

面向延迟感知的车联网缓存优化策略

王亚丽^{1,2}, 韩晓培²

(1.苏州城市学院 计算科学与人工智能学院,江苏 苏州 215104;2.河南师范大学 计算机与信息工程学院,河南 新乡 453007)

摘要:车联网技术的发展催生了各种数据密集型和延迟敏感型的车载应用.将内容缓存在用户侧,可为车载应用提供高效的内容交付.研究了车联网中以用户为中心的内容交付问题,提出了一种基于延迟感知的内容缓存优化策略,通过协同路边单元和车载单元,在不确定请求内容的情况下最小化传输时延.首先基于多准则决策收集缓存车辆特征信息,确定最佳缓存车辆;然后提出了一种基于流行度的动态内容缓存算法,获取最优内容缓存决策.仿真实验表明,与其他方案相比,所提策略可有效提高缓存命中率,减少时延.

关键词:边缘缓存;移动边缘计算;车联网;协作缓存

中图分类号:TP391

文献标志码:A

文章编号:1000-2367(2025)06-0074-10

随着信息技术飞速发展,车联网(Internet of Vehicles, IoV)融合了智能计算、无线通信等特点,不仅为驾驶员提供实时导航、紧急救援等服务,还能实现车辆间的协同控制和自主感知^[1],成为支撑未来交通工具和信息通信系统发展的基础平台^[2-3].在这一背景下,多功能车载应用对数据资源的实时性和鲁棒性等方面要求更严格^[4-5].车辆的车载单元(On Board Unit, OBU)具有一定的计算和存储能力,相较于日益复杂的应用需求显得捉襟见肘.借助云计算服务成为一种切实可行的解决方案^[6],但在车辆密度大的交通高峰时段可能会导致主干网络堵塞,难以保证内容交付的服务质量(Quality of Service, QoS).

为应对如上挑战,边缘缓存技术应用到 IoV 中^[7-8],将关键和流行的内容放在网络边缘接近终端用户的地方,减少降低延迟和回程链路的负担^[9].一方面,移动车辆通过与路边单元(the Roadside Unit, RSU)等基础设施通信(vehicle-to-infrastructure, V2I)来共享和获取信息,既保护隐私,还为车载设备提供可靠低延迟智能服务^[10].另一方面,借助 OBU 之间的车对车通信(vehicle-to-vehicle, V2V)拓宽信息传播的渠道,提升信息获取的灵活性和实时性.在实际应用中,传统的单一缓存方式常常面临着存储空间不足和资源不合理分配导致的浪费等问题^[11].因此,制定高效缓存策略至关重要.RSU 和 OBU 的协同合作缓存可显著降低对核心网络基础设施的依赖,有效减少传输时延.

在实际应用中,RSU 和 OBU 协同缓存必须考虑路况和内容受欢迎程度,管理通信进程并协调车辆通信资源,避免信道冲突.基于此,本文研究了 IoV 中以用户为中心的内容交付问题,提出了一种基于延迟感知的内容缓存优化算法,在不确定请求内容的情况下最小化传输时延,保证 QoS.本文的创新点总结如下:

1)考虑实际路况中高速移动的车辆,为避免因超过通信范围而传输中断产生不必要的信令开销,设计了一种基于多准则决策(multiple-criteria decision making, MCDM)的最佳缓存车辆选择算法,获取最佳缓存车辆.

收稿日期:2024-07-12; **修回日期:**2024-07-31.

基金项目:国家自然科学基金(62072159;61902112).

作者简介(通信作者):王亚丽(1979—),女,河南三门峡人,苏州城市学院副教授,博士,研究方向为边缘计算、软件定义网络, E-mail: yaliwan_g@163.com.

引用本文:王亚丽,韩晓培.面向延迟感知的车联网缓存优化策略[J].河南师范大学学报(自然科学版),2025,53(6):74-83.(Wang Yali, Han Xiaopei. An optimization strategy for latency-aware internet of vehicle caching[J]. Journal of Henan Normal University(Natural Science Edition), 2025, 53(6): 74-83. DOI: 10.16366/j.cnki.1000-2367.2024.07.12.0001.)

2)研究了以用户为中心的系统最小化传输时延的问题,协同 OBU 和 RSU 缓存,设计了一种基于延迟感知内容缓存优化(delay-aware content caching optimization algorithm based,DACO)算法,最小化内容传输时延。

1 相关工作

随着用户对即时信息需求的日益增长,网络拥塞现象日益严重,内容交付延迟显著增加,用户体验受到严重影响.传统的方案难以解决上述挑战,车联网缓存技术逐步成为解决数据传输和处理效率问题的重要手段.文献[12]提出一种基于边缘缓存和免疫克隆策略的车联网内容分发方法,主动缓存高热度内容,降低通信成本.文献[13]将数据孪生技术和人工智能融入 IoV 中,采用深度强化学习中的分布式调度计算任务卸载和边缘资源分配.文献[14]提出一种协作式内容缓存策略,解决了传统缓存系统中低效率问题.文献[15]提出一种高效的边云系统缓存策略,降低任务执行过程中的数据迁移的延迟.文献[16]提出一种预测移动性的协作缓存方案,降低内容访问延迟.文献[17]设计了一种基于车辆轨迹交通相关性的 RSU 间协作机制,优化车辆服务的网络延迟.文献[18]共同考虑 peer-to-peer(P2P)车载网络的内容传播、缓存更新和公平性问题,开发了一种分布式异构媒体服务方案,提高系统命中率,降低内容传输延迟.文献[19]提出一种社交感知车载边缘计算架构,通过使用网络中的某些车辆作为边缘服务器来解决内容交付问题。

