

# 不可靠通信下的联邦抗干扰模型优化方案

李中捷<sup>abc</sup>, 郭海榕<sup>a\*</sup>, 邱凡<sup>ab</sup>

(中南民族大学 a. 电子信息工程学院; b. 智能无线通信湖北省重点实验室; c. 智能物联技术湖北省工程研究中心,  
武汉 430074)

**摘要** 针对蜂窝车联网(Cellular-Vehicle to Everything, C-V2X)通信场景下无线信道受到干扰导致通信过程中可能存在信息丢失的情况,通过对联邦分布式随机梯度下降(Federated Learning-Distributed Stochastic Gradient Descent, FL-DSGD)进行抗干扰模型更新机制的优化以减少上述通信链路不可靠情况的影响.该方案首先建立车辆与基站的通信链路及传输模型参数;然后在通信链路不可靠导致模型参数在传输过程中部分缺失的情况下,根据链路可靠性混合权重矩阵,利用车辆上存储的本地模型以及基站存储的全局模型参与当前轮次联邦学习的模型更新,以填充丢失的模型参数.仿真结果表明:在通信链路不可靠的情况下,FL-DSGD方案达到90%的训练准确率以及85%的测试准确率所需的通信轮次约为分布式基线方案所需通信轮次的50%.

**关键词** 联邦学习;车联网;随机梯度下降

中图分类号 TN929.5 文献标志码 A 文章编号 1672-4321(2025)06-0826-07

doi:10.20056/j.cnki.ZNMDZK.20250611

## Optimization scheme of federated anti-interference model with unreliable communication

LI Zhongjie<sup>abc</sup>, GUO Hairong<sup>a\*</sup>, QIU Fan<sup>ab</sup>

(South-Central Minzu University, a. College of Electronic and Information Engineering; b. Hubei Key Laboratory of Intelligent Wireless Communications; c. Hubei Engineering Research Center of Intelligent Internet of Things Technology, Wuhan 430074, China)

**Abstract** Aiming at the possible information loss in the communication process caused by interference to the wireless channel in the Cellular Vehicle to Everything (C-V2X) communication scenario, the impact of the above unreliable communication link is reduced by optimizing the anti-interference model update mechanism of the Federated Learning Distributed Stochastic Gradient Descent (FL-DSGD). Firstly, the communication link between the vehicle and the base station and the transmission model parameters are established; Then, when the communication link is unreliable, leading to partial loss of model parameters in the transmission process, according to the link reliability mixed weight matrix, the local model stored on the vehicle and the global model stored in the base station are used to participate in the model update of the current round of federated learning to fill in the missing model parameters. Simulation results show that when the communication link is unreliable, the communication rounds required for FL-DSGD scheme to achieve 90% training accuracy and 85% test accuracy are about 50% of the communication rounds required for the distributed baseline scheme.

**Keywords** federated learning; vehicle to everything; stochastic gradient descent

随着道路上联网车辆的不断增多,车载通信已成为一个重要的研究领域<sup>[1]</sup>.为了满足车联网(Vehicle to Everything, V2X)通信的严格要求,C-V2X需要共享频谱资源,从而实现车对基础设施

收稿日期 2024-04-18 \* 通信作者 郭海榕,研究方向:无线通信,E-mail:892781753@qq.com

作者简介 李中捷(1974-),男,教授,博士,研究方向:无线通信,E-mail:lizhongjie@mail.scuec.edu.cn

基金项目 国家自然科学基金资助项目(61379028);湖北省自然科学基金资助项目(2022CFB905);中央高校基本科研业务费专项资金资助(CZY23027)

(Vehicle to Infrastructure, V2I)和车对车(Vehicle to Vehicle, V2V)的同步通信.因此,如何在有限的频谱资源内减少干扰并同时保证V2I链路的信道容量和V2V通信的可靠性,成为V2X通信亟待解决的一个重要问题.

传统的优化方法已经被用来应对V2X通信中的干扰问题<sup>[2-5]</sup>.在文献[2]中,提出了一种无线电资源管理(Radio Resource Management, RRM)算法,保证了基于端到端(Device to Device, D2D)的V2X系统的延迟和可靠性要求.基于类似的V2X框架,文献[3]提出了一个优化问题,在只考虑缓慢变化的大规模衰减信道的情况下为D2D车辆系统设计一个高效的频谱和功率分配算.此外,在文献[4-5]中研究了排队延迟对吞吐量和可靠性的影响.

