

## 双因子更新的车联网双层异步联邦学习研究

王力立\* 吴守林 杨妮 黄成

(南京理工大学自动化学院 南京 210094)

**摘要:** 针对车联网(IoV)中节点资源异构、拓扑结构动态变化等特点, 该文建立了一个双因子更新的双层异步联邦学习(TTAFL)框架。考虑到模型版本差和车辆参与联邦学习(FL)次数对局部模型更新的影响, 提出基于陈旧因子和贡献因子的模型更新方案。同时, 为了避免训练过程中, 车辆移动带来路侧单元切换的问题, 给出考虑驻留时间的节点选择方案。最后, 为了减少精度损失与系统能耗, 利用强化学习方法优化联邦学习的本地迭代次数与路侧单元局部模型更新次数。仿真结果表明, 所提算法有效提高了联邦学习的训练效率和训练精度, 降低了系统能耗。

**关键词:** 车联网; 联邦学习; 异步训练; 深度强化学习

中图分类号: TN929.5

文献标识码: A

文章编号: 1009-5896(2022)YU-0001-08

DOI: 10.11999/JEIT230918

## A Study of Two-layer Asynchronous Federated Learning with Two-factor Updating for Vehicular Networking

WANG Lili WU Shoulin YANG Ni HUANG Cheng

(Institute of Automation, Nanjing University of Science and Technology, Nanjing 210094, China)

**Abstract:** In response to the characteristics of heterogeneous node resources and dynamic changes in the network topology in the Internet of Vehicles (IoV), a Two-layer Asynchronous Federated Learning with Two-factor updating (TTAFL) framework is established in this paper. Considering the impact of model version differences and the number of times that vehicles participate in Federated Learning (FL) on server model updates, a model update scheme based on staleness factor and contribution factor is proposed. Furthermore, to avoid the problem of roadside unit switching caused by vehicle mobility during the training process, a node selection scheme considering the residence time is given. Finally, in order to reduce the accuracy loss and system energy consumption, a reinforcement learning method is used to optimize the number of local iterations of FL and the number of local model updates of roadside units. Simulation results show that the proposed algorithm effectively improves the training efficiency and training accuracy of federated learning and reduces the system energy consumption.

**Key words:** Internet of Vehicles (IoV); Federated Learning (FL); Asynchronous training; Deep reinforcement learning

### 1 引言

车辆在移动过程中会产生大量数据, 车辆之间的数据共享对驾驶安全与增强车载服务起着重要的作用。例如, 搭载摄像头的车辆可以采集行驶过程中的实时数据, 经过训练得到不同行车服务的模型, 能够实现交通标志识别、停车信息查询、路线规划和导航等功能<sup>[1, 2]</sup>。

联邦学习(Federated Learning, FL)是一种分布式机器学习方法, 能调度客户端利用自身数据在

本地进行模型训练, 然后将模型参数上传服务器聚合得到全局模型, 缓解了分布式现场学习的模型泛化能力差的问题<sup>[3, 4]</sup>。联邦学习可以促使车辆实现协同训练机器学习模型, 因车辆仅上传模型参数而非原始数据, 大幅增强了行车用户的隐私性<sup>[5]</sup>, 适合在车联网(Internet of Vehicles, IoV)中采用<sup>[6]</sup>。

根据参数服务器的位置, 可以将联邦学习分为基于边缘的联邦学习与基于云的联邦学习<sup>[7]</sup>。在车联网环境下, 分层联邦学习是一种更合适的训练方式, 具有低延迟和高可靠性<sup>[8]</sup>。文献<sup>[9]</sup>提出了一种基于车辆—路边设备—基站的分层联邦学习框架, 并且在车辆和路边设备之间设计了多领导者多玩家

的Stackelberg博弈。通过分层学习,克服了云联邦学习与边缘联邦学习的不足,但车联网中的分层联邦学习还存在诸多挑战。首先,由于车联网中参与车辆资源、计算能力的异构性,同步更新会导致快节点花费较长时间等待慢节点,这既造成了资源的浪费也降低了训练效率。第二,现有联邦学习大多采用联邦平均算法聚合模型,但在车联网动态复杂的环境中,不同车辆训练的本地模型对全局模型的贡献并不相同,若忽略该影响对模型采取平均聚合显然有失公平。第三,在双层联邦学习中,有两个参数对模型的训练效果有重要影响,分别是本地模型训练的迭代次数与局部模型的更新次数。一般来说,它们的值越大模型的精度越高,但也会造成更大的能耗。因此,有必要确定本地训练的迭代次数与局部模型的更新次数,以降低系统能耗与训练损失。此外,训练速度慢或训练效果差的节点会降低预测精度<sup>[10]</sup>,为保障联邦学习的可靠性并提高训练效率,可以在进行训练前对参与节点进行选择<sup>[11-13]</sup>。如文献[11]提出了基于信誉评分的联邦学习环境的最佳用户选择方法,可以消除性能欠佳的本地模型贡献,从而改善移动环境中的联邦学习。但车辆的机动性会导致链路不稳定,使参与车辆在训练过程中落后,严重影响联邦学习的训练性能<sup>[14]</sup>。

综上所述,本文构建了双因子更新的双层异步联邦学习(Two-layer Asynchronous Federated Learning with Two-factor updating, TTAFL)框架,在本联邦学习框架中,各路侧单元先对提出训练请求的车辆进行筛选,筛选通过后在本地训练。本文采取的异步方案是当路侧单元一旦接收到车辆

完成训练后提交的本地模型参数,便在路侧单元对局部模型进行更新。当所有的路侧单元都累计进行了一定次数更新后,基站便对各局部模型进行全局聚合。此外,为了优化联邦学习过程中本地训练的迭代次数与局部模型的更新次数,设置了基于双深度Q网络(Double Deep Q Network, DDQN)的联邦学习参数优化算法,有效降低了能耗与模型精度损失。

