x_n) \text{sign}(\mathbf{v}_m - \mathbf{v}_n)] / |\mathbf{v}_m - \mathbf{v}_n| \quad (7)$$

其中, R 表示固定传输功率下的V2V传输范围, $\text{sign}(\cdot)$ 是符号函数, 当 $\cdot > 0$ 时, $\text{sign}(\cdot) = 1$, 表示两个车辆彼此间正在远离。则任务车辆 m 与服务车辆 n 间的通信时延为

$$t_{m,n}^{l,\text{trans}} = D_m^l / V_{m,n} \quad (8)$$

3.4 计算模型

若车辆 m 选择本地执行, 任务的计算时延仅取决于本地处理器分配给任务的频率 $F_m^{k_l}$, 总时延为

$$t_{m,m}^l = t_{m,m}^{l,\text{comp}} = D_m^l C_m^l / F_m^{k_l} \quad (9)$$

若车辆 m 选择将任务卸载到基站, 假设基站有无限大的计算能力同时服务多个车辆, 并为车辆提供固定大小为 F_0 的计算能力, 则任务在边缘服务器上的计算时延为

$$t_{m,0}^{l,\text{comp}} = D_m^l C_m^l / F_0 \quad (10)$$

由于结果通常很小, 本文忽略反馈时延, 仅考虑传输和计算过程, 因此选择卸载到基站的总时延为

$$t_{m,0}^l = t_{m,0}^{l,\text{trans}} + t_{m,0}^{l,\text{comp}} \quad (11)$$

若车辆 m 选择将任务卸载到服务车辆 n , 用 $F_n^{k_l}$ 表示服务车辆分配给任务的频率, 计算时延为

$$t_{m,n}^{l,\text{comp}} = D_m^l C_m^l / F_n^{k_l} \quad (12)$$

因此选择卸载到服务车辆 n 的总时延为

$$t_{m,n}^l = t_{m,n}^{l,\text{trans}} + t_{m,n}^{l,\text{comp}} \quad (13)$$

综上, 任务的完成时延可统一表示为

$$t_m^l = \sum_{d=0}^M a_{k_m^l, d} t_{m,d}^l \quad (14)$$

3.5 能耗模型

若车辆 m 选择本地执行, 车辆端每个CPU周期的能耗为 $\kappa (F_m^{k_l})^2$, 其中, κ 受芯片架构影响, 则本地执行任务消耗的绿色能源为

$$E_m^l = \kappa (F_m^{k_l})^2 D_m^l C_m^l \quad (15)$$

若车辆 m 选择将任务卸载到基站, 任务车辆为传输任务消耗的绿色能源为

$$E_m^l = t_{m,0}^{l,\text{trans}} P_m \quad (16)$$

假设边缘服务器处理一个CPU周期的能耗为 u , 则服务器处理任务消耗的电网电力为 $u D_m^l C_m^l$ 。

若车辆 m 选择将任务卸载到服务车辆 n , 任务车辆为传输任务消耗的绿色能源为

$$E_m^l = t_{m,n}^{l,\text{trans}} P_m \quad (17)$$

而服务车辆为执行任务消耗的绿色能源为

$$E_n^l = \kappa (F_n^{k_l})^2 D_m^l C_m^l \quad (18)$$

3.6 任务效用模型

一般来说, 车辆任务可以分为高优先级和低优先级两类。高优先级任务是指一类与安全密切相关, 具有严格延迟约束的任务, 如导航、路况感知等。低优先级任务是一类容忍延迟的任务, 如车载娱乐服务等。如果高优先级任务在截止时间前完成, 则任务的收益为非负值, 并取决于完成时间。如果完成时间超过截止时间而失败, 则收益为负常值作为严厉惩罚。因此高优先级任务的收益函数定义为

$$U_{k_m^l}^H = \begin{cases} \ln(1 + \alpha_m^l - t_m^l), & t_m^l \leq \alpha_m^l \\ -C^H, & t_m^l > \alpha_m^l \end{cases} \quad (19)$$