考虑到 RSU 和 OBU 存储空间有限,很多学者致力于协同 RSU 和 OBU 的高效利用缓存空间的研究.文献[20]充分利用道路中的闲置资源构建协作缓存系统,提出一种面向 IoV 的社会感知分散式协作缓存,减少传输延迟.文献[21]提出一种基于内容请求预测的协同缓存策略,以减少内容延迟.文献[10]提出一种车辆与 RSU 之间的协同缓存方案,降低传输延迟.文献[22]提出一种基于联邦学习的边缘缓存方案,提高缓存命中率,减少回程负载.文献[23]提出一种联合智能车辆和小型基站以减少内容检索延迟的缓存方案.文献[24]从全局优化的角度出发提出一种 V2I 和 V2V 的联合调度方案,提高传输效率.文献[25]提出一种基于 RSU 协作的自适应调度算法,减轻和平衡双向道路场景中 RSU 的工作负载.文献[26]提出一种基于弹性段的 V2I 和 V2V 协同策略应对车联网拓扑结构的动态变化,降低车辆密度和对可实现吞吐量的影响.文献[27]提出一种由传统和专用两种 RSU 组成的新型车辆云作为运营骨干,动态配置网络中托管的服务及数据转发信息,最小化服务延迟.以上文献虽然考虑了 RSU 与 OBU 的协同存储,但对车载设备的 QoS 关注较少。

2 系统模型

2.1 系统框架

系统框架模型由一个云服务器、一个宏基站(macro base station,MBS)、多个 RSU 以及若干移动车辆组成,如图 1 所示.MBS 具有一定存储和调度能力,管理覆盖范围内的 RSU 集群和存储其内容目录.设 $C = \{1, 2, \dots, C\}$ 为云服务器存储的所有内容文件集合,每个文件内容大小为 S_c .设 $N = \{1, 2, \dots, N\}$ 为 RSU 集合,每个 RSU 存储空间为 D .RSU 覆盖一个特定的区域,可以处理覆盖范围的车辆请求. $V_R = \{1, 2, \dots, M\}$ 为请求车辆(requested vehicles,RV)集合,在行驶过程中发送内容请求. $V_C = \{1, 2, \dots, K\}$ 为缓存车辆(caching vehicles,CV)集合,车辆的 OBU 存储部分内容,向请求车辆发送传输内容.服务时间 $l_T = \{1, 2, \dots, T\}$ 划分为 T 个离散时隙,车辆每个时隙只能请求一个内容.规定:如果请求内容不能在 l_T 内完成传输,则传输内容直接从云服务器传输。

车辆所请求内容可以通过 V2V 链路从缓存车辆的 OBU 中获取,也可通过 V2I 链路从 RSU 或云服务器中获取.避免通信链路频繁切换导致信令开销增加,每一时刻只能选择一条链路.车辆请求发送后,先从 OBU 中查询,查询成功直接传输;查询失败从 RSU 中查询,若查询成功直接传输,否则从云服务器传输。

2.2 车辆移动模型

车辆可通过 V2V 链路和 V2I 链路获取内容,链路传输性能会受到车辆的高速移动和通信范围的影响,因此需要分为两种情况进行分析。

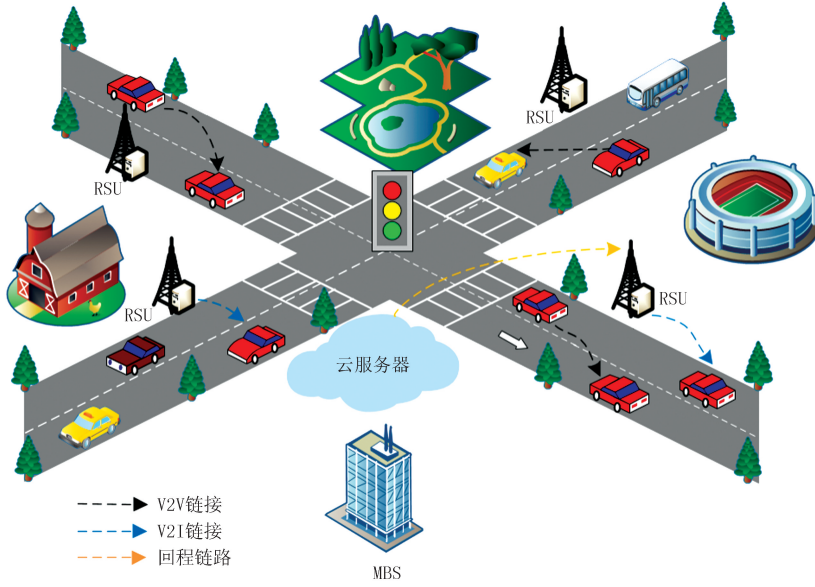


图1 系统框架模型

Fig.1 The schematic framework model

(1)情况 1:当车辆 $V_{R_m} \in V_R$ 通过与 V2V 链路与缓存车辆 $V_{C_k} \in V_C$ 通信,V2V 链路的连接时间为 $\Delta t_{m,k}$,表示为 $|\Delta s_{m,k} + v_m \Delta t_{m,k} - v_k \Delta t_{m,k}| = s$,其中, v_m 是 V_{R_m} 的速度, v_k 是 V_{C_k} 的速度,车距为 $\Delta s_{m,k}$, s 为链路的通信范围.

(2)情况 2:当 V_{R_m} 与 $RSU N_i \in N$ 通过 V2I 链路连接,最大连接时间为 $\Delta t_{m,N_i}$,表示为 $\Delta s_0 + v_m \Delta t_{m,N_i} = 2r$,其中, Δs_0 为 V_{R_m} 起始位置, r 为 N_i 覆盖半径.

