

·车联网通信性能优化与安全技术专题·

# 车联网中基于联邦和强化学习的边缘缓存策略\*

张良 张国栋 卢剑伟 雷夏阳 程浩

(合肥工业大学,合肥 230009)

**【摘要】**为解决车联网中传统内容流行度预测方法无法准确捕获车辆请求特性,导致缓存命中率较低的问题,提出了一种基于联邦学习和强化学习的边缘协同缓存策略。该策略将车辆请求概率更高的内容预缓存存在其他车辆或路侧单元中,以提高缓存命中率和降低平均内容获取延时。采用联邦学习方法利用分布在多个车辆上的私有数据进行训练并预测内容流行度,然后使用强化学习算法求解目标函数,获得流行内容的最佳缓存位置。结果表明,所提出的策略在缓存命中率和平均内容获取延时方面均优于其他对比缓存策略,有效提升了车联网边缘缓存性能。

**关键词:**智能交通 边缘缓存 车联网 联邦学习 强化学习

**中图分类号:**U463.6;TP181 **文献标志码:**A **DOI:** 10.19620/j.cnki.1000-3703.20230389

## Edge Caching Strategy of Internet of Vehicles Based on Federated and Reinforcement Learning

Zhang Liang, Zhang Guodong, Lu Jianwei, Lei Xiayang, Cheng Hao

(Hefei University of Technology, Hefei 230009)

**【Abstract】**In order to solve the problem that the traditional content popularity prediction method in the Internet of Vehicles cannot accurately capture the vehicle request characteristics and leads to the low cache hit rate, an edge collaborative caching strategy based on federated learning and reinforcement learning is proposed. This strategy pre-caches content with a higher probability of vehicle requests in other vehicles or roadside units to improve the cache hit ratio and reduce the average content acquisition delay. The federated learning method is used to train and predict the content popularity using private data distributed across multiple vehicles, and then the reinforcement learning algorithm is used to solve the objective function to obtain the best cache location for the popular content. The results show that the proposed strategy is better than other caching strategies in terms of cache hit ratio and average content acquisition delay, which effectively improves the performance of the edge cache of the Internet of Vehicles.

**Key words:** Intelligent transportation, Edge caching, Internet of Vehicles, Federated learning, Reinforcement learning

**【引用格式】**张良,张国栋,卢剑伟,等.车联网中基于联邦和强化学习的边缘缓存策略[J].汽车技术,2024(10):49-55.

ZHANG L, ZHANG G D, LU J W, et al. Edge Caching Strategy of Internet of Vehicles Based on Federated and Reinforcement Learning[J]. Automobile Technology, 2024(10): 49-55.

## 1 前言

为了解决车联网(Internet of Vehicles, IoV)远程云服务器的访问延时较高,难以满足低延时和多样化应用需求的问题,研究人员提出了内容缓存技术,通过将内容预先缓存到边缘节点上,以减少回程链路上的数据流量并降低服务延时<sup>[1]</sup>。然而,边缘节点的存储容量有限,因此缓存策略必须优先缓存最受车辆用户关注的流

行内容。缓存策略主要分为反应式缓存和主动式缓存两类<sup>[2]</sup>。反应式缓存依赖于用户请求后再进行缓存操作,如先进先出和最近最少使用策略。但这种方式仅在内容被请求后才会缓存,无法提前缓存未请求过的内容。主动式缓存通过预测内容流行度,在用户请求到达之前预先缓存流行内容,更适用于车联网场景。

在主动式缓存中,机器学习广泛应用于预测内容流行度,以提高缓存效率。文献[3]提出了基于联邦学

\*基金项目:国家重点研发计划项目(2021YFE0116600)。

通信作者:张国栋,硕士,研究方向为智能汽车技术,niandaiyijiu@163.com。

习的上下文感知主动缓存策略,通过对抗性自动编码器预测内容流行度。文献[4]利用齐夫(Zipf)模型计算内容流行度,并根据车辆和路侧单元(Road Side Unit, RSU)的内容获取延时决定内容的获取位置。然而,Zipf模型基于历史请求数据进行计算,难以准确捕捉内容流行度的时变特性<sup>[5]</sup>。为了解决这一问题,文献[6]采用长短期记忆网络预测时间序列中的内容请求,并将其结果作为Zipf模型的输入,以利于更准确地预测未来的内容请求。然而,识别不流行的内容仍然具有挑战性,因为未知和未请求的内容常常混淆在一起<sup>[7]</sup>。基于上述内容缓存策略的不足,本文构建一种城市交通边缘缓存场景,并对边缘节点的缓存决策以及网联车辆通信场景进行了建模与假设,在此基础上提出了一种基于联邦和强化学习的协同缓存策略(Edge Cooperative Caching based on Federated and Reinforcement learning, ECCFR)。然后以缓存命中率和内容传输延时为评价指标,通过Python平台进行仿真,系统地测试分析缓存总容量、缓存方式对所提出的协同缓存策略的影响。