由于车辆的高速移动性导致快速变化的信道以及多样化的服务需求,传统的优化算法难以利用数学方法进行精确建模并且会产生大量的计算开销.因此,机器学习被认为是一种有前途的技术,可以解决传统优化方法中的挑战.

联邦学习作为一种分布式机器学习范式,可以使多个车辆共同训练一个全局模型,而不需要在车辆间或与服务器交换数据<sup>[6-10]</sup>.文献[6]通过聚合所有智能体的本地模型,为所有车辆学习一个单一的全局模型来降低通信开销.文献[7]通过在本地图标中引入额外的近端项,以限制本地模型的数量从而提高了联邦学习的稳定性,文献[8-10]为本地模型学习额外的个性化模型来提高本地模型的效率和性能.

然而,上述文献都假设通信链路处于完美的情况,而在实际通信系统中,由于恶劣的无线信道会引入噪声、衰落和干扰,导致丢包和传输错误,以至于模型无法达到最优效果.

本文在C-V2X通信场景下,针对模型传输过程中数据缺失的情况,提出了一种联邦抗干扰模型更新方案,以减少通信链路不可靠的影响.该方案根据链路可靠性混合权重矩阵,对车辆和基站接收到的模型中缺失的部分,利用车辆和基站存储的模型填充缺失的模型,以进行更新模型.最后仿真实验评估了本算法的收敛性,分析不同程度的链路可靠性对系统性能的影响.

## 1 FL-DSGD通信链路优化方案设计

本文考虑到车辆与基站在行驶过程建立V2I链路进行模型参数传输的过程中,双方只接收到部分模型参数消息的情况下进行共识更新,并根据链路可靠性矩阵 $P^r$ 优化混合权重矩阵 $W$ .

### 1.1 不可靠通信链路问题描述

在FL-DSGD方案中,车辆 $n$ 维护并训练自己的本地参数 $w_n \in \mathbb{R}^d$ , $d$ 为训练参数的维度.车辆根据从本地数据集 $\mathcal{D}_n$ 采样的样本 $\xi_n$ 计算本地梯度 $[g_{gd}]_n = \nabla F_n(w_n, \xi_n)$ ,式中, $F_n(w_n, \xi_n)$ 为车辆的本地损失函数,之后通过V2I链路与基站服务器 $J$ 传输模型参数.每次训练迭代的两个步骤分别为SGD更新和联邦共识更新:

$$w_n^{(t+\frac{1}{2})} = w_n^{(t)} - \mu [g_{gd}]_n^{(t)}, \quad (1)$$

$$w_J^{(t+1)} = \sum_{i=1}^N w_{n,J} w_J^{(t+\frac{1}{2})}, \quad (2)$$

式中: $w_n^{(t)}$ 表示车辆 $n$ 在 $t$ 时刻的参数, $\mu$ 为学习率, $w_{n,J}$ 是混合权重矩阵的第 $(n, J)$ 项, $W \in \mathbb{R}^{N \times N}$ ,如果第 $n$ 个车辆向基站服务器 $J$ 传输模型参数失败,则 $w_{n,J} = 0$ .在车辆通过本地训练进行SGD更新时,训练的收敛速度很大程度上取决于混合矩阵 $W$ ,为了保证收敛性,将 $W$ 设置为对称双随机,即 $W^T = W$ , $W\mathbf{1} = \mathbf{1}$ ,其中 $\mathbf{1}$ 表示一个全为1的 $N$ 维向量.在实际的通信系统传输过程中模型参数暴露在不可靠的通信条件下导致容易出现丢包和传输错误.且由于没有数据重传这样的可靠性增强机制,因此无法保证交付.本文令矩阵 $p^r = [p_{n,J}^r] \in \mathbb{R}^{N \times N}$ 表示通信网络中的链路可靠性程度,式中 $p_{n,J}^r$ 表示从第 $n$ 个车辆成功传输模型至基站服务器的概率.为了模拟随机通信网络,在每次车辆和基站服务器通信前产生一个随机掩码 $m_{n,J}$ ,以确定基站获取车辆数据的部分,第 $n$ 个车辆向基站服务器传输模型参数 $w_n$ 的数据包被随机接收,出现在基站服务器上接收到的数据 $z_n \rightarrow J$ 可能只包括 $w_n$ 的一部分.这一过程可表示为:

$$z_n \rightarrow J = m_{n,J} \odot w_n, \quad (3)$$

式中: $\odot$ 表示逐元素相乘法,如果基站服务器 $J$ 成功接收到第 $n$ 个车辆传输的第 $h$ 个参数,则 $m_{n,J}(h) = 1$ ,否则 $m_{n,J}(h) = 0$ .