2.3 系统缓存模型

定义二进制矩阵 $A_{K,C}$ 表示缓存车辆 V_{C_k} 是否缓存内容 $c \in [1, C]$,其中 $a_{k,c} \in \{0, 1\}$. $a_{k,c} = 1$ 表示 V_{C_k} 缓存 c , $a_{k,c} = 0$ 表示 V_{C_k} 未缓存 c .类似地, $A_{N_i,C}$ 表示是否缓存 c ,其中 $a_{N_i,c} \in \{0, 1\}$. $a_{N_i,c} = 1$ 表示 N_i 缓存 c ,表示 N_i 未缓存 c .为合理利用缓存资源,减少冗余,每辆缓存车辆的车载单元存储不同内容.RSU 存储资源有限,任何时隙缓存数据量不能违反 RSU 可用容量约束:

$$\sum_{c=1}^C a_{k,c} \leq 1, \quad (1)$$

$$\sum_{c=1}^C a_{N_i,c} \leq D_i. \quad (2)$$

2.4 内容请求模型

车辆请求随机到达,其到达独立同分布.设 Q 为请求内容集合, q_c 表示时隙 t 时车辆请求状态,其中 $q_c \in \{0, 1\}$. $q_c = 1$ 表示有车辆请求 c , $q_c = 0$ 表示无车辆请求 c .时隙 t 车辆请求内容集合 Q 为: $Q = \{q_1, q_2, \dots, q_c, \dots, q_C\}$.将来自移动车辆的内容请求的到达建模为泊松过程, λ 为来自车辆请求内容的到达率, η 表示 Zipf 分布的偏度参数^[10].时隙 t 车辆请求内容 c 的概率为:

$$\lambda_c(t) = \frac{c^{-\eta}}{\sum_{c=1}^C c^{-\eta}}. \quad (3)$$

2.5 通信模型

传输请求内容的方式不同,产生不同传输时延.为防止通信链路频繁切换带来的信令开销增加,在每次传输过程中,只能选择一条传输链路获取所需内容.此外,V2V 链路和 V2I 链路分配了不同频谱资源,两条链路之间不存在干扰.

2.5.1 V2V 链路传输

请求车辆通过 V2V 链路从缓存车辆的 OBU 中获取内容时,通信链路信噪比为^[21]:

$$\gamma_m^k = \frac{P_k^t h_{k,m}}{\sigma^2 + \sum_{m=1, m \neq k}^M \Phi_{k,m} P_k^t h_{k,m}}. \quad (4)$$

其中, P_k^t 为发射功率, $h_{k,m}$ 为缓存车辆 V_{C_k} 与请求车辆 V_{R_m} 之间的信道功率增益, σ^2 为噪声功率, $\sum_{m=1, m \neq k}^M \Phi_{k,m} P_k^t h_{k,m}$ 表示是否有其他车辆干扰通信链路, $\Phi_{k,m}$ 为频谱分配指标, B_m^k 是 V_{R_m} 与 V_{C_k} 之间的信道带宽, V_{R_m} 通过 V2V 链路获取内容的延迟为:

$$T_k = \frac{S_c^k}{B_m^k \log_2(1 + \gamma_m^k)}. \quad (5)$$

2.5.2 V2I 链路传输

车辆通过 V2I 链路从 RSU 中获取内容时,通信链路信噪比为:

$$\gamma_m^{N_i} = \frac{P_{N_i}^t h_{N_i,m}}{\sigma^2 + \sum_{i=1}^N \Phi_{N_i,m} P_{N_i}^t h_{N_i,m}}. \quad (6)$$

其中, $P_{N_i}^t$ 为发射功率, $h_{N_i,m}$ 为 N_i 与 V_{R_m} 之间的信道功率增益, $\sum_{i=1}^N \Phi_{N_i,m} P_{N_i}^t h_{N_i,m}$ 表示是否有其他 V2I 通信链路干扰, $\Phi_{N_i,m}$ 为频谱分配指标, $B_m^{N_i}$ 是 V_{R_m} 与 N_i 之间的信道带宽, V_{R_m} 通过 V2I 链路获取内容的延迟为:

$$T_{N_i} = \frac{S_c^{N_i}}{B_m^{N_i} \log_2(1 + \gamma_m^{N_i})}. \quad (7)$$

2.6 问题形式化

车联网系统对延迟十分敏感,需要考虑整个网络的内容获取延迟.若缓存车辆和 RSU 均无车辆请求的内容,传输内容需要从云服务器传输.上述场景问题中,系统的传输时延为:

$$T = \sum_{m=1}^M (\sum_{k=1}^K T_k + \sum_{i=1}^N T_{N_i} + T_C). \quad (8)$$

其中 T_C 为云服务器传输内容的时延.系统传输时延的优化问题表示为:

$$\min(\sum_{m=1}^M (\sum_{k=1}^K T_k + \sum_{i=1}^N T_{N_i} + T_C)) \quad (9)$$

$$\text{s.t.} \begin{cases} a_{k,c} + a_{N_i,c} \leq 1, & (a) \\ \sum_{c=1}^C a_{k,c} \leq 1, & (b) \\ \sum_{c=1}^C a_{N_i,c} \leq D_i, & (c) \\ T_k \leq \Delta t_{m,k}, T_{N_i} \leq \Delta t_{m,N_i}, & (d) \\ a_{k,c} \in \{0,1\}, a_{N_i,c} \in \{0,1\}, & (e) \end{cases} \quad (10)$$

其中,式(10a-10c)保证合理利用缓存资源;式(10d)保证车辆传输内容必须在服务器覆盖范围内;式(10e)保证传输的可能性.V2V 链路传输时延小,V2I 链路次之,云服务器最大.系统需要优化缓存决策,最小化传输时延.

3 基于延迟感知的缓存优化算法

车辆的高移动性、请求内容的实时性和不确定性决定了 OBU 和 RSU 存储内容是时变的,增加了在每个时隙中求解问题的难度,需合理分配存储内容.由于缓存决策变量是二进制变量,该问题是一个 0-1 型整数线性规划问题.为降低计算的复杂度,将优化问题分解为两个子问题:缓存车辆选择和内容缓存决策优化.