## 2 双因子更新的双层异步联邦学习

### 2.1 系统模型

本文所提“云-边缘-移动设备”双层异步联邦学习框架,由基站、 $M$ 个车辆 $V = \{v_1, \dots, v_m, \dots, v_M\}$ ,以及 $N$ 个路侧单元(RoadSide Unit, RSU)  $R = \{R_1, \dots, R_n, \dots, R_N\}$ 构成。其中,车辆 $v_m$ 可以在行驶过程中从环境中收集数据作为本地数据,记作 $D_m$ 。如图1所示,进入 $R_n$  ( $R_n \in R$ )覆盖范围的车辆 $v_m$ 可以在任意时刻向 $R_n$ 提出训练请求,经筛选后从 $R_n$ 上下载当前最新的局部模型然后利用本地数据进行训练。一旦完成本地训练,车辆便立即将模型参数上传到 $R_n$ 以完成对局部模型的更新。当所有的RSU都累计收到了车辆完成本地训练后上传的 $\kappa$ 个训练结果,并由此对RSU上的局部模型完成 $\kappa$ 次异步更新后,基站便对RSU上的局部模型执行全局聚合操作,得到的全局模型再下发给各RSU作为下一轮的局部模型。虽然车辆是动态的,但车辆请求的服务模型是从所属RSU上直接获取的,因此通过把联邦学习得到高精度的全局模型进一步部署到路侧单元,可以在长时间内满足不同来往车辆的服务请求。

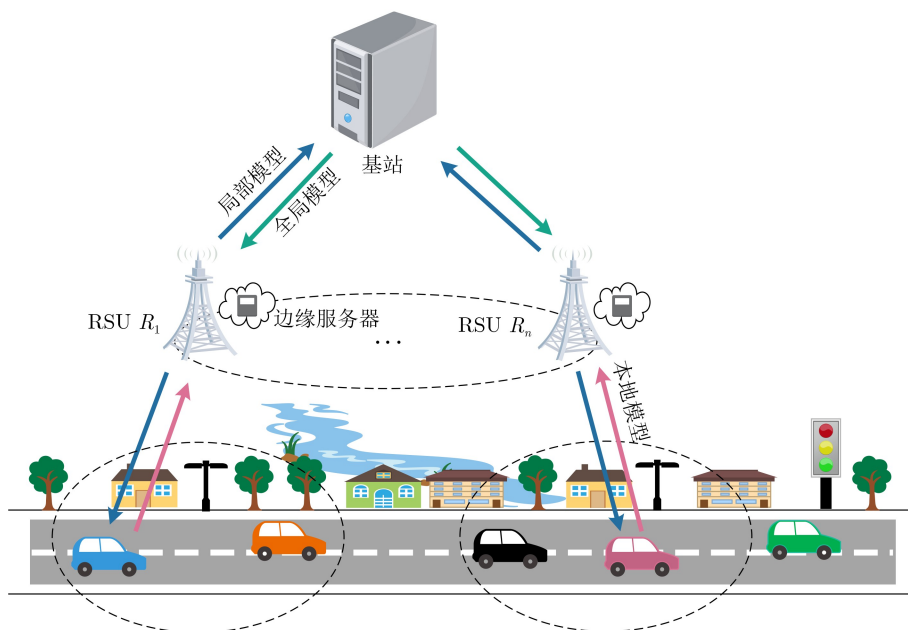


图1 车联网中的双层异步联邦学习框架

## 2.2 考虑驻留时间的车辆选择

在联邦学习过程中，由于车辆的持续移动导致RSU之间的服务切换，有的车辆在利用当前RSU  $R_n$  的局部模型进行本地训练时，可能还未完成就已经进入了下一个RSU  $R_{n+1}$  区域，导致本地模型无法上传给  $R_n$ 。鉴于此，在以模型精度等因素进行车辆选择的基础上，同时考虑车辆驻留时间非常有必要。只有训练时间及通信时间的总和小于行驶时间的车辆才可以参加训练。具体地，根据车辆  $v_m$  提出训练请求时在该路段的剩余里程数  $L_m$ ，与在该路段车辆通行的平均速度  $\vartheta$ ，得到  $v_m$  的驻留时间

$$T_{\text{stay}}(m) = \frac{L_m}{\vartheta} \quad (1)$$

车辆  $v_m$  在数据为  $D_m$  时完成1次本地训练的时间为

$$T_L(m) = \frac{|D_m| \cdot \beta_m}{\xi_m} \quad (2)$$

其中， $|D_m|$  是车辆  $v_m$  本地训练数据的大小， $\beta_m$  是在1个单位数据上训练模型所需的CPU周期数， $\xi_m$  (CPU周期/s) 是车辆在某一时间段内执行计算任务的可用计算资源。

将数据传输中的总带宽记为  $B$ ，对于车辆  $v_m$ ，其上传模型参数到RSU的数据传输速率为

$$r_m = B \cdot \log_2 \left( 1 + \frac{h_m p_m}{N_0} \right) \quad (3)$$

其中， $N_0$  是噪声强度， $p_m$  是车辆的发射功率， $h_m$  是车辆与上行RSU之间的无线信道增益。车辆在完成本地训练后将模型参数  $\omega_m^\theta$  上传到RSU所需的通信时间则为

$$T_C(m) = \frac{|\omega_m^\theta| \cdot p_m}{r_m} \quad (4)$$

其中， $|\omega_m^\theta|$  是车辆传输模型参数的大小。

结合式(2)和式(4)，车辆  $v_m$  在1轮训练中所需要的总时间成本可表示为

$$T_{\text{sum}}(m) = T_L(m) + T_C(m) \quad (5)$$

被选中参与联邦学习的车辆必须满足条件为

$$T_{\text{stay}}(m) > T_{\text{sum}}(m) \quad (6)$$

## 2.3 本地训练

各RSU首先从基站上下载初始化的全局模型  $\omega_G^0$ ，进入其覆盖范围内的车辆  $v_m$  若满足式(6)，则从RSU上下载当前最新版本的局部模型。然后通过梯度下降方法对收集到的来自周围环境的数据集  $D_m$  进行训练，每训练1次视作完成1轮本地迭代，本地迭代的次数影响了模型训练的精度。将  $D_m$  中的数