如果低优先级任务在截止时间前完成，则任务的收益为正常值。如果完成时间超过截止时间，结果仍然被认为是可用的，但收益随着超过截止期限时间的增加而下降。因此低优先级任务的收益函数定义为

$$U_{k_m^l}^L = \begin{cases} C^L, & t_m^l \leq \alpha_m^l \\ C^L e^{-(t_m^l - \alpha_m^l)}, & t_m^l > \alpha_m^l \end{cases} \quad (20)$$

任务的完成效用定义为任务收益与付出成本的差值，表示为

$$\left. \begin{aligned} U_{k_m^l} &= \mathbb{I}(\beta_m^l = \beta^H) U_{k_m^l}^H + \mathbb{I}(\beta_m^l = \beta^L) \\ &U_{k_m^l}^L - \mathbb{I}(a_{k_m^l,0} \neq 1) p_{m,n}^l \\ &F_n^{k_m^l} t_{m,n}^{l,\text{comp}} - \lambda \mathbb{I}(a_{k_m^l,0} = 1) u D_m^l C_m^l \end{aligned} \right\} \quad (21)$$

其中， $\mathbb{I}(\cdot)$ 是指示函数，当 \cdot 为真时， $\mathbb{I}(\cdot) = 1$ ， β^H 和 β^L 分别表示高、低优先级任务， $p_{m,n}^l$ 表示任务车辆 m 为卸载任务到服务车辆 n 支付的价格，当任务在本地执行时， $p_{m,m}^l = 0$ ， λ 为系统能耗的重要性权重。

3.7 定价模型

服务车辆要同时执行本地任务和任务车辆的卸载任务。考虑到有限的计算能力，计算资源的分配首先要保证本地任务优先完成。考虑服务车辆 n 有本地任务 k_n^l 表征为 $\{D_n^l, C_n^l, \alpha_n^l, \beta_n^l\}$ ，则本地任务所需的最低频率 $F_n^{\min} = D_n^l C_n^l / \alpha_n^l$ 。如果车辆拒绝所有来自其他车辆的卸载任务，将所有计算资源分配给本地任务，则有 $F_n^{\max} = F_n$ 。如果车辆决定接受来自车辆 m 的卸载请求并将频率 $F_n^{k_m^l}$ 分配给卸载任务，本地任务的效用将下降，卸载任务的服务价格应补偿本地任务效用的损失。为任务支付的价格应满足

$$p_{m,n}^l F_n^{k_m^l} t_{m,n}^{l,\text{comp}} = U_{k_n^l} \Big|_{F_n^{k_n^l} = F_n} - U_{k_n^l} \Big|_{F_n^{k_n^l} = F_n - F_n^{k_m^l}} \quad (22)$$

从式(22)可以看出，如果任务车辆想要使用服务车辆更多的计算资源，则应向其支付更高的价格。

4 问题描述

本文的目标是在基站的控制辅助下，调整相应的任务执行策略以最大化一段时间内的社会福利，即在最大化所有车辆平均任务完成效用的同时高效

利用绿色能源，减少电网电力的使用。在每个时隙，基站要确定任务执行方式和卸载到服务车辆时对应的价格。最优化问题由式(23)给出

$$\left. \begin{aligned} &\max_{a_{k_m^l,d}, p_{m,n}^l} \sum_{l=0}^{l_{\max}} \frac{1}{M} \sum_{m=1}^M U_{k_m^l} \\ \text{s.t. C1: } &0 < p_{m,n}^l \\ &\leq \left(U_{k_n^l} \Big|_{F_n^{k_n^l} = F_n^{\max}} - U_{k_n^l} \Big|_{F_n^{k_n^l} = F_n^{\min}} \right) \\ &/[(F_n - F_n^{\min}) t_{m,n}^{l,\text{comp}}] \\ \text{C2: } &a_{k_m^l,d} \in \{0, 1\} \\ \text{C3: } &\sum_{d=0}^M a_{k_m^l,d} = 1 \\ \text{C4: } &\mathbb{I}(a_{k_m^l,n} = 1) \rho_{m,n} \geq t_m^l \\ \text{C5: } &\sum_{m=1}^M F_n^{k_m^l} \leq F_n \end{aligned} \right\} \quad (23)$$