## 2 城市交通边缘缓存场景构建

本文构建的城市边缘缓存场景为三层式架构,底层是网联车辆,中间层包含多个RSU,用于缓存附近车辆可能请求的内容,顶层是多个具有缓存能力的基站(Base Station, BS),用于存储其覆盖范围内所有RSU的缓存内容列表并管理其缓存资源,如图1所示。

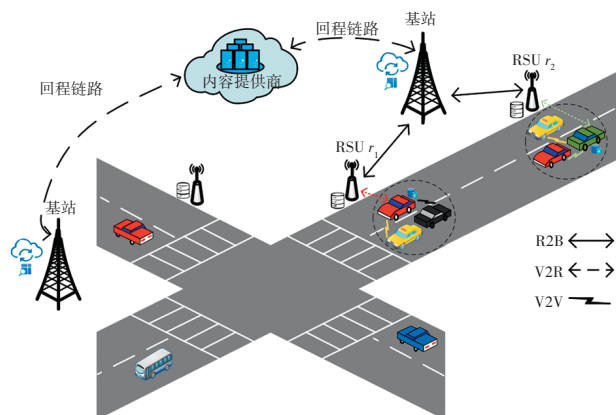


图1 边缘缓存场景

在BS的覆盖区域内,RSU  $r_s$  ( $s=1,2,3,\dots,S$ )以距离  $D_s$  等距放置在道路一侧,任意两个RSU之间没有重叠的覆盖区域。车辆  $v_i$  ( $i=1,2,3,\dots,I$ )会穿过多个RSU的覆盖区域<sup>[8]</sup>,假设所有车辆沿同一方向行驶,车辆的速度  $u_i$  ( $i=1,2,3,\dots,I$ )是独立同分布的,并且在每次通信过程中保持其指定速度不变,在  $t$  时间内进入每个RSU范围

内的车辆数量为  $V(t)$ ,遵循参数为  $\lambda$  的泊松过程<sup>[3]</sup>。因此,在时间  $t$  内进入RSU范围内的车辆数量  $g$  的概率可以表示为:

$$P(V(t) = g) = \frac{(\lambda t)^g}{g!} e^{-\lambda t} \quad (1)$$

当车辆经过某个RSU覆盖区域时,移动车辆可以连接到RSU并向其发送内容请求。如果请求的内容已缓存,即缓存命中。否则,请求车辆必须从互联网获取所请求的内容,即缓存未命中<sup>[9-10]</sup>。

## 3 缓存决策与通信时延模型构建

### 3.1 缓存决策模型

本文将RSU和用户车辆进行协作缓存,以便更准确地满足车辆用户的内容请求<sup>[11-14]</sup>。RSU会在每一时间段内根据车辆速度和位置选择多个连接车辆协作训练共享的全局模型,并在车群中选取一个车辆  $v_{cv}$  作为辅助缓存车辆,且辅助缓存车辆与车群中其他车辆仅存在一跳信号传输。

假设每个RSU  $r_s$  的缓存容量大小为  $C_{rsu}$ ,辅助缓存车辆  $v_{cv}$  的缓存容量大小为  $C_{cv}$ ,通常  $C_{cv} < C_{rsu} < N$ ,其中  $N$  为预测的前  $N$  个流行度较高的内容。为了表示每轮的请求内容  $q$  ( $q=1,2,3,\dots,Q$ )是否缓存在RSU和  $v_{cv}$  中,定义一个二进制矩阵  $Y_Q$  表示缓存决策:

$$Y_Q = \begin{bmatrix} y_{r_s,q} \\ y_{cv,q} \end{bmatrix} \quad (2)$$

$$y_{r_s,q}, y_{cv,q} \in \{0,1\} \quad (3)$$

式中:  $y_{r_s,q}=1$  表示将内容  $q$  缓存在RSU  $r_s$  中,  $y_{r_s,q}=0$  表示不缓存,  $y_{cv,q}=1$  表示将内容缓存在辅助缓存车辆  $v_{cv}$  中,  $y_{cv,q}=0$  表示不缓存。

### 3.2 通信时延模型

每辆车在每一轮通信中保持相同的通信模型,并在不同轮中更改其通信模型<sup>[15]</sup>。当通信轮次为  $d$  时,第  $i$  辆车  $v_i^d$  的信道增益建模为:

$$h_i^d(\text{dis}(x, v_i^d)) = l_i^d(\text{dis}(x, v_i^d)) g_i^d(\text{dis}(x, v_i^d)) \quad (4)$$

式中:  $x$  为与车辆  $v_i^d$  进行通信的边缘节点,当  $x=r_s$  时表示RSU  $r_s$ ,当  $x=v_{cv}$  时表示辅助缓存车辆  $v_{cv}$ ;  $\text{dis}(x, v_i^d)$  为辅助缓存车辆  $v_{cv}$  和车辆  $v_i^d$  之间的距离;  $l_i^d(\text{dis}(x, v_i^d))$  为辅助缓存车辆  $v_{cv}$  和车辆  $v_i^d$  之间的路径损耗;  $g_i^d(\text{dis}(x, v_i^d))$  为辅助缓存车辆  $v_{cv}$  和车辆  $v_i^d$  之间的阴影衰落,它们都遵循对数正态分布。