3.1 缓存车辆选择

移动车辆的速度、行驶方向以及加速度不同,可能出现超过 V2V 链路通信范围传输中断的情况,产生不必要的信令开销.采用 MCDM 选择符合相似特征的缓存车辆^[28].具体过程如下:

首先,收集请求车辆和缓存车辆相关特征信息 $[\omega_1, \omega_2, \Delta s_{m,k}, e]$,其中, ω_1 计算车辆的相对速度, ω_2 判

断车辆行驶方向是否一致, $\Delta s_{m,k}$ 表示车辆的相对行驶距离, e 为车辆行驶的加速度.

$$\omega_1(V_{R_m}, V_{C_k}) = \frac{\max(\nu_m, \nu_k) - \nu_k}{\max(\nu_m, \nu_k) - \min(\nu_m, \nu_k)}, \quad (11)$$

$$\omega_2(V_{R_m}, V_{C_k}) = \begin{cases} 1, & \text{同向,} \\ 0, & \text{其他.} \end{cases} \quad (12)$$

其次, 根据整合 $[\omega_1, \omega_2, \Delta s_{m,k}, e]$ 可得车辆缓存矩阵 $\mathbf{\Gamma}$:

$$\mathbf{\Gamma} = \begin{matrix} & \omega_1 & \omega_2 & \Delta s_{m,k} & e \\ \begin{matrix} V_{C_1} \\ V_{C_2} \\ \vdots \\ V_{C_k} \end{matrix} & \begin{bmatrix} \tau_{1,1} & \tau_{1,2} & \tau_{1,3} & \tau_{1,4} \\ \tau_{2,1} & \tau_{2,2} & \tau_{2,3} & \tau_{2,4} \\ \vdots & \vdots & \vdots & \vdots \\ \tau_{k,1} & \tau_{k,2} & \tau_{k,3} & \tau_{k,4} \end{bmatrix} \end{matrix}, \quad (13)$$

然后, 由 $\mathbf{\Gamma}$ 计算缓存车辆综合评分, 计算 $\mathbf{\Gamma}$ 中每个元素 τ_j 的加权指标值 $y_{i,j}$:

$$x_{i,j} = \frac{\tau_{i,j}}{\sqrt{\sum_{i=1}^k \tau_{i,j}^2}}, \quad i = 1, 2, \dots, k; j = 1, 2, \dots, 4. \quad (14)$$

$$y_{i,j} = \omega_j \times x_{i,j}.$$

其中, ω 为参数的权重, $\sum_{j=1}^4 \omega_j = 1$. 获取所有 $y_{i,j}$ 后可得综合评分矩阵 $\mathbf{R}$, 计算 $\mathbf{R}$ 中每个元素 r_i 的值:

$$r_i = \frac{\sqrt{\sum_{j=1}^4 (y_{i,j} - \max(y_{i,j}))^2}}{\sqrt{\sum_{j=1}^4 (y_{i,j} - \max(y_{i,j}))^2} - \sqrt{\sum_{j=1}^4 (y_{i,j} - \min(y_{i,j}))^2}}, \quad i = 1, 2, \dots, k. \quad (15)$$

最后, 根据计算综合评分 r_i 并降序排列, 获得缓存车辆队列 $\Lambda = (r_1, r_2, \dots, r_k)$, 具体算法如算法 1 所示.

算法 1 缓存车辆选择算法

输入 CV 集合 V_C 、RV 集合 V_R

输出 最佳缓存车辆队列

收集请求车辆和缓存车辆信息;

for $k = 1, 2, \dots, K$ do

 根据式(11)计算 ω_1 , 式(12)计算 ω_2 , 收集 $[\omega_1, \omega_2,$

$\Delta s_{m,k}, e]$ 创建矩阵 $\mathbf{\Gamma}$;

end for

for $i = 1, 2, \dots, k$ do

 for $j = 1 : 4$ do

 根据式(14)计算 τ_j 加权指标值 $y_{i,j}$;

 end for

end for

for $i = 1, 2, \dots, k$ do

 根据式(15)计算 r_i , 获得综合评分矩阵 $\mathbf{R}$;

end for

计算综合评分 r_i 降序排列, 获取综合评分队列 $\mathbf{\Gamma}$.

3.2 基于延迟感知内容缓存优化算法

车辆可以通过 V2V 链路从缓存车辆的 OBU 获取内容, 也可以通过 V2I 链路从 RSU 获取. 因此, 缓存决策需要根据车辆所请求内容的动态变化而制定. 从 OBU 获取内容的延迟通常小于从 RSU 获取的, 若某缓存车辆的 OBU 中存储了被请求的内容, 请求车辆会优先选择从缓存车辆处获取. 此外, RSU 中的缓存内容会受到请求车辆的请求和缓存车辆的存储内容的影响. 鉴于缓存车辆的存储内容是不确定的且无法预测, 需要设计一种能够应对内容流行度变化的动态内容缓存算法, 以更有效地更新 RSU 缓存内容.

内容 c 在 RSU N_i 中的命中率为 $u_{N_i} = \frac{\sum \mathbf{r}_{N_i}}{\sum q_{N_i}} \cdot \lambda_c$, 其中, $\mathbf{r}_{N_i}$ 为 N_i 内容请求命中次数. 设 c 在第 t 次迭

代时缓存在 N_i 的概率为 $p_{i,c}^t$, 则 N_i 缓存向量 $\mathbf{P}_{N_i}^t$ 可表示为 $\mathbf{P}_{N_i}^t = [p_{N_i,1}^t, p_{N_i,2}^t, \dots, p_{N_i,N}^t]$. N_i 中缓存内容相互独立, 内容缓存概率分布函数为:

$$f(a_{N_i,c}^t, p_{N_i,c}^t) = \prod_{c=1}^C p_{N_i,c}^t a_{N_i,c}^t (1 - p_{N_i,c}^t)^{1-a_{N_i,c}^t}. \quad (16)$$