在上述问题中，约束条件C1保证支付价格为正且不超过最大值，约束C2表示任务采用整体卸载的形式，约束C3确保任务只能在本地端、边缘端、服务车辆端中采取一种执行方式，约束C4表示为任务选择的服务车辆必须可用，约束C5表示服务车辆提供的计算资源不超过自身限制。

5 解决方案

TD3算法克服了DDPG算法存在的超参数和其他微调鲁棒性不足的缺陷。在处理最优决策问题时，TD3无需知道完整定义的系统模型即可通过与环境的交互感知环境变换，从而实现从环境特征到策略的映射，更适用于优化动态变换环境下的实时控制决策问题，可对任务卸载过程中计算资源的分配进行细粒度的频率控制。基于以上考虑，本文基于TD3设计了车辆协同任务卸载方法。

5.1 马尔可夫决策过程

由于系统动态特性对于基站是未知的，在每一时隙基站中的智能体观测基站覆盖范围内的车辆环境，并收集环境参数作为观测状态。时隙 l 的状态向量表示为： $\mathbf{s}^l = \{Q_1^l, \dots, Q_M^l, D_1^l, C_1^l, \alpha_1^l, \beta_1^l, \dots, D_M^l, C_M^l, \alpha_M^l, \beta_M^l\}$ 。

智能体依据策略进行决策，时隙 l 产生的动作向量表示为： $\mathbf{a}^l = \{a_1^l, a_2^l, \dots, a_M^l\}$ ，其中 a_m^l 是一个确定的实数，用于决定任务执行方式及服务车辆应向任务车辆提供的计算资源。注意，由式(22)可知，给定可获得的计算资源 $F_n^{k_m^l}$ 后，应付价格 $p_{m,n}^l$ 是唯一确定的。因此动态定价问题可等价转化为计算资源分配问题。

在每一时隙中, 智能体观察状态 s^l 并执行动作 a^l , 获得即时奖励 R^l , 其值与式(23)中定义的单时隙的优化目标一致, l_{\max} 时隙内的累积奖励用以表征系统的长期性能。

5.2 基于TD3的车辆协同任务卸载方法

如图2所示, 方法包含3个组件: actor网络、critic网络和经验回放池。actor网络包含主actor网络 μ^θ 和目标actor网络 $\mu^{\theta'}$ 。critic网络包含主critic网络 $1Q^{\phi_1}$ 、主critic网络 $2Q^{\phi_2}$ 、目标critic网络 $1Q^{\phi'_1}$ 和目标critic网络 $2Q^{\phi'_2}$ 。目标网络具有与主网络相同的结构, 参数周期性地从主网络复制。在每个时隙, 环境状态的转变、智能体进行的动作和即时奖励形成经验 (s^l, a^l, R^l, s^{l+1}) 存储在池中。主critic网络使用近似动作-价值的 Q 值函数评估所选行动。用神经网络进行函数近似存在不精确性, 同时使用后续状态的估计来更新值函数的估计进一步夸大了这种不精确性。在每一次更新策略中错误被累加, 使得不好的状态被高估, 策略无法被优化到最优。TD3使用双重critic网络来评估 Q 值, 选取较小的 Q 值更新可以缓解 Q 值高估的现象。当更新critic网络时, 使用确定性策略的学习目标极易导致目标估计的方差大。这种诱导方差可通过目标策略平滑正则化来减少, 即在计算目标值时, 在下一个状态的动作上加入扰动, 使学到的价值函数在动作维度上更平滑, 价值评估更准确。添加的噪声服从正态分布, 随训练进行逐渐衰退, 对噪声进行裁剪以保持与原始动作相似。在每个训练步, 从池中随机抽样一小批经验作为样本集 $\psi = \{(s^j, a^j, R^j, s^{j+1})\}$ 。依据贝尔曼方程递归关系, 通过最小化时间差分误差更新主critic网络, 表示为

$$L(\Phi_i) \approx \frac{1}{|\psi|} \sum_{j \in \psi} (y - Q^{\Phi_i}(s^j, a^j))^2 \quad (24)$$

$$y = R^j + r \min_{i=1,2} Q^{\Phi'_i}(s^{j+1}, a') \quad (25)$$

$$a' = \mu^{\theta'}(s^{j+1}) + \varepsilon', \varepsilon' \sim \text{clip}(N(0, \delta), -c, c) \quad (26)$$