每个RSU通过车对路侧单元(Vehicle to RSU, V2R)链路与其覆盖区域内的车辆进行通信,而同一个车群中车辆通过车对车(Vehicle to Vehicle, V2V)链路

进行通信<sup>[16]</sup>。RSU  $r_s$ 和辅助缓存车辆  $v_{cv}$ 在不同的通信轮次中与车辆  $v_i^d$ 之间的距离不同,因此在不同通信轮次中以不同的速率进行传输,V2R链路的传输速率为<sup>[17]</sup>:

$$R_{r_s,i}^d = B \log_2 \left( 1 + \frac{P_{r_s} h_i^d (dis(r_s, v_i^d))}{\sigma_c^2} \right) \quad (5)$$

式中: $B$ 为可用带宽, $P_{r_s}$ 为RSU  $r_s$ 的发射功率, $\sigma_c^2$ 为噪声功率。

类似的,V2V链路的传输速率计算如下:

$$R_{v_{cv},i}^d = B \log_2 \left( 1 + \frac{P_{v_{cv}} h_i^d (dis(v_{cv}, v_i^d))}{\sigma_c^2} \right) \quad (6)$$

式中: $P_{v_{cv}}$ 为  $v_{cv}$ 的发射功率。

基于上述分析,可以建立内容传输时延模型,当内容被缓存在不同位置时,获取所请求内容  $q$ 的内容传输延时的不同的。如果内容  $q$ 缓存在本地RSU  $r_s$ 中,车辆  $v_i^d$ 通过V2R链路从RSU  $r_s$ 处获取内容的延时可以定义为:

$$T_{r_s,i,q}^d = \frac{s_q}{R_{r_s,i}^d} \quad (7)$$

式中: $s_q$ 为内容  $q$ 的大小。

通过V2V链路从辅助缓存车辆  $v_{cv}$ 处获取内容的延时为:

$$T_{v_{cv},i,q}^d = \frac{s_q}{R_{v_{cv},i}^d} \quad (8)$$

## 4 基于联邦和强化学习的协同缓存策略

该策略主要包含2个部分,即基于联邦学习的内容流行度预测和基于强化学习的协作式缓存算法,如图2所示。

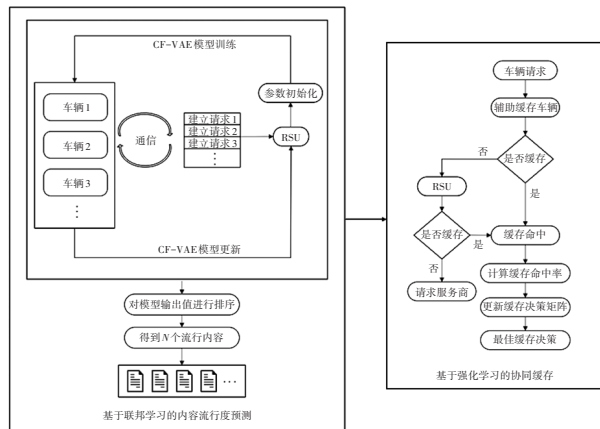


图2 基于联邦和强化学习的协同缓存策略

### 4.1 基于联邦学习的内容流行度预测

#### 4.1.1 联邦学习

联邦学习(Federated Learning, FL)模型训练会执行多轮通信,每轮通信分为5步,如图3所示。

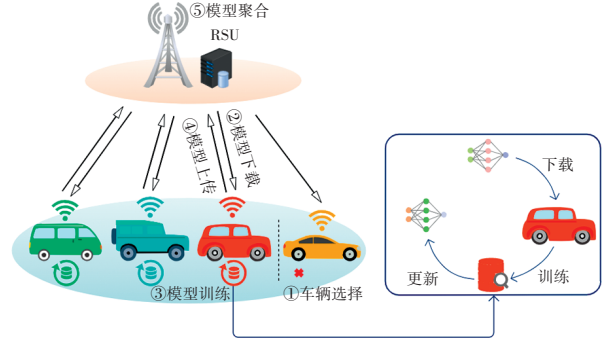


图3 联邦学习模型

a. 车辆选择。车辆选择过程将考虑车辆在RSU区域内的停留时间,这主要取决于联网车辆的位置和速度。假设第  $d$  轮通信第  $i$  个连接车辆的速度为  $u_i^d$ , ( $u_{\min} \leq u_i^d \leq u_{\max}$ ), 且  $u_i^d$  遵循截断高斯分布<sup>[3]</sup>, 则  $u_i^d$  的概率密度函数表示为:

$$f(u_i^d) = \begin{cases} \frac{e^{-\frac{1}{2\sigma^2}(u_i^d - \mu)^2}}{\sqrt{2\pi\sigma^2} \left( \text{erf}\left(\frac{u_{\max} - \mu}{\sigma\sqrt{2}}\right) - \text{erf}\left(\frac{u_{\min} - \mu}{\sigma\sqrt{2}}\right) \right)}, & u_{\min} \leq u_i^d \leq u_{\max} \\ 0, & \text{其他} \end{cases} \quad (9)$$

