文章编号:1674-2974(2025)04-0079-12 DOI:10.16339/j.cnki.hdxbzkb.2025268

多层分布式车联网边缘计算任务动态卸载策略

巨涛†,张宇斐,马雅玲,火久元
(兰州交通大学 电子与信息工程学院,甘肃 兰州 730070)

摘 要:针对车联网计算任务动态卸载成功率和数据传输效率低的问题,基于多智能体深度强化学习设计了多层分布式车联网边缘计算任务动态卸载策略.首先融合软件定义网络和移动边缘计算设计了多层分布式车联网边缘计算系统模型,实现在不同层次上的协同调度优化,更好地满足移动车辆资源动态分配和任务实时处理的需求;之后从车辆计算任务卸载成功率和数据卸载速率两方面考虑,提出了一种多智能体深度强化学习算法框架,利用多智能体系统的协作学习,使车载边缘系统自主选择最优任务卸载决策;同时引入动作空间搜索优化和优先经验回放机制,进一步提升动作空间的有效搜索,提高任务卸载决策的稳定性和准确性;最终在以上算法框架和优化机制的基础上,设计了多层分布式车辆任务卸载决策优化算法,保证车辆能根据当前网络状态和任务大小,以最小的任务传输时间和高效的卸载成功率完成计算任务卸载.仿真结果表明,与已有的卸载方法相比,本文所提方法在计算任务卸载成功率方面提高了5%~20%,在数据传输效率方面平均提高了17.8%.

关键词:车联网;移动边缘计算;任务卸载;深度强化学习;软件定义网络
中图分类号:TP311 **文献标志码**:A

A Multi-layer Distributed Edge Computing Task Dynamic Offloading Strategy in Internet of Vehicles

JU Tao†,ZHANG Yufei,MA Yaling,HUO Jiuyuan
(School of Electronic and Information Engineering, Lanzhou Jiaotong University, Lanzhou 730070, China)

Abstract: To address the challenges of low offloading success rates and inefficient data transmission in the internet of vehicles (IoV), this paper proposes a multi-layer distributed dynamic offloading strategy for edge computing tasks in IoV based on multi-agent deep reinforcement learning. Firstly, a multi-layer distributed internet of vehicles edge computing system model is designed by integrating software defined network and mobile edge computing. The system model can realize collaborative scheduling optimization at different levels, which can better meet the needs of dynamic allocation of mobile vehicle resources and real-time processing of tasks. Then, considering the success rate of offloading and data transmission rate of vehicle computing tasks, a multi-agent deep reinforcement learning algorithm framework is proposed. The algorithm framework uses collaborative learning of

* **收稿日期**:2024-04-12
基金项目:国家自然科学基金资助项目(61862037,62262038),National Natural Science Foundation of China(61862037,62262038);甘肃省科技计划项目(23CXGA0028),Gansu Provincial Science and Technology Plan Project(23CXGA0028);兰州市人才创新创业项目(2021-RC-40),Lanzhou Talent Innovation and Entrepreneurship Project(2021-RC-40)
作者简介:巨涛(1980—),男,甘肃永登人,兰州交通大学副教授,博士
† 通信联系人,E-mail:jutao@mail.lzjtu.cn

multi-agent systems to enable the vehicle edge system to independently select the optimal task offloading decision. At the same time, the optimization of the action space search and the priority experience replay mechanism were introduced to further improve the effective search of the action space and the stability and accuracy of the task offloading decision. Finally, based on the above algorithm framework and optimization mechanism, a multi-layer distributed vehicle task offloading decision optimization algorithm is proposed. The algorithm can ensure that the vehicle can complete the computing task offloading with the minimum task transmission time and effective offloading success rate according to the current network status and task size. Simulation results show that, compared with the existing offloading methods, the proposed method improves the success rate of computing task offloading by 5%~20% and the efficiency of data transmission by 17.8% on average.

Key words: internet of vehicles (IoV); mobile edge computing; task offloading; deep reinforcement learning; software defined network

随着自动驾驶和辅助驾驶等技术融入智能交通系统,车联网应用变得越来越多样化.在相关的车载边缘计算场景中,由于车辆的高速移动性使得车载边缘系统网络拓扑快速变化,对资源分配造成了很大的不确定性.为了解决此类问题,目前已有研究通过利用软件定义网络(software defined network,SDN)和移动边缘计算(mobile edge computing,MEC)来提升车联网的性能.SDN 技术通过解耦控制层和数据层,提高了对车辆网络的管理和扩展,而 MEC 则将计算资源和存储资源推向车辆附近的边缘,减少了数据传输的时延,可以更好地满足车联网应用对低延迟和高计算能力的要求[1-2].

在已有的工作中,文献[3]提出一种基于多步深度 Q 网络的深度强化学习算法,在不考虑复杂传输信道的情况下,将车联网边缘计算任务卸载整个过程建模为马尔可夫决策过程,充分考虑了车联网计算卸载过程中的时延、能耗和通信质量问题,结果显示所提卸载算法在处理计算密集型任务时能显著降低车联网计算时延和计算能耗.文献[4]为了解决由于车辆高移动性在网络中引入的不确定性问题,开发了一个相互协作的通用三层分散式车辆辅助多接入边缘计算网络,使用基于多智能体深度强化学习的匈牙利算法求解车联网中动态任务卸载问题.文献[5]利用有向无环图获取任务调度的优先级,结合分布式深度强化学习策略确定任务卸载决策,最大限度地减少了能源和时间开销.文献[6]提出一种预测性车辆任务卸载方法,通过车对车通信和车对基础设施通信,将计算任务卸载到边缘服务器,利用遗传算法进行多目标优化,采用加权算法和多准则决策求解最优卸载策略,以有效减少车辆计算卸载的能耗和时延.文献[7]提出一种云边协同分层智能驱动的车载边缘计算网络架构,利用云中心、聚合服务器和 MEC 服务器的异构计算能力,实现网络资源的综合协同和智能管理.

以上研究工作,都致力于解决车联网边缘计算任务卸载问题,但由于任务卸载的动态性,车辆通常处于快速移动状态,在高速移动或网络拓扑不稳定的情况下,任务传输的连接状态可能变得不稳定,从而影响任务卸载的成功率[8].同时,随着车联网规模不断扩大,会有大量任务同时传输,而这些任务之间的相互干扰会对任务传输质量产生影响[9].为了解决上述问题,本文针对车载边缘计算动态任务卸载优化问题,设计了一种包含多车辆、多路边单元、多边缘服务器的多层分布式 SDN 边缘计算系统模型,该模型通过多层次资源管理、网络流量控制,在不同层次上进行协同调度和优化,可以更好地满足车辆移动性、资源动态分配和任务实时处理等方面的需求.同时基于所提的系统模型,从车辆计算任务卸载成功率和数据卸载速率两方面考虑,将优化问题转化为多智能体马尔可夫决策过程,提出了一种多智能体深度强化学习算法框架.该算法框架利用多智能体系统的协作学习能力,实现车载边缘系统最优任务卸载决策.为了进一步提升对动作空间的有效搜索,提高任务卸载决策的稳定性和准确性,引入了动作空间搜索优化和优先经验回放机制,以充分利用已有的知识加快深度强化学习的训练速度,提升边

缘计算任务卸载决策效率.最终在以上算法框架和优化机制的基础上,提出了多层分布式车辆任务卸载决策及资源分配优化算法SP-MAD3DQN.该算法可保证车辆能够根据当前的网络状态和任务大小,以最小的任务传输时间和高效的卸载成功率完成计算任务卸载.