据样本记作  $d_j = \{\mathbf{x}_j, y_j\}$ ，其中  $\mathbf{x}_j$  是模型的输入向量， $y_j$  是  $\mathbf{x}_j$  的实际标签值。本文使用随机梯度下降法对车辆的联邦学习模型进行迭代更新。而且，为了提高联邦学习中模型训练效率，在计算模型梯度值时使用小批量数据样本。

假设车辆  $v_m$  在局部模型迭代到  $t$  次时加入训练，其下载此时RSU上最新版本的局部模型  $\omega_{R_n}^t$  然后在本地进行随机梯度下降。当车辆处于第  $k$  轮迭代时，它在上一轮本地模型  $\omega_m^{k-1}$  的基础上利用本地数据继续训练，本地模型更新为

$$\omega_m^k = \omega_m^{k-1} - \eta_m \nabla g(\omega_m^{k-1}; B_{m,k}) \quad (7)$$

其中， $\omega_m^k$  是车辆  $v_m$  第  $k$  轮迭代时的本地模型， $\eta_m$  是  $v_m$  的学习速率， $B_{m,k} \in D_m$  是  $v_m$  第  $k$  轮迭代时使用的小批量数据样本。 $\nabla g(\omega_m^{k-1}; B_{m,k})$  是下降的梯度，可表示为

$$\nabla g(\omega_m^{k-1}; B_{m,k}) = \frac{1}{|B_{m,k}|} \sum_{j=1}^{|B_{m,k}|} \frac{\partial L(f(\omega_m^{k-1}, \mathbf{x}_j), y_j)}{\partial \omega_m^{k-1}} \quad (8)$$

其中， $f(\omega_m^{k-1}, \mathbf{x}_j)$  表示第  $k$  轮迭代的本地模型预测值，损失函数  $L(\omega_m^k)$  表示模型的预测值与真实值之间的差异程度，可由具体任务定义，如交叉熵损失函数和均方误差损失函数。车辆在训练数据集为  $B_{m,k}$  时的损失函数可表示为

$$L(\omega_m^k) = \frac{1}{|B_{m,k}|} \sum_{j=1}^{|B_{m,k}|} L(f(\omega_m^k, \mathbf{x}_j), y_j) \quad (9)$$

训练的目标是使本地模型损失函数最小化，即

$$\omega_m^k = \arg \min L(\omega_m^{k-1}) \quad (10)$$

然后经过  $\theta$  轮随机梯度下降过程，车辆  $v_m$  的本地模型更新为

$$\omega_m^\theta = \omega_m^{t-1} - \sum_{k=1}^{\theta} \eta_m \nabla g(\omega_m^{k-1}; B_{m,k}) \quad (11)$$

## 2.4 双因子更新的局部模型

与同步联邦学习不同，异步联邦学习不必等候所有的车辆都完成本地训练轮次后才对局部模型更新，而是一旦有车辆完成训练就立即将参数上传RSU。由于是异步更新，车辆可能会在任意时间下载当前时刻最新的局部模型，并利用自身数据进行训练。每个参与训练的车辆从下载局部模型到上传训练完成的本地模型过程中，都可能会有其他车辆的上传操作穿插其中。

### 2.4.1 陈旧因子

在异步联邦学习中，车辆在不同时刻从RSU下载的局部模型版本可能是不同的。一般来说，下载

的模型版本越新训练的模型越好,故要记录版本号加以区分。此外,车辆 $v_m$ 从下载模型到利用自身数据完成训练并将其上传至RSU的过程中,可能已经有其它多个车辆提交了本地模型并对局部模型进行了数次更新。车辆下载的局部模型版本以及训练完成时局部模型的版本差,在一定程度上反映了模型参数的陈旧情况。利用车辆下载的局部模型版本和版本差,可以将陈旧因子定义为

$$s_c = \frac{b1 \cdot \ln(\tau + b2)}{1 + \exp(t - \tau - b3)} \quad (12)$$

其中,  $b1, b2, b3$ 是调节参数,  $t$ 是当前模型的版本,  $\tau$ 是车辆提出训练请求时下载的用于本地训练的局部模型的版本,  $t - \tau$ 是模型的版本差。陈旧因子将作为一个重要的权重,调节不同车辆在时间尺度上传的本地模型对局部模型的影响。

#### 2.4.2 贡献因子

有的车辆可能在同一个RSU区域内频繁提交更新,由于数据样本变化小,后续训练出来的模型贡献也较小,有必要降低其上传的本地模型的权重。为此设置车辆本地模型贡献因子,贡献因子用于局部模型更新中,淡化了车辆在同一路段频繁上传本地模型造成局部模型无意义更新的影响,可表示为

$$s_g = \frac{1}{1 + \exp(h - c)} \quad (13)$$

其中,  $h$ 表示该车辆在同一RSU覆盖区域进行联邦学习训练的次数,  $c$ 为调节参数,  $\exp$ 是以自然常数 $e$ 为底的指数函数。故可以将车辆 $v_m$ 提交更新的综合权重表示为

$$\rho_m = (1 - \beta) \cdot s_c + \beta \cdot s_g \quad (14)$$

其中,  $\beta$ 是用于调节陈旧因子和贡献因子权重的参数。RSU只要接收到车辆上传的本地模型便对局部模型进行一次更新,则 $R_n$ 上局部模型的第 $t$ 次更新表示为

$$\omega_{R_n}^t = (1 - \rho_m) \cdot \omega_{R_n}^{t-1} + \rho_m \cdot \omega_m^\theta \quad (15)$$