主actor网络旨在基于当前状态产生一个使 Q 值最大化的动作, 其最后一层为tanh层, 有 M 个神经元, 每个神经元的输出 $\{a_1^j, a_2^j, \dots, a_M^j\}$ 分别对应每个车辆。将区间 $(-1, 1)$ 均等划分为 $M+1$ 个小区间, 若 a_m^j 落在小区间 $(-1, -1+2/(M+1))$, 代表任务车辆 m 选择将任务卸载到基站; 若 a_m^j 落在小区间 $(-1+2(n-1)/(M+1), -1+2n/(M+1))$, 代表任务车辆 m 选择将任务卸载到服务车辆 n , 从而满足约束条件C2和C3。若 $m=n$, 即任务在车辆 m 本地处理, 若没有其他车辆将其作为服务车辆, 则为任务分配最大频率, 否则为任务分配所需的最低频率。若 $m \neq n$, 为了满足约束条件C1和C5, 服务车辆 n 为任务车辆 m 分配的计算频率在经正则化操作后得出, 表示为

$$F_n^{k_m^j} = (F_n - F_n^{\min}) \frac{(a_m^j - (-1+2(n-1)/(M+1)))}{\sum_{i=1}^M \mathbb{I}(a_{k_i^j, n} = 1) (a_i^j - (-1+2(n-1)/(M+1)))} \quad (27)$$

当卸载过程中发生链路中断时, 不再考虑结果传递, 高优先级任务效用与超过截止期限时相同, 低优先级任务因 $t_m^j = \infty$, 效用为0, 从而通过环境反馈施加的惩戒使智能体倾向于满足约束条件C4。智能体通过策略梯度方法更新主actor网络, 损失函数的梯度表示为

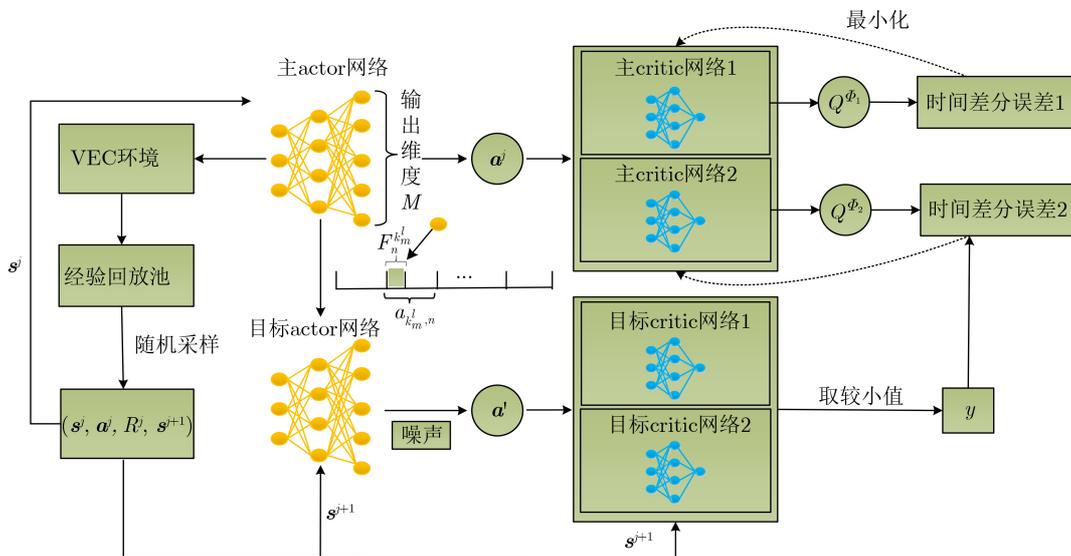


图2 基于TD3的车辆协同任务卸载方法

$$\nabla_{\theta} J(\theta) = \frac{1}{|\psi|} \sum_{j \in \psi} \nabla_{\theta} \mu^{\theta}(\mathbf{s}^j) \nabla_{\mathbf{a}^j} Q^{\Phi_1}(\mathbf{s}^j, \mathbf{a}^j) \Big|_{\mathbf{a}^j = \mu^{\theta}(\mathbf{s}^j)} \quad (28)$$

