

车联网智能协同缓存策略: 车辆匹配、 内容推荐与资源优化

梁 琰¹, 张海君^{1,2*}, 任 超¹, 马 旭¹, 任语铮¹

(1. 北京科技大学计算机与通信工程学院, 北京 100083; 2. 北京科技大学人工智能学院, 北京 100083)

摘要: 自动驾驶、高清地图更新及车载娱乐等车联网应用的规模化落地引发数据量爆发式增长, 对缓存任务调度、内容处理推荐及网络资源分配提出了更高要求。协同缓存作为缓解车联网核心网络负载、降低内容获取端到端时延的关键技术, 其性能高度依赖车辆匹配、内容推荐与资源优化三个核心环节的深度协同, 而现有方案普遍存在车辆匹配稳定性差、推荐适配性不足、资源优化效率偏低等问题, 难以适配高动态车联网环境的需求。为了应对这些挑战, 本文从通信匹配、内容适配、资源优化三个维度展开研究, 提出一套车联网智能协同缓存策略。首先, 针对车辆高移动性导致的通信链路稳定性差、频繁切换等问题, 设计短期车道预测驱动的车辆协同匹配算法, 综合挖掘车道流量时空相关性、车辆通信速率及节点资源占用率等多维度特征, 实现需求车辆与服务车辆、路边单元的协同匹配, 为内容推荐与资源优化构建稳定的通信基础。其次, 针对传统缓存推荐策略对新老用户群体适配性不足、个性化指标优化不充分等问题, 构建融合协同过滤与内容特征的混合推荐算法, 针对新用户设计流行度-多样性双维度推荐策略解决冷启动问题, 针对老用户依据历史行为动态调整协同过滤与内容特征的推荐权重, 有效提升缓存内容与用户需求的契合度, 较基准算法使推荐命中率提升 21.11%, F1 分数等推荐指标提升 20.94%, 减少冗余数据传输与对远程云端资源的依赖。最后, 面向协同匹配关系与内容推荐分布形成的动态环境状态, 提出基于软演员评论家的资源优化算法, 通过最大化策略熵实现缓存任务分解与带宽资源分配的自适应动态决策, 较 TD3 (Twin Delayed Deep Deterministic Policy Gradient) 和 DDPG (Deep Deterministic Policy Gradient) 算法使缓存任务成本分别降低 5.80% 和 13.27%。仿真实验结果表明, 所提智能协同缓存策略在通信效率、推荐指标、资源调度经济性等方面均优于现有基准算法, 能有效适配车联网高动态特性, 为“车路云一体化”场景下的协同缓存系统设计提供了有效理论支撑。

关键词: 车联网; 协同缓存; 推荐算法; 深度强化学习; 资源优化

基金项目: 国家自然科学基金 (No.62225103, No.U22B2003, No.U2441227, No.U24A20211); 北京市自然科学基金 (No.L241008); 中央高校基本科研业务费 (No.FRF-TP-22-002C2); 小米青年学者

中图分类号: TN929.5 **文献标识码:** A **文章编号:** 0372-2112(XXXX)XX-0001-16

电子学报 URL: <http://www.ejournal.org.cn>

DOI: 10.12263/DZXB.20260065

Intelligent Cooperative Caching Strategy of Internet of Vehicles: Vehicle Matching, Content Recommendation and Resource Optimization

LIANG Yan¹, ZHANG Haijun^{1,2*}, REN Chao¹, MA Xu¹, REN Yuzheng¹

(1. School of Computer and Communication Engineering, University of Science and Technology Beijing, Beijing 100083, China;

2. School of Artificial Intelligence, University of Science and Technology Beijing, Beijing 100083, China)

Abstract: The large-scale deployment of Internet of Vehicles (IoV) applications such as autonomous driving, high-definition map updates and in-vehicle entertainment has triggered an explosive growth in data volume, imposing higher requirements on cache task scheduling, content processing and recommendation, as well as network resource allocation. As a key technology to alleviate the network load of IoV and reduce the end-to-end delay of content acquisition, cooperative caching's performance highly relies on the in-depth collaboration of three core links: vehicle matching, content recommendation and resource optimization. However, existing schemes generally suffer from poor stability of vehicle matching, insufficient recommendation adaptability and low resource optimization efficiency, making it difficult to adapt to the demands of the highly dynamic IoV environment. To address these challenges, this paper conducts research from three dimensions of communication matching, content adaptation and resource optimization, and proposes an intelligent cooperative caching strategy for IoV. Firstly, aiming at the problems of poor communication link stability and frequent handovers caused by the high mobility of vehicles, a short-term lane prediction-driven vehicle cooperative matching (SLPVC) algorithm is de-

signed. This algorithm comprehensively explores multi-dimensional features including spatiotemporal correlation of lane flow, vehicle communication rate and node resource occupancy rate, to realize the collaborative matching between demand vehicles, service vehicles and roadside units, thus establishing a stable communication foundation for content recommendation and resource optimization. Secondly, in view of the insufficient adaptability of traditional cache recommendation strategies to new and existing user groups and the inadequate optimization of personalized indicators, a hybrid recommendation based on collaborative filtering and content features(HRCF2) algorithm is constructed. For new users, a dual-dimensional recommendation strategy of popularity and diversity is designed to solve the cold-start problem; for existing users, the recommendation weights of collaborative filtering and content features are dynamically adjusted according to historical behaviors, which effectively improves the matching degree between cached content and user demands. Compared with the benchmark algorithms, this algorithm increases the recommendation hit rate by 21.11% and improves recommendation indicators such as F1-score by 20.94%, while reducing redundant data transmission and the dependence on remote cloud resources. Finally, facing the dynamic environmental state formed by collaborative matching relationships and content recommendation distribution, a soft Actor-Critic based vehicle caching tasks and resources optimization(SVCTRO) algorithm is proposed. By maximizing the policy entropy, this algorithm achieves adaptive and dynamic decision-making for cache task decomposition and bandwidth resource allocation, reducing the cache task costs by 5.80% and 13.27% respectively compared with the Twin Delayed Deep Deterministic Policy Gradient(TD3) and Deep Deterministic Policy Gradient(DDPG) algorithms. Simulation experimental results show that the proposed intelligent cooperative caching strategy outperforms the existing benchmark algorithms in terms of communication efficiency, recommendation indicators and resource scheduling economy, and can effectively adapt to the highly dynamic characteristics of IoV. It provides an effective theoretical support for the design of cooperative caching systems in the vehicle-road-cloud integration scenario.

Keywords: Internet of Vehicles; cooperative caching; recommendation algorithm; deep reinforcement learning; resource optimization

Foundation Item(s): National Natural Science Foundation of China (No.62225103, No.U22B2003, No.U2441227, No.U24A20211); Beijing Natural Science Foundation (No.L241008); Fundamental Research Funds for the Central Universities (No.FRF-TP-22-002C2); Xiaomi Fund of Young Scholar

0 引言

车联网(Internet of Vehicles, IoV)作为智能交通体系核心技术,正推动全球交通向智能化、网联化演进。根据市场研究权威机构 Mordor Intelligence 的数据,2025 年全球车联网市场规模预计达到 1 880 亿美元,至 2030 年将达到 4 010 亿美元,复合年增长率为 16.38%^[1]。车联网的应用场景日益丰富,其中自动驾驶技术的发展尤为引人注目,通过车联网技术实现车辆(Vehicle to Vehicle, V2V)、基础设施(Vehicle to Infrastructure, V2I)等通信节点间的实时通信,能够提供危险预警、交通拥堵信息及突发状况的实时反馈,可显著降低碰撞风险,提升应急响应速度与导航安全性,为保障交通安全发挥了重要作用^[2]。

在全球车联网市场蓬勃发展背景下,中国车联网产业作为国家战略性新兴产业,正在经历从技术验证到规模化应用的关键转型期,展现出强劲的发展趋势。根据中研普华研究院《2025-2030 年中国车联网行业竞争分析及发展前景预测报告》,中国车联网行业正处于高速发展期,2024 年市场规模已达 5 000 亿元,预计到 2030 年将突破 1.2 万亿元,年复合增长率达 15.8%^[3]。随着《智能网联汽车准入和上路通行试点实施指南》等相关政策措施陆续出台,中国车联网

产业已进入“车路云一体化”发展的新阶段。2024 年中国搭载 L2 级及以上自动驾驶功能的新车渗透率已达 42.4%,据中研普华预测到 2030 年中国新能源汽车渗透率将超 75%,智能网联功能将成为新车标配^[4]。

自动驾驶等 IoV 应用的数据量激增,对网络带宽、时延和数据处理能力提出了更严格的要求,使内容缓存机制成为支撑车联网大规模应用、保障用户体验的关键技术之一。对此,科研学者围绕协同缓存展开大量研究。西安电子科技大学空地一体化综合业务网全国重点实验室团队、北京邮电大学和西安邮电大学等联合研究团队提出车载边缘网络协同缓存方案,利用路边单元(Roadside Unit, RSU)等多边缘缓存节点协同,促进缓存内容及时交付,降低了传输时延和缓存成本^[5-6]。在此基础上,南京邮电大学等研究团队考虑车载自组织网络缓存资源有限,依托车辆与 RSU 协同构建缓存系统,进一步优化了缓存命中率和时延^[7-8]。由此可知,现有研究较多将车辆与 RSU 纳入缓存协同系统,以提升时延、成本、缓存命中率等性能。协同缓存的高效实现依赖于车辆协同匹配、内容推荐与资源优化三个环节:通过匹配机制实现需求车辆与服务车辆、RSU 的动态关联;通过内容推荐机制提升缓存内容与用户偏好的契合度;通过资

源优化降低任务成本并提升系统效率。这三个关键环节有效协同可显著减轻核心网络负担、降低内容获取时延,从而为自动驾驶实时决策、高清地图更新、多媒体娱乐服务等场景提供技术支持^[9-11]。

在车联网协同匹配方面,车辆移动性导致其与RSU或服务车辆等单个通信节点的通信窗口短、链路稳定性差,易引发频繁切换和传输中断问题。为此,需要引入多节点协同匹配机制,使需求车辆同时与RSU、服务车辆建立协同缓存,保障通信稳定性、降低服务时延。现有研究主要聚焦于两类匹配场景:一是车路协同,清华大学、东南大学等研究团队通过感知融合与通信协同实现车辆与RSU的匹配,但指出资源占用率与通信稳定性仍有不足^[12-13];二是车车协同,美国与加拿大等研究学者分别利用分布式启发式算法、双边拍卖方法,实现了车与车能量共享稳定匹配^[14-15]。北京邮电大学、武汉大学等团队从车辆用户关系度、连接概率及车距等因素探索匹配策略^[16-17]。尽管这些工作在匹配机制设计上取得一定进展,但现有算法多侧重于车辆距离或能量等单一因素,尚未充分融合车道时空特性、资源状态与用户需求,导致通信稳定性不足,难以满足高动态车联网环境下的实时性与可靠性要求。

在车联网内容推荐方面,随着网联汽车技术的快速发展,用户对娱乐内容的需求逐渐向个性化和定制化转变。为满足这一需求,服务器端需要根据用户兴趣动态调整缓存策略,实现更精准内容推送^[18]。当边缘RSU无法提供请求内容时,可推荐类似缓存内容,以快速交付并减轻回程负担^[19-20]。在车辆缓存内容推荐中的研究,主要围绕内容特征和用户偏好两个方面展开。一方面,基于内容特征的推荐通过分析主题、类型、格式等属性,推荐高相似度缓存资源给用户。中国、韩国、美国和英国的研究者引入了内容的流行价值矩阵,提出了时变机制来预测高度动态环境中的内容受欢迎程度,根据动态流量中的流行数据提高缓存命中率,从而降低系统成本和数据传输时延^[21-23]。另一方面,基于用户偏好的推荐更注重通过分析请求记录、交互反馈等用户历史行为,构建个性化偏好模型,以捕捉不同场景需求。文献[24-25]通过感知用户历史搜索数据与兴趣点区域类型两类要素对用户进行个性化的内容推荐,减小了需求响应时延。现有研究多侧重于优化时延、能耗等系统指标,缺乏对新老用户差异化需求的适配性,需进一步分析召回率、准确度及F1分数等个性化指标以提升车辆用户体验。