根据缓存概率生成数量为 Z 的样本集合 $Z = \{1, 2, \dots, z\}$, 样本 $z \in Z$ 的形式为:

$$z = [h_{N_i,1}^t(z), h_{N_i,2}^t(z), \dots, h_{N_i,C}^t(z)]. \quad (17)$$

计算 z 的样本效用:

$$U_{N_i}^t(z) = \sum_c \left(\frac{\sum r_{N_i}}{\sum q_{N_i}} \cdot \lambda_c \cdot h_{N_i,c}^t(z) \right), \quad (18)$$

然后,将所有样本效用按非递增顺序排列: $U_{N_i}^t(1) \geq U_{N_i}^t(2) \geq \dots \geq U_{N_i}^t(z)$.紧接着,选择样本数量为 $\lceil (1-\rho) \cdot z \rceil$ 的 Z 部分子集,计算缓存 c 的概率:

$$p_{N_i,c}^{i+1} = \frac{\sum_{z=1}^z X_{U_{N_i}^t(z) \geq \lceil (1-\rho) \cdot z \rceil} \cdot h_{N_i,c}^t(z)}{\sum_{z=1}^z X_{U_{N_i}^t(z) \geq \lceil (1-\rho) \cdot z \rceil}}. \quad (19)$$

其中,样本效用 $U_{N_i}^t(z) \geq U_{N_i}^t(\lceil (1-\rho) \cdot z \rceil)$ 时, $X_{U_{N_i}^t(z) \geq \lceil (1-\rho) \cdot z \rceil} = 1$; 否则为 0.

如上所述,RSU 可以根据内容的流行程度制定动态的内容缓存策略.RSU 基于流行度动态内容缓存算法如算法 2 所示.

算法 2 RSU 基于流行度动态内容缓存算法

```

输入  迭代次数  $t_d$ , 观察周期  $\Delta t$ , 请求集合  $Q = \{q_1, q_2, \dots, q_c, \dots, q_C\}$ , 优势比例  $\rho$ 
输出  RSU 缓存内容策略
初始化  $N_i$  每个内容  $c$  的缓存概率  $p_{i,c}^{t+1}$ ;
for  $i=1, 2, \dots, N$  do
    根据  $Q$  统计内容请求概率, 计算 RSU 存储内容  $c$  的流行度;
end for
while  $\Delta t \geq 0$  do
    for  $i=1, 2, \dots, N$  do
        统计  $N_i$  每个内容  $c$  请求次数, 根据流行度更新  $N_i$ 
    end for
    缓存内容;
     $\Delta t = \Delta t - 1$ ;
end while
for  $t=1, 2, \dots, t_d$  do
    for  $i=1, 2, \dots, N$  do
        根据式(15)计算内容分布概率, 获取样本集合  $Z$ ;
        for  $z=1, 2, \dots, z$  do
            根据式(18)计算  $z$  效用;
        end for
        按照非递增的顺序排列样本;
        选择数量为  $\lceil (1-\rho) \cdot z \rceil$  样本;
    end for
    for  $i=1, 2, \dots, N$  do
        for  $z=1, 2, \dots, z$  do
            根据式(18)和(19)更新  $p_{i,c}^{t+1}$ ;
        end for
    end for
end for

```

DACO 算法分为两个阶段进行.第一阶段是选择符合特征的缓存车辆队列,算法复杂度为 $O(K^2)$, 其中 K 是缓存车辆数量.第二阶段应对内容流行度变化的动态内容缓存算法设计,这一阶段的算法复杂度为 $O(N)$, 其中 N 是 RSU 数量.

4 仿真实验与数值分析

4.1 仿真设置

本文仿真实验数据集是 2023 年 Elevated Highway Vehicle Trajectory Dataset 数据集^[29],实现 100% 的车辆识别和跟踪.路段全长 352 m,南北走向,日均车流量适中,涵盖多种类型的车辆(如轿车、公交车等),较好地反映车联网中车辆行驶状况.仿真实验中设置了 4 个 RSU,RSU 存储空间为 [20, 50] MB, OBU 的为 [1.0, 2.5] MB.每个内容大小遵循 [0.1, 0.5] MB 均匀分布,传输速率介于 [1, 3] Mbps 之间, OBU 间传输速率介于 [5, 10] Mbps,所有内容请求独立同分布.考虑命中率 and 传输时延两个性能指标验证算法性能.除本文算法之外,还考虑另外 3 种缓存替换算法:随机缓存算法(Random)、最不经常使用(least frequently Used, LFU)缓存算法和最近最少缓存算法(least recently used, LRU).

1) Random: 存储内容随机更新,每个内容被选中的概率相等.即每个缓存设备在每个时间段随机选择一个内容更新.

2) LFU: 根据请求内容频率更新内容,避免长时间内没被访问的“僵尸”条目占用缓存空间.本文中,RSU

在每个时间段将请求次数最少的内容进行更新.

3)LRU:基于缓存内容最近被访问时间进行替换决策,即更新最近访问时间最早的缓存内容.本文中,RSU 在每个时间段将最久未被访问的内容进行替换.

4.2 算法性能及结果分析

为验证协算法的有效性,对比相同条件下缓存车辆的 OBU 与 RSU 协同传输内容(DACO)、仅 RSU 传输内容(RSU Transport Contents,RTC)、仅缓存车辆传输内容(caching vehicles transport contents,CTC)下的命中率和传输时延,以“响应时间”衡量算法实时处理能力.为提高结果的准确性和可信度,按照相同实验设置,进行 10 次独立的重复实验,计算平均值,实验结果如表 1 所示.

表 1 实验结果

Tab. 1 Experimental Result

算法名称	DACO	LRU	LFU	Random
响应时间/ms	61	170	185	258

由表 1 可知,DACO 在做出缓存策略决策时,所需响应时间最小,响应时间远小于 Random,LRU 和 LFU 的响应时间相差不大.258 ms 的延迟可能会导致行驶车辆或 RSU 无法响应其他车辆请求,也可能导致车辆无法及时响应环境变化从而影响车辆的安全行驶.结合实验结果,DACO 算法能够高效、快速地制定缓存决策,低延迟响应请求,满足实时车载应用的需求.