为了稳定 Q 值，减少一些错误的更新，采用策略延迟更新。首先目标网络与主网络不同步更新，在主网络更新多次后，再对目标网络进行更新。同样地，actor网络与critic网络不同步更新，在critic网络更新多次后，再对actor网络进行更新，即策略网络以低于价值网络的频率更新。这样一方面减少了不必要的重复更新，另一方面减少了多次更新中累积的误差。基于TD3的任务卸载方法实现如算法1所示。

6 性能评估

6.1 仿真设置

本节通过仿真实验来评估本文方法在最大化社会福利方面的性能。表1给出了实验中典型参数值。

为了验证本文方法，引入以下4种方法进行比较：

(1)完全边缘执行(Naive Edge Execution, NEE)：所有车辆的任务全部卸载到边缘服务器进行处理。

(2)随机原则执行(Random Principle Execution, RPE)：随机选择任务的执行方式和以V2V方式卸载任务时服务车辆应向任务车辆提供的计算资源。

(3)GPE：贪心地选择具有最多剩余绿色能源的车辆作为服务车辆，并优先为高优先级任务分配使任务能在最大延迟容忍内完成的计算资源。

(4)DDPG：该方法使用与上文中定义的状态空间、动作空间和奖励函数。

6.2 仿真分析

首先本文将一段时间内实现的累计社会福利作为评估标准，探索不同学习率对本文方法性能的影响。如图3所示，方法在训练迭代次数达400次左右时达到收敛，不同学习率实现的社会福利之间的差异很小，这表明本文方法对学习率参数不太敏感。

算法1 基于TD3的车辆协同任务卸载方法

初始化VEC环境，包括：基站的位置信息 $(0, h)$ ，所有车辆的计算能力 F_m ，车辆的移动模式 $\{(x_m, 0), \mathbf{v}_m\}$ 和EH模块初始能源存储 $Q_m^0, m \in M$ 。

for 迭代次数 $v = 1, 2, \dots, V$ do

初始化一个随机过程用于动作探索，即产生一个随机噪声 $\varepsilon \sim N(0, \delta)$ 。

for $l = 0, 1, \dots, l_{\max}$ do

基站从车辆处接收任务请求，并从环境收集状态信息 \mathbf{s}^l ，依据当前策略及探索噪声生成动作 $\mathbf{a}^l = \mu^{\theta}(\mathbf{s}^l) + \varepsilon$ ，发送任务执行和资源分配信息给车辆，执行动作获得环境反馈的即时奖励 R^l 及新状态 \mathbf{s}^{l+1} ，添加经验 $(\mathbf{s}^l, \mathbf{a}^l, R^l, \mathbf{s}^{l+1})$ 到回放池中。

if $v \geq$ 训练开始阈值 then

从池中随机抽取一小批经验作为样本集 $\psi = \{(\mathbf{s}^j, \mathbf{a}^j, R^j, \mathbf{s}^{j+1})\}$ 。

根据式(24)更新主critic网络。

if l_{mod} 更新频率 = 0 then

根据式(28)更新主actor网络。

以软更新方式更新目标critic网络和目标actor网络。

end for

end for

表1 仿真参数设置

| 参数 | 值 |
|----------------------|-------------------------|
| 车辆速度(km/h) | 40~80 |
| 车辆计算能力(GHz) | 5~10 |
| EH模块初始能源(J) | 50~4000 |
| 任务大小(kbit) | 200~1000 |
| 任务所需计算资源(cycles/bit) | 500~1000 |
| 最大延迟容忍(ms) | 高优先级: 200 低优先级: 1000 |
| 最大V2V链路传输范围(m) | 100 |
| 边缘服务器处理一个CPU周期能耗(W) | 4×10^{-9} |