### 1.2 FL-DSGD算法设计

联邦学习通过车辆一定的通信能力达成了协作学习的效果,每个车辆用自己的数据来训练和优化模型参数,并通过基站服务器实现模型信息的交

互以达成共识<sup>[11-12]</sup>.

在实际的通信场景下,车辆上传  $w_n^{(t+\frac{1}{2})}$  到基站服务器,基站服务器接收到车辆发送的信息表示为:

$$z_{n \rightarrow j}^{(t+\frac{1}{2})} = m_{n \rightarrow j}^{(t)} \odot w_n^{(t+\frac{1}{2})}. \quad (4)$$

本文设计用车辆的本地参数以及基站服务器的全局模型参数填补对应传输过程中缺失的数据.即在车辆进行本地梯度下降过程中,第  $n$  个车辆使用  $w_n$  中的参数来填补  $z_{j \rightarrow n}$  中缺失的参数,填补后的信息  $\hat{z}_{j \rightarrow n}$  可以表示为:

$$\hat{z}_{j \rightarrow n}^{(t+\frac{1}{2})} = z_{j \rightarrow n}^{(t+\frac{1}{2})} + (1 - m_{j \rightarrow n}^{(t)}) \odot w_n^{(t+\frac{1}{2})}. \quad (5)$$

基站服务器进行联邦共识更新可以表示为:

$$\begin{aligned} w_j^{(t+1)} &= w_j w_j^{(t+\frac{1}{2})} + \sum_{n=1}^N w_{n,j} \hat{z}_{n \rightarrow j}^{(t+\frac{1}{2})} \\ &= w_j^{(t+\frac{1}{2})} + \sum_{n=1}^N w_{n,j} m_{n \rightarrow j}^{(t)} \odot (w_n^{(t+\frac{1}{2})} - w_j^{(t+\frac{1}{2})}) \end{aligned} \quad (6)$$

由于通信网络的随机性,共识更新步骤具有随机性.本文设置了两个重要的矩阵  $\bar{W}$  和  $\bar{W}^2$ , 分别表示更新参数的一阶矩阵和二阶矩阵.在期望中,  $\bar{W}$  是一个双随机矩阵,不可靠通信网络下的共识更新等价于以  $\bar{W}$  为权重矩阵的可靠共识更新.平均混合权重  $\bar{W}$  不仅取决于  $W$ , 还取决于链路可靠性矩阵  $P$ . 其中矩阵  $p^r = [p_{n,j}^r] \in \mathbb{R}^{N \times N}$  表示通信网络中的链路可靠性程度,  $p_{n,j}^r$  表示从第  $n$  个车辆成功传输模型至基站服务器的概率.

根据矩阵  $P^r$  的可靠性,本文采用两种方法来选择混合矩阵  $W$ : (1) 若链路可靠性矩阵  $P^r$  未知,则对每条链路采用平均混合权重,即  $W = W^T/N$ ; (2) 如果链路可靠性矩阵  $P^r$  已知,则可以通过优化  $W$  以加快收敛速度.FL-DSGD 方案的收敛性取决于矩阵  $\bar{W}^2 - W^T/N$  的最大特征值.因此,本文的优化目标为最小化  $W$  矩阵的最大特征值:

$$\begin{aligned} \min_W \quad & \mu_{\max}(\bar{W}^2 - W^T/N) \\ \text{s.t.} \quad & 0 \leq w_{n,j} \leq 1, W^T = W, W\mathbf{1} = \mathbf{1} \end{aligned} \quad (7)$$

在 FL-DSGD 方案下各车辆与基站服务器进行联邦学习交互信息过程如图 1 所示,车辆的模型参数交换包括以下步骤:

(1) 车辆以及基站服务器初始化模型.