1　系统模型

1.1　多层分布式SDN边缘计算系统模型架构

单一软件定义车联网边缘系统存在单点故障的风险,一旦控制器发生故障,整个车联网络就会受到影响.图1为单一软件定义车联网系统架构图.

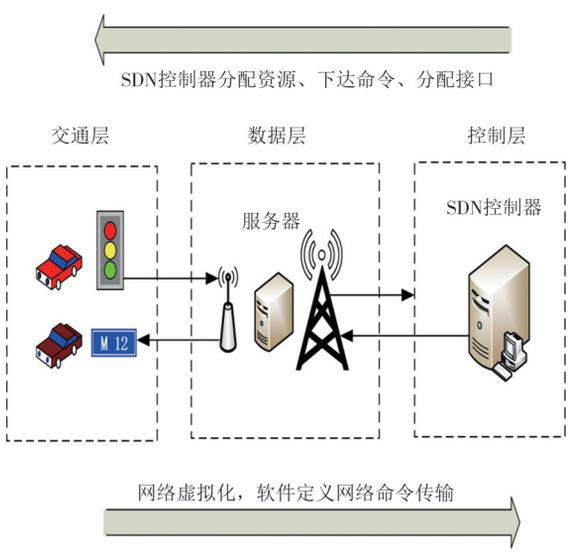


图1　单一软件定义车联网系统

Fig.1　Single software-defined internet of vehicle system

为了给车辆任务卸载提供更高效、灵活、可靠的边缘计算环境,在图1所示的单一软件定义车联网系统模型架构的基础上,本文提出一种多层分布式SDN车载边缘系统模型,如图2所示.该模型通过对软件定义车联网基础模型架构进一步优化和扩展,以满足车载边缘系统中多层次资源管理、网络流量控制等需求[10-11].多层分布式SDN系统可以实现更细粒度的资源调度和任务卸载.通过在不同层次上进行协同调度和优化,可更好地满足车辆移动性、资源动态分配和任务处理实时性等方面的需求,提高边缘计算任务卸载效率和计算性能[12].

图中车载边缘计算环境中存在多个层次的资源,例如车间级、车载级和边缘云级.通过采用多层分布式网络结构,可更好地管理和组织这些资源,实现灵活的资源分配和调度.在该模型中,SDN采用分

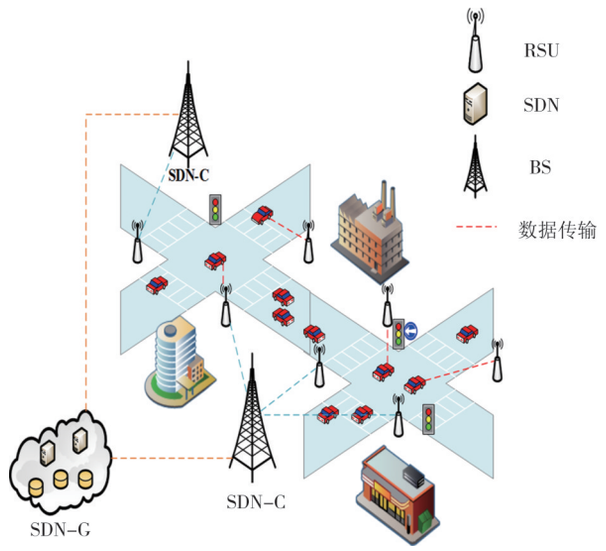


图2　多层分布式SDN车载边缘系统模型

Fig.2　Multi-layer distributed SDN vehicular edge system model

布式控制思想,将控制层和数据层分离,以提高车载边缘系统的可扩展性和灵活性.其中控制层负责系统的集中管理和全局资源控制;数据层负责处理和转发车载边缘系统中的各种数据[13-14].整个系统模型由数据层、本地控制层和中央控制层三层构成.

1.1.1　数据层

数据层分为固定层和移动层两个层级.固定层由一系列部署在道路两旁的路边单元(RSU)构成,这些节点具备数据处理的能力[15].移动层则由不同类型的车辆组成.

1.1.2　本地控制层

与现有大多数SDN方案不同,本文采用了多层分布式SDN的控制架构.下层由多个具备SDN功能的本地边缘控制器(SDN-C)组成,这些控制器集成在位于道路边缘的路边单元(RSU)或基站(BS)中,一般部署在道路两侧的地理中心,以最大限度地减少传输延迟[16].本地控制器可以定期收集其覆盖范围内车辆的状态信息、环境信息,车辆也可以主动将信息传输给本地控制器.

1.1.3　中央控制层

位于上层的是覆盖范围更大的中央控制器(SDN-G),可以对其覆盖范围内的本地控制器进行控制.同时,中央控制器可以定期向本地控制器发送全局信息状态,以同步车载边缘计算系统的全局信息.此外,SDN-G通过无线通信与边缘数据中心和云数据中心进行连接,为车载边缘系统提供更强大的计算、存储能力[17].具体多层分布式SDN车载边缘系统模型架构如图3所示.

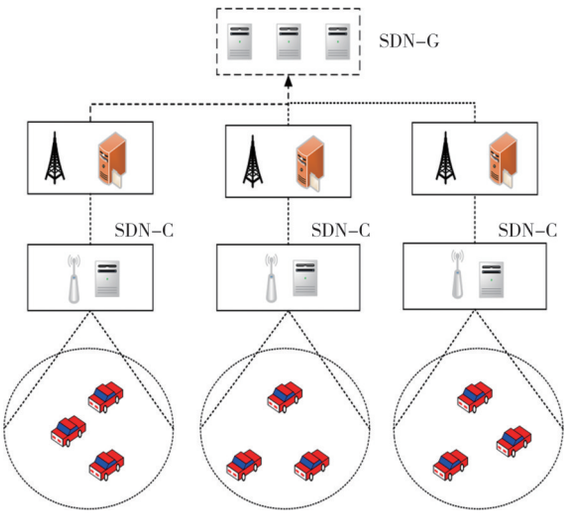


图 3 多层分布式SDN车载边缘系统模型架构

Fig.3 Architecture of multi-layer distributed SDN vehicular edge system model

在高速动态移动环境中,为车辆提供稳定的卸载是一项复杂的任务.为了提高车辆计算卸载的成功率,本文将SDN-C部署到更靠近车辆的RSU或BS中,使得SDN-C更接近车辆,更有效地与车辆进行交互,达到高效管理本地计算资源、根据车辆的需求和环境变化做出合理计算卸载决策的目的.同时,中央控制器位于更高层次的控制层面,负责全局的协调和管理,以确保卸载决策的一致性和高效性.本地控制器根据实时情况快速做出决策,而中央控制器提供更全面的网络视图和资源状态信息,进一步支持本地控制器做出更好的任务卸载策略[18-19].