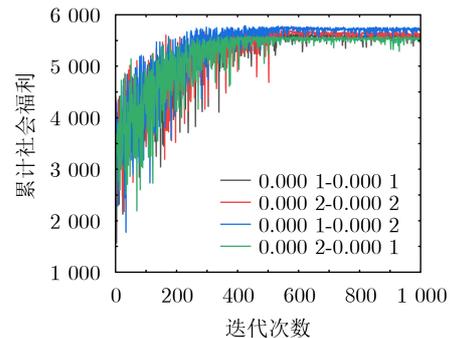


图3 不同学习率下本文方法实现的累计社会福利

当actor网络学习率 $\delta_a = 0.0001$, critic网络学习率 $\delta_c = 0.0002$ 时,方法的收敛性能最好,因为critic网络更快的收敛能更好地指导actor网络的更新。

然后本文比较了5种方法在不同交通密度下实现的平均社会福利。假设在一段时间内没有车辆的加入和离开,如图4所示,在不同的交通密度下,每辆车获得的平均社会福利维持基本稳定。当车辆增多时,状态空间和输出动作维度增加可能导致维数灾难,但本文方法性能没有出现明显下降,说明方法具有很好的扩展性,可适应中等规模的车辆环境。且本文方法实现的平均社会福利均高于其他方法,相比DDPG和GPE方法分别有7.34%和37.47%的提升,而NEE方法的表现最差。这是因为在NEE方法中,绿色能源仅用作通信开销,丧失了车辆协同减少电力使用的优势。GPE方法目标仅是使每个时隙绿色能源的使用效率最大化,无法获得一段时间内社会福利的最大化。尽管DDPG方法的优化目标也是最大化预期长期回报,但DDPG实现的平均社会福利略低于本文方法,这是因为DDPG虽然借鉴使用了两个critic网络,但在实际过程中仍存在高估Q值的情况,而在TD3中利用两套critic网络缓解了这一问题,并通过调整actor网络的更新频率避免了因盲目迭代被困在次优值。此外DDPG可能出现错误Q估值引导下的错误策略,TD3在动作一定范围内随机选择来实现策略平滑,从而摆脱错误峰值的影响。这些策略使得TD3在选择动作时更准确、更鲁棒,可以探索更合适的行为。

本文模拟了当车辆数量为15时,可收集绿色能源最大值对社会福利的影响。如图5所示,随着可收集绿色能源的增加,除NEE方法外,其他方法实现的社会福利都有提升,这验证了更多的绿色能源可以为车辆协同任务卸载的优化提供更多潜力的事实。本文方法和DDPG方法实现的社会福利的增长幅度明显高于GPE方法和RPE方法。这是因为GPE方法基于贪心原则优先选择具有最多剩余绿色能源的车辆作为服务车辆,而RPE方法不考虑这些,随机选择任务的执行方式,所以对富余绿色能源的使用效率较低。但GPE方法同样缺乏对最佳资源分配的考虑,为保障计算资源优先向高优先级任务分配,任务车辆支付给服务车辆的价格是根据预定义的法则得出的,这种固定模式无法使任务获得最大完成效用。而本文方法通过动态定价,在确保高优先级任务可通过支付更高的价格获取更多计算资源的同时,充分发挥了共享资源按需分配的优势,提高了计算资源的利用效率。此外随着可收集绿色能源的增加,社会福利的增速有所放缓,这意味着绿