在车联网资源优化方面,自动驾驶决策、高清地图实时更新、车载多媒体流等资源密集型应用的激增对网络带宽、计算能力和存储资源提出了严苛要求,

车辆移动性与RSU覆盖受限进一步加剧资源供需失衡,这使得资源优化成为提升系统性能的重要环节。为解决动态环境适应性差、通信时延敏感性强等问题,科研学者围绕资源优化展开研究。中国台湾、中国香港和德国的研究团队针对动态车辆环境的时延敏感任务调度问题,提出了基于多网络深度强化学习实现资源分配优化^[26]。进一步地,通过内容分割技术缓解服务器负载^[27]。英国、韩国等科研人员提出基于联邦强化学习的车联网通信资源分配算法,在最大功率和信干噪比的约束下,优化车辆吞吐量^[28-29]。为激发车辆共享其计算、存储、带宽等闲置资源的积极性,WU等人^[30]提出定价驱动资源分配算法,权衡车辆收入与能源消耗,以实现高效资源分配。现有针对车联网资源优化的研究,仍缺乏对车辆高动态性所引发的环境不确定性、任务负载波动以及经济激励约束的综合建模能力,导致传统优化方法在部署中存在收敛慢、适配性差等局限。

尽管现有研究在车联网协同缓存、内容推荐与资源优化方面已取得一定进展,但仍存在以下关键问题尚未解决:(1)在车辆协同匹配方面,现有匹配算法多侧重如通信速率或距离等单一因素,未充分融合车道时空相关性与资源状态,导致通信稳定性不足;(2)在内容推荐方面,传统缓存推荐策略对新老用户群体的适配性较差,召回率、精确率、F1分数等个性化指标优化不足,难以满足车辆用户体验要求;(3)在资源优化方面,传统方法难以应对车辆高动态性带来的环境不确定性,任务分解与带宽分配效率有限,缺乏对经济激励与成本优化等多维约束的综合考量。

出于上述动机,本文在城市车联网的“车路云一体化”协同缓存发展背景下,构建了短期车道预测通信模型、车辆缓存任务时延模型与动态定价推荐满意度模型,聚焦于高动态环境下车辆协同缓存系统成本高、用户体验差的关键问题,主要贡献如下:

(1)引入短期车道轨迹预测机制,融合车道流量相关性、车辆通信速率与资源占用率,实现需求车辆与服务车辆、RSU的最佳匹配,提升通信效率与稳定性;

(2)针对新用户采用流行度与多样性推荐策略解决冷启动问题,针对老用户动态调整协同过滤与内容特征的推荐权重,结合用户历史评分与内容时空属性优化推荐列表,实现召回率、精确率、F1分数和命中率个性化性能的提升;

(3)将软演员评论家深度强化学习框架引入车辆缓存场景,通过最大化策略熵实现任务分解与带宽分配的动态优化,较TD3和DDPG对比算法任务成本分别降低5.80%和13.27%,收敛稳定性显著提升。

1 基于短期车道预测匹配的车辆协同缓存系统模型

聚焦于城市区域内的一个典型十字路口场景,该路口各主要道路均设有双向多车道,且整体交通流量处于较高运行状态。在此场景下,车辆广泛分布于道路及停车场中。同时,沿线部署配备边缘服务器的道路侧单元(Roadside Units, RSUs)为车辆提供了通信与缓存服务,系统场景如图1所示,本文主要符号含义如表1所示。

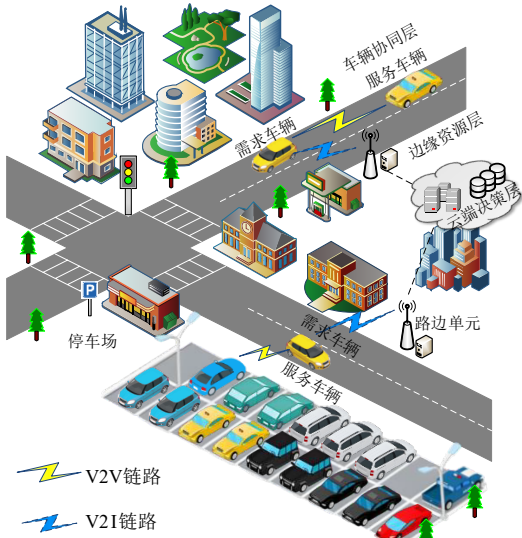


图1 车辆协同缓存系统场景示意图

Figure 1 Vehicle collaborative caching system scenario

表1 符号含义

Table 1 Symbol meaning

符号	含义
M	需求车辆数量
N	服务车辆数量
K	RSU数量
T	时隙总数
$Lc_{a,b}$	第 a 和第 b 条车道的平均相关性
$Vc_{i,j}$	DV $_i$ 与SV $_j$ 的车辆协同度
$Rc_{i,k}$	DV $_i$ 与RSU $_k$ 的资源占用率与通信速率比
$Tsi(ta)$ 、 $Tpc(ta)$	缓存任务 ta 文件的大小、处理复杂度
$R_{i,jk}(t)$	时隙 t 内DV $_i$ 与SV $_j$ /RSU $_k$ 的传输速率
bw^{occ}	带宽资源占用率

根据车辆在资源共享过程中的不同作用,将车辆划分为服务车辆(Service Vehicles, SVs)与需求车辆(Demand Vehicles, DVs)两类。其中,SV的范畴涵盖了正在道路上行驶的出租车、网约车,以及停放在停车场内愿意提供缓存服务的车辆。DV利用车对车

(Vehicle to Vehicle, V2V)通信方式向SV发起请求,以获取缓存的资源文件。此外,DV可以通过车对路侧基础设施发起(Vehicle to Roadside unit, V2R)通信,向RSU提出缓存请求,其配备的处理单元与存储资源比DV更为强大。若SV和RSU的服务器均没有DV所需资源文件,则需向云端发起请求,这将消耗更多的时延与能源。

本车辆协同缓存系统架构可分为以下三层。

车辆协同层:涵盖 $\mathcal{M}=\{1,2,\dots,i,\dots,M\}$ 个DV s 以及 $\mathcal{N}=\{1,2,\dots,j,\dots,N\}$ 个SV s 。其预先缓存了一些流行资源文件,使DV s 能够对这些缓存资源发起请求。

边缘资源层:由 $\mathcal{K}=\{1,2,\dots,k,\dots,K\}$ 个RSUs组成。各RSU依据各区域的缓存流行偏好来存储不同的资源,边缘侧还决定DV s 与SV s 间的成对缓存关系。此外,通过推荐算法,RSU能够引导DV缓存相似资源。

云端决策层:包含所有的缓存资源。依据区域中的时间和空间相关性偏好,作出缓存资源的放置和更新的全局性决策。

1.1 短期车道预测通信模型

在时间段 t ,设采样车辆数量为 $M+N$ 、平均到达值为 $\varepsilon(t)$,区域网络内车辆用户的用户数服从泊松分布,则概率密度函数为

$$Pb\{X=(M+N)\} = \frac{\varepsilon(t)^{(M+N)}}{(M+N)!} e^{-\varepsilon(t)} \quad (1)$$

在道路上行驶的车辆具有车道属性,车道属性可以显示车辆在接近交叉路口时的短期轨迹。在为SV匹配DV时,优先考虑具有相同或相似短期轨迹的车辆,这可以增强用于缓存任务的车辆间连接的持续时间和稳定性。

分析区域中在第 t 个时隙中第 a 条车道的相关性,它们形成的矩阵 C_r 为

$$C_r = \begin{bmatrix} l_1^1 & l_1^2 & \dots & l_1^T \\ l_2^1 & l_2^2 & \dots & l_2^T \\ \vdots & \vdots & \ddots & \vdots \\ l_A^1 & l_A^2 & \dots & l_A^T \end{bmatrix} \quad (2)$$

其中, $l_a^t, t \in [1, T], a \in [1, A]$ 代表在第 t 个时隙中第 a 条车道的交通流量。

交通流数据中可能存在不同的计算单位,这些不同的计算单位很难得出正确一致的结论。为了消除这种对数据的不利影响,采用零均值归一化法对问题进行无量纲化。处理后的矩阵 C'_r 为

$$C'_r = \frac{C_r - C_{r_c}}{C_{r_m}} = \begin{bmatrix} l_1^1 & l_1^2 & \dots & l_1^T \\ l_2^1 & l_2^2 & \dots & l_2^T \\ \vdots & \vdots & \ddots & \vdots \\ l_A^1 & l_A^2 & \dots & l_A^T \end{bmatrix} \quad (3)$$

其中, $C_{r_c}=[v_1, \dots, v_A]$ 和 $C_{r_m}=[m_1, \dots, m_A]$ 分别是每个车

道的交通数据的方差和平均值。经处理的交通流数据的总样本方差值为1,平均值为0。

将车道数据 $l_a = [l_a^1, \dots, l_a^T]$ 作为参考序列,车道数据 $l_b = [l_b^1, \dots, l_b^T]$ 用作比较序列。在时隙 t 中任意两条车道 l_a 和 l_b 间的相关系数 ϑ 表示为

$$\vartheta_{a,b}(t) = \frac{\min_b \min_t \Delta(t) + \alpha \max_b \max_t \Delta(t)}{\Delta(t) + \alpha \max_b \max_t \Delta(t)} \quad (4)$$

$$\Delta(t) = |l_a^t - l_b^t| \quad (5)$$

其中, $\alpha \in (0, 1)$ 是用于区分不同序列之间相关程度的分辨率系数。

任意两车道 l_a 和 l_b 间的相关度是在每个时隙计算的相关系数 $\vartheta_{a,b}(t)$ 的平均值,则它们的平均相关性 $Lc_{a,b}$ 可表示为

$$Lc_{a,b}(t) = \frac{1}{T} \sum_{t=1}^T \vartheta_{a,b}(t) \quad (6)$$

行驶在道路上的 DV*i* 需要缓存资源,停车场的 SV*j* 和 RSU*k* 被分配为其服务。通过在平面交通地图上创建笛卡尔坐标系,地图上 i 和 j/k 间的欧几里得距离为

$$d_{i,j/k}(t) = \sqrt{(x_{j/k}(t) - x_i(t))^2 + (y_{j/k}(t) - y_i(t))^2} \quad (7)$$

进一步地,结合车道流量历史数据计算两车所在车道的时空相关性,综合表征车辆间协同度的表达式为

$$Vc_{i,j}(t) = \frac{Lc_{a,b}(t)}{d_{i,j}(t)} \quad (8)$$

其中,分子 $Lc_{a,b}(t) \in [0, 1]$ 表示车道相关性,值越接近1,表示车道相关性越高。分母 $d_{i,j}(t)$ 表示距离的加权因子,值越大,相关性越小,但是不能超过距离阈值 d_{th} 。 $\omega_c(\gamma_{i,j}(t))$ 表示以 DV*i* 和 SV*j* 之间信号干扰加噪声比 (Signal Interference plus Noise Ratio, SINR) 为自变量的权重。鉴于车辆能够配对传输任务的时间整体较短,这种车辆协同相关性有利于找到与 DV 行程路线类似的 SV,提高通信稳定性。