式中: $\text{erf}()$ 为  $u_i^d$ 在均值  $\mu$ 和方差  $\sigma^2$ 下的高斯误差函数; $u_{\min}$ 和  $u_{\max}$ 分别为联网车辆的最小车速和最大车速,则每辆车  $v_i^d$ 在当前RSU覆盖区域内停留的时间为:

$$T_{\text{std}}^{i,d} = \frac{(D_r - P_i^d)}{u_i^d} \quad (10)$$

式中: $D_r$ 为RSU覆盖区域的直径, $P_i^d$ 为  $v_i^d$ 在RSU覆盖范围内经过的距离。

假设每一轮通信的平均训练时间为  $T_{\text{tm}}$ 并且推理时间为  $T_{\text{inf}}$ ,若  $T_{\text{std}}^{i,d} > T_{\text{tm}} + T_{\text{inf}}$ ,则表示车辆满足停留时间的要求,并被选择参与FL训练,反之忽略<sup>[3]</sup>。

b. 模型下载。在第  $d$  轮通信中,RSU基于条件过滤变分自动编码器(Conditional Filtering based Variational AutoEncoder, CF-VAE)初始化全局模型  $\omega^d$ 。在之后每一轮中,RSU都更新全局模型,并将更新后的全局模型发送给所有选定车辆。

c. 模型训练。每个车辆将从RSU处下载的全局模型  $\omega^d$ 设置为初始本地模型,并利用车辆的本地数据进行训练迭代该模型。设第  $k$  轮迭代的本地模型为  $\omega_{i,k}^d$ , 则车辆  $v_i^d$ 在本地模型  $\omega_{i,k}^d$ 下的局部损失函数为:

$$l(\omega_{i,k}^d) = \frac{1}{|Q_{i,k}^d|} \sum_{j \in Q_{i,k}^d} l_j(\omega_{i,k}^d) \quad (11)$$

式中: $Q_{i,k}^d$ 为车辆  $v_i^d$ 从训练集中随机选取的数据,  $l_j(\omega_{i,k}^d)$ 为第  $k$  轮迭代时选取的数据中第  $j$  个数据的预测损失。

第 $k$ 轮迭代结束后,车辆 $v_i^d$ 再次随机采样一些训练数据并开始第 $(k+1)$ 次迭代。在 $(k+1)$ 次迭代中的本地模型为:

$$\omega_{i,k+1}^d \leftarrow \omega_i^d - \eta \nabla l(\omega_{i,k}^d) \quad (12)$$

式中: $\eta$ 为固定学习率, $\nabla l(\omega_{i,k}^d)$ 为 $l(\omega_{i,k}^d)$ 的梯度。

重复以上过程,当迭代次数达到设定的最大值时,车辆 $v_i^d$ 完成本地训练。

d. 模型上传。将更新后的本地模型 $\omega_i^{d+1}$ 从车辆上传到RSU服务器。

e. 模型聚合。车辆上传模型后,RSU将所有本地模型进行加权和计算,生成新的全局模型 $\omega^{d+1}$ ,下一轮通信训练的全局模型更新为:

$$\omega^{d+1} = \sum_{i=1}^I \frac{q_i^d}{q^d} \gamma_i \omega_i^{d+1} \quad (13)$$

式中: $q_i^d$ 为车辆 $v_i^d$ 中的本地数据量, $q^d$ 为选定车辆的总本地数据量, $\omega_i^{d+1}$ 为车辆 $v_i^d$ 更新后的本地模型, $\gamma_i = P_i^d/D_r$ 为车辆 $v_i^d$ 的聚合权重。

#### 4.1.2 流行内容预测

用于联邦训练的模型是CF-VAE,主要由推理神经网络和生成神经网络组成。分别将矩阵 $X$ 和矩阵 $\hat{c}$ 输入推理神经网络以学习潜在在表征 $z$ ,然后将这些获得的表征组合重构后输出矩阵 $\tilde{X}$ 。用作输入数据的 $X$ 是车辆用户的内容检索历史,若用户之前请求过某个内容,则将这个内容标记1,反之则标记为0。但将所有未请求的内容都标记0是不合理的,因为不感兴趣的内容中包含了未知内容和未被请求的内容。因此,为了区别未知内容,使用随机抽样标记未知内容,对应于矩阵中空白的位罝。 $\hat{c}$ 为车辆用户上下文信息矩阵, $\tilde{X}$ 为内容流行度矩阵。将输出值进行排序后,输出中得分最高的 $N$ 个内容将被选作流行内容,如图4所示。