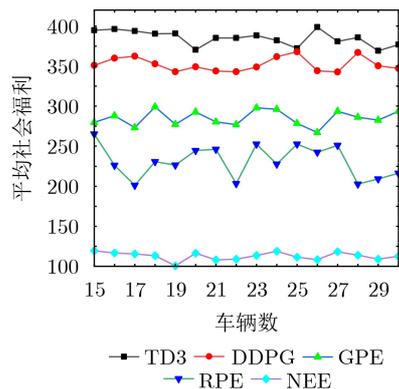


图4 不同交通密度下的平均社会福利

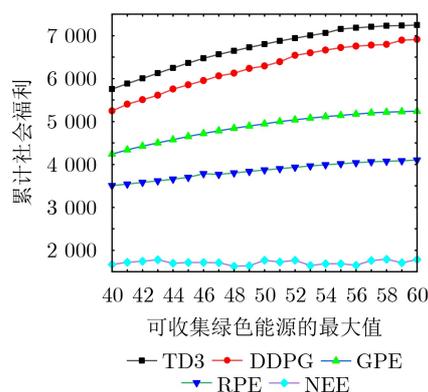


图5 不同可收集绿色能源最大值下的累计社会福利

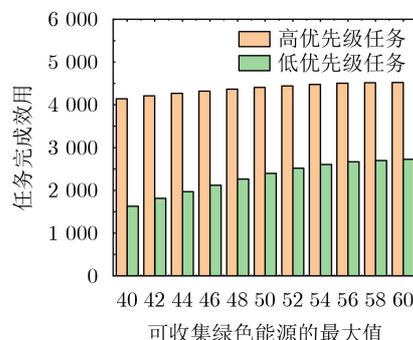


图6 不同可收集绿色能源最大值下的任务完成效用

色能源足够使任务在车辆端执行,社会福利更多地取决于任务的完成效用。

最后,如图6所示,随着可收集绿色能源的增加,低优先级任务完成效用的增速比高优先级任务快,说明本文方法在可收集绿色能源较少时优先确保了高优先级任务以尽可能大的效用完成,同时使低优先级任务以相对适当的效用完成。因此,本文方法在区分执行不同优先级的任务时具有较好的性能。

7 结束语

为了提高VEC系统中边缘服务器能效,实现绿色计算,本文构建了“绿色能源-电网”混合能源供应模式,考虑多个配备能源收集设备的车辆之

间任务协同卸载问题，通过动态定价解决了任务车辆与服务车辆之间存在的利益冲突，促进了二者间的资源共享。同时提出了一种动态VEC环境下基于TD3的任务卸载方法，通过引入双重网络、策略延迟更新、目标策略平滑等策略，本文方法在选择任务执行方式和资源分配方面更具有准确性。最后实验结果验证了本文所提方法的优越性能，相比DDPG和GPE性能上分别提升了7.34%和37.47%。未来会考虑跨越边缘服务器边界的任务卸载以及协同过程中车辆个人隐私保护。