其中,  $\omega_{R_n}^{t-1}$ 为当前 $R_n$ 上最新版本的局部模型,  $\omega_m^\theta$ 为车辆 $v_m$ 完成 $\theta$ 轮本地训练提交给 $R_n$ 的模型。

综上所述,RSU进行完 $\kappa$ 次本地更新后,局部模型的损失函数可以表示为

$$\begin{aligned} L(\omega_{R_n}^\kappa) &= \frac{1}{\kappa} \sum_{m=1}^{\kappa} \rho_m \cdot L(\omega_m^\theta) \\ &= \frac{1}{\kappa} \sum_{m=1}^{\kappa} \frac{\rho_m}{|D_m|} \sum_{j=1}^{|D_m|} L(f(\omega_m^\theta, \mathbf{x}_j), y_j) \end{aligned} \quad (16)$$

## 2.5 全局聚合

当所有RSU都对局部模型完成了 $\kappa$ 次更新后,基站就对各RSU执行全局聚合操作,得到的全局模型将用于下一轮的联邦学习。通常不同RSU经过 $\kappa$ 轮更新后的局部模型精度有所差异。精度高的局部模型对全局模型的贡献应该更大,相反精度低的局部模型则应在全局聚合时占据较小的比重以减小影响。为了评估局部模型的精度,路侧单元 $R_n$ 利用自身采集到的环境数据集合 $D_{R_n}$ 对更新所得的局部模型进行精度评估,评估得分 $\gamma_n$ 表示为

$$\gamma_n = \sum_{j=1}^{|D_{R_n}|} \frac{1}{|D_{R_n}|} \left( 1 - \left| \frac{f(\omega_{R_n}^\kappa, \mathbf{x}_j) - y_j}{y_j} \right| \right) \quad (17)$$

基于该精度评估得分,基站聚合各局部模型得到第 $t$ 轮全局模型 $\omega_G^t$

$$\omega_G^t = \frac{1}{N} \sum_{n=1}^N \gamma_n \omega_{R_n}^\kappa \quad (18)$$

联邦学习训练的目标是使全局模型损失函数最小化,全局模型损失函数为

$$L(\omega_G^t) = \frac{1}{N} \sum_{n=1}^N \gamma_n L(\omega_{R_n}^\kappa) \quad (19)$$

综上所述,本文提出的双因子更新的双层异步联邦学习完整流程为:在训练开始时,首先对基站上的全局模型进行初始化,然后各RSU从基站下载全局模型,作为初始的局部模型。对系统中的车辆进行选择,满足驻留时间大于训练时间的车辆才可以参加联邦学习。RSU计算车辆上传的本地模型的新旧因子与贡献因子,得到综合权重后对局部模型更新。当所有RSU的局部模型都完成了 $\kappa$ 次更新,基站就按照各局部模型评估的精度作为权重进行全局聚合,聚合完成后的模型作为基站新一轮的全局模型。

## 3 基于DDQN的联邦学习参数优化

### 3.1 联邦学习参数优化问题

在本地模型训练和局部模型更新中,增加迭代次数和更新次数通常可以提高模型精度,但也会增大系统能耗,因此需要找到联邦学习参数 $\theta$ 和 $\kappa$ 值的最优解,以保证较低的系统能耗和训练精度损失。将 $v_m$ 在一个单位数据上训练模型所需的CPU周期数记为 $\beta_m$ ,则进行1轮本地训练所需的CPU周期数为 $\beta_m |D_m|$ ;将车辆 $v_m$ 给模型训练分配的CPU频率记作 $f_m$ ,则进行1轮本地训练的能耗为

$$E_m^{\text{cmp}} = \frac{\alpha_m}{2} f_m^2 \beta_m |D_m| \quad (20)$$

其中,  $\alpha_m/2$ 表示车辆 $v_m$ 计算芯片组的有效电容系数。

车辆 $v_m$ 的传输功率记为 $p_m$ , 其数据传输速率为 $r_m$ , 则 $v_m$ 向RSU传递参数过程中的能耗为

$$E_m^{\text{com}} = \frac{|\omega_m^\theta| \cdot p_m}{r_m} \quad (21)$$

所以 $v_m$ 完成 $\theta$ 轮本地训练并上传到RSU的总能耗为

$$E_m^{\text{total}} = \theta \cdot E_m^{\text{cmp}} + E_m^{\text{com}} \quad (22)$$

RSU对局部模型完成了 $\kappa$ 次更新后, 便将第 $\kappa$ 次更新后的局部模型上传到基站参与全局聚合。则在一轮全局模型更新过程中, 系统的总能耗为

$$E^{\text{total}} = \sum_{n=1}^N \sum_{m=1}^{\kappa} (\theta \cdot E_m^{\text{cmp}} + E_m^{\text{com}}) \quad (23)$$

系统能耗与精度损失可以用式(24)联合表示

$$\varphi = \sigma \cdot E^{\text{total}} + (1-\sigma) \cdot L(\omega_G^t) \quad (24)$$