1.2 通信模型

在基于SDN的车载边缘卸载系统中,任务会被卸载到本地控制器进行策略选择,并通过OFDM(正交频分复用)技术[20]进行车辆间的数据传输.在评估车辆卸载链路的质量时,通常使用信道功率增益进行综合评价.式(1)用来计算车辆卸载链路m的信道功率增益g_m.通过将小尺度衰落和大尺度衰落的影响结合在一起计算卸载链路质量.其中参数a_m表示与频率无关的大尺度衰落增益,即阴影衰落和路径损耗,h_m表示与频率相关的小尺度衰落增益.

$$g_m = a_m h_m \tag{1}$$

当对多个车辆同时进行计算卸载任务时,需考虑车辆之间的信道增益以及干扰信道,以解决传输过程中相互干扰的问题.

设将k号车辆进行卸载传输时产生的干扰信道增益表示为$g'_m(k)$,信道增益表示为$g_m(k)$.同时n号车辆进行计算卸载时产生的干扰信道增益表示为$g'_m(n)$,信道增益表示为$g_m(n)$.则k号车辆在计算任务卸载时产生的信干噪比r_k的计算如式(2)所示.其中σ^2为噪声功率,$I_m(k)$为k号车辆受到的干扰功率大小,P_k为车辆的发射功率大小.

$$r_k = \frac{P_k g_m(k)}{\sigma^2 + I_m(k)} \tag{2}$$

k号车辆在进行计算任务卸载时产生的功率计算如式(3)所示.

$$I_m(k) = P_k g'_m(k) + \rho_m(k) P_n g'_m(n) \tag{3}$$

式中:P_n表示n号车辆发射功率大小;$\rho_m(k)$是一个布尔变量,表示k号车辆是否在车载边缘系统中进行任务卸载,如果是则$\rho_m(k)=1$,否则$\rho_m(k)=0$.

通过式(3),可以在考虑车辆之间相互干扰的情况下,评估每个车辆的计算卸载执行质量.通过考虑车辆之间的相互干扰,优化管理分配卸载资源,进而提高整个车载边缘卸载系统的性能和效果.同时根据香农公式,车载边缘计算系统中卸载信道容量的大小可通过式(4)来计算,其中$C_m(k)$表示k号车辆进行任务卸载时的信道容量大小,W表示卸载信道的带宽.

$$C_m(k) = W\log_2(1 + r_k) \tag{4}$$

根据式(4),则在约束时间T内,k号车辆计算卸载时任务的卸载速率可表示为式(5).

$$V(k) = \sum_{t=1}^{T} \rho_{m,k}(t) C_{m,k}(t) \tag{5}$$

1.3 优化目标

本文目标是通过找到一个合适的卸载和资源分配方案,以达到优化任务卸载成功率和降低车辆任务卸载时延.因此目标函数设置为最大化车辆卸载链路的容量和最大化车辆卸载任务的卸载速率,目标函数具体定义如公式(6)所示.

$$\text{Maximize:} f(x,y) = \sum_{i=1}^{n} C_i + \sum_{i=1}^{z} V(i) \tag{6}$$

在式(6)中参数C_i表示第i条通信链路的传输容量,$V(i)$表示i号车辆卸载任务的传输速率,n表示正在执行卸载任务的通信链路的数量,z表示正在系统中进行卸载的车辆数量.

2 计算任务卸载策略

在车载边缘计算场景之中,由于车辆的高速移动性,常会使任务传输出现中断,从而导致卸载任务传输失败.为了应对车辆的动态变化给车辆边缘计

算任务卸载带来的巨大挑战,本文将车辆边缘计算卸载问题转化为多智能体马尔可夫决策过程,提出一种多智能体深度强化学习算法框架.该算法框架利用多智能体系统的协作学习,使车载边缘系统能够选择最优的任务卸载方案.同时为了进一步提升对动作空间的有效搜索,提高任务卸载决策的稳定性和准确性,引入动作空间搜索优化和优先经验回放机制,以充分利用已有的知识加快深度强化学习的训练速度,提升边缘计算任务卸载决策效率.最终在以上算法框架和优化机制的基础上,提出了多层分布式车辆任务卸载决策及资源分配优化算法SP-MAD3DQN.通过SP-MAD3DQN算法,车辆能够根据当前网络状态和任务特点,实现高效的计算任务的卸载,最小化任务传输时间,提高任务卸载的成功率和边缘计算资源的利用率.

2.1　多智能体马尔可夫决策过程

首先将车辆任务卸载决策问题,转化为一个多智能体马尔可夫决策过程,然后基于多智能体深度强化学习进行问题求解.在整个求解过程中,每个本地控制器作为一个智能体与未知环境进行交互获取经验,并从中学习最优策略.多个智能体共同探索环境,根据环境状态变化优化控制策略.为提高神经网络的全局性能,所有智能体使用相同的奖励函数.多智能体强化学习可以被定义为一个离散时间的随机过程,使用四元组(S,A,R,P)来描述.其中元组S表示环境的状态集,A表示智能体可以执行的动作集,R表示智能体对完成某个动作的回报,P描述动作A从状态S过渡到状态S'的概率[21].多智能体马尔可夫决策过程设计如下.

2.1.1　状态空间设计

对于本地控制器而言,仅能获取自己所处覆盖范围内的车辆状态信息S_t^k和动作$A(t)$,而全局信息和其他智能体的动作都是未知的.当车辆进行计算卸载时,单个智能体观测到的局部信息有:卸载车辆自身的信道增益$g_m(k)$,来自其他卸载车辆产生的干扰信道增益$g'_m(n)$、车辆卸载的数据量大小B、车辆的行驶速度V,以及车辆的位置信息X_k,Y_k.则智能体k所观察到的状态空间可表示为:

$$S_t^k=\left\{B,V,g_m(k),\ g'_m(n),X_k,Y_k\right\} \tag{7}$$

2.1.2　动作空间设计

本文车辆计算任务卸载优化策略动作决策,主要确定将车辆任务卸载到哪个边缘服务器上.具体动作空间如公式(8)所示,其中a_i^{edge}表示任务要卸载到n个边缘服务器中的第i个边缘服务器进行计算.

$$A(t)=\left(a_1^{\text{edge}},a_2^{\text{edge}},\cdots,a_n^{\text{edge}}\right) \tag{8}$$

2.1.3　奖励函数设计

在车辆卸载决策过程中,本地控制器之间需要充分协作.每个时间步都要根据当前状态和所采取的行动获得具体的奖励.本文目标是在最小化车辆任务卸载时延和提高车载任务卸载成功率之间找到最佳平衡,以实现最优的控制策略.在设计奖励函数时,将车辆卸载成功率的奖励和车辆卸载时延的奖励结合到总预期回报中,引入权重参数λ,取值范围在0到1之间,使计算卸载策略倾向于不同的回报,并通过考虑车辆卸载链路总容量和车辆卸载任务传输成功率的折中来平衡整个奖励回报.具体的奖励函数如式(9)所示.

$$R_t=\lambda\sum_k C_m(k,t)+(1-\lambda)\sum_k L_k(t) \tag{9}$$

式中:$L_k(t)$为k号车辆在t时间内任务卸载成功率,其值为t时刻内经边缘服务器成功处理的车辆任务数量与t时刻内车辆产生的总任务数量的比值;$C_m(k,t)$表示k号车辆在t时刻车辆卸载信道的总容量,m代表车辆卸载链路,信道总容量越大,系统的带宽能力越强,车辆任务的卸载时延就会越短.