设定 cg_0 是指初始参考距离 d_0 处的信道增益, ω 代表具有指数分布的随机变量, α^{pl} 代表信道路径损耗指数。对于远处停车场的 RSU 和 SV 而言, DV*i* 和 SV*j* 或 RSU*k* 在时隙 t 的信道增益表达式为

$$cg_{i,j/k}(t) = cg_0 \omega^2 d_{i,j/k}(t)^{-\alpha^{pl}} \quad (9)$$

当 DV*i* 与 RSU*k* 通信时,需要考虑同时与它通信的其他车辆用户的信道干扰。设 pw_i 和 pw_b 分别代表 DV*i* 和干扰车辆 b 的发射功率, $cg_{i,k}$ 和 $cg_{b,k}$ 分别是 DV*i* 和其他车辆 b 到 RSU*k* 的信道增益, σ^2 代表高斯白噪声。DV*i* 和 RSU*k* 之间的 SINR 的表达式为

$$\gamma_{i,k}(t) = \frac{pw_i(t)cg_{i,k}(t)}{\sum_{b=1}^B pw_b(t)cg_{b,k}(t) + \sigma^2} \quad (10)$$

相似地, DV*i* 与 SV*j* 的 SINR 为

$$\gamma_{i,j}(t) = \frac{pw_i(t)cg_{i,j}(t)}{\sigma_v^2} \quad (11)$$

由此,它们之间的传输速率为

$$R_{i,j/k}(t) = W_{i,j/k} \log_2(1 + \gamma_{i,j/k}(t)) \quad (12)$$

其中, DV*i* 与 SV*j* 间的带宽,以及 DV*i* 与 RSU*k* 间的带宽分别为 $W_{i,j}$ 和 $W_{i,k}$ 。

1.2 车辆缓存任务时延模型

接下来,计算 DV*i* 在协同系统内完成缓存任务 ta 的时延。

(1) 在车道上行驶的 DV*i* 产生缓存需求,向 RSU*k* 请求资源,该过程中产生的请求时延表示为

$$t_{i,k}^{req}(ta) = \frac{Tpc(ta)Tsi(ta)}{R_{i,k}(t)} \quad (13)$$

其中, Tpc 和 Tsi 分别代表任务的处理复杂度和文件大小。

(2) 若配对的 RSU 或 SV 没有相应的资源,系统可以向 DV 推荐类似的缓存资源。推荐系统算法的处理时延为

$$t_{i,k}^{proc}(ta) = \frac{Tpc(ta)Tsi(ta)}{Pr_k} \quad (14)$$

其中, Pr_k 表示 RSU*k* 的 CPU 处理能力。

(3) 从 SV、RSU 或远端到 DV 的下载时延。为了降低对远程云资源下载的依赖,采用 Zipf 模型来优化缓存策略,提高缓存系统的命中率。引入 Zipf 模型,可根据文件的流行度分布来优化缓存中的数据分布,使缓存系统更有效地存储和快速访问热门数据,更适合大众需求。在时隙 t 中,内容 c 的 Zipf 表达式为

$$Z_c(t) = \frac{s_c^{-\varphi}}{\sum_{c=1}^c s_c^{-\varphi}} \quad (15)$$

其中, s_c 表示内容 c 在整个内容库中的请求频率排序索引, $\varphi \in [0, 1]$ 为流行度参数, φ 值越大,则内容 c 被请求的概率越高。此外,设计相关推荐系统算法,通过分析用户行为和内容特征,能够进一步优化缓存资源的分配,提高系统的整体性能。

资源文件从上层传输到 DV 的时延,需要根据具体情况进行讨论。若 SV 或 RSU 有 DV 需要的资源文件,或者 DV 接受了推荐的资源,就会出现三种情况。第一种情况,只有 SV*j* 向 DV*i* 提供缓存资源,时延表达式为

$$t_{j,i}^{dl}(ta) = \frac{Tsi(ta)}{R_{j,i}(t)} \quad (16)$$

第二种情况,仅 RSU*k* 为 DV*i* 提供了缓存资源,则时延表达式为

$$t_{k,i}^{dl}(ta) = \frac{Tsi(ta)}{R_{k,i}(t)} \quad (17)$$

第三种情况, SV j 和 RSU k 都为 DV i 提供了缓存资源, 时延取决于两者较长的一方, 表达式为

$$t_{jk,i}^{dl}(ta) = \max \left\{ \frac{S_j Tsi(ta)}{R_{j,i}(t)}, \frac{S_k Tsi(ta)}{R_{k,i}(t)} \right\} \quad (18)$$

其中, S_j 和 S_k 代表从 SV 和 RSU 下载的文件片段。

此外, 若 SV j 和 RSU k 均没有 DV i 所需的资源文件, 并且 DV 不接受推荐的资源文件, 则会向远程云请求资源。RSU 从云端下载并传输到 DV, 这会产生较高时延, 其表达式为

$$t_{cloud,i}^{dl}(ta) = \frac{Tsi(ta)}{R_{cloud,k}(t)} + t_{k,i}^{dl}(ta) \quad (19)$$

(4) DV 完成了缓存的文件段后, 需要在本地合并。文件段合并的时延表达式为

$$t_i^{mge}(ta) = \frac{Tpc(ta)Tsi(ta)}{Pr_i} \quad (20)$$

其中, Pr_i 表示 DV i 的 CPU 处理能力。

因此, 整个缓存任务由四个部分组成: (1) DV 提交缓存请求和初始服务器处理; (2) 推荐系统处理; (3) 将资源下载到 DV; (4) DV 在本地合并文件。则 DV i 完成缓存任务 ta 的总时延为

$$t_i^{total}(ta) = t_{i,k}^{req}(ta) + \zeta t_{i,k}^{proc}(ta) + t_i^{dl}(ta) + t_i^{mge}(ta) \quad (21)$$

其中, $\zeta \in [0, 1]$ 表示系统推荐的文件服务的推荐指数。

1.3 动态定价和推荐满意度模型

综合分析 SV 和 RSU 等缓存服务提供方的贡献, 为其构建了资源定价模型, 保障其基本收益和服务意愿。在缓存服务的过程中, 主要占用服务方的通信资源, 考虑资源占用率以及带宽资源单价, 定义定价函数表达式为

$$P_{bw}^{unit} = P^{init} PW_{bw} \quad (22)$$

$$PW_{bw} = \begin{cases} \varpi_a \frac{bw^{occ}}{OR_{th}}, & \text{if } bw^{occ} \leq OR_{th} \\ \varpi_a \left(\frac{bw^{occ} - OR_{th}}{1 - OR_{th}} + 1 \right), & \text{if } bw^{occ} > OR_{th} \end{cases} \quad (23)$$

其中, P^{init} 为资源基准初始单价, PW_{bw} 为带宽资源动态定价权重函数。 $bw^{occ} \in [0, 100\%]$ 是带宽资源的占用率, $OR_{th} \in (0, 100\%)$ 是占用率阈值, ϖ_a 是定价调节参数, 这里默认取值为 1。通过两段函数的动态定价机制, 可实现以阈值 OR_{th} 为临界点, 占用率 $bw^{occ} \in [0, OR_{th}]$ 时, 定价权重 PW_{bw} 取值范围为 $[0, \varpi_a]$; 占用率 $bw^{occ} \in (OR_{th}, 100\%]$ 时, PW_{bw} 取值范围为 $(\varpi_a, 2\varpi_a]$; 总体取值范围为 $[0, 2\varpi_a]$ 。这样, 若某节点资源出现闲置状况, 较低的单价吸引 DV 优先选用该节点; 反之, 繁忙状态时较高的单价会促使 DV 选择其他节点, 有利于系统的负载平衡。此外, 该模

型确保 SV 和 RSU 在提供缓存服务的同时, 获得合理的经济回报, 提高服务积极性。

根据带宽资源占用率和任务持续时间, SV j 和 RSU k 的收益函数定义为

$$BF_j(ta) = (P_{bw}^{unit} W_j - P_j^e) t_{j,i}^{dl}(ta) \quad (24)$$

$$BF_k(ta) = (P_{bw}^{unit} W_k - P_k^e) t_{k,i}^{dl}(ta) \quad (25)$$

其中, W_{jk} 是带宽大小, P_{jk}^e 是资源成本。

为确保推荐内容的用户体验, 采用召回率 (Recall)、精确率 (Precision)、F1 分数和命中率 (Hit-Rate) 等指标来评估用户对推荐资源的满意度。召回率衡量推荐系统找到用户感兴趣的相关内容的的能力, 即推荐系统能够推荐的用户感兴趣的所有资源文件的百分比。则 DV i 的召回率表达式为

$$Rc_i = \frac{recommended_interest_num_i}{all_interest_num_i} \quad (26)$$

精确率是指推荐给用户的资源文件中, 用户真正感兴趣的文件所占的比例。则 DV i 的精确率表达式为

$$Prec_i = \frac{interest_num_i}{all_recommended_num_i} \quad (27)$$

F1 分数是精确率和召回率的调和平均值, 是两者之间的折中值。则 DV i 的 F1 表达式为

$$F1_i = \frac{2Rc_i Prec_i}{Rc_i + Prec_i} \quad (28)$$

命中率表示推荐算法给用户的推荐内容列表中至少有一个用户感兴趣内容的比例。则算法的命中率表达式为

$$HR = \begin{cases} 1, & interest_num \geq 1 \\ 0, & interest_num = 0 \end{cases} \quad (29)$$

则系统内完成车辆缓存任务的带宽资源成本最小化问题的表达式为

$$\begin{aligned} \mathbf{P1:} \quad & \min_{S_{jk}, W_{jk}} \sum_{i=1}^M \sum_{ta=1}^{Ta} (P_j^{unit} t_{j,i}^{dl}(ta) + P_k^{unit} t_{k,i}^{dl}(ta)) \\ \text{s.t.} \quad & \mathbf{C1:} \quad S_j + S_k = 1, S_{jk} \in [0, 1] \\ & \mathbf{C2:} \quad W_{i,j} + W_{i,k} \leq W_i^{total}, \forall i \in M, \forall j \in N, \forall k \in K \\ & \mathbf{C3:} \quad t_i^{total}(ta) \leq Td, \forall ta \in Ta \\ & \mathbf{C4:} \quad Tsi(ta) \leq disk_{ijk}^{max}, \forall i \in M, \forall j \in N, \forall k \in K \\ & \mathbf{C5:} \quad OR_{jk} \leq OR_{th}, \forall j \in N, \forall k \in K \\ & \mathbf{C6:} \quad BF_{jk}(ta) \geq BF^{base}, \forall ta \in Ta \end{aligned} \quad (30)$$

其中, C1 确保来自 SV S_j 和 RSU S_k 文件分段的完整性。C2 表示当 DV i 的同时接收 SV j 和 RSU k 的缓存服务时, 总带宽是有限的。C3 表示完成任务的总时延不能超过可容忍时延, C4 和 C5 分别表示缓存容量、资源占用率有上限。C6 确保 SV j 和 RSU k 都有不小于基本值的回报, 维持其服务意愿。

2 车辆智能协同缓存成本优化策略

分析问题 P1 表达式得知,有两种减少成本的关键途径:一是降低带宽资源单价 P^{unit} ,二是缩短缓存文件的下载时延 t^{dl} 。一方面,降低带宽资源单价的核心在于优先选择带宽资源占用率较低的 RSU 和 SV,从而在后续资源分配上实现更高的性价比。另一方面,通过内容推荐算法降低远端云下载带来的高时延,并提出有效的任务分解和资源分配算法,以优化 RSU 和 SV 的下载时延,实现缓存系统中的时延有效控制。

2.1 短期车道预测驱动的车辆协同匹配

由短期车道预测通信模型可知,以式(2)交通流量数据矩阵 C_r 作为输入,记录了区域内各个车道在不同时隙的交通流量情况。为了消除不同计算单位对数据的影响并进行无量纲化处理,采用零均值归一化方法对矩阵 C_r 进行处理,得到了均值为 0,方差为 1 的式(3)归一化矩阵 C_r' 。