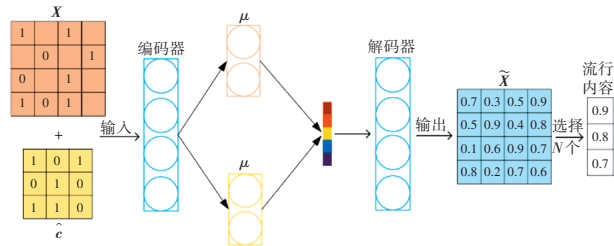


图4 基于CF-VAE的流行内容预测

通常,CF-VAE假设对于每个 $x \in X$ ,会对应1组或多组 $z \sim p(z)$ ,其中 $p(z)$ 是潜在变量 $z$ 的概率分布。目标是最大化生成每个输入数据 $x$ 的概率,它可以定义成:

$$p(x) = \int p(x|z)p(z)dz \quad (14)$$

虽然先验分布 $p(z)$ 和似然估计 $p(x|z)$ 都可以精确表示,但后验分布 $p(z|x)$ 需要在潜在空间上进行难以处理

的积分。因此,CF-VAE选择利用推理神经网络 $q(z|x)$ 来近似真实的后验分布 $p(z|x)$ ,从而提供更可能产生输入数据 $x$ 的潜在变量的分布,且近似后验分布 $q(z|x)$ 遵循高斯分布 $N(\mu, \sigma^2)$ , $\mu$ 为均值, $\sigma^2$ 为方差。这是通过最小化 $q(z|x)$ 和 $p(z|x)$ 之间的散度(Kullback-Leibler, KL)来实现的,将其表示为:

$$D[q(z|x)||p(z|x)] = E_{z \sim q(z|x)}[\log q(z|x) - \log p(z|x)] \quad (15)$$

应用贝叶斯推理,可以得到:

$$D[q(z|x)||p(z|x)] = E_{z \sim q(z|x)}[\log q(z|x) - \log p(z|x)] + \log p(x) \quad (16)$$

为了最小化 $D[q(z|x)||p(z|x)]$ ,将式(16)简化为:

$$\log p(x) \geq E_{z \sim q(z|x)}[\log p(x|z)] - D[q(z|x)||p(z)] \quad (17)$$

公式(17)右侧为CF-VAE的变分下界,生成网络 $p(x|z)$ 和推理网络 $q(z|x)$ 通过最大化其变分下界来进行训练。

最后,通过训练好的模型来预测内容的流行度。完整的内容流行度预测算法在算法1中进行概述,如图5所示。

```

输入:  $X, \hat{c}, P_i^d, D_r$ 
1: 初始化全局模型  $\omega^d$ 
2: for each 通信轮次  $d=1,2,3,\dots$  do:
3:   for each 请求车辆  $v_i^d$  do:
4:      $T_{std}^{i,d} = (D_r - P_i^d)u_i^d$ 
5:     if  $T_{std}^{i,d} > T_{th} + T_{inf}$  then
6:        $v_i^d$  被选择参加 FL 训练
7:     end if
8:   end for
9: end for
10: for each 被选择的车辆  $v_i^d$  do:
11:    $\omega_i^{d+1} \leftarrow$  车辆本地训练( $\omega^d, i$ )
12:   上传本地模型  $\omega_i^{d+1}$  到 RSU
13: end for
14:  $\omega^{d+1} \leftarrow \sum_{i=1}^I \frac{q_i^d}{q^d} \gamma_i \omega_i^{d+1}$ 
15: 预测流行内容并排序后选择前  $N$  个
16: return  $\omega^{d+1}, N$ 
车辆本地运行:
输入:  $\omega^d$ 
1: function 车辆本地训练( $\omega^d, i$ ):
2:   for each  $k = 1, 2, 3, \dots$  do:
3:     随机从数据集采样一部分数据  $Q_{i,k}^d$  用于训练
4:      $\omega_{i,k+1}^d \leftarrow \omega_i^d - \eta \nabla l(\omega_{i,k}^d)$ 
5:   end for
6:   return  $\omega_i^{d+1}$ 
7: end function
    
```

图5 算法1:基于联邦学习的内容流行度预测算法

## 4.2 基于强化学习的协同缓存

强化学习(Reinforcement Learning, RL)是一种自适应算法,对于主动式缓存,RL是一种较好的求解方法,故本文提出了一种基于Q-learning<sup>[18]</sup>的协作式内容缓存算法。Q-learning会将状态与动作对应的奖励值记录在

$Q$ 矩阵中,在强化学习每次迭代的过程中, $Q$ 矩阵会不断更新其中的元素,迭代一定的次数后, $Q$ 矩阵将趋于稳定。最后,在测试阶段,智能体将选择最高的奖励及其相应的动作来执行给定系统状态的缓存决策。完整的协同缓存算法在算法2中进行了概述,如图6所示。

```

1: 初始化  $Q$  矩阵、缓存决策矩阵  $Y_0$ 
2: 输入:第  $d$  轮通信本地 RSU 中所有车辆请求的前  $N$  个流行内容
3: for  $episode = 1, 2, 3 \dots$  do:
4:   for 时隙  $slot = 1, 2, 3 \dots$  do:
5:     观察系统状态  $s$ 
6:     从给定状态  $s$  的动作集中随机选取动作  $a$ 
7:     根据动作表示的缓存决策向量  $\alpha$  计算缓存命中率  $Hr$ 
8:      $R(s, a) = Hr$ 
9:      $Q(s, a) \leftarrow Q(s, a) + \alpha [R(s, a) + \lambda \max_{a'} Q(s', a') - Q(s, a)]$ 
10:     $s \leftarrow s'$ 
11:   end for
12: end for
13: return  $Q$  矩阵
14: 根据系统状态找到最佳动作,即最佳缓存决策矩阵  $\alpha_{best}$ 
    
```