2.2　传统D3DQN车联网任务卸载控制框架

D3DQN算法网络结构是在DDQN算法网络的基础上,通过引入Dueling Network来解决强化学习中的动作选择问题[22].该算法由状态值网络$V(s)$和动作优势网络$A(s,a_i)$两个子网络构成.其中动作优势网络$A(s,a_i)$的任务是估计在给定状态下每个动作相对于其他动作的优势.该网络接受状态和动作作为输入,并输出每个动作的优势估计.优势估计用于衡量每个动作的相对优劣.状态值网络$V(s)$任务是估计在给定状态下每个可能动作的价值.该网络接受状态s作为输入,并输出每个可能动作的价值估计,价值估计用于衡量每个动作的优劣[23].通过组合$V(s)$状态值网络和动作优势网络$A(s,a_i)$,可以得出作用值函数估计值,从而更准确地评估每个动作的Q值.计算公式可表示为:

$$Q(s,a_i)=V(s)+A(s,a_i) \tag{10}$$

图4为本文基于D3DQN算法的车联网计算任务卸载控制策略框架.首先,车辆通过车联网通信网络将自身任务需求上传至本地控制器.每个本地控制器会维护一个状态动作对,其中s表示当前时刻的状

态信息,而动作a表示所采取的措施.本地控制器为了估算在状态s下执行某个动作a得到的预期回报,通过状态值网络和动作优势网络进行计算.状态值网络利用神经网络估算当前状态下的预期回报.动作优势网络则用于衡量执行某个动作相比其他动作更优的概率.通过组合状态值网络和动作优势网络获得Q值$Q(s,a_i)$,最后通过使用argmax(a')策略从在线网络中选取动作,并利用目标网络计算获得目标Q值$Q'(s,a_i)$,同时本地控制从网络环境中获得动作奖励γ.随后,本地控制器将采取的行动a发送给其他控制器,并且通过最小化损失函数来更新网络参数.

损失函数定义如下:

$$L(\theta)=\frac{1}{N}\left[y_i-Q(s_i,a_i;\ \theta)\right]^2 \tag{11}$$

式中:$L(\theta)$表示损失函数;N表示样本数量;y_i表示目标Q值.

最终,将优化后的参数信息上传到中央控制端,中央控制器依据收集到的状态信息进行任务和资源的合理分配.在整个过程中,本地控制器通过不断学习,持续更新状态值网络和动作优势网络的参数,以此提高卸载策略质量和任务卸载效率.

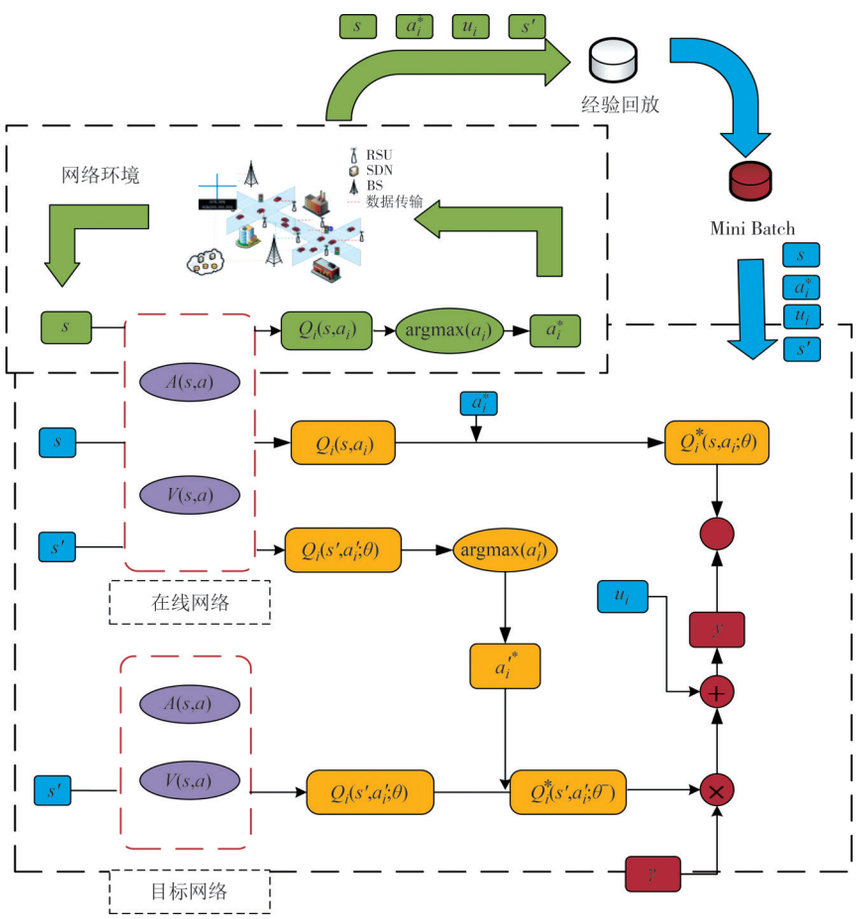


图4 基于D3DQN算法的车联网计算任务卸载控制框架

Fig.4 Control framework for task offloading in internet of vehicles based on D3DQN algorithm

2.3 SP-MAD3DQN算法框架

车载边缘计算任务卸载是一个复杂的过程,要同时考虑车辆自身的移动性,网络资源的动态分配方式,以及任务处理的实时性要求等多个方面.传统强化学习算法D3DQN,通常根据局部信息进行优化和决策,而在车载边缘计算环境中,任务卸载和网络资源管理不仅需要关注局部情况,更需要考虑整个系统的全局优化目标,包括系统的整体性能和效率.D3DQN由于无法有效实现全局优化,在进行车载边缘计算任务卸载优化时会导致局部最优,影响整个计算性能.多智能体深度强化学习框架,能够把决策过程拆解为多个智能体,通过多个智能体之间的协同合作,管控车载边缘计算环境中可能存在的多个边缘服务器和控制器.

本文为了有效应对传统强化学习算法 D3DQN 存在的不足,设计了基于多层分布式SDN架构的多智能体深度强化学习的车辆边缘计算任务卸载框架 SP-MAD3DQN.通过多个智能体之间的协作,实现分布式的车辆边缘计算任务卸载决策,从而更好地利用边缘系统的计算资源,提高任务处理的效率.本文 SP-MAD3DQN算法框架如图5所示.

多智能体算法在集中式学习训练阶段,每个智能体获取车辆的局部状态信息,并将这些信息作为算法的输入.然后本地控制器根据获取的信息,通过SP-MAD3DQN算法框架得到状态-动作值函数,确定当前节点上需要卸载的任务或资源分配情况.当边缘节点执行动作后,将动作信息传输至中央控制器,中央控制器根据全局信息进行全局资源协调与车辆任务卸载决策.最终每个智能体节点根据自身的状态信息、动作和奖励信号共同更新经验池,并利用 MAD3DQN 算法不断学习,优化任务卸载策略.