参 考 文 献

- [1] BAI Shengxi and LIU Chunhua. Overview of energy harvesting and emission reduction technologies in hybrid electric vehicles[J]. *Renewable and Sustainable Energy Reviews*, 2021, 147: 111188. doi: [10.1016/j.rser.2021.111188](https://doi.org/10.1016/j.rser.2021.111188).
- [2] LIU Lei, CHEN Chen, PEI Qingqi, et al. Vehicular edge computing and networking: A survey[J]. *Mobile Networks and Applications*, 2021, 26(3): 1145–1168. doi: [10.1007/s11036-020-01624-1](https://doi.org/10.1007/s11036-020-01624-1).
- [3] 马惠荣, 陈旭, 周知, 等. 绿色能源驱动的移动边缘计算动态任务卸载[J]. *计算机研究与发展*, 2020, 57(9): 1823–1838. doi: [10.7544/jissn1000-1239.2020.20200184](https://doi.org/10.7544/jissn1000-1239.2020.20200184).
MA Huirong, CHEN Xu, ZHOU Zhi, et al. Dynamic task offloading for mobile edge computing with green energy[J]. *Journal of Computer Research and Development*, 2020, 57(9): 1823–1838. doi: [10.7544/jissn1000-1239.2020.20200184](https://doi.org/10.7544/jissn1000-1239.2020.20200184).
- [4] SHI Jinming, DU Jun, WANG Jingjing, et al. Priority-aware task offloading in vehicular fog computing based on deep reinforcement learning[J]. *IEEE Transactions on Vehicular Technology*, 2020, 69(12): 16067–16081. doi: [10.1109/TVT.2020.3041929](https://doi.org/10.1109/TVT.2020.3041929).
- [5] LIN Yan, ZHANG Yijin, LI Jun, et al. Popularity-aware online task offloading for heterogeneous vehicular edge computing using contextual clustering of bandits[J]. *IEEE Internet of Things Journal*, 2022, 9(7): 5422–5433. doi: [10.1109/JIOT.2021.3109003](https://doi.org/10.1109/JIOT.2021.3109003).
- [6] FUJIMOTO S, HOOF H, and MEGER D. Addressing function approximation error in actor-critic methods[C]. *The 35th International Conference on Machine Learning*, Stockholm, Sweden, 2018: 1587–1596.
- [7] WANG Haipeng, LV Tiejun, LIN Zhipeng, et al. Energy-delay minimization of task migration based on game theory in MEC-assisted vehicular networks[J]. *IEEE Transactions on Vehicular Technology*, 2022, 71(8): 8175–8188. doi: [10.1109/TVT.2022.3175238](https://doi.org/10.1109/TVT.2022.3175238).
- [8] LUO Qu Yuan, LI Changle, LUAN T H, et al. Collaborative data scheduling for vehicular edge computing via deep reinforcement learning[J]. *IEEE Internet of Things Journal*, 2020, 7(10): 9637–9650. doi: [10.1109/JIOT.2020.2983660](https://doi.org/10.1109/JIOT.2020.2983660).
- [9] 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).
- [10] MA Huirong, HUANG Peng, ZHOU Zhi, et al. GreenEdge: Joint green energy scheduling and dynamic task offloading in multi-tier edge computing systems[J]. *IEEE Transactions on Vehicular Technology*, 2022, 71(4): 4322–4335. doi: [10.1109/TVT.2022.3147027](https://doi.org/10.1109/TVT.2022.3147027).
- [11] LEE J and KO H. Neighbor-aware distributed task offloading algorithm in energy-harvesting internet of things[J]. *IEEE Internet of Things Journal*, 2023, 10(10): 8744–8753. doi: [10.1109/JIOT.2022.3232710](https://doi.org/10.1109/JIOT.2022.3232710).
- [12] KAZMI S M A, DANG T N, YAQOOB I, et al. A novel contract theory-based incentive mechanism for cooperative task-offloading in electrical vehicular networks[J]. *IEEE Transactions on Intelligent Transportation Systems*, 2022, 23(7): 8380–8395. doi: [10.1109/TITS.2021.3078913](https://doi.org/10.1109/TITS.2021.3078913).
- [13] LIWANG Minghui, DAI Shijie, GAO Zhibin, et al. A truthful reverse-auction mechanism for computation offloading in cloud-enabled vehicular network[J]. *IEEE Internet of Things Journal*, 2019, 6(3): 4214–4227. doi: [10.1109/JIOT.2018.2875507](https://doi.org/10.1109/JIOT.2018.2875507).
- [14] 孙慧婷, 范艳芳, 马孟晓, 等. VEC中基于动态定价的车辆协同计算卸载方案[J]. *计算机科学*, 2022, 49(9): 242–248. doi: [10.11896/jsjcx.210700166](https://doi.org/10.11896/jsjcx.210700166).
SUN Huiting, FAN Yanfang, MA Mengxiao, et al. Dynamic pricing-based vehicle collaborative computation offloading scheme in VEC[J]. *Computer Science*, 2022, 49(9): 242–248. doi: [10.11896/jsjcx.210700166](https://doi.org/10.11896/jsjcx.210700166).
- [15] XU Huiying, QIU Xiaoyu, ZHANG Weikun, et al. Privacy-preserving incentive mechanism for multi-leader multi-follower IoT-edge computing market: A reinforcement learning approach[J]. *Journal of Systems Architecture*, 2021, 114: 101932. doi: [10.1016/j.sysarc.2020.101932](https://doi.org/10.1016/j.sysarc.2020.101932).

张红霞：女，博士，副教授，研究方向为边缘计算等。

吕智豪：男，硕士生，研究方向为边缘计算、强化学习算法等。

席诗语：男，硕士生，研究方向为空地一体网络等。

刘佳敏：女，硕士生，研究方向为未来网络人工智能、深度学习算法等。

郭加树：男，博士，研究方向为机器学习算法等。

张培颖：男，博士，副教授，研究方向为未来网络人工智能、网络虚拟化技术、深度学习算法等。

责任编辑：余蓉