图 2 展示 3 种算法的系统命中率和传输时延的表现.条件相同时 DACO 的命中率最高,时延最小,是因为 DACO 充分利用 OBU 和 RSU 的存储空间.随着内容种类增加 3 种方式命中率均降低,是因为随着内容种类增多,车辆请求内容变得更加多样化,而存储空间受限,缓存未命中风险增加.图 2(b)中随内容种类增加 RTC 的时延从高于 CTC 逐渐低于 CTC 的情况,这是合理的,因为虽然 RTC 的命中率比 CTC 高,但 RSU 距离请求车辆相较于缓存车辆较远,时延增加.

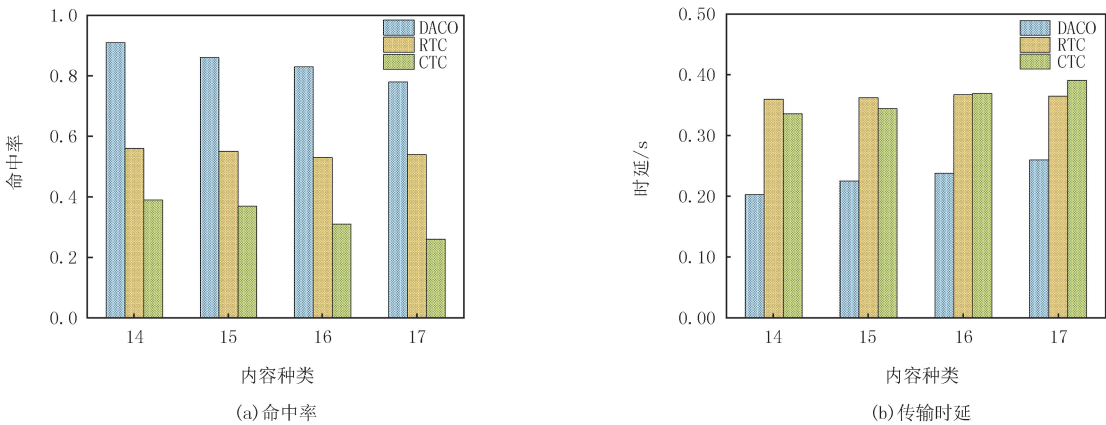


图2 缓存协作方式的性能表现

Fig.2 Performance of the cache collaboration approach

图 3 展示内容种类从 14 到 17 时命中率和传输时延表现.图 3(a)中随着内容种类增加,4 种算法命中率均呈下降趋势.因为内容种类增多导致需要缓存更多内容满足请求需求,而存储容量有限,命中率降低.内容种类相同时 DACO 命中率最高,体现了本文所提算法的优越性.LRU 通过替换每个时间段内容最久未被访问内为流行内容腾空间,但效果不及 DACO.图 3(b)中随着内容种类增多 4 种算法传输时延均有增加,因为内容种类增多,OBU 和 RSU 存储内容更难覆盖所有需求,导致更多内容从云服务器传输,传输时延增加.内容种类相同时本文算法传输时延最低,因为其能动态调整 RSU 存储内容,使更多请求内容从 RSU 直接传输降低传输时延.

图 4 展示缓存车辆的 OBU 存储容量在[1.0,2.5] MB 变化时命中率和传输时延表现,4 种算法的命中率均随 OBU 容量的增加而增加,传输时延随 OBU 容量的增加而减少.这是因为 OBU 存储容量增加相当于整体的缓存空间变大,可缓存更多内容,减少从云服务器获取内容.条件相同时 DACO 命中率最高,是因为它能够有

效地更新更符合请求“偏好”.Random 是随机选择内容进行更新,不确定性因素多,命中率较低.云服务器传输时延相较于其他两种传输方式更长.LFU 基于内容的访问频率做出替换决策,表现介于 DACO 和 LRU 之间.