其中,  $\sigma$ 用以调节系统能耗与训练精度损失之间的权重。

问题的目标是要找到最优的本地迭代次数 $\theta$ 值和局部模型更新次数 $\kappa$ 值, 以达到系统能耗与训练精度的综合最优值。根据分析, 可以将其建模成如式(25)的优化问题P1

$$\left. \begin{array}{l} \text{P1: } Q(\theta, \kappa) = \arg \min \varphi \\ \text{s.t. C1: } T_{\text{stay}}(m) > T_{\text{sum}}(m) \\ \text{C2: } v_m \in V, R_n \in R \\ \text{C3: } \theta \cdot E_m^{\text{cmp}} + E_m^{\text{com}} \leq E_m^{\text{limit}} \\ \text{C4: } \sum_{m=1}^{\kappa} (\theta \cdot E_m^{\text{cmp}} + E_m^{\text{com}}) \leq E^{\text{R}} \\ \text{C5: } \sum_{n=1}^N \sum_{m=1}^{\kappa} (\theta \cdot E_m^{\text{cmp}} + E_m^{\text{com}}) \leq E^{\text{max}} \\ \text{C6: } \theta, \kappa, N \in \{1, 2, \dots\} \end{array} \right\} \quad (25)$$

上述约束条件中, C1表示车辆的驻留时间要大于完成本地训练并发送到RSU所需的预计时间; C2表示车辆 $v_m$ 与路侧单元 $R_n$ 的所属集合; C3表示车辆进行本地训练并上传到路侧单元的能耗, 其不能超过车辆自身可用资源的最大值 $E_m^{\text{limit}}$ ; C4表示RSU进行了 $\kappa$ 次更新时, 所消耗的能量不能超过RSU要求资源的最大值 $E^{\text{R}}$ ; C5表示完成1轮全局模型的更新, 所消耗的总能耗不能超过系统资源的最大值 $E^{\text{max}}$ ; C6表示本地训练的迭代次数、RSU局部模型更新次数及路侧单元个数都为正整数。

### 3.2 MDP模型

进一步将问题P转化为马尔可夫决策过程(Markov Decision Process, MDP), 即 $\{S, A, P, R\}$ , 这4个元

组分别表示状态空间、动作空间、状态转移概率和奖励函数。将 $t$ 时刻系统的状态定义为

$$s_t = (E_t, T_t, F(w^t)) \in S \quad (26)$$

其中,  $E_t$ 表示完成当前轮的总能耗,  $T_t$ 表示完成当前轮的总时延,  $F(w^t)$ 表示该轮全局模型的损失函数。

动作空间 $A$ 为训练参数的取值集合, 利用强化学习得出最优的本地迭代次数和RSU局部模型更新次数, 即在 $t$ 时刻的动作 $a_t = (\theta, \kappa) \in A$ ; 策略以概率分布 $\pi(a|s)$ 表示,  $\pi(a_t|s_t) \in [0, 1]$ 表示在状态 $s_t$ 下采取动作 $a_t = (\theta, \kappa)$ 的概率。

当动作被应用后, Agent会从环境中获得奖励 $r_t$ , 可以根据联邦学习训练精度的损失与能耗构建奖励函数来评估学习策略的好坏。具体的奖励函数为

$$r_t = - \left( \lambda \cdot E_t + (1-\lambda) \cdot \frac{F(w^t) - F(w^*)}{F(w^*)} \right) \quad (27)$$

其中,  $\lambda$ 用以调节系统能耗与训练精度损失之间的权重,  $F(w^*)$ 为目标损失值。

### 3.3 基于DDQN的参数优化方法

深度Q网络(Deep Q Network, DQN)算法是深度强化学习中处理离散动作空间的一种常用的方法, 但其容易产生过度估计问题。为解决这一问题, DDQN算法应运而生, 其整体结构与DQN算法大致相同, 只不过在目标Q值对应动作的选择和目标Q值计算时有所差异。

DDQN算法有两个结构相同的Q网络, 其中Target Q网络是每经过一定回合才进行更新, 更新的方式是将Q网络的参数复制过去。在DQN算法中, 其目标Q值进行计算的方法为

$$y_{t+1} = r_t + \gamma Q(s_{t+1}, a_{t+1}) \quad (28)$$

而DDQN中, 是先选择出最大Q值对应的动作, 然后将该动作输入Target Q网络, 找到对应的Target Q网络的输出, 即目标Q值。动作的选择与Target Q网络的输出为

$$a_{t+1} = \arg \max_a Q(s_{t+1}, a) \quad (29)$$

$$y_{t+1} = r_t + \gamma Q^{\text{Target}}(s_{t+1}, a_{t+1}) \quad (30)$$

然后, 以 $Q(s_t, a_t)$ 作为网络的预测值, 以 $y_{t+1}$ 作为实际值, 通过误差反向传播进行Q网络的更新。如果每次从经验回放池中随机取出 $h$ 个样本进行训练, 则DDQN中Q网络的训练损失函数为

$$\text{loss} = \frac{1}{h} \sum_{i=1}^h (y_{t+1} - Q^{\text{Target}}(s_{t+1}, a_{t+1}))^2 \quad (31)$$

图2为基于DDQN的联邦学习参数优化算法流程图, 算法1则是其伪代码。首先初始化联邦学习



状态, 得到足够的训练样本(1~7行)。当样本数达到 $P_1$ 后以 $P_2$ 为频率进行DDQN训练, 每次从经验回放池采样 $h$ 个样本将误差反向传播更新 $Q$ 网络参数(8~10行)。Target  $Q$ 网络的参数每隔 $H$ 步从 $Q$ 网络复制, 并且每次迭代随机探索率都会衰减(11~17行)。