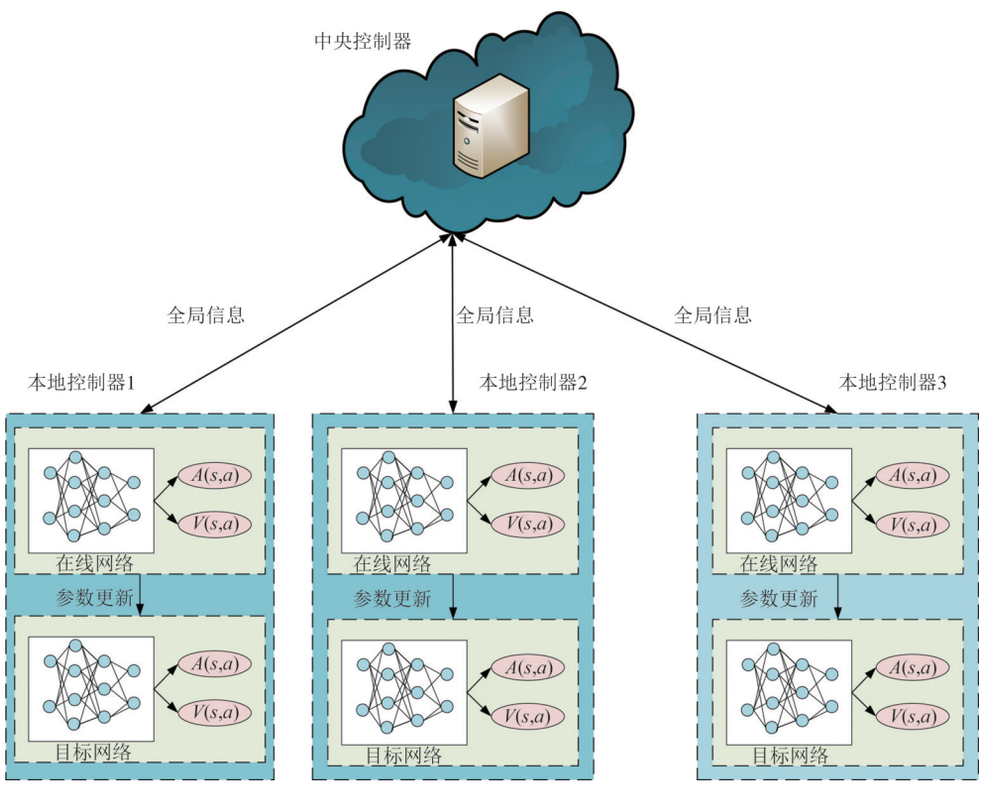


图 5　SP-MAD3DQN 算法框架

Fig.5　SP-MAD3DQN algorithm framework

2.4　动作空间搜索优化

在车载边缘计算场景中,智能体需要在不同的车辆状态下探索动作空间,包括选择合理的卸载方式和适当的边缘服务器进行卸载任务.这涉及动作空间的选择问题,传统的算法通常采用随机策略进行动作探索,但是随机性设置过高或过低会影响探索效率和已有知识的利用.

为了有效地探索动作空间,结合Softmax动作空间搜索优化策略,使智能体在探索动作空间时更注重估计价值较高的动作,倾向于选择在当前状态下更有利的计算卸载和资源分配方式,提高任务执行效率和系统性能.

将 Softmax 动作空间搜索优化策略融入 SP-MAD3DQN算法框架中,通过使用Softmax()函数计算神经网络输出动作的价值概率,从而更有效地训练神经网络,使算法可以学习到在不同状态下,哪些动作会有更好的任务执行效果,让智能体在探索中更加有针对性地选择动作,从而得出最优任务卸载和资源分配方案.图6为融入了Softmax策略的内部 Q 网络结构.

通过反复训练和调整参数,系统可以逐渐寻找到最优的策略,实现车载边缘系统计算资源的高效利用和任务的高效执行.对于每个状态-动作对 (s,a),使用Softmax()函数将 Q 值转换为一个概率值 $\mathrm{Softmax}\big(Q(s,a)\big)$,如式(12)所示.

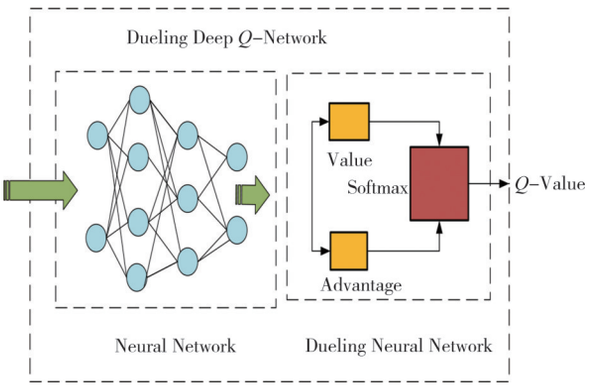


图6 融合Softmax策略的SP-MAD3DQN内部Q网络结构

Fig.6 Internal Q network structure of SP-D3DQN integrated with Softmax

$$\mathrm{Softmax}\big(Q(s,a)\big)=\frac{\exp\big[\beta Q(s,a)\big]}{\sum\exp\big[\beta Q(s,a')\big]} \tag{12}$$

在式(12)中,β是一个温度参数,控制了探索程度和已有知识之间的平衡.当温度参数较小时,概率分布更集中,智能体更倾向于选择具有高Q值的动作;而当温度参数较大时,概率分布更均匀,智能体更倾向于随机选择动作.通过反复训练和调整参数,系统可以逐渐寻找到最优的任务卸载策略.

2.5 优先经验回放机制

在车载边缘计算环境中,智能体需要不断适应网络状况、车辆状态等变化做出决策.简单地根据当前状态选择动作可能导致决策的不稳定.为了进一步提高任务卸载决策的稳定性和准确性,本文引入TD-error优先经验回放机制,使智能体利用已有知识数据,根据经验数据优先级权重,更有效地选择任务卸载对象,加速学习过程,提升决策性能.

具体结合TD-error优先经验回放机制策略的核心思想是:对训练有价值的数据赋予更高的优先级权重,在算法对经验池进行采样时,优先级高的数据被更频繁地选择出来,进而提高训练效率和收敛速度[24].经验池中第j个训练数据,其经验优先级定义如公式(13)所示.

$$P_j=|\partial|+\delta \tag{13}$$

式(13)中δ为一个很小的正数,用于避免数据的优先级计算为0;而参数∂被定义为TD误差,如公式(14)所示:

$$\partial=r(s_t,a_t)+Q'(s_{t+1},a_{t+1})-Q(s_t,a_t) \tag{14}$$

式(14)中Q'、Q分别表示从在线网络和目标网络中计算获得的Q值;$r(s_t,a_t)$表示智能体在状态s_t和动作a_t的共同作用下变为后续状态s_{t+1}后获得的奖励.在利用优先经验回放对样本进行采样时,学习经验j被抽中的概率定义为$P(j)$,如式(15)所示.

$$P(j)=\frac{p_j^{\lambda}}{\sum_{i=1}^{J}p_i^{\lambda}} \tag{15}$$

式中:J为经验池的容量,λ为优先级指数.该机制通过调节λ的大小,保留优先级更高的转移样本.通过优先经验回放机制,可以更好地利用有价值的数据,提高训练的效率和准确性,优化车载边缘计算任务的卸载决策.

2.6 多层分布式SP- MAD3DQN算法

在以上多智能体深度强化学习车辆边缘计算任务卸载框架、动作空间搜索优化及优先经验回放机制的基础上,基于D3DQN算法设计了本文多层分布式车辆任务卸载决策及资源分配优化算法SP-MAD3DQN.该算法通过多层分布式SDN架构为车辆提供实时网络状态信息和流量数据,控制器通过感知车辆计算任务的计算需求和资源状态,对本地资源进行管理和调度,并且通过分布式控制结构减少传输延迟,提高任务卸载的成功率.