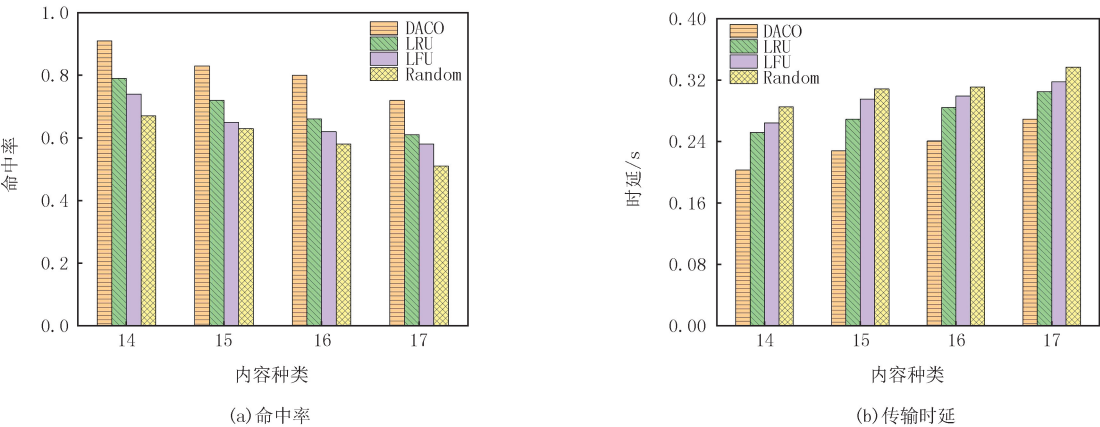


图3 内容种类变化时的表现
Fig.3 Performance when varying the types of content

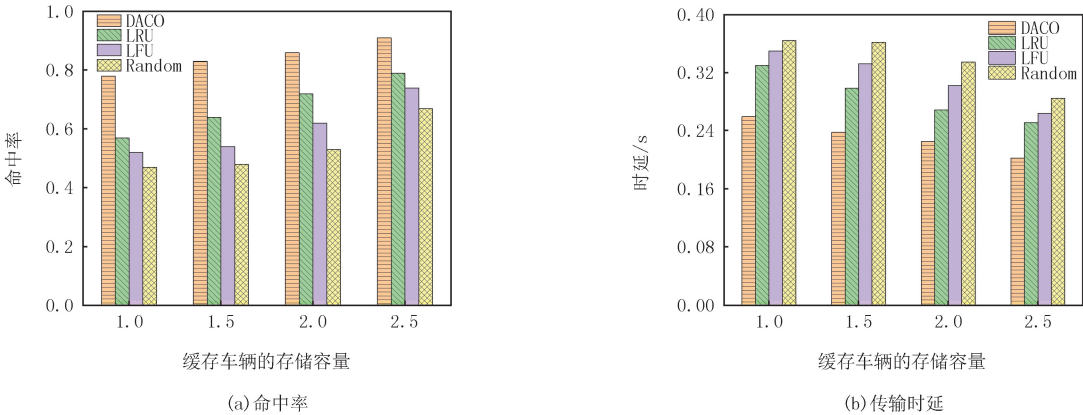


图4 缓存车辆存储容量变化时的表现
Fig.4 Performance when varying storage capacity of cache vehicle

图 5 显示 RSU 存储容量在[20,50] MB 范围内变化时,命中率和传输时延的表现.图 5(a)显示 4 种算法命中率与 RSU 存储容量正相关,原因是 RSU 存储容量增加能满足更多请求,提高命中率.其他条件相同时,DACO 命中率最高,原因是它可以根据内容的动态变化更新 RSU 存储内容.图 5(b)显示 4 种算法时延与 RSU 存储容量呈负相关,原因是更多车辆请求可从 RSU 或缓存车辆得到满足,无需通过云服务器进行长距离传输.RSU 和缓存车辆与车辆之间的传输距离相较于云服务器更短,传输时延也相应减少.

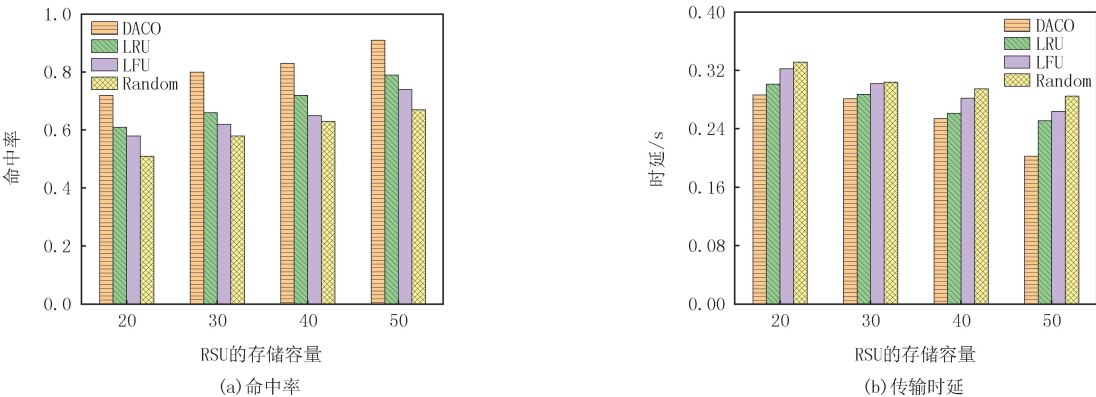


图5 RSU 的存储容量变化时的表现
Fig.5 Performance when varying storage capacity of RSU

5 结 论

本文研究了以用户为中心的车联网中内容交付问题,其目标是在不确定请求内容的情况下最小化传输时延.考虑到 IoV 实际场景,联合 RSU 和 OBU 存储空间提出了一种基于延迟感知内容缓存优化算法,优化传输时延.实验结果表明,相比于仅考虑 RSU 或仅考虑 OBU 存储的策略以及其他基准算法,本文所提出策略在性能上有显著的优势.下一步将深化对 RSU 和 OBU 合作中缓存策略研究,进一步提升内容传输的效率和质量,优化整个车联网系统性能.

参 考 文 献

[1] ZHOU H,JIANG K,HE S B,et al.Distributed deep multi-agent reinforcement learning for cooperative edge caching in Internet-of-vehicles[J].IEEE Transactions on Wireless Communications,2023,22(12):9595-9609.

[2] ZHANG D G,NIC H,ZHANG J,et al.New method of vehicle cooperative communication based on fuzzy logic and signaling game strategy[J].Future Generation Computer Systems,2023,142:131-149.

[3] ZHANG T,ZHANG D G,YAN H R,et al.A new method of data missing estimation with FNN-based tensor heterogeneous ensemble learning for Internet of vehicle[J].Neurocomputing,2021,420:98-110.

[4] MUHAMMAD K,ULLAH A,LLORET J,et al.Deep learning for safe autonomous driving;current challenges and future directions[J].IEEE Transactions on Intelligent Transportation Systems,2021,22(7):4316-4336.

[5] YANG J N,DING M,MAO G Q,et al.Optimal base station antenna downtilt in downlink cellular networks[J].IEEE Transactions on Wireless Communications,2019,18(3):1779-1791.

[6] JUNG S,KIM J,LEVORATO M,et al.Infrastructure-assisted on-driving experience sharing for millimeter-wave connected vehicles[J].IEEE Transactions on Vehicular Technology,2021,70(8):7307-7321.

[7] ZHAO J H,LI Q P,GONG Y,et al.Computation offloading and resource allocation for cloud assisted mobile edge computing in vehicular networks[J].IEEE Transactions on Vehicular Technology,2019,68(8):7944-7956.

[8] DAI P L,SONG F,LIU K,et al.Edge intelligence for adaptive multimedia streaming in heterogeneous Internet of vehicles[J].IEEE Transactions on Mobile Computing,2021:1.