(2) 深度神经网络的参数通常无法在一个数据包中传输,本方案通过将参数随机分组为多个数据包并独立传输.

(3) 模型通过 V2I 链路传输到基站服务器进行联邦共识更新.由于通信链路的不可靠,传输信息

在接收时可能被随机丢弃或污染.在这种情况下,部分模型数据被宣布丢失并被丢弃,否则数据包将被成功接收.对于缺失部分,基站服务器利用存储的全局模型以及混合权重矩阵进行联邦共识更新.

(4) 全局模型反馈至本地车辆并进行梯度下降.若全局模型在传输过程中缺失,车辆利用存储的本地模型以及混合权重矩阵进行模型更新.

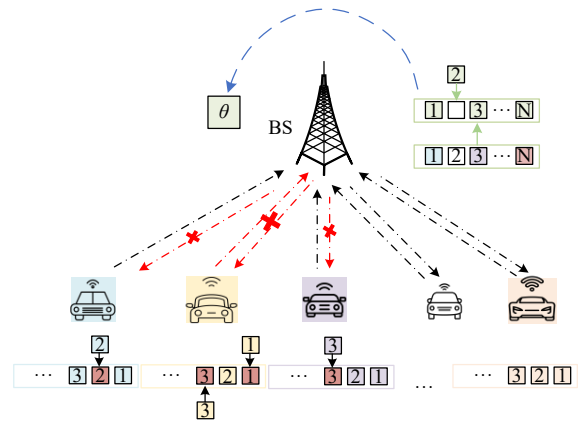


图 1 FL-DSGD 通信流程

Fig. 1 FL-DSGD communication process

FL-DSGD 优化通信链路方案的流程如表 1 所示:

表 1 基于 FL-DSGD 的优化通信链路方案

Tab. 1 Optimization of communication link scheme based on FL-DSGD

步骤	FL-DSGD 优化方案流程
初始化:	初始化本地模型 $\{w_n\}$ , 学习率 $\mu$ , 混合权重矩阵 $W$
1:	for episode $t = 0, 1, \dots, E$ :
2:	for all vehicles $n$ :
3:	从本地数据集 $\mathcal{D}_n$ 中采样一个小批量 $\xi_n$
4:	用 $g_n^t = \nabla F_n(w_n^{(t)}, \xi_n)$ 计算本地梯度.
5:	本地 SGD 更新: $w_n^{(t+\frac{1}{2})} = w_n^{(t)} - \mu [g_{gd}^t]_n^{(t)}$
6:	上传 $w_n^{(t+\frac{1}{2})}$ 到基站服务器 基站服务器从车辆接收消息:
7:	$z_{n \rightarrow j}^{(t+\frac{1}{2})} = m_{n \rightarrow j}^{(t)} \odot w_n^{(t+\frac{1}{2})}$ 用陈旧数据替换缺失值:
8:	$\hat{z}_{n \rightarrow j}^{(t+\frac{1}{2})} = z_{n \rightarrow j}^{(t+\frac{1}{2})} + (1 - m_{n \rightarrow j}^{(t)}) \odot w_j^{(t+\frac{1}{2})}$ 共识更新:
9:	$w_j^{(t+1)} = w_j w_j^{(t+\frac{1}{2})} + \sum_{n=1}^N w_{n,j} \hat{z}_{n \rightarrow j}^{(t+\frac{1}{2})}$ 基站服务器反馈全局模型:
10:	$\hat{z}_{j \rightarrow n}^{(t+\frac{1}{2})} = z_{j \rightarrow n}^{(t+\frac{1}{2})} + (1 - m_{j \rightarrow n}^{(t)}) \odot w_n^{(t+\frac{1}{2})}$

## 2 收敛性分析

假设每个车辆本地损失函数 $f_n(\mathbf{w})$ 是光滑的且具有L-Lipschitzian梯度,即存在常数 $L > 0$ ,使得

$$\|\nabla f_n(\mathbf{w}) - \nabla f_n(\mathbf{w}')\| \leq L\|\mathbf{w} - \mathbf{w}'\|, \quad (8)$$

式中,对于 $\forall \mathbf{w}, \mathbf{w}' \in \mathbb{R}^d, \|\cdot\|$ 为欧式范数.