图6 算法2:基于Q-learning的协同缓存算法对强化学习在ECCFR中的作用,作如下解释:

a. 状态。将RSU  $r_s$ 和辅助缓存车辆  $v_{cs}$ 的总缓存容量视为系统状态。

b. 动作。为了获得相应的奖励,智能体首先需要基于观测到的系统状态执行动作,即缓存决策。因此,本文将缓存决策矩阵  $Y_0$  视为强化学习中的动作集。

c. 奖励。在根据系统状态执行动作后,将生成相应的奖励,用于在训练  $Q$  矩阵的过程中更新元素值,在测试阶段将直接执行最大奖励值对应的动作。因此,本文将最大化缓存命中率设计为奖励函数,缓存命中率表示为:

$$Hr = Pr_{\zeta} \times \alpha_r + Pr_{\zeta} \times \alpha_{cv} \quad (18)$$

式中: $Pr_{\zeta} = [p_1, p_2, p_3, \dots, p_N]$ 为下一时间段内请求的每个内容的概率向量, $\alpha_r$ 为RSU相应的缓存决策向量, $\alpha_{cv}$ 为辅助缓存车辆的缓存决策向量。

每一状态动作组合对应一个奖励值  $R(s, a)$ ,即缓存命中率值,智能体会根据该奖励值通过更新公式对  $Q$  矩阵中第  $s$  行、第  $a$  列的元素值进行更新。 $Q(s, a)$ 的更新公式如下所示:

$$Q(s, a) = Q(s, a) + \alpha [R(s, a) + \lambda \max_{a'} Q(s', a') - Q(s, a)] \quad (19)$$

式中: $s'$ 和  $a'$ 为下一时刻的状态和动作; $\alpha$ 为学习率,是小于1的常数; $\lambda$ 为折扣因子,表示对未来奖励的衰减,若  $\lambda$  值接近0,表示智能体更关注直接的奖励, $\lambda$  值接近1,则表示智能体更注重未来的奖励。

## 5 仿真试验与分析

### 5.1 仿真环境设置与数据集

本文模拟了一个城市区域中的车辆边缘缓存场  
2024年 第10期

景,包括1个BS,2个RSU和15台仿真车辆。车辆和RSU、MBS之间的通信采用第三代合作伙伴计划(3rd Generation Partnership Project, 3GPP)蜂窝V2X(Cellular Vehicle to Everything, C-V2X)架构,其中参数设置参考3GPP标准<sup>[19]</sup>。仿真环境为Python3.6,深度学习框架为Pytorch。在试验中使用的数据来自MovieLens网站收集的真实数据集MovieLens 1M。该数据集包含了来自3884部电影的6041个匿名用户的1000210个评分,以及评分用户的上下文信息,例如性别、年龄、职业、邮编地址。每辆车从本地数据中随机选择80%的数据作为训练集,20%的数据作为测试集。为了模拟车辆请求内容的过程,假设用户对一部电影进行1次评分算为1次内容请求过程。

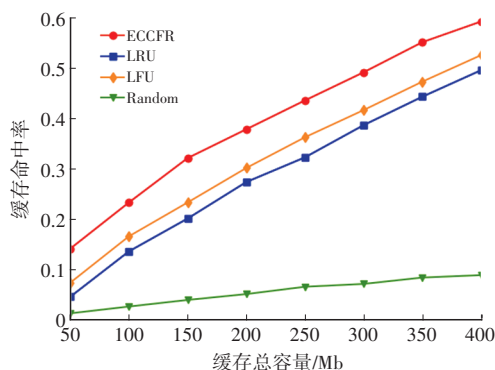
### 5.2 评价指标与试验对比

本文选择缓存命中率和内容传输延时作为性能指标来评估ECCFR策略,并使用3种传统缓存策略与ECCFR进行比较,3种比较策略为:随机缓存(Random),即随机缓存内容在RSU和辅助缓存车辆中;最近最少使用(Least Recently Used, LRU),即当达到缓存容量限制时,它首先删除缓存中最近使用次数最少的内容。最不常使用(Least Frequently Used, LFU),即只要缓存容量已满,缓存中最不常用的内容就会被删除。