然后,以归一化矩阵 C_r' 为基准,根据式(4)和式(5),计算任意两个车道 l_a 和 l_b 在每个时隙 t 的相关性系数 $\rho_{a,b}(t)$,该系数通过量化两车道流量差值的相对程度来衡量两车道间的相似性。通过平均处理所有时隙内的相关性系数,得到式(6)表示的车道间平均相关性 $Lc_{a,b}$,实现量化两车道间的整体相关性。

在此基础上,基于车辆位置信息建立平面交通笛卡尔坐标系,根据式(7)计算出的 DV_i 与 SV_j 或 RSU_k 之间的欧几里得距离 $d_{i,j/k}(t)$ 。结合车道流量历史数据计算两车所在车道的时空相关性,综合表征式(8)所示的车辆间协同度 $Vc_{i,j}(t)$,选取值最大的车辆作为匹配的 SV。

此外,RSU 的匹配目标是降低成本。在问题 P1 中,成本是由资源单价与时延相乘得出,实质上是 RSU 服务器的资源占用率与传输速率的倒数之积。该值越小,成本越低,选择取值最小者为匹配的 RSU。则 DV_i 与匹配的 RSU_k 资源占用率与通信速率比例表达式为

$$Rc_{i,k}(t) = \frac{bw_k^{\text{occ}}(t)}{R_{i,k}(t)} \quad (31)$$

综上所述,短期车道预测驱动的车辆协同匹配(Short-term Lane Prediction-driven Vehicle Cooperative Matching, SLPVCM)算法如算法 1 所示。

2.2 基于协同过滤的车辆缓存内容推荐

如果 DV 匹配的 RSU 服务器缺乏所需的缓存内容,就必须从远端云中寻找资源。该过程会导致任务时延的显著增加,影响车辆用户体验。为解决这个问题,提出了基于协同过滤和内容特征的混合推荐(Hybrid Recommendation based on Collaborative Filtering

and Content Features, HRCF2)算法,综合了车辆用户画像和资源内容特征,更好地满足用户体验要求,如算法 2 所示。

算法 1 短期车道预测驱动的车辆协同匹配(SLPVCM)算法

输入:区域各车道在 T 个时隙内流量数据组成的矩阵 C_r

输出:为 DV 序列 $\{DV_1, DV_2, \dots, DV_M\}$ 匹配的 SV 序列

$\{SV_1, SV_2, \dots, SV_M\}$ 、RSU 序列 $\{RSU_1, RSU_2, \dots, RSU_M\}$

1. 初始化式(2)矩阵 C_r ,存储区域各车道在 T 个时隙内流量数据

2. 根据式(3),对矩阵 C_r 进行零均值归一化处理得到矩阵 C_r'

3. FOR $a = 1$ to $A - 1$

4. FOR $b = a + 1$ to A

5. 根据式(4)和式(5),计算车道间的相关系数矩阵 $\rho_{a,b}(t)$

6. 根据式(6),计算车道间的平均相关度 $Lc_{a,b}(t)$

7. 对称赋值,填充空值: $Lc_{b,a}(t) \leftarrow Lc_{a,b}(t)$

8. END FOR

9. END FOR

10. 在平面交通图上建立笛卡尔坐标系,获取 t 时隙下 DV_i 、 SV_j 和 RSU_k 的坐标

11. FOR $i = 1$ to M

12. 根据式(7),计算 DV_i 和 SV_j/RSU_k 之间的欧几里得距离

13. 根据式(8),计算 DV_i 和 SV_j 的车辆间协同指标 $Vc_{i,j}(t)$,选取协同指标最大的车辆作为 DV_i 匹配的 SV

14. 根据式(31),计算 DV_i 和 RSU_k 的匹配指标 $Rc_{i,j}(t)$,取值最小者为 DV_i 匹配的 RSU

15. END FOR

算法 2 基于协同过滤和内容特征的混合推荐算法(HRCF2)算法

输入:DV 数量 M 、内容库 C

输出:为 DV 推荐的内容序列 $RList$

1. 初始化 DV 和内容信息

2. FOR $i = 1$ to M

3. IF DV_i 是新用户 THEN

4. 利用流行度和多样性策略选择访问量排名前列的内容资源

5. 从每个主类型中选择最高频的数个资源

6. 根据式(32),生成推荐列表 $RList_i$

7. ELSE

8. 根据式(33)和(34),计算混合推荐算法的动态权重 w_{cf} 、 w_{cb}

9. 由式(36)和式(37),计算预测得分 $CFps_{ui}^c$ 、 $CBps_{ui}^c$

10. 根据式(38),计算用户给内容的综合得分 $ComSco_i^c$

11. 根据式(39),生成推荐列表 $RList_i$

12. END IF

13. END FOR

首先,在服务器数据库中初始化用户和内容的时、位置和分数。接着,为当前区域 RSU、SV 通信覆盖范围内的每个 DV,创建内容的时间、位置、主题和类型等内容特征的偏好字典。对于在历史时隙中对内容资源的请求次数较少的新用户 DV_i ,利用流行度和多样性的推荐策略。选择访问量排名前列的内容

资源,进一步从每个主类型中选择最高频的数个资源,有效解决冷启动问题。则为新用户 DV_i 推荐的 x 条内容的表达式为

$$RList_i = \text{sort}\{\text{type1}^{\text{top}}, \text{type2}^{\text{top}}, \dots, \text{typex}^{\text{top}}\} \quad (32)$$

若 DV_i 是内容资源的请求次数较多的老用户,计算基于协同过滤和内容特征的混合推荐算法的动态权重为

$$w_{cf} = \min\left(0.7, 0.3 + \frac{\text{avgs}_i}{10}\right) \quad (33)$$

$$w_{cb} = 1 - w_{cf} \quad (34)$$

其中, $\text{avgs}_i \in [0, 5]$ 是 DV_i 的历史平均评分, 这样使 $w_{cf} \in [0.3, 0.7]$, 用户历史评分越高, 协同过滤权重越大, 表明对该 DV 的个性化推荐更依赖协同过滤。

在用户协同过滤推荐中, 采用余弦相似度, 通过基于邻居评分偏差的加权平均预测 DV 偏好。具体来说, 通过比较 DV_i 与邻居 DV_{i+1} 的兴趣向量之间的夹角大小, 即两个向量的内积与各自模积的比值, 以量化用户之间的相似程度。设两个 DV 的兴趣评分向量分别为 \mathbf{u}_i 和 \mathbf{u}_{i+1} , 则相似度的计算表达式为

$$\text{sim}(\mathbf{u}_i, \mathbf{u}_{i+1}) = \frac{\mathbf{u}_i \cdot \mathbf{u}_{i+1}}{\|\mathbf{u}_i\| \|\mathbf{u}_{i+1}\|} \quad (35)$$

选取与 DV_i 最相似的 I 个邻居用户, 设 $ps_{u_i}^c$ 是用户 \mathbf{u}_i 对内容 c 的预测得分, $hs_{u_{i+1}}^c$ 是用户 \mathbf{u}_{i+1} 对内容 c 的历史得分, $\text{sim}(\mathbf{u}_i, \mathbf{u}_{i+1})$ 是两个用户之间的相似度。以 DV_i 的历史平均评分 avgs_i 为基准项, 减少不同用户的评分标准的系统性偏差, 则用户协同过滤评分预测得分 $CFps_{u_i}^c$ 表达式为

$$CFps_{u_i}^c = \text{avgs}_i + \frac{\sum_{i=1}^I (hs_{u_{i+1}}^c - \text{avgs}_{i+1}) \cdot \text{sim}(\mathbf{u}_i, \mathbf{u}_{i+1})}{\sum_{i=1}^I \text{sim}(\mathbf{u}_i, \mathbf{u}_{i+1})} \quad (36)$$

在基于时间位置等偏好的内容推荐中, 内容 c 特征矩阵包括时间 st 、地点 sl 、类型 sty 、主题 sto 等属性, 采用余弦相似度预测用户 \mathbf{u}_i 对内容 c 的预测得分 $CBps_{u_i}^c$, 表达式为

$$CBps_{u_i}^c = \frac{\mathbf{u}_i \cdot \mathbf{c}}{\|\mathbf{u}_i\| \|\mathbf{c}\|} \quad (37)$$

结合用户协同过滤推荐和基于时间位置等偏好的内容推荐, 计算得出加权综合得分, 则 DV_i 给内容 c 的综合得分表达式为

$$\text{ComSco}_i^c = w_{cf} CFps_{u_i}^c + w_{cb} CBps_{u_i}^c \quad (38)$$

根据综合得分对内容进行排序, 为 DV_i 推荐得分最高的 x 条内容, 表达式为

$$RList_i = \text{sort}\{rc_1, rc_2, \dots, rc_x\} \quad (39)$$

该推荐算法优化了用户的时间位置偏好加权策略, 允

许用户接受相似信息推荐并跳远程下载过程, 消除了远端云的下载时延。

2.3 基于SAC的车辆缓存任务资源优化

通过短期车道预测驱动的车辆协同匹配算法和基于时间位置偏好的车辆内容推荐算法, 分别从降低带宽资源单价和跳远程下载时延两个方面, 有效地减少了车辆协同缓存系统的总体成本。为了进一步优化RSU和SV的下载时延, 实现车辆协同缓存系统的下载途径、带宽资源用量控制, 需要提出一种有效的任务分解和资源分配算法。

在车辆缓存场景中, 车辆的动态性、缓存资源的有限性以及通信环境的复杂性使得传统的优化方法难以有效应对这种动态多变的环境。为了克服这些挑战, 引入深度强化学习(Deep Reinforcement Learning, DRL)技术, 利用DRL的自适应学习能力和环境交互决策能力, 能够动态地调整策略以应对复杂环境中的不确定性, 实现任务分解和资源分配的优化。

首先, 将车辆协同缓存系统建模为一个包含 $\langle S, A, R \rangle$ 的马尔可夫决策过程(Markov Decision Process, MDP), 其中 S 是系统代理的状态空间, A 是代理每回合的行动空间, R 是行动完成后环境给予系统代理的回报。

代理: 该车辆协同缓存系统中的 DVs 被视为一个代理。

状态空间: 定义为时隙比例和剩余缓存任务量比例。其中, 时隙比例是当前时隙 t^{cur} 与最大时隙 t^{max} 的比值; 剩余缓存任务量比例是每个 DV 的剩余缓存任务量 $csize^{\text{res}}$ 与总缓存任务量 $csize^{\text{total}}$ 的比值。使用归一化将时隙、任务量缩放到 $[0, 1]$ 范围, 使得之具有相同的量纲, 有助于DRL算法更稳定地处理状态信息, 避免因变量量纲差异导致的学习困难。如果时隙比例升为1, 说明用完时间没有在可容忍时延内完成缓存任务, DRL训练循环终止。在此过程中需要考虑问题P1的时延约束C3和缓存容量约束C4, 存在 M 个 DV 参与缓存, 则状态空间的表达式为

$$S = \left\{ \frac{t_1^{\text{cur}}}{t_1^{\text{max}}}, \frac{csize_1^{\text{res}}}{csize_1^{\text{total}}}, \dots, \frac{t_i^{\text{cur}}}{t_i^{\text{max}}}, \frac{csize_i^{\text{res}}}{csize_i^{\text{total}}}, \dots, \frac{t_M^{\text{cur}}}{t_M^{\text{max}}}, \frac{csize_M^{\text{res}}}{csize_M^{\text{total}}} \right\} \quad (40)$$

动作空间: 定义为每个时隙每个 DV 的RSU任务分配比例、SV任务分配比例、RSU带宽分配比例。其中, RSU和SV任务分配比例是每个 DV 分别从RSU和SV获取缓存任务的比例, 范围为 $[0, 1]$, 两者总量需要满足文件分段的完整性约束C1。RSU带宽分配比例范围为 $[0, 1]$, 满足带宽总量约束C2。需要注意的是, 缓存任务需要多个时隙长期完成, 每个时隙内RSU