假设存在常数 $\sigma > 0, \zeta > 0$ ,且 $\sigma^2$ 和 $\zeta^2$ 分别限制每个车辆上SGD的方差以及数据分布的差异:

$$\mathbb{E}\{\|\nabla F_n(\mathbf{w}, \xi) - \nabla f_n(\mathbf{w})\|\} \leq \sigma^2, \quad (9)$$

$$\frac{1}{N} \sum_{n=1}^N \|\nabla f_n(\mathbf{w}) - \nabla f(\mathbf{w})\| \leq \zeta^2, \quad (10)$$

每个车辆获得的随机梯度是对本地目标的真实梯度的无偏估计<sup>[13]</sup>:

$$\mathbb{E}\{[\mathbf{g}_{gd}]_n\} = \nabla f_n(\mathbf{w}_n). \quad (11)$$

为了分析FL-DSGD方案的收敛性,本文首先研究了不可靠通信网络下的共识算法.假设设备初始化为 $\{\mathbf{w}_n^{(0)} \in \mathbb{R}^d, n = 1, \dots, N\}$ ,每次迭代中只执行共识的更新步骤,即:

$$\mathbf{w}_j^{(t+1)} = \mathbf{w}_j^{(t)} + \sum_{n=1}^N w_{n,j} m_{n \rightarrow j}^{(t)} \odot (\mathbf{w}_j^{(t)} - \mathbf{w}_n^{(t)}). \quad (12)$$

按照共识更新规则,每一次梯度下降的期望受 $\mathbf{w}_i$ 方差的限制.其中 $\kappa$ 为建立通信链路的确定性.如果通信链路是确定的,即 $p_{n,j}^r$ 等于0或1, $\kappa$ 值则为0.

$$\mathbb{E}\{\|\bar{\mathbf{w}}_{t+1} - \bar{\mathbf{w}}_t\|^2\} \leq \frac{\kappa}{N^2} \sum_{n=1}^N \|\mathbf{w}_n^{(t)} - \bar{\mathbf{w}}_t\|^2. \quad (13)$$

$$\kappa = 2 \max_n \sum_{n=1}^N w_{n,j}^2 p_{n,j}^r (1 - p_{n,j}^r)$$

FL-DSGD方案的收敛程度取决于共识更新的收敛速度.不可靠通信链路的共识收敛率和期望分别为:

$$\frac{1}{N} \sum_{n=1}^N \|\mathbf{w}_n^{(t)} - \bar{\mathbf{w}}_t\|^2$$

$$\mathbb{E}\left\{\frac{1}{N} \sum_{n=1}^N \|\mathbf{w}_n^{(t)} - \bar{\mathbf{w}}_t\|^2\right\} \leq \varpi^t \sum_{n=1}^N \|\mathbf{w}_n^{(0)} - \bar{\mathbf{w}}_0\|^2, \quad (14)$$

式中: $\varpi = \max(\bar{\mathbf{W}}^2 - \mathbf{W}^T/N)$ ,如果通信网络是可靠的,共识更新规则具有指数级的收敛速度<sup>[14-15]</sup>.

FL-DSGD方案中假设所有本地模型都用 $\mathbf{w}_n^{(0)} \in \mathbb{R}^d$ 进行初始化.其中学习率满足 $\mu L \leq \min\{1, \sqrt{\varpi^{-1}} - n\}$ ,则在 $T$ 次迭代之后:

$$\frac{1}{T} \sum_{i=1}^T \|\nabla f(\bar{\mathbf{w}}_i)\|^2 \leq \left( \frac{\mathbb{E}[f(\bar{\mathbf{w}}_i)] - \mathbb{E}[f(\bar{\mathbf{w}}_i)]}{\mu T} \right.$$

$$\left. + \frac{\mu L}{N} \sigma^2 + \frac{2\gamma L \kappa}{N} \sigma^2 + \frac{6L\kappa\mu\zeta^2}{N} \right) \frac{1-D}{1-2D}$$

$$+ \left( L^2 + \frac{2L\kappa}{\mu N} + \frac{2(3N+1)L^3\mu\kappa}{N} \right), \quad (15)$$

$$\left( \frac{2\mu^2\sigma^2\varpi}{1-\varpi} + \frac{6\mu^2\zeta^2\varpi}{(1-\sqrt{\varpi})^2} \right) \frac{1}{1-2D}$$