Softmax()函数将每个卸载动作的值转换为概率分布,智能体通过概率分布选择动作,在车辆任务卸载决策过程中更全面地考虑动作空间的选择,有效地探索决策空间,提高决策准确性.

优先经验回放机制根据经验数据的优先级权重选择训练数据,确保重要的训练数据被有效地选择和学习,以加速训练效率和收敛速度.最终智能体能够更有效地利用已有的知识数据,提高决策的准确性和效率.

本文SP-MAD3DQN算法采用分布式执行方式,每个智能体都可以独立做出决策和执行,大大提高了执行效率.在分布式执行阶段,每个智能体只需要局部观察,可在不了解其他智能体环境信息的情况下获得自己的动作.每个智能体可以同时进行决策和执行,不需要进行全局的信息交流,以提高执行效率.具体实现如算法1所示.

3 实验仿真及结果分析

3.1 实验环境及参数设置

本文车联网任务卸载传输场景,根据文献[25]中定义的城市中车联网模拟器进行设置,包括车辆、车道和无线通信网络模型.使用Python 3.7、TensorFlow 1.13.1、CUDA10.0等工具进行模拟实验.具体的硬件环境为NVUDIA 1050Ti显卡,24 GB内存,CPU

算法 1:多层分布式车辆卸载决策及资源分配优化算法 SP-MAD3DQN

输入:$S_t^k=\{B,V,g_m(k),g'_m(n),X_k,Y_k\}$

输出:车辆任务卸载策略$A(t)$

1. 初始化在线网络策略μ、评估网络策略μ'及其权重为θ、θ',经验池大小D,随机噪声N_t,
2. **for** 每个训练集 **do**
3. 初始化观测状态s
4. **for** 每个时间步t **do**
5. **for** 每个智能体m **do**
6. 智能体根据状态选择动作$A(t)$
7. 智能体执行动作s后获得下一时刻的状态s'
8. **end for**
9. 所有智能体执行完动作,获取共同奖励
10. **for** 每个智能体m **do**
11. **if** 经验缓冲池D未满
12. 在D中存储状态转移信息 (s,a,r,s')
13. 从D中根据式(13)抽取数据大小为K的批量训练数据
14. **else**
15. 更新经验缓冲池D
16. 根据式(10)获得目标值Q
17. 使用梯度下降法更新动作优势网络参数
18. 使用式(11)最小化损失函数值更新状态值网络
19. 更新目标网络参数,更新速率为τ
20. **end if**
21. **end for**
22. **end for**
23. **end for**

为 CORE i7 8th .激活函数为 ReLu.为了验证所提 SP-MAD3DQN 算法在进行车载任务卸载时的有效性,和以下四种方法进行比较.

1)随机优化算法(Random):在每个时间步,车辆以一种均匀分布的随机方式卸载计算任务.

2)单智能体深度强化学习算法(D3DQN):该算法只有一个智能体,智能体具有在线网络和目标网络双网络结构,智能体只能获取单个车辆的局部信息和奖励.

3)引入 TD-error 机制和 Softmax 策略的单智能体深度强化学习算法(SP-D3DQN):该算法在 D3DQN 算法的基础上加入了 TD-error 优先经验回放机制和 Softmax 动作空间搜索优化策略.

4)引入 TD-error 机制和 Softmax 策略的多智能体深度强化学习方法(SP-MADQN):该算法中单个智能体遵循 DQN 算法,即每个智能体基于当前车联网环境选择离散的动作.算法的超参数学习率和折扣因子与本文所提算法一致.

3.2　结果分析

3.2.1　收敛性分析

为了验证本文方法的收敛性,将本文方法与随机优化、D3DQN、SP-D3DQN、SP-MADQN 四种方法训练过程的奖励回报值进行了对比.从图 7 可以看出,随着训练的持续迭代,奖励值逐渐增加.经过 500 轮的训练后,四种深度强化学习算法逐渐开始收敛.相比于其他强化学习算法,本文所提算法收敛速度快约 2%~10%.随机优化算法因没有设置针对特定目标或最优情况的调整和优化,在实验中没有出现高奖励或低奖励情况,得到的奖励值的变化范围相对较窄.在整个算法的收敛过程中,本文方法表现出较小的波动性,因为本文算法引入了优先经验回放机制,能够确保模型关注对性能提升关键的经验,而不被其他次要经验所干扰,能够更好地控制训练样本,降低训练过程中的奖励波动.

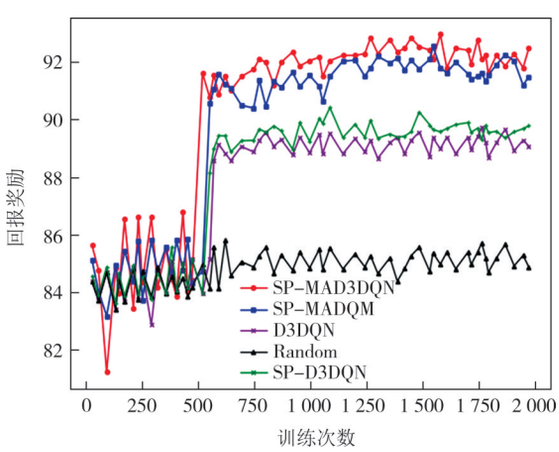


图 7　回报奖励对比

Fig.7　Reward comparison

3.2.2　最大卸载速率分析

为了验证本文方法在数据卸载速率方面的性能,比较了在不同任务数据量情况下车载计算任务卸载时的数据传输时延.从图 8 可以看出,随着任务数据量的增加,所有优化算法的车辆任务传输时延明显上升.而本文方法和其他四种方法相比,传输时延随任务数据量的增大变化较为缓慢,其原因是本文考虑了传输数据量增大时车辆卸载任务之间的相互干扰问题,通过优化车载边缘系统信道资源分配来减少彼此之间的干扰,以此来降低任务传输过程中因相互干扰所导致的传输时延,进而提高任务卸载速率.

3.2.3　卸载成功率分析

为了验证本文方法在车载任务卸载成功率方面

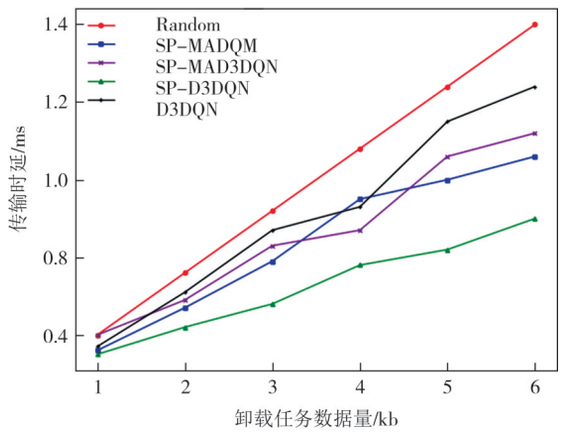


图8 不同负载下车辆任务传输时延

Fig.8 Task transmission delay under different load sizes

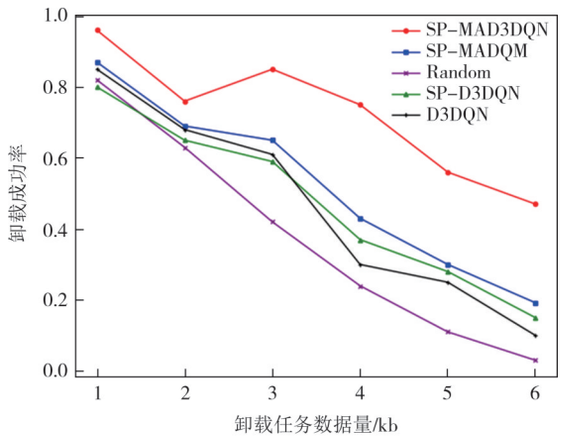


图9 不同负载下车辆任务卸载成功率对比

Fig.9 Comparison of success probability of vehicle task transmission under different load sizes