和SV任务分配比例之和不为1,全部时隙的两者累计之和才为1,故需要在动作空间单独设计两者的比例;每个时隙的带宽分配比例总和为1,无需设计两者的分配比例。则动作空间的表达式为

$$A = \{rtask_1, stask_1, rbw_1, \dots, rtask_i, stask_i, rbw_i, \dots, rtask_M, stask_M, rbw_M\} \quad (41)$$

奖励函数:环境会根据代理的当前状态和执行的行动提供反馈奖励函数。一般地,奖励函数越大越好。鉴于优化目标为降低系统缓存任务的成本,如果缓存任务在可容忍的时延内完成,且满足资源占用率约束C5和基本回报约束C6,则将奖励的正反馈设为优化目标的倒数,再乘以一个正数放大奖励;否则,设置一个较大负值作为惩罚函数。则奖励函数为

$$R = \begin{cases} \frac{100}{\sum_{i=1}^M \sum_{td=1}^{Ta} (P_j^{unit} t_{j,i}^{dl}(ta) + P_k^{unit} t_{k,i}^{dl}(ta))}, & \text{满足约束} \\ -1000, & \text{不满足约束} \end{cases} \quad (42)$$

设置状态空间序列为 $\{S_1, S_2, \dots, S_t, \dots, S_T\}$,动作空间序列为 $\{A_1, A_2, \dots, A_t, \dots, A_T\}$ 。奖励值用于评估在状态 S_t 下,执行行动 A_t ,到达下一个状态 S_{t+1} 的有利程度。引入折扣因子 φ 来表示当前时刻与未来时刻的不同奖励权重,如果其值接近0,则更关注当前奖励,反之则更关注未来奖励。则累积奖励函数表示为

$$R_t^{cum} = R_{t+1} + \varphi R_{t+2} + \dots + \varphi^{T-t-1} R_T \quad (43)$$

传统优化方法难以应对车辆协同缓存系统的动态性、状态空间高维性及复杂优化目标与约束。深度强化学习具有自适应学习和环境交互决策能力,能够应对复杂环境中的不确定性。引入软演员评论家(Soft Actor-Critic, SAC)算法,以实现任务分解和资源分配的优化。相较于传统的DRL算法,SAC算法通过最大化累积奖励的同时最大化策略熵,能够更有效地探索策略空间,避免陷入局部最优解,并在复杂的动态环境中展现出更优异的稳定性和收敛性,更适用于具有连续状态和动作空间的车联网场景。

SAC算法框架主要包含策略(Actor)网络和双价值(Critic)网络、目标网络和温度参数。其中,Actor网络负责根据车辆的缓存需求、通信环境状态生成获取资源途径的动作,Critic网络用于评估缓存任务的完成时延、资源分配的合理性以及系统的整体成本策略的优劣。同时,引入目标网络保持对Actor网络和Critic网络的平滑跟踪,缓解车辆动态环境中网络更新带来的振荡问题,确保训练过程的稳定性。温度参数用来控制策略的探索性,避免算法过早陷入局部最优解。

所提出基于SAC的车辆缓存任务资源优化(SAC-

based Vehicle Caching Tasks and Resources Optimization, SVCTRO)算法,如算法3所示。其核心目标为最大化累积奖励与策略熵的加权和,表达式为

$$J(\pi) = \sum_{t=0}^T \mathbb{E}_{(S_t, A_t) \sim p_\pi} [\varphi R(S_t, A_t) + \tau \mathcal{H}(\pi(\cdot|S_t))] \quad (44)$$

其中, τ 为温度参数,控制策略熵 $\mathcal{H}(\pi)$ 的权重,确保策略在收敛过程中保持足够的探索性。

算法3 基于SAC的车辆缓存任务资源优化(SVCTRO)算法

输入:DV数量、缓存任务量等车辆环境参数

输出:优化后的任务分解和带宽分配比例序列

$\{\dots, rtask^*_i, stask^*_i, rbw^*_i, \dots\}$

1. 初始化Actor网络、Critic网络、目标网络和温度参数
 2. FOR *episode* = 1 to max_ep
 3. 由DV数量、缓存任务量等参数初始化当前状态 S_0
 4. FOR *t* = 1 to *T*
 5. 从Actor网络采样任务分解、带宽分配比例等动作
 6. 根据式(45)给动作添加探索噪声
 7. 执行动作 A_t ,获取下一个状态 S_{t+1} ,根据式(42)计算奖励
 8. 将经验数据 S_t, A_t, R_t, S_{t+1} 存入经验池 D 中
 9. IF D 存满 THEN
 10. 从 D 中随机采样批量 S_t, A_t, R_t, S_{t+1}
 11. 根据式(46)和式(47),计算Critic网络目标 Q 值损失函数
 12. 根据式(48),通过梯度下降更新Critic网络的 ϕ_1, ϕ_2
 13. 根据式(49)和式(50),计算Actor网络策略熵、损失函数
 14. 根据式(51),通过梯度下降更新Actor网络的 θ
 15. 根据式(52)和式(53),计算温度损失,更新温度参数 τ
 16. 根据式(54)和式(55),软更新目标网络参数 $\theta', \phi'_1, \phi'_2$
 17. END IF
 18. 更新当前状态 $S_t \leftarrow S_{t+1}$
 19. END FOR
 20. END FOR
-

首先,初始化Actor网络参数 $\pi_\theta(A|S_t)$ 、双Critic网络参数 $Q_{\phi_1}(S_t, A_t), Q_{\phi_2}(S_t, A_t)$ 和目标网络参数 $\pi'_\theta(A|S_t), Q'_{\phi_1}(S_t, A_t), Q'_{\phi_2}(S_t, A_t)$ 、温度参数 τ ,以及经验回放池 D 。

然后,在每个训练周期内,初始化当前状态 S_0 。在每个时隙 t 内从Actor网络采样动作,添加探索噪声 δ 后执行动作,获取下一个状态 S_0 并计算奖励 R_t 。其中,添加探索噪声 ϵ 后,动作的表达式为

$$A_t \sim \pi_\theta(A|S_t) + \epsilon \quad (45)$$

每时隙将经验数据 S_t, A_t, R_t, S_{t+1} 存入经验池 D 中,当经验池存满后,从池中随机采样一批数据。计算Critic网络的目标 Q 值和损失函数,表达式分别为

$$Y_t = R_t + \varphi' \min(Q_{\phi_1}(S_{t+1}, A_{t+1}), Q_{\phi_2}(S_{t+1}, A_{t+1})) \quad (46)$$

$$L_{\phi_x} = \frac{1}{N} \sum_{t=1}^N [Y_t - Q_{\phi_x}(S_t, A_t)]^2, x=1, 2 \quad (47)$$

设 η_ϕ 为 Critic 网络的学习率, 则通过梯度下降更新 Critic 网络的参数 ϕ_1, ϕ_2 表达式为

$$\phi_x \leftarrow \phi_x - \eta_\phi \nabla L_{\phi_x}, x=1, 2 \quad (48)$$

同时, 计算 Actor 网络的策略熵和损失函数, 表达式分别为

$$\mathcal{H}(\pi_\theta) = -\mathbb{E}_{A \sim \pi_\theta} [\log \pi_\theta(A|S_t)] \quad (49)$$

$$L_\theta = \frac{1}{N} \sum_{t=1}^N [\tau \log \pi_\theta(A|S_t) - \min(Q_{\phi_1}(S_t, A_t), Q_{\phi_2}(S_t, A_t))] \quad (50)$$

设 η_θ 为 Actor 网络的学习率, 通过梯度下降更新 Actor 网络的参数 θ :

$$\theta \leftarrow \theta - \eta_\theta \nabla L_\theta \quad (51)$$

计算温度损失函数为

$$L_\tau = \frac{1}{N} \sum_{t=1}^N [\tau (\mathcal{H}_0 - \mathcal{H}(\pi_\theta))] \quad (52)$$

设 η_τ 为温度参数的学习率, 更新温度参数 τ :

$$\tau \leftarrow \tau - \eta_\tau \nabla L_\tau \quad (53)$$

最后, 设为 l 软更新因子, 更新目标网络参数 $\theta', \phi'_1, \phi'_2$:

$$\theta' \leftarrow l\theta + (1-l)\theta' \quad (54)$$

$$\phi'_x \leftarrow l\phi_x + (1-l)\phi'_x, x=1, 2 \quad (55)$$

更新当前状态 $S_t \leftarrow S_{t+1}$ 。重复以上步骤, 直到最大迭代次数 \max_ep 停止训练。

3 车辆智能协同缓存仿真实验分析

本节通过搭建车联网仿真环境对所提算法进行验证, 车辆通信特征及缓存内容流行度分布等具体参数设置如表 2 所示^[31-32]。通过开展参数配置分析实验, 探究学习率、折扣因子等关键参数对算法收敛性能及优化效果的影响, 明确各参数的最优配置范围。进一步地, 将所提算法与现有典型算法进行多维度对比, 从命中率、任务成本等核心指标验证所提算法在动态车辆缓存场景下的有效性, 为车辆协同缓存系统提供理论与实验支撑。

3.1 车辆缓存优化算法参数分析

设定这些在城市车联网系统的车辆 DVs、SVs 发射 100 mW 功率。每个 RSU 的发射功率为 10 W, 覆盖半径为 500 m, 且进入 RSU 覆盖范围的车辆数量遵循泊松分布。单个缓存任务的可容忍时延为 $[0, 10]$ s, 任务处理复杂度为 $[0.01, 0.2]$ 。单个 RSU 和车辆的总带宽分别为 200 MHz 和 100 MHz。本实验中使用的内容库包含约 2 000 个内容项, 涵盖了视频、音频、文档、软件包等多种类型资源, 每个内容项的大小分布范围为 100 kB 至 2 GB, 以适应车联网场景中多样化

表 2 仿真参数

Table 2 Simulation parameter

参数	取值
车辆数量 $M \& N$	[5, 30]
RSU 数量 K	[10, 50]
车辆、RSU 传输功率	30, 50 dBm
高斯白噪声 σ^2	-118 dBm
车辆、RSU 带宽	100 MHz, 200 MHz
RSU 覆盖范围	500 m
内容库规模	2 000 项
内容类型分布	视频、音频、文档、软件包等
内容项的大小分布	[100 kB, 2 GB]
Actor 网络学习率 α_A	[0.00005, 0.005]
Critic 网络学习率 α_C	[0.0005, 0.05]
温度参数 α_τ	[0.00005, 0.005]
折扣因子 ϕ	[0.9, 0.99]
软更新参数 l	[0.001, 0.1]
经验池大小	10 000

的内容需求。

在基于 SAC 的车辆缓存任务资源优化算法中, 学习率 α 、折扣因子 ϕ 和软更新参数 l 的配置对算法性能有重要影响。为了保障算法在车辆协同缓存环境中能够发挥其最优性能, 需要校准这三个关键参数。

学习率 α 决定了参数更新的步长, 较小学习率能让算法在参数空间精细探索、提升收敛稳定性, 但会减缓收敛速度、增加时间成本; 较大学习率可加速迭代以快速接近最优解, 但过高易导致优化过程跳跃过大, 难以捕捉最优解甚至出现振荡, 影响算法稳定性与性能。需选择合适学习率, 平衡收敛速度与稳定性。在 SAC 神经网络中, 主要有 Actor 网络 α_A 、双 Critic 网络 α_C 、温度参数 α_τ 三种学习率。