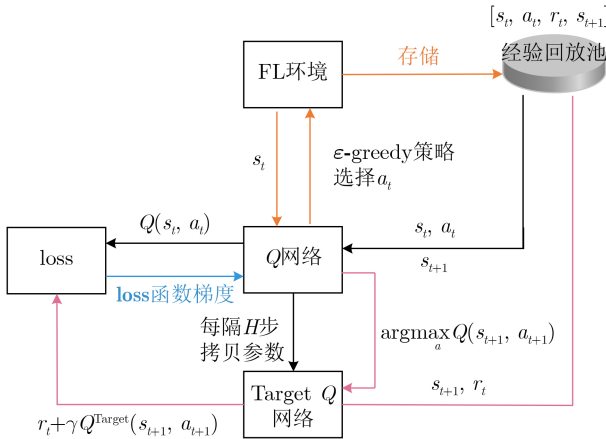


图2 DDQN算法流程图

#### 算法1 基于DDQN的联邦学习参数优化算法

输入: 迭代轮数 $T$ , 动作集 $A$ , 衰减因子 $\gamma$ , 探索率 $\varepsilon$ , 探索衰减decay, 当前 $Q$ 网络 $Q$ , Target  $Q$ 网络 $Q'$ , 批量梯度下降的样本数 $h$ , Target  $Q$ 网络参数更新频率 $H$ , 神经网络开始训练时系统当前轮次为 $P_1$ , 神经网络训练的频率为 $P_2$

输出: 最优动作 $a = (\theta, \kappa)$

- (1) 初始化当前 $Q$ 网络与Target  $Q$ 网络, 将经验回放池Memory清空。
- (2) **for** epoch=1 to  $T$  **do**
- (3) 初始化双层异步联邦学习的状态 $s$
- (4) 将 $s$ 输入 $Q$ 网络, 得到 $Q$ 网络所有动作对应的 $Q$ 值输出, 并利用 $\varepsilon$ -贪婪策略选出动作 $a = (\theta, \kappa)$
- (5) 根据动作更新联邦学习参数 $\theta$ 和 $\kappa$ 并进行训练, 得到新的状态 $s_{t+1}$ 和奖励 $r_t$
- (6) 将 $\{s_t, a_t, r_t, s_{t+1}\}$ 存入经验回放池
- (7) 令 $s_t = s_{t+1}$
- (8) **if** epoch >  $P_1$  && epoch %  $P_2 == 0$
- (9) 从经验回放池中采样 $h$ 个样本, 并计算Target  $Q$ 的值 $y_{t+1}$
- (10) 将误差 $\text{loss} = \frac{1}{h} \sum_{i=1}^h (y_{t+1} - Q^{\text{Target}}(s_{t+1}, a_{t+1}))^2$ 反向传播以更新 $Q$ 网络参数
- (11) **if** epoch %  $H == 0$
- (12) 将 $Q$ 网络参数复制给Target  $Q$ 网络
- (13) **end if**
- (14) **end if**
- (15) epoch = epoch + 1
- (16)  $\varepsilon = \varepsilon \cdot \text{decay}$
- (17) **end for**

## 4 仿真分析

本文采用Python3.8实现仿真环境, 依靠Pytorch 1.11.0深度学习框架, 并使用MNIST<sup>[15]</sup>数据集来模拟交通图像识别的任务场景。针对该图像识别问题, 在训练中使用卷积神经网络作为训练模型, 该模型包括两个卷积层、两个最大池化层和两个全连接层。

基于DDQN的联邦学习参数优化过程包含离线阶段和在线阶段。离线阶段, 利用云服务器先训练一个通用的DDQN模型, 获得初始策略网络; 在线阶段, 将离线阶段生成的策略网络部署到基站作为初始策略, 然后使用实时的数据在线确定每轮的两个训练参数。将预训练的模型部署到基站作为初始DDQN策略, 而不是从零开始训练, 避免了重复大批量计算, 加速收敛。另外, DDQN的训练周期可以根据环境变化指标动态调整, 稳定时扩大周期, 剧变时缩短周期, 这可以在很大程度上减少引入强化学习带来的额外计算开销与能耗。

图3是DDQN与DQN训练的奖励, 图4为训练损失, 训练的奖励是由联邦学习的精度与完成该轮联邦学习的能耗决定的。在进行强化学习时, 为了获得足够数量的样本, 需要不断初始化状态并依据 $\varepsilon$ -greedy策略选择动作, 本文设定初始的 $\varepsilon$ 为1, 每轮迭代都会乘以0.99的衰减率。在训练前期, 强化学习的动作以随机动作为主, 当进行了足够数量的

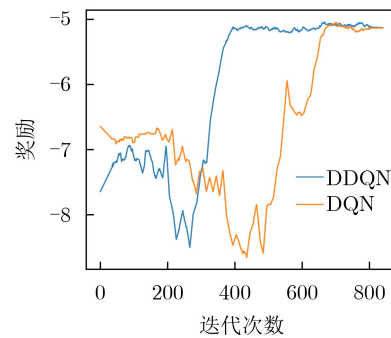


图3 强化学习的训练奖励

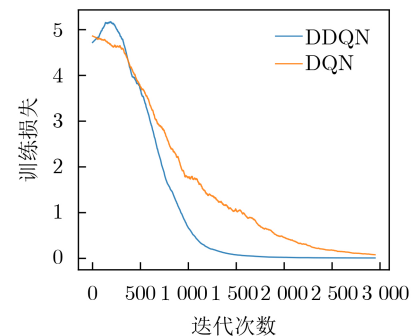


图4 强化学习的训练损失

探索之后，Agent倾向于选择最大 $Q$ 值对应的动作。图3中蓝色波形为DDQN奖励值的变化，橙色波形为DQN奖励值的变化，在训练前期两条曲线的值都较低且有较大波动，此时强化学习还没有进行足够的探索。随着训练轮次的增加，神经网络逐渐学习到好的策略此时奖励值迅速上升并趋于收敛，这表明强化学习已经通过大量的训练得到最优策略并根据该策略选择出最优动作—— $\theta$ 和 $\kappa$ 。通过比较我们可以看到，DDQN算法的收敛速度比DQN算法要快，波动也更小。根据上述基于DDQN的联邦学习参数的优化结果，得到最优的车辆本地训练次数为2，局部模型更新次数为10。为车辆随机分配的样本量区间为[600,1 200]，每个车辆随机设置更新延迟时间1~5 s。