[9] GU H X,WANG H J.A distributed caching scheme using non-cooperative game for mobile edge networks[J].IEEE Access,2020,8:142747-142757.

[10] SU Z,HUI Y L,XU Q C,et al.An edge caching scheme to distribute content in vehicular networks[J].IEEE Transactions on Vehicular Technology,2018,67(6):5346-5356.

[11] QIN J M,XUN Y J,LIU J J.CVMIDS:cloud-vehicle collaborative intrusion detection system for Internet of vehicles[J].IEEE Internet of Things Journal,2024,11(1):321-332.

[12] ZHANG D G,WANG S,ZHANG J,et al.A content distribution method of Internet of vehicles based on edge cache and immune cloning strategy[J].Ad Hoc Networks,2023,138:103012.

[13] ZHANG K,CAO J Y,ZHANG Y.Adaptive digital twin and multiagent deep reinforcement learning for vehicular edge computing and networks[J].IEEE Transactions on Industrial Informatics,2022,18(2):1405-1413.

[14] RUI L L,SONG D,CHEN S Y,et al.Content collaborative caching strategy in the edge maintenance of communication network;a joint download delay and energy consumption method[J].IEEE Transactions on Parallel and Distributed Systems,2022,33(12):4148-4163.

[15] ZENG F,ZHANG K W,WU L,et al.Efficient caching in vehicular edge computing based on edge-cloud collaboration[J].IEEE Transactions on Vehicular Technology,2023,72(2):2468-2481.

[16] YAO L,CHEN A L,DENG J,et al.A cooperative caching scheme based on mobility prediction in vehicular content centric networks[J].IEEE Transactions on Vehicular Technology,2018,67(6):5435-5444.

[17] LING C,ZHANG W Z,FAN Q Y,et al.Cooperative service caching in vehicular edge computing networks based on transportation correlation analysis[J].IEEE Internet of Things Journal,2024,11(12):22754-22767.

[18] ZHOU L,ZHANG Y,SONG K,et al.Distributed media services in P2P-based vehicular networks[J].IEEE Transactions on Vehicular Technology,2011,60(2):692-703.

[19] AUNG N,DHELIM S,CHEN L M,et al.VeSoNet:traffic-aware content caching for vehicular social networks using deep reinforcement learning[J].IEEE Transactions on Intelligent Transportation Systems,2023,24(8):8638-8649.

[20] WU H H,FAN Y Z,JIN J C,et al.Social-aware decentralized cooperative caching for Internet of vehicles[J].IEEE Internet of Things Journal,2023,10(16):14834-14845.

[21] WANG R Y,KAN Z W,CUI Y P,et al.Cooperative caching strategy with content request prediction in Internet of vehicles[J].IEEE In-

ternet of Things Journal,2021,8(11):8964-8975.

[22] WANG P F,YAN Z H,OBAIDAT M S,et al.Edge caching with federated unlearning for low-latency V2X communications[J].IEEE Communications Magazine,2024,62(10):118-124.

[23] WANG X J,ZHU H L,NING Z L,et al.Blockchain intelligence for Internet of vehicles:challenges and solutions[J].IEEE Communications Surveys & Tutorials,2023,25(4):2325-2355.

[24] SU L,NIU Y,HAN Z,et al.Content distribution based on joint V2I and V2V scheduling in mmWave vehicular networks[J].IEEE Transactions on Vehicular Technology,2022,71(3):3201-3213.

[25] KO B,LIU K,SON S H,et al.RSU-assisted adaptive scheduling for vehicle-to-vehicle data sharing in bidirectional road scenarios[J].IEEE Transactions on Intelligent Transportation Systems,2021,22(2):977-989.

[26] LIU X Y,XU Z L,MENG Y,et al.An elastic-segment-based V2V/V2I cooperative strategy for throughput enhancement[J].IEEE Transactions on Vehicular Technology,2022,71(5):5272-5283.

[27] ALI SALAHUDDIN M,AL-FUQAHA A,GUIZANI M.Software-defined networking for RSU clouds in support of the Internet of vehicles[J].IEEE Internet of Things Journal,2015,2(2):133-144.

[28] KHAN M T R,JEMBRE Y Z,SAAD M M,et al.Proactive content retrieval based on value of popularity in content-centric Internet of vehicles[J].IEEE Transactions on Intelligent Transportation Systems,2024,25(8):8514-8526.

[29] 冯汝怡,李志斌,吴启范,等.航拍视频车辆检测目标关联与时空轨迹匹配[J].交通信息与安全,2021,39(2):61-69.

FENG R Y,LI Z B,WU Q F,et al.Association of vehicle object detection and the time-space trajectory matching from aerial videos[J].Journal of Transport Information and Safety,2021,39(2):61-69.

An optimization strategy for latency-aware internet of vehicle caching

Wang Yali^{1,2}, Han Xiaopei²

(1. Computing Science and Artificial Intelligence College, Suzhou City University, Suzhou 215104, China;
2. College of Computer and Information Engineering, Henan Normal University, Xinxiang 453007, China)

Abstract: The development of vehicular networking technology has given rise to various data-intensive and latency-sensitive in-vehicle applications. Caching content at the user side can enable efficient content delivery for these applications. This paper investigates the user-centric content delivery problem in vehicular networks and proposes a latency-aware content caching optimization strategy, which minimizes transmission delays by coordinating roadside units(RSUs) and onboard units(OBUs) under uncertain content requests. First, basing on multi-criteria decision-making, we collects caching vehicle characteristics to determine the optimal caching vehicle. Then, a popularity-based dynamic content caching algorithm is proposed to derive the optimal caching strategy. Simulation results demonstrate that, compared with other schemes, the proposed strategy effectively improves cache hit rates while reducing content delivery latency.

Keywords: edge caching; mobile edge computing; internet of vehicles; collaborative cache

[责任编辑 陈留院 杨浦]