的性能,在不同卸载任务负载下,将本文方法和其他四种方法的车载计算任务卸载成功率进行了对比.从图9可以看出,随着负载的增加,所有优化方法的车载计算任务卸载成功率会逐渐降低.对比五种方法,在相同的卸载任务数据量下,本文SP-MAD3DQN算法相对于其他方法表现出了更优的车载任务卸载成功率.随着卸载任务负载的增加,SP-MAD3DQN算法和其他四种方法在卸载成功率之间的差距逐渐扩大,和其余对比算法相比本文算法卸载成功率最低提高了5%,最高提高了20%,表现出了较好的计算任务卸载性能.本文所提算法,可以根据任务的要求和边缘服务器的可用资源情况,智能地选择最佳的边缘服务器进行任务卸载,可以确保任务在具备足够计算能力和存储空间的服务器上卸载执行,从而提高车载任务卸载的成功率.此外,还考虑了传输时延和数据传输的可靠性,算法会评估边缘服务器之间的传输时延,为车辆分配最合适的边缘服务器和卸载资源,降低额外的传输开销,以此来提高任务卸载的成功率.

3.2.4 卸载效率分析

为了验证本文方法在处理多个车辆任务时的卸载效率,将本文方法和随机优化算法在处理3个车辆的任务请求时的卸载时间和任务卸载速率进行了比较.图10和图11分别为本文方法和随机优化算法在处理3个车辆计算任务卸载时的卸载时间和任务卸载速率变化.本文方法和随机优化算法相比,平均任务卸载效率提高了17.8%.

在任务卸载速率方面,从图10可以看出,系统会根据车辆任务优先级分别为1~3号车辆提供卸载服务,首先为1号车辆优先分配资源,确保该车辆在初始卸载过程中具有较高的卸载速率,以提高整个车辆系统任务卸载的效率,减少任务传输时间.随着1号车辆完成卸载,系统会根据请求任务的优先级及计算资源情况,将剩余资源分配给后续卸载车辆,实现系统资源的合理分配和利用.本文方法根据车辆任务的优先级,通过在卸载过程中合理分配和利用车载边缘计算系统的资源,可使获得卸载服务的车辆持续处于高速卸载状态,高效地利用边缘系统计算资源保证优先级高的任务及时得到处理,同时协同不同车辆计算任务特点进行最优的计算资源分配,从而提高整个车载边缘计算系统的计算任务卸载效率.

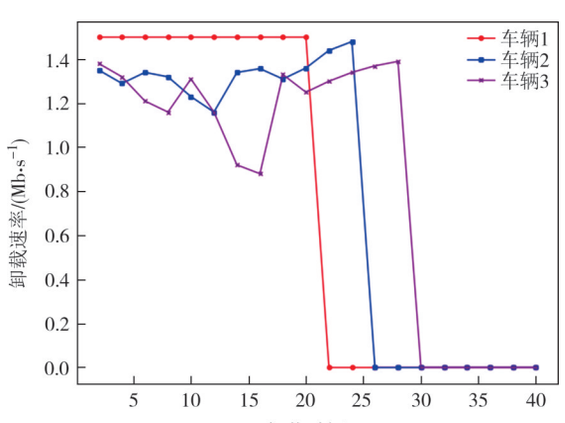


图10 SP-MAD3DQN算法任务卸载时间和卸载速率

Fig.10 Task offloading time and transmission rate of SP-MAD3DQN algorithm

如图11所示,和本文算法相比,随机优化算法由于没有考虑车辆计算任务的特点,无法将计算任务特点和边缘计算资源特点结合,无法有效协调和优化资源分配,导致车辆在卸载过程中卸载速率的波动范围较大,从而影响了整个车辆计算任务的卸

载速率和卸载效率.特别是在复杂的车辆卸载场景中,随机优化算法无法实现高效的任务卸载.

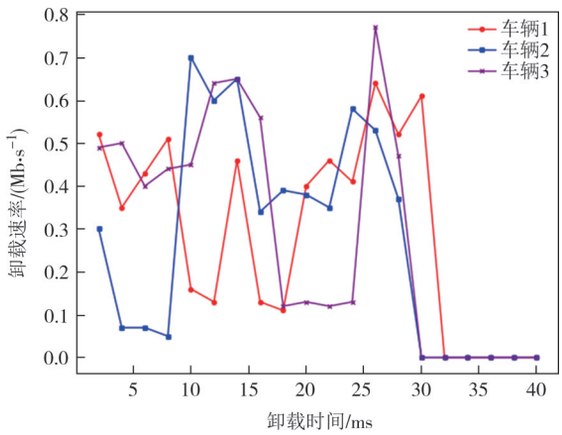


图 11　随机优化算法任务卸载时间和传输速率

Fig.11　Task offloading time and transmission rate of random optimization algorithm

4　结 论

本文针对车载边缘计算任务卸载优化问题,以最大化任务卸载成功率和最小化车辆任务卸载时延为目标,设计了一种面向多车辆、多路边单元、多边缘服务器,融合软件定义网络和移动边缘计算的车联网系统模型,以提供更可靠的车联网边缘计算环境.通过引入SDN控制器收集全局环境信息,控制全局网络状态,为车辆分配相应的资源,以实现最佳的资源利用和任务执行效果.基于多智能体深度强化学习算法设计了动态车辆计算任务卸载框架,以提高车辆计算任务的卸载成功率和数据卸载速率,同时为了进一步提升算法的收敛性和计算性能,加入了动作空间搜索优化机制和优先经验回放策略,充分利用已有的知识加快深度强化学习的训练速度.最终在所提出的动态车辆计算任务卸载框架和优化机制的基础上设计实现了多层分布式车辆任务卸载决策及资源分配优化算法.实验验证结果表明,本文所提方法在相同的实验环境下,与其他卸载方法相比能够更好地适应车辆移动和任务卸载的不确定性,可以提升计算任务卸载成功率和数据传输效率,提高边缘服务器资源利用率并减小任务传输时延.

参考文献

[1] YUAN T T, DA ROCHA NETO W, ROTHENBERG C E, et al. Dynamic controller assignment in software defined internet of vehicles through multi-agent deep reinforcement learning [J]. IEEE Transactions on Network and Service Management, 2021, 18(1): 585-596.