图 2 展示了四种学习率组合方案对基于 SAC 的车辆缓存任务资源优化算法累积奖励的影响, 横坐标为训练轮次, 纵坐标为累积奖励。低学习率方案 ($\alpha_A = \alpha_\tau = 0.00005, \alpha_C = 0.0005$) 更新保守, 探索推进迟缓, 导致收敛效率不足。中等偏低学习率方案 ($\alpha_A = \alpha_\tau = 0.0001, \alpha_C = 0.001$) 收敛速度和奖励攀升动力不足, 体现学习率提升对收敛效率的促进作用, 但仍未达到最优平衡。中等学习率方案 ($\alpha_A = \alpha_\tau = 0.001, \alpha_C = 0.01$) 在收敛速度与稳定性上取得最佳平衡, 前期累积奖励快速上升, 60 轮后进入平稳阶段, 后期波动极小。高学习率 ($\alpha_A = \alpha_\tau = 0.005, \alpha_C = 0.05$) 则加速收敛, 但易因跳跃过大而导致振荡, 影响稳定性。因此, 选择中等学习率方案, 既能高效更新参数, 又能避免过度振荡, 确保算法在车辆协同缓存环境中稳定运行。

折扣因子 ϕ 影响代理对未来奖励的重视程度及长期规划能力。 ϕ 接近 1 时, 算法侧重未来奖励, 采取

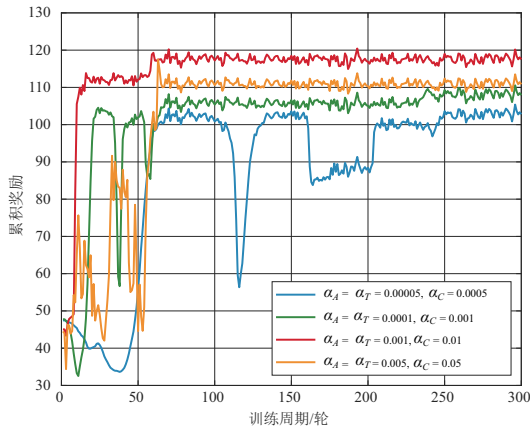


图2 学习率对SVCTRO算法的影响

Figure 2 Impact of learning rate on SVCTRO algorithm

前瞻性策略,有利于长期规划; φ 趋近于0时,算法关注即时奖励,适合即时反馈场景。但过度偏重短期或长期利益均可能导致局部最优解。因此,需选择合适的折扣因子以平衡当下与未来回报,确保在动态环境中实现高效且稳健的决策。图3展示了低折扣因子、中等折扣因子及高折扣因子等三种方案对基于SAC的车辆缓存任务资源优化算法累积奖励的影响,横坐标为训练轮次,纵坐标为累积奖励。低折扣因子方案($\varphi = 0.9$)初期累积奖励飙升快,后期忽略长期策略迭代的价值积累,导致性能受限;高折扣因子方案($\varphi = 0.99$)前期累积奖励攀升慢,最终奖励水平未达最优,体现过度前瞻对学习效率的抑制。中等折扣因子方案($\varphi = 0.95$)在收敛速度与性能上实现平衡,故被选为算法配置方案。

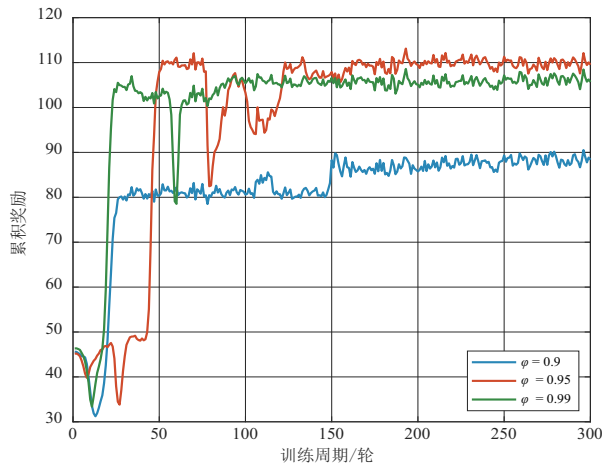


图3 折扣因子对SVCTRO算法的影响

Figure 3 Impact of discount factor on SVCTRO algorithm

软更新参数 τ 控制目标网络的平滑更新过程,影响算法的稳定性和收敛速度。 τ 值较小时,目标网络更新缓慢,有助于防止参数剧烈波动,提高算法稳定

性,但可能导致目标网络滞后,限制适应能力和学习效率; τ 值较大时,目标网络能更快适应环境变化和策略更新,加速收敛,但可能引入过多噪声和不稳定性。故需选择合适的软更新参数以平衡目标网络的平滑性与响应速度。

图4展示了低软更新参数、中等软更新参数及高软更新参数等三种方案对基于SAC的车辆缓存任务资源优化算法累积奖励的影响,横坐标为训练轮次,纵坐标为累积奖励。低软更新参数方案($\tau = 0.001$)初期收敛慢,最终累积奖励低, τ 值过小导致目标网络更新保守,当前网络的改进无法及时反馈到目标网络,落后的策略迭代削弱了学习效率;高软更新参数方案($\tau = 0.1$)初期累积奖励剧烈震荡,当前网络的短期波动被快速传递,破坏策略学习的稳定性,导致后期波动频繁且奖励水平低。中等软更新参数方案($\tau = 0.01$)在前50轮累积奖励快速上升,随后稳定收敛,波动小,平衡了收敛效率与性能,故被选为算法配置方案。

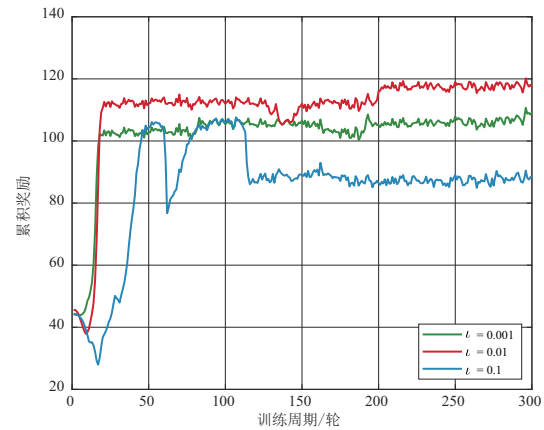


图4 软更新参数对SVCTRO算法的影响

Figure 4 Impact of soft update parameter on SVCTRO algorithm

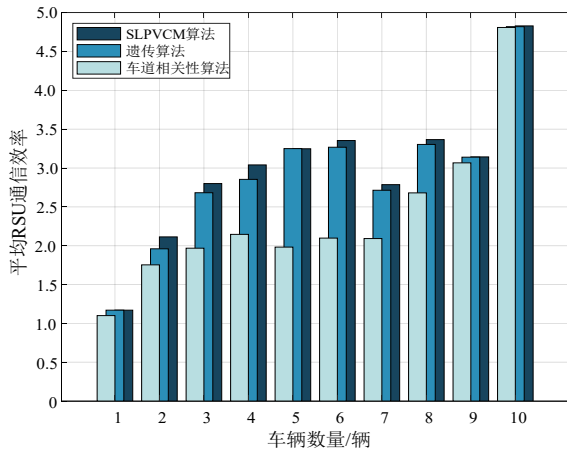
3.2 车辆缓存优化算法性能比较

从车辆协同匹配、缓存内容推荐和任务资源优化三个方面对所提短期车道预测驱动的车辆协同匹配(SLPVCM)算法、基于协同过滤和内容特征的混合推荐(HRCF2)算法以及基于SAC的车辆缓存任务资源优化(SVCTRO)算法进行了综合评估与性能比较分析。

首先,在相同道路状态、通信条件下,比较车道相关性算法、遗传算法和所提出的SLPVCM算法在车辆协同匹配方面的通信效率,以及DV与匹配的RSU资源占用率与通信速率比例等性能。其中,车道相关性算法优先选择与目标DV处于同一车道,相关性最高的SV,选择通信速率最高的RSU。遗传算法选择与目标DV的通信速率最高的SV进行匹配;考虑通信速率和带宽资源占用率,选择车辆间协同度最高的RSU。所提的SLPVCM算法,考虑了车道相关性、通

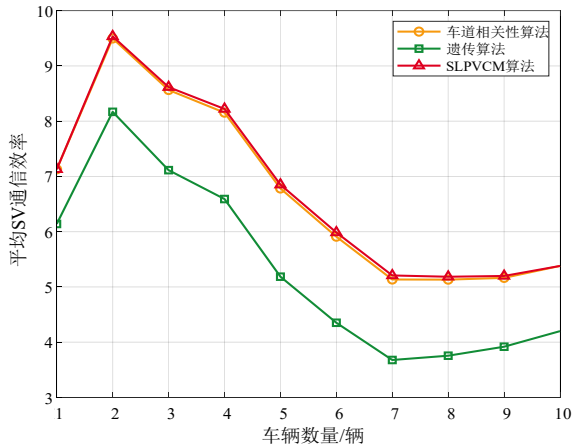
信速率和带宽资源占用率,选择与目标DV的车辆间协同度最高的SV、资源占用率与通信速率比例最低的RSU进行匹配。

图5(a)和图5(b)分别展示了三种算法在RSU和SV通信效率方面的性能。考虑到优化目标的成本与带宽资源占用率有关,定义通信效率(Communication Efficiency, CE)为DV与SV/RSU之间的数据传输和带宽资源的综合效率,表示节点间的考虑资源空闲率的通信质量。设 w_{ce} 为权重调节因子,则CE的表达式为 $CE_{i,j/k}(t) = w_{ce} R_{i,j/k}(t) (1 - bw_{i,j/k}^{occ}(t))$,平均CE的表达式为 $CE^{avg} = \sum_{i=1}^M \sum_{t=1}^T CE_{i,j/k}(t)$ 。分别用柱状图和曲线,来表示RSU和SV的平均通信效率。



(a) 平均RSU通信效率随车辆数量的变化

(a) Variation of average RSU communication efficiency with number of vehicles



(b) 平均SV通信效率随车辆数量的变化

(b) Variation of average SV communication efficiency with number of vehicles

图5 RSU、SV通信效率性能对比

Figure 5 Comparison of RSU and SV communication efficiency

具体来说,在RSU平均通信效率方面,遗传算法和SLPVCM算法性能较好;在SV平均通信效率方面,车道相关性算法和SLPVCM算法性能较好。一方面,RSU是固定基础设施,匹配核心是资源高效调度。车辆数增多时,遗传算法可调度更多RSU分摊负载,SLPVCM也考虑多因素适配实现资源均衡,故效率持续上升;车道相关性算法忽略了RSU资源占用率,效率远低于前两者。另一方面,SV是动态移动节点,匹配核心是通信稳定性。车辆数增多时,SLPVCM、车道相关性算法考虑车辆协同性,通信稳定,效率相对平稳;遗传算法因解空间迅速增长,效率下降严重。所提的SLPVCM算法比车道相关性算法的RSU平均通信效率提升25.97%,比遗传算法的SV平均通信效率提升26.76%,展现出对固定RSU资源的调度能力与对动态SV通信的稳定能力。

考虑到不同RSU资源占用率差异较大影响成本优化,分析如图6所示的三种算法在DV与匹配的RSU资源占用率与通信速率比例方面的性能,该比值越低越好。随着车辆数量增加,RSU需为更多DV提供服务,资源占用率上升,通信速率下降,导致三种算法的比例均呈上升趋势。SLPVCM算法因综合考虑车道相关性、通信速率和带宽资源占用率,能选择最优RSU,故性能最好;遗传算法次之,组合解空间指数级膨胀,匹配效率下降;车道相关性算法因忽略资源占用率,仅优先选择高通信速率RSU,导致比例最高。相较于车道相关性和遗传算法,SLPVCM算法分别降低了10.03%和26.11%的RSU资源占用率与通信速率比例。