图5对车辆总数 $M$ 分别为30、50和80、以及相应RSU数量 $N$ 为5和10的情况下对训练精度的影响进行了分析。由图5可知，随着车辆总数的增加，模型的精度也呈现上升趋势。这是因为随着车辆总数的增加，高质量节点用于训练的数量也相应增加。此外，增加路侧单元数量 $N$ 对精度有所改善，其影响相对更加显著。这是因为增加RSU数量会提高参与全局数据聚合的局部模型样本数量，从而更有利于模型训练以提高精度。

选择RSU数量 $N = 5$ ，当全局模型的迭代次数达到15轮时，每个RSU上的局部模型进行了150次更新。从图6可以看出，随着训练的进行这5个RSU局部模型的精度迅速上升，但由于不同RSU覆盖范围内参与训练的车辆资源异构，所以它们的精度曲线表现出一定的差异，但总体趋势相同。这5个RSU上的局部模型经过150轮更新后基本达到稳定，识别精度在90%左右。

为了验证本文改进算法的有效性，本文将同步联邦学习算法(Synchronous Federated Learning, SFL)、自适应异步联邦学习算法(Adaptive Asynchronous Federated Learning, AAFL)与本文所改进的双因子更新的异步联邦学习算法TTAFL比较。

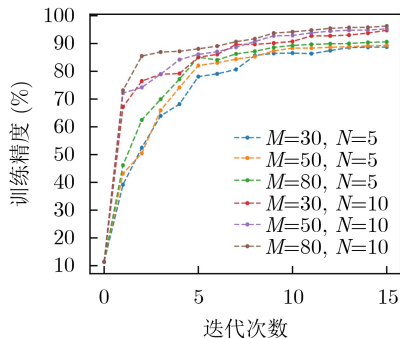


图5 不同车辆总数及路侧单元数对精度的影响

由图7可以看出，SFL算法在训练前期时，训练精度上升速度比异步算法慢，但收敛后的准确率比AAFL算法略高。而本文提出的TTAFL算法既能保持较高的收敛速度，又能达到更高的准确率，且从不同学习率下的训练精度可以看出， $\alpha = 0.01$ 时具有更好的表现。

此外，为了验证车辆选择策略的有效性，统计了SFL算法和TTAFL算法分别在随机车辆选择策略和本文提出的考虑驻留时间的车辆选择策略下的系统能耗与时延，如图8所示。本文考虑驻留时间的车辆选择方案可以在一定程度上减小系统能耗与时延，这是因为与随机车辆选择方案相比，其减少了车辆驶出当前RSU时，车辆在服务器间切换及模型

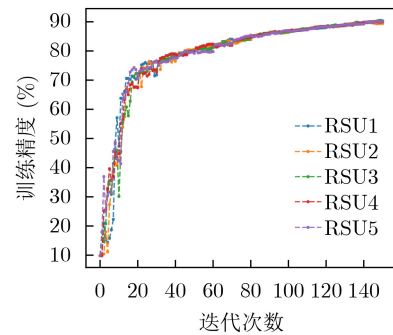


图6 各RSU局部模型精度

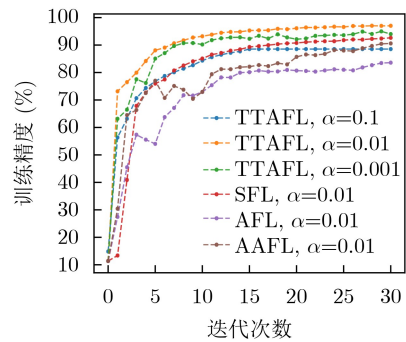


图7 训练精度对比

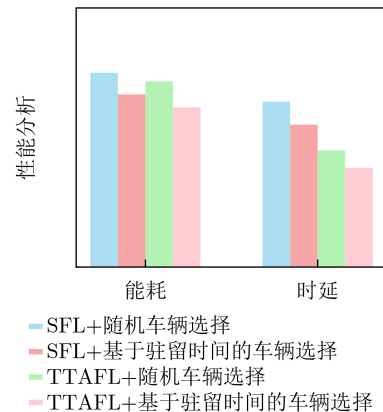


图8 基于不同选择策略的能耗与时延分析

传输产生的能耗与时延。并且在收敛到相同准确率时,TTAFL算法比SFL算法的训练时延减少了45.5%。

## 5 结论

本文设计并构建了一个车联网中的双层异步联邦学习框架,综合考虑版本差与训练次数提出了双因子更新的方法。考虑车辆在当前RSU的驻留时间对参与训练的车辆进行选择,避免了训练过程中服务器间切换带来的困扰。此外,本文使用DDQN算法优化联邦学习过程中的本地迭代次数和局部模型更新次数,在时延和资源约束下,有效提高了训练效率,减小了资源消耗与精度损失。仿真结果表明,所提双层异步算法比传统算法具有更快的收敛速度与精度,在车联网的异构环境下具有较好的性能和应用前景。