[2] 李国燕,薛翔,刘毅,等. 改进TD3的SDN车联网边缘计算卸载策略[J]. 计算机集成制造系统,2023,29(5):1627-1634.
LI G Y, XUE X, LIU Y, et al. Improved TD3 edge computing offloading strategy for software defined networking internet of vehicles[J]. Computer Integrated Manufacturing Systems, 2023, 29(5): 1627-1634.(in Chinese)

[3] HAN S D, CHEN Y Q, CHEN G H, et al. Multi-step reinforcement learning-based offloading for vehicle edge computing [C]//2023 15th International Conference on Advanced Computational Intelligence (ICACI). Seoul, Korea, IEEE, 2023: 1-8.

[4] ALAM M Z, JAMALIPOUR A. Multi-agent DRL-based Hungarian algorithm (MADRLHA) for task offloading in multi-access edge computing internet of vehicles (IoVs) [J]. IEEE Transactions on Wireless Communications, 2022, 21(9): 7641-7652.

[5] LIU H Q, ZHAO H B, GENG L W, et al. A distributed dependency-aware offloading scheme for vehicular edge computing based on policy gradient [C]//2021 8th IEEE International Conference on Cyber Security and Cloud Computing (CSCloud)/2021 7th IEEE International Conference on Edge Computing and Scalable Cloud (EdgeCom). Washington, D C, USA. IEEE, 2021: 176-181.

[6] SUN Y L, WU Z Y, SHI D Y, et al. Task offloading method of internet of vehicles based on cloud-edge computing [C]//2022 IEEE International Conference on Services Computing (SCC). Barcelona, Spain. IEEE, 2022: 315-320.

[7] WANG S S, XIN N, LUO Z Y, et al. An efficient computation offloading strategy based on cloud-edge collaboration in vehicular edge computing [C]//2022 International Conference on Computing, Communication, Perception and Quantum Technology (CCPQT). Xiamen, China. IEEE, 2022: 193-197.

[8] SOUA A, TOHME S. Multi-level SDN with vehicles as fog computing infrastructures: a new integrated architecture for 5G-VANETs [C]//2018 21st Conference on Innovation in Clouds, Internet and Networks and Workshops (ICIN). Paris, France. IEEE, 2018: 1-8.

[9] LIN C, HAN G J, JIANG J F, et al. Underwater pollution tracking based on software-defined multi-tier edge computing in 6G-based underwater wireless networks[J]. IEEE Journal on Selected Areas in Communications, 2023, 41(2): 491-503.

[10] WANG Y Q, HU X, GUO L J, et al. Research on V2I/V2V hybrid multi-hop edge computing offloading algorithm in IoV environment [C]//2020 IEEE 5th International Conference on Intelligent Transportation Engineering (ICITE). Beijing, China. IEEE, 2020: 336-340.

[11] ZHAO N, LIANG Y C, NIYATO D, et al. Deep reinforcement learning for user association and resource allocation in heterogeneous cellular networks [J]. IEEE Transactions on Wireless Communications, 2019, 18(11): 5141-5152.

[12] SHARMA H, KUMAR N, TEKCHANDANI R. Mitigating jamming attack in 5G heterogeneous networks: a federated deep reinforcement learning approach [J]. IEEE Transactions on Vehicular Technology, 2023, 72(2): 2439-2452.

[13] LIU Y B, MAO B M. On a novel content edge caching approach based on multi-agent federated reinforcement learning in internet of vehicles[C]//2023 32nd Wireless and Optical Communications Conference (WOCC). Newark, NJ, USA. IEEE, 2023: 1-5.

[14] GONG Y Q, HAO F, WANG L, et al. A socially-aware dependent tasks offloading strategy in mobile edge computing[J]. IEEE Transactions on Sustainable Computing, 2023, 8(3): 328-342.

[15] 刘雷,陈晨,冯杰, 等. 车载边缘计算卸载技术研究综述[J]. 电子学报, 2021, 49(5): 861-871.
LIU L, CHEN C, FENG J, et al. A survey of computation offloading in vehicular edge computing networks [J]. Acta Electronica Sinica, 2021, 49(5): 861-871.(in Chinese)

[16] 朱思峰,赵明阳,柴争义. 边缘计算场景中基于粒子群优化算法的计算卸载[J]. 吉林大学学报(工学版),2022,52(11): 2698-2705.
ZHU S F, ZHAO M Y, CHAI Z Y. Computing offloading scheme based on particle swarm optimization algorithm in edge computing scene[J]. Journal of Jilin University (Engineering and Technology Edition), 2022, 52(11): 2698-2705.(in Chinese)

[17] GAUR K, GROVER J. Exploring VANET using edge computing and SDN[C]//2019 Second International Conference on Advanced Computational and Communication Paradigms (ICACCP). Gangtok,India. IEEE,2019:1-4.

[18] WANG B X, LIU L, WANG J. Multi-agent deep reinforcement learning for task offloading in vehicle edge computing[C]//2023 IEEE International Symposium on Broadband Multimedia Systems and Broadcasting (BMSB). Beijing, China. IEEE, 2023: 1-6.

[19] 吕品, 许嘉, 李陶深, 等. 面向自动驾驶的边缘计算技术研究综述[J]. 通信学报, 2021, 42(3): 190-208.
LYU P, XU J, LI T S, et al. Survey on edge computing technology for autonomous driving [J]. Journal on Communications, 2021, 42(3): 190-208.(in Chinese)

[20] 巨涛, 王志强, 刘帅, 等. D3DQN-CAA:一种基于DRL的自适应边缘计算任务调度方法[J]. 湖南大学学报(自然科学版), 2024, 51(6): 73-85.
JU T, WANG Z Q, LIU S, et al. D3DQN-CAA: a DRL-based adaptive edge computing task scheduling method [J]. Journal of Hunan University (Natural Sciences), 2024, 51(6): 73-85.(in Chinese)

[21] 朱思峰, 胡家铭, 杨诚瑞, 等. 物联网边缘计算场景下基于优先级任务的卸载决策优化[J]. 吉林大学学报(工学版), 2024, 54(11): 3338-3350.
ZHU S F, HU J M, YANG C R, et al. Optimization of priority task based unloading decision in the context of edge computing of the internet of things[J]. Journal of Jilin University (Engineering and Technology Edition): 2024, 54(11): 3338-3350.(in Chinese)

[22] TANG S H, CHEN B, IWEN H, et al. VECFrame: a vehicular edge computing framework for connected autonomous vehicles [C]//2021 IEEE International Conference on Edge Computing (EDGE). Chicago, IL, USA. IEEE, 2021: 68-77.

[23] CHEN X W, ZHANG Q Y, JIN Z G, et al. Research on intelligent vehicle infrastructure cooperative system based on 5G mobile edge computing[C]//2021 6th International Conference on Transportation Information and Safety (ICTIS). Wuhan, China. IEEE, 2021:21-27.

[24] TAO X, HAFID A S. DeepSensing: a novel mobile crowdsensing framework with double deep Q-network and prioritized experience replay[J]. IEEE Internet of Things Journal,2020, 7(12): 11547-11558.

[25] 方维维, 王云鹏, 张昊, 等. 基于多智能体深度强化学习的车联网通信资源分配优化[J]. 北京交通大学学报, 2022, 46(2): 64-72.
FANG W W, WANG Y P, ZHANG H, et al . Optimization of communication resource allocation in connected vehicles based on multi agent deep reinforcement learning [J]. Journal of Beijing Jiaotong University, 2022, 46(2): 64-72.(in Chinese)