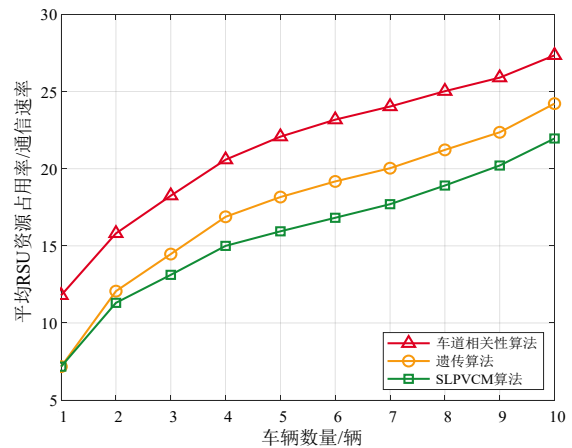


图6 DV与匹配的RSU资源占用率与通信速率比例性能对比

Figure 6 Comparison of resource occupancy rate and communication rate ratio between DV and matched RSU

其次,比较流行度推荐算法、内容特征推荐算法和所提出的HRCF2算法在车辆缓存内容推荐方面的召回率、精确率、F1分数和命中率等性能。其中,流行度

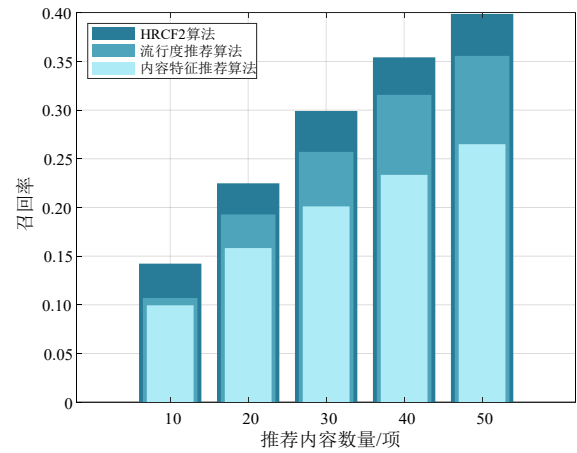
推荐算法主要通过统计内容被请求的次数,取频率高者推荐。内容推荐算法通过计算用户画像与物品内容特征之间的余弦相似度,取排名高者推荐。所提出的HRCF2算法,考虑用户请求资源次数,通过结合用户协同过滤和时空偏好内容推荐,取内容排名高者推荐。

图7(a)与图7(b)分别为三种算法在召回率与精确率上的性能对比。随着推荐内容数量增加,推荐列表覆盖范围扩大,召回率上升;但为填充列表,算法会纳入更多低相关性内容,导致精确率下降。整体训练矩阵密度较低,使得三种算法的召回率和精确率普遍偏低。图中显示,流行度推荐和内容特征推荐算法的精确率在推荐数量约35时相交于0.128。推荐数量较小时,内容特征算法因精准匹配而精确率更高;随着数量增加,流行度算法精确率快速下降,而内容特征算法下降较慢,最终两者相交。本文所提HRCF2算法结合新老用户特征,融合用户协同过滤相似度与内容特征属性。高分用户更依赖协同过滤,低分用户更依赖内容特征。这使得HRCF2在覆盖范围、个性化精度和场景适应性上均优于对比算法。

图8(a)和图8(b)分别对比了三种算法的命中率与F1分数性能。命中率随推荐数量增加而上升并趋于平稳,因为推荐列表覆盖范围扩大,更易命中用户潜在感兴趣内容。F1分数作为召回率和精确率的调和平均,受精确率下降影响,先升后趋于平稳。流行度推荐和内容特征推荐的F1分数在推荐数量约15时相交于0.137,凸显了单一算法的性能偏向性。相比之下,HRCF2算法通过动态调整权重,针对高活跃用户加强内容特征,针对低活跃用户加强流行度,在召回率和精确率间取得更优平衡,F1分数整体更高且更稳定。HRCF2的命中率较流行度推荐和内容推荐算法分别提升8.29%和21.11%,F1分数分别提升12.63%和20.94%。

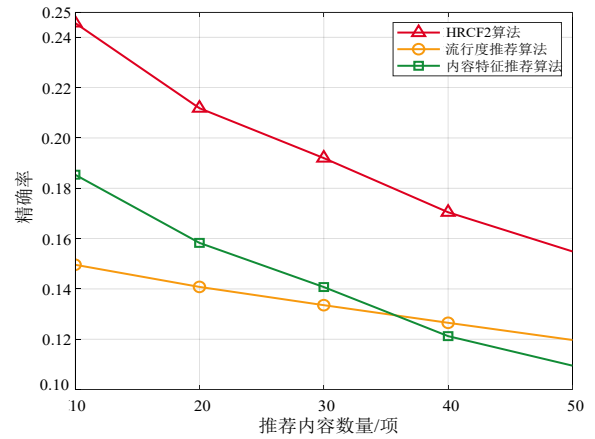
最后,比较深度确定性策略梯度(Deep Deterministic Policy Gradient, DDPG)算法、双延迟确定性策略梯度(Twin Delayed Deep Deterministic Policy Gradient, TD3)算法和所提出的SVCTRO算法在车辆任务资源优化方面的累积奖励、任务成本等性能。其中,DDPG算法通过经验回放的随机采样,将强化学习中的序列相关数据转化为近似独立同分布的数据,加强其训练稳定性。TD3算法是DDPG的改进版,引入了Critic网络中的双Q网络来减少价值函数的高估问题,并采用延迟策略更新和目标策略平滑的方法来提高算法的稳定性和收敛性。所提的SVCTRO算法是基于SAC的最大熵的DRL算法,使用双Q网络解决过估计问题,并通过自动温度参数调整平衡探索与利用,从而提高学习效率。

如图9所示,展示了三种算法在累积奖励的性能



(a) 召回率随推荐内容数量的变化

(a) Variation of recall with recommended content quantity



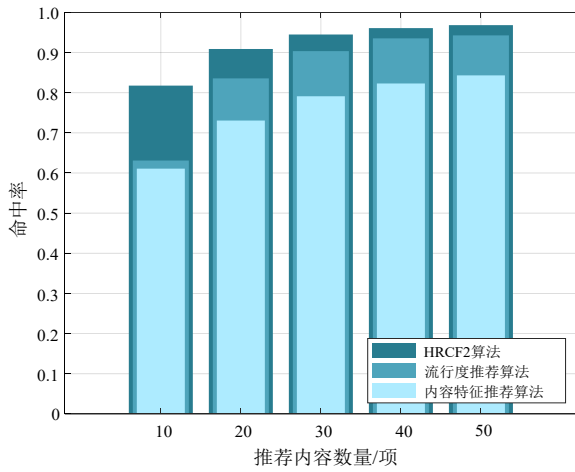
(b) 精确率随推荐内容数量的变化

(b) Variation of precision with recommended content quantity

图7 召回率和精确率的性能对比

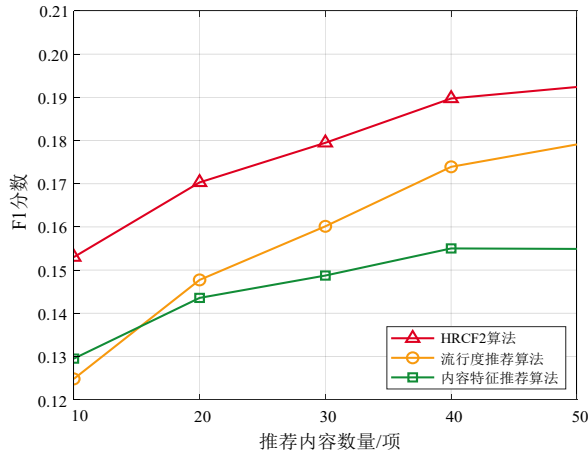
Figure 7 Performance comparison of recall and precision

对比。算法核心目标是优化车辆通过RSU和SV下载缓存任务的成本,根据式(44)奖励与成本成反比,奖励越大性能越好。三个算法均是前期探索波动,后期逐渐收敛稳定。具体来说,DDPG算法初期收敛速度较慢,但后期逐渐稳定并最终反超TD3。这是因为DDPG策略更新频繁,在环境动态变化较快的后期能更及时调整策略。TD3虽引入双Q网络和延迟策略更新机制以减少估计误差和提升稳定性,但在本环境中后期性能被DDPG超越,原因在于延迟更新导致策略滞后,以及保守估计策略在动作空间收窄时可能过度压缩Q值,导致策略偏向次优动作。所提SVCTRO算法通过最大化策略熵鼓励探索,能够更快发现低成本的缓存调度方案,并通过熵正则化机制降低陷入局部最优的风险,较TD3与DDPG算法分别平均提高了4.54%和7.56%的累积奖励。



(a) 命中率随推荐内容数量的变化

(a) Variation of hit rate with recommended content quantity



(b) F1分数随推荐内容数量的变化

(b) Variation of F1-score with recommended content quantity

图8 命中率和F1分数的性能对比

Figure 8 Performance comparison of hit rate and F1-score

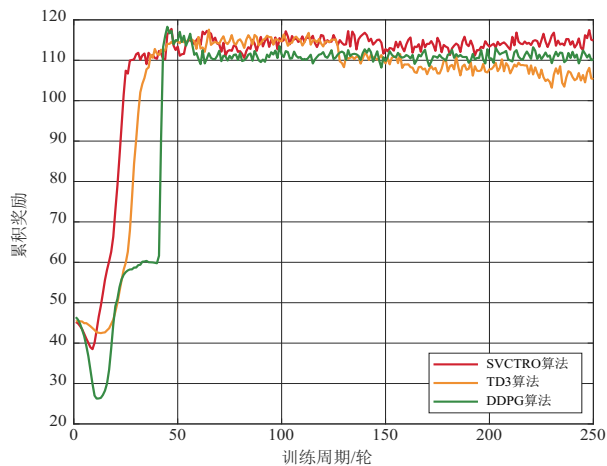


图9 累积奖励性能对比

Figure 9 Performance comparison of cumulative reward

如图10所示,展示了三种算法在任务成本的性能对比。根据式(42),任务成本是放大系数除以累积奖励还原得到,成本越低性能越好。性能分析同累积奖励,所提出的SVCTRO算法在整个训练过程中表现最优,与TD3和DDPG算法相比,分别平均降低了5.80%和13.27%的任务成本。

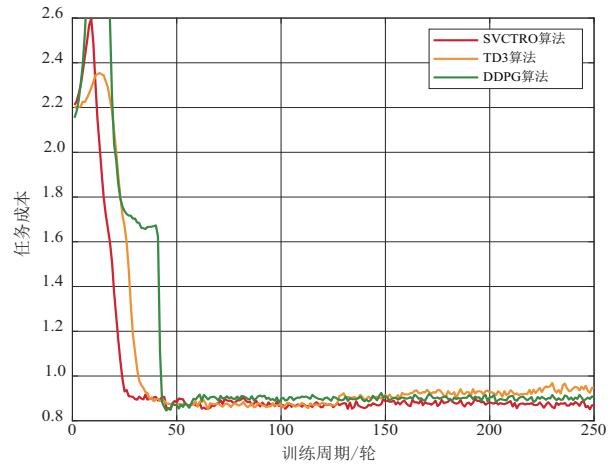


图10 任务成本性能对比

Figure 10 Performance comparison of task cost

4 结束语

针对城市车联网场景中车辆匹配稳定性差、内容推荐个性化不足、资源优化动态适应性差等问题,构建了短期车道预测通信模型、车辆缓存任务时延模型、动态定价和推荐满意度模型等系统模型,提出了车辆协同缓存成本优化策略。具体来说,首先,通过SLPVCM算法综合分析车道时空相关性,实现了SV、RSU与DV的协同匹配,有效提升了通信效率。其次,提出HRCF2算法,针对新老用户动态调整推荐权重,通过优化缓存内容推荐策略减少了对远端云资源的依赖,提升了推荐精度与用户满意度。最后,提出SVCTRO算法,通过深度强化学习动态调整任务分解与带宽分配策略,进一步优化了系统缓存成本。仿真实验表明,所提算法在通信效率、命中率、F1分数及任务成本等方面均优于对比算法。未来研究工作聚焦方向包括:一是细化车辆动态行为模型,结合实时路况与驾驶意图优化协同匹配策略;二是探索多业务并发场景下的缓存资源调度机制,提升系统对异构任务的适配能力。

参考文献

- [1] Mordor Intelligence. Internet of vehicle market size & share analysis - growth trends & forecasts (2025-2030)[EB/OL]. [2026-01-05]. <https://www.mordorintelligence.com/industry-reports/internet-of-cars-market-industry>.