## 参考文献

- [1] ZHOU Xiaokang, LIANG Wei, SHE Jinhua, *et al.* Two-layer federated learning with heterogeneous model aggregation for 6G supported internet of vehicles[J]. *IEEE Transactions on Vehicular Technology*, 2021, 70(6): 5308–5317. doi: [10.1109/TVT.2021.3077893](https://doi.org/10.1109/TVT.2021.3077893).
- [2] XU Chenhao, QU Youyang, LUAN T H, *et al.* An efficient and reliable asynchronous federated learning scheme for smart public transportation[J]. *IEEE Transactions on Vehicular Technology*, 2023, 72(5): 6584–6598. doi: [10.1109/TVT.2022.3232603](https://doi.org/10.1109/TVT.2022.3232603).
- [3] MCMAHAN B, MOORE E, RAMAGE D, *et al.* Communication-efficient learning of deep networks from decentralized data[C]. The 20th International Conference on Artificial Intelligence and Statistics, Fort Lauderdale, USA, 2017: 1273–1282.
- [4] ABDULRAHMAN S, TOUT H, OULD-SLIMANE H, *et al.* A survey on federated learning: The journey from centralized to distributed on-site learning and beyond[J]. *IEEE Internet of Things Journal*, 2021, 8(7): 5476–5497. doi: [10.1109/JIOT.2020.3030072](https://doi.org/10.1109/JIOT.2020.3030072).
- [5] ZHU Hangyu, ZHANG Haoyu, and JIN Yaochu. From federated learning to federated neural architecture search: A survey[J]. *Complex & Intelligent Systems*, 2021, 7(2): 639–657. doi: [10.1007/s40747-020-00247-z](https://doi.org/10.1007/s40747-020-00247-z).
- [6] SUN Feng, ZHANG Zhenjiang, ZEADALLY S, *et al.* Edge computing-enabled internet of vehicles: Towards federated learning empowered scheduling[J]. *IEEE Transactions on Vehicular Technology*, 2022, 71(9): 10088–10103. doi: [10.1109/TVT.2022.3182782](https://doi.org/10.1109/TVT.2022.3182782).
- [7] LIM W Y B, LUONG N C, HOANG D T, *et al.* Federated learning in mobile edge networks: A comprehensive survey[J]. *IEEE Communications Surveys & Tutorials*, 2020, 22(3): 2031–2063. doi: [10.1109/COMST.2020.2986024](https://doi.org/10.1109/COMST.2020.2986024).
- [8] YANG Zhigang, ZHANG Xuhua, WU Dapeng, *et al.* Efficient asynchronous federated learning research in the internet of vehicles[J]. *IEEE Internet of Things Journal*, 2023, 10(9): 7737–7748. doi: [10.1109/JIOT.2022.3230412](https://doi.org/10.1109/JIOT.2022.3230412).
- [9] CHAI Haoye, LENG Supeng, CHEN Yijin, *et al.* A hierarchical blockchain-enabled federated learning algorithm for knowledge sharing in internet of vehicles[J]. *IEEE Transactions on Intelligent Transportation Systems*, 2021, 22(7): 3975–3986. doi: [10.1109/TITS.2020.3002712](https://doi.org/10.1109/TITS.2020.3002712).
- [10] SAPUTRA Y M, HOANG D T, NGUYEN D N, *et al.* Dynamic federated learning-based economic framework for internet-of-vehicles[J]. *IEEE Transactions on Mobile Computing*, 2023, 22(4): 2100–2115. doi: [10.1109/TMC.2021.3122436](https://doi.org/10.1109/TMC.2021.3122436).
- [11] WANG Yuwei and KANTARCI B. A novel reputation-aware client selection scheme for federated learning within mobile environments[C]. 2020 IEEE 25th International Workshop on Computer Aided Modeling and Design of Communication Links and Networks, Pisa, Italy, 2020: 1–6. doi: [10.1109/CAMAD50429.2020.9209263](https://doi.org/10.1109/CAMAD50429.2020.9209263).
- [12] XIAO Huizi, ZHAO Jun, PEI Qingqi, *et al.* Vehicle selection and resource optimization for federated learning in vehicular edge computing[J]. *IEEE Transactions on Intelligent Transportation Systems*, 2022, 23(8): 11073–11087. doi: [10.1109/TITS.2021.3099597](https://doi.org/10.1109/TITS.2021.3099597).
- [13] 贺文晨, 郭少勇, 邱雪松, 等. 基于DRL的联邦学习节点选择方法[J]. *通信学报*, 2021, 42(6): 62–71. doi: [10.11959/j.issn.1000-436x.2021111](https://doi.org/10.11959/j.issn.1000-436x.2021111).  
HE Wenchen, GUO Shaoyong, QIU Xuesong, *et al.* Node selection method in federated learning based on deep reinforcement learning[J]. *Journal on Communications*, 2021, 42(6): 62–71. doi: [10.11959/j.issn.1000-436x.2021111](https://doi.org/10.11959/j.issn.1000-436x.2021111).
- [14] YANG Peng, YAN Mengjiao, CUI Yaping, *et al.* FedDD: Federated double distillation in IoV[C]. 2022 IEEE 96th Vehicular Technology Conference, London, United Kingdom, 2022: 1–5, doi: [10.1109/VTC2022-Fall57202.2022.10012798](https://doi.org/10.1109/VTC2022-Fall57202.2022.10012798).
- [15] LECUN Y, BOTTOU L, BENGIO Y, *et al.* Gradient-based learning applied to document recognition[J]. *Proceedings of the IEEE*, 1998, 86(11): 2278–2324. doi: [10.1109/5.726791](https://doi.org/10.1109/5.726791).

王力立: 女, 副教授, 研究方向为深度强化学习及无线传感网络。  
吴守林: 男, 硕士生, 研究方向为车联网中的联邦学习、任务卸载。  
杨妮: 女, 硕士生, 研究方向为联邦学习、知识交易。  
黄成: 男, 讲师, 研究方向为机器学习及检测技术。

责任编辑: 余蓉