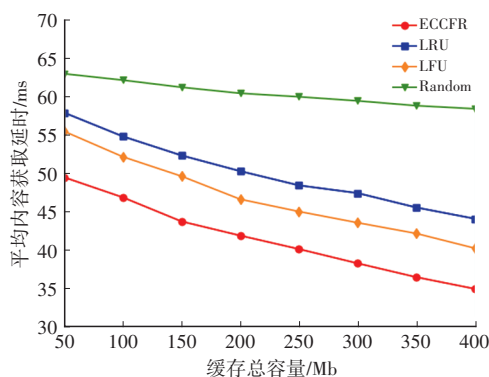
### 5.3 试验结果与分析

图7为不同缓存总容量下不同缓存策略性能对比。如图7a所示,各种缓存策略的缓存命中率随总缓存容量的增加而呈现上升趋势。这是因为更大的缓存容量意味着可以缓存更多的内容,辅助缓存车辆和RSU可以满足来自车辆用户的更多请求。此外,随机缓存策略Random的缓存命中率最低,因为该策略仅随机选择内容,而不考虑内容流行度。在缓存容量为350Mb时,Random、LRU、LFU和ECCFR的缓存命中率分别为8.32%、44.32%、47.33%和55.18%。与其他缓存策略相比,ECCFR的缓存命中率分别超过了46.86%、10.86%和7.85%。ECCFR缓存命中率更高的原因是因为采用了基于CF-VAE的内容请求预测,使得内容流行度更加准确,缓存了车辆用户在下一个时间段请求的可能性更大的内容。如图7b所示,在不同缓存容量下,所有缓存策略的平均内容传输延时随着缓存容量的增加而减小。这是因为,随着缓存容量的增加,缓存命中率将提高,较高缓存命中率使得车辆用户可以直接从边缘节点获取请求内容,从而避免通过延时更高的回程链路从远程内容提供商处获取请求的内容。在缓存容量为350Mb时,Random、LRU、LFU和ECCFR的内容获取平均延时

分别为 58.78 ms、45.50 ms、42.10 ms 和 36.41 ms,这意味着与其他策略相比,ECCFR 的平均内容获取延时分别节省了 22.37 ms、9.09 ms 和 5.69 ms。



(a)缓存命中率



(b)平均内容获取延时

图 7 不同缓存总容量下不同缓存策略的性能指标

图 8 为当缓存总容量为 100 时,ECCFR 中 RL 的迭代次数和缓存命中率及内容传输延时的关系。随着迭代次数的增加,缓存命中率逐渐增加,内容传输延时逐渐降低,这是因为辅助缓存车辆和 RSU 在前 10 次迭代中逐渐缓存了适当的流行内容。当迭代次数达到 10 次左右时,缓存命中率和内容传输延时开始收敛,这是因为智能体在第 10 次迭代时开始学习执行最佳协作缓存的策略。

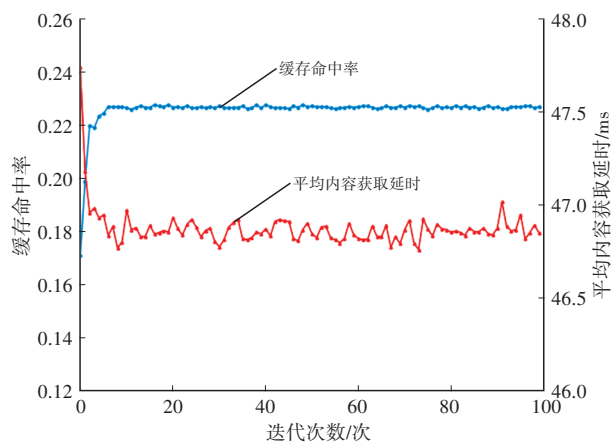
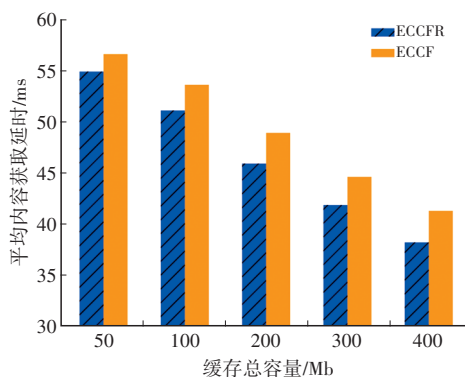
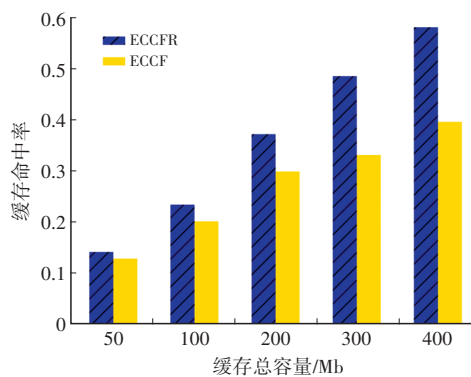


图 8 迭代次数与缓存命中率和平均内容获取延时的关系

图 9 为在不同缓存容量下,使用了 RL 的 ECCFR 缓存策略与没有使用 RL 的基于联邦学习的协同缓存策略 (Edge Cooperative Caching based on Federated learning, ECCF) 的缓存命中率和平均内容获取延时。如图 9a 所示,ECCFR 的缓存命中率优于无强化学习的 ECCF,这是因为 ECCFR 可以通过强化学习自适应地根据预测的流行内容来确定最佳协作缓存,从而可以在边缘节点中缓存更合适的流行内容。如图 9b 所示,ECCFR 的平均内容获取延时小于没有强化学习的 ECCF,这是因为 ECCFR 的缓存命中率优于没有强化学习的 ECCF,从而车辆用户能够以相对较低的延时从辅助缓存车辆和 RSU 处获取所请求的内容。