- [2] Clancy J, Mullins D, Deegan B, et al. Wireless access for V2X communications: Research, challenges and opportunities[J]. *IEEE Communications Surveys & Tutorials*, 2024, 26(3): 2082-2119.
- [3] 中研普华中国产业研究院. 2025-2030年中国车联网行业竞争分析及发展前景预测报告[R/OL]. 2025-11-07. <https://www.chinairm.com/report/20251107/155426613.html?id=1922815&name=LiBo2>.
China Industry Research Institute, Zero Power Intelligence Group. China Automotive Manufacturing Industry Market Analysis and Development Prospect Forecast Report(2025-2030) [R/OL]. 2025-11-07. <https://www.chinairm.com/report/20251107/155426613.html?id=1922815&name=LiBo2>. (in Chinese)
- [4] 工业和信息化部,公安部,住房和城乡建设部,等. 智能网联汽车准入和上路通行试点实施指南[EB/OL]. 2023-11-17. https://www.miit.gov.cn/zwgk/zcwj/wjfb/tz/art/2023/art_be04a39345354b1ba4624262f86e7cfb.html.
Ministry of Industry and Information Technology, Ministry of Public Security, Ministry of Housing and Urban-Rural Development, et al. Implementation Guide for Pilot Access and Road Use of Intelligent Connected Vehicles[EB/OL]. 2023-11-17. https://www.miit.gov.cn/zwgk/zcwj/wjfb/tz/art/2023/art_be04a39345354b1ba4624262f86e7cfb.html. (in Chinese)
- [5] Cai Xuelian, Zheng Jing, Fu Yuchuan, et al. Cooperative content caching and delivery in vehicular networks: A deep neural network approach[J]. *China Communications*, 2023, 20(3): 43-54.
- [6] Wang Chaowei, Wang Ziye, Xu Lexi, et al. Collaborative caching in vehicular edge network assisted by cell-free massive MIMO[J]. *Chinese Journal of Electronics*, 2023, 32(6): 1218-1229.
- [7] 陈清秋,朱琦. VANET中基于车辆分簇的协作缓存策略[J]. *计算机技术与发展*, 2022, 32(11): 214-220.
Chen Qingqiu, Zhu Qi. Cooperative caching strategy based on vehicle clustering in VANET[J]. *Computer Technology and Development*, 2022, 32(11): 214-220. (in Chinese)
- [8] 吴红海,王白冰,马华红,等. 移动车载边缘网络中基于递归深度强化学习的协作缓存接力算法[J]. *通信学报*, 2024, 45(11): 277-286.
Wu Honghai, Wang Baibing, Ma Huahong, et al. Recursive deep reinforcement learning-based collaborative caching relay algorithm in mobile vehicular edge network[J]. *Journal on Communications*, 2024, 45(11): 277-286. (in Chinese)
- [9] 许小龙,杨威,杨辰翊,等. 车联网边缘计算环境下基于流量预测的高效任务卸载策略研究[J]. *电子学报*, 2025, 53(2): 329-343.
Xu Xiaolong, Yang Wei, Yang Chenyi, et al. Efficient task offloading based on traffic prediction in IoV-enabled edge computing[J]. *Acta Electronica Sinica*, 2025, 53(2): 329-343. (in Chinese)
- [10] 林志坚,侯映,曹晓晓,等. 车联网中基于信息年龄价值的边缘缓存策略[J]. *电子学报*, 2023, 51(12): 3410-3421.
Lin Zhijian, Hou Ying, Cao Xiaoxiao, et al. Edge caching scheme based on value of information age in the Internet of vehicles[J]. *Acta Electronica Sinica*, 2023, 51(12): 3410-3421. (in Chinese)
- [11] 沈俊杰,彭江,郭坤银,等. 车联网中基于位置信息映射和相关性评估的进化多任务优化算法[J]. *电子学报*, 2025, 53(5): 1661-1676.
Shen Junjie, Peng Jiang, Guo Kuiyin, et al. Location mapping and correlation assessment based evolutionary multi-task optimization algorithm in the Internet of vehicles[J]. *Acta Electronica Sinica*, 2025, 53(5): 1661-1676. (in Chinese)
- [12] 张新钰,卢毅果,高鑫,等. 面向智能网联汽车的车路协同感知技术及发展趋势[J]. *自动化学报*, 2025, 51(2): 233-248.
Zhang Xinyu, Lu Yiguo, Gao Xin, et al. Vehicle-road collaborative perception technology and development trend for intelligent connected vehicles[J]. *Acta Automatica Sinica*, 2025, 51(2): 233-248. (in Chinese)
- [13] 伊笑莹,芮一康,冉斌,等. 车路协同感知技术研究进展及展望[J]. *中国工程科学*, 2024, 26(1): 178-189.
Yi Xiaoying, Rui Yikang, Ran Bin, et al. Vehicle-infrastructure cooperative sensing: Progress and prospect[J]. *Strategic Study of CAE*, 2024, 26(1): 178-189. (in Chinese)
- [14] Yassine A, Hossain M S. Match maximization of vehicle-to-vehicle energy charging with double-sided auction[J]. *IEEE Transactions on Intelligent Transportation Systems*, 2023, 24(11): 13250-13259.
- [15] Chatterjee P, Majumder P, Das S K. V2VDisCS: Vehicle to vehicle distributed charge sharing in intelligent transportation systems[J]. *IEEE Transactions on Intelligent Transportation Systems*, 2025, 26(4): 4960-4974.
- [16] Cao Tengfei, Yi Jie, Wang Xiaoying, et al. Interaction trust-driven data distribution for vehicle social networks: A matching theory approach[J]. *IEEE Transactions on Computational Social Systems*, 2024, 11(3): 4071-4086.
- [17] Liu Shuohan, Ni Qiang, Cao Yue, et al. A reservation-based vehicle-to-vehicle charging service under constraint of parking duration[J]. *IEEE Systems Journal*, 2023, 17(1): 176-187.
- [18] Zhang Chuan, Zhao Mingyang, Zhang Weiting, et al. Bilateral content recommendation for intelligent connected vehicles: Challenges and applications[J]. *IEEE Network*, 2024, 38(6): 309-315.
- [19] Ahani G, Yuan Di. Optimal content caching and recommendation with age of information[J]. *IEEE Transactions on Mobile Computing*, 2024, 23(1): 689-704.
- [20] Zhao Yi, Yu Zhanwei, Yuan Di. Caching with personalized and incumbent-aware recommendation: Modeling and optimization[J]. *IEEE Transactions on Mobile Computing*, 2024, 23(10): 9595-9613.
- [21] Khan M T R, Jembre Y Z, Saad M M, et al. Proactive con-

- tent retrieval based on value of popularity in content-centric internet of vehicles[J]. IEEE Transactions on Intelligent Transportation Systems, 2024, 25(8): 8514-8526.
- [22] Ashraf M W A, Raza A, Singh A R, et al. Intelligent caching based on popular content in vehicular networks: A deep transfer learning approach[J]. IEEE Transactions on Intelligent Transportation Systems, 2024, 25(12): 20643-20656.
- [23] 崔亚平, 石宏吉, 吴大鹏, 等. 内容新鲜度保障的车联网多智能体缓存分发策略[J]. 通信学报, 2025, 46(1): 52-66.
Cui Yaping, Shi Hongji, Wu Dapeng, et al. Multi-agent caching distribution strategy for content freshness guarantee in IoV[J]. Journal on Communications, 2025, 46(1): 52-66. (in Chinese)
- [24] Tian Shujuan, Zou Song, Tan Yi, et al. Dynamic content cache strategy based on content prediction in the internet of vehicles[C]//Proceedings of the 2023 19th International Conference on Mobility, Sensing and Networking. Piscataway: IEEE, 2023: 262-269.
- [25] 秦鹏, 和昊婷, 赵雄文, 等. 基于停放车辆路边单元环境感知的车联网资源高效分配[J]. 通信学报, 2022, 43(7): 113-125.
Qin Peng, He Haoting, Zhao Xiongwen, et al. Efficient resource allocation with context-awareness for parked car road side unit-based Internet of vehicles[J]. Journal on Communications, 2022, 43(7): 113-125. (in Chinese)
- [26] Hazarika B, Singh K, Li C P, et al. RADiT: Resource allocation in digital twin-driven UAV-aided Internet of vehicle networks[J]. IEEE Journal on Selected Areas in Communications, 2023, 41(11): 3369-3385.
- [27] Zhu Sifeng, Tian Xiaohua, Zhang Zonghui, et al. Content placement and edge collaborative caching scheme based on deep reinforcement learning for Internet of vehicles[J]. IEEE Transactions on Intelligent Transportation Systems, 2025, 26(6): 8050-8064.
- [28] Xu Kaidi, Zhou Shenglong, Li G Y. Federated reinforcement learning for resource allocation in V2X networks[J]. IEEE Journal of Selected Topics in Signal Processing, 2024, 18(7): 1210-1221.
- [29] Bhardwaj S, Kim D H, Kim D S. Federated learning-based resource allocation for V2X communications[J]. IEEE Transactions on Intelligent Transportation Systems, 2025, 26(1): 382-396.
- [30] Wu Sijun, Yang Liang, Li Junjie, et al. DRL-based pricing-driven for task offloading and dynamic resource in vehicle edge computing[J]. IEEE Transactions on Mobile Computing, 2025, 24(10): 10389-10404.
- [31] Liang Yan, Zhang Haijun, Ma Hui, et al. Learning-based cooperative cache strategy integrating content recommendation in IoV[J]. IEEE Transactions on Vehicular Technology, 2025, 74(2): 2080-2093.
- [32] Ren Yuzheng, Xie Renchao, Yu F R, et al. Connected and autonomous vehicles in Web3: An intelligence-based reinforcement learning approach[J]. IEEE Transactions on Intelligent Transportation Systems, 2024, 25(8): 9863-9877.

作者简介



梁 琰 男, 1998年1月出生于湖北省孝感市。现为北京科技大学计算机与通信工程学院博士研究生。主要研究方向为未来智能网络、车联网。
E-mail: liangy188@126.com



马 旭 男, 1996年7月出生于山东省菏泽市。现为北京科技大学计算机与通信工程学院博士研究生。主要研究方向为6G移动通信、空天地一体化网络。
E-mail: maxu@xs.ustb.edu.cn



张海君 男, 1986年10月出生于辽宁省朝阳市。现为北京科技大学人工智能学院教授、博士生导师。主要研究方向为智能通信、无人机通信、卫星通信。中国电子学会会员编号: E190020833S。
E-mail: zhanghaijun@ustb.edu.cn



任语铮 女, 1995年3月出生于河北省保定市。现为北京科技大学计算机与通信工程学院副教授。主要研究方向为未来网络、智能内生网络。中国电子学会会员编号: E190157261M。
E-mail: renyuzheng@ustb.edu.cn



任 超 1989年4月出生于山东省菏泽市。现为北京科技大学计算机与通信工程学院副教授。主要研究方向为无线多模态通信、未来移动网络、智能通信。中国电子学会会员编号: E190157259M。
E-mail: chaoren@ustb.edu.cn