(a)平均内容获取延时



(b)缓存命中率

图 9 不同缓存总容量下 ECCFR 与 ECCF 策略对比

## 6 结束语

为了进一步提高 IoV 中的内容缓存命中率,减少内容获取延时,本文提出了一种基于联邦和强化学习的协作缓存策略 ECCFR。该策略基于条件过滤变分自动编码器来估计内容流行度,并使用强化学习来最大化缓存命中率。仿真结果表明,相比于其他缓存策略,ECCFR 在缓存命中率和平均内容传输延时方面表现更优。考虑到车辆高速移动时会导致频繁的网络拓扑结构变化,下一步将详细分析车辆的移动性对缓存决策的影响,研

究具有自主缓存决策能力的边缘节点之间的协作缓存决策,提高方法的可用性。

参 考 文 献

- [1] YU Z X, HU J, MIN G Y, et al. Proactive Content Caching for Internet-of-Vehicles Based on Peer-to-Peer Federated Learning[C]// 2020 IEEE 26th International Conference on Parallel and Distributed Systems (ICPADS). Hong Kong, China: IEEE, 2020.
- [2] YU Z X, HU J, MIN G Y, et al. Privacy-Preserving Federated Deep Learning for Cooperative Hierarchical Caching in Fog Computing[J]. IEEE Internet of Things Journal, 2021, 9(22): 22246–22255.
- [3] YU Z X, HU J, MIN G Y, et al. Mobility-Aware Proactive Edge Caching for Connected Vehicles Using Federated Learning[J]. IEEE Transactions on Intelligent Transportation Systems, 2020, 22(8): 5341–5351.
- [4] YE H, LI G Y, JUANG B H F. Deep Reinforcement Learning Based Resource Allocation for V2V Communications[J]. IEEE Transactions on Vehicular Technology, 2019, 68(4): 3163–3173.
- [5] 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.
- [6] WANG R Y, KAN Z W, CUI Y P, et al. Cooperative Caching Strategy with Content Request Prediction in Internet of Vehicles[J]. IEEE Internet of Things Journal, 2021, 8(11): 8964–8975.
- [7] MANIAS D M, SHAMI A. Making A Case for Federated Learning in the Internet of Vehicles and Intelligent Transportation Systems[J]. IEEE Network, 2021, 35(3): 88–94.
- [8] CUI Y P, DU L J, WANG H G, et al. Reinforcement Learning for Joint Optimization of Communication and Computation in Vehicular Networks[J]. IEEE Transactions on Vehicular Technology, 2021, 70(12): 13062–13072.
- [9] ZHAO L, LI H X, LIN M W, et al. Intelligent Content Caching Strategy in Autonomous Driving Toward 6G[J]. IEEE Transactions on Intelligent Transportation Systems, 2021, 23(7): 9786–9796.
- [10] ZHANG M, WANG S, GAO Q. A Joint Optimization Scheme of Content Caching and Resource Allocation for Internet of Vehicles in Mobile Edge Computing[J]. Journal of Cloud Computing, 2020, 9(1): 1–12.
- [11] ZHANG K, CAO J, MAHARJAN S, et al. Digital Twin Empowered Content Caching in Social-Aware Vehicular Edge Networks[J]. IEEE Transactions on Computational Social Systems, 2021, 9(1): 239–251.
- [12] ZHANG Y, WANG R, HOSSAIN M S, et al. Heterogeneous Information Network-Based Content Caching in The Internet of Vehicles[J]. IEEE Transactions on Vehicular Technology, 2019, 68(10): 10216–10226.
- [13] ZHANG K, LENG S P, HE Y J, et al. Cooperative Content Caching in 5G Networks with Mobile Edge Computing[J]. IEEE Wireless Communications, 2018, 25(3): 80–87.
- [14] CUI Y P, DU L J, HE P, et al. Multi-Vehicle Intelligent Collaborative Computing Strategy for Internet of Vehicles[C]// 2022 IEEE Wireless Communications and Networking Conference (WCNC). Austin, TX, USA : IEEE, 2022: 1647–1652.
- [15] CHEN S Z, HU J L, SHI Y, et al. Vehicle-to-Everything (V2X) Services Supported by LTE-Based Systems and 5G[J]. IEEE Communications Standards Magazine, 2017, 1(2): 70–76.
- [16] LIANG L, YE H, LI G Y. Spectrum Sharing in Vehicular Networks Based on Multi-Agent Reinforcement Learning[J]. IEEE Journal on Selected Areas in Communications, 2019, 37(10): 2282–2292.
- [17] CHEN J Y, WU H Q, YANG P, et al. Cooperative Edge Caching with Location-Based and Popular Contents For Vehicular Networks[J]. IEEE Transactions on Vehicular Technology, 2020, 69(9): 10291–10305.
- [18] WATKINS C J C H. Learning from Delayed Rewards[J]. Robotics & Autonomous Systems, 1989, 15(4): 233–235.
- [19] CHEN S Z, HU J L, SHI Y, et al. Vehicle-to-Everything (V2X) Services Supported by LTE-based Systems and 5G[J]. IEEE Communications Standards Magazine, 2017, 1(2): 70–76.

(责任编辑 王 一)

修改稿收到日期为2023年6月12日。