式中: $D = \frac{6\mu^2 L^2 \varpi}{(1-\sqrt{\varpi})^2}$ , $\varpi$ 为式(4)的最大特征值.收敛

边界在很大程度上取决于 $\varpi$ ,如果学习率 $\mu = \sqrt{\frac{N}{T}}$ ,在总共 $T$ 次迭代之后:

$$\frac{1}{T} \sum_{i=1}^T \|\nabla f(\bar{\mathbf{w}}_i)\|^2 = \mathcal{O}\left(\frac{1}{\sqrt{NT}}\right) + \mathcal{O}\left(\frac{N}{T}\right), \quad (16)$$

式中: $\mathcal{O}$ 为常数汇总,因此具有不可靠通信链路的FL-DSGD方案可以实现收敛.

## 3 仿真结果及分析

本文设置所有车辆在仿真场景内初始化距离相同,在60次训练迭代后学习率衰减10倍.为了验证FL-DSGD方案的效率,本文在模拟实验中采用了两种对比算法:(1)完美通信链路下的联邦学习方案(perfect-FL):假设所有车辆都与基站服务器建立连接并且通信信道处于可靠状态下,不会出现信息丢失情况,车辆在训练本地模型后将本地模型上传至基站,并由基站服务器聚合后反馈给每个车辆.(2)传统分布式学习方案(DSGD):每个车辆统一分配数据独自训练模型.

信道系统在Pytorch中实现,数据集为CIFAR-10,其中包含50000个训练样本和10000个测试样本.网络模型为ResNet-20,激活函数为Tanh,表2给出模型训练的超参数.

图2给出了DSGD方案以及FL-DSGD方案在通

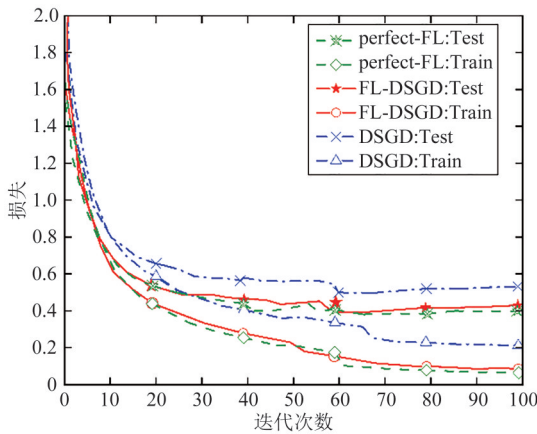
表2 训练超参数

Tab. 2 Training hyperparameters

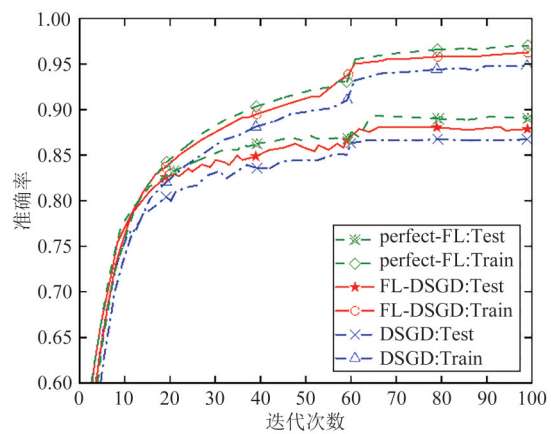
参数	值
迭代次数	100
学习率	0.1
Batch-size	32
权重衰减	0.0001
车辆数	4, 8, 12, 16
可靠性程度	0.3, 0.5, 0.7, 0.9

信链路可靠性程度  $k=0.7$  的情况下的收敛性对比. DSGD 方案与 FL-DSGD 方案相比, 准确率收敛所需的通信轮次更多. 从图中可以看出, DSGD 方案达到 90% 训练准确率以及 85% 测试准确率所需的通信轮次约为 FL-DSGD 方案所需通信轮次的两倍, 这是因为在保证所有训练数据完全被车辆接收的情况下, 需要另外花费更多通信轮次来维持稳定的通信链路. 同时, FL-DSGD 方案的训练准确率高于测试准确率, 这是由于车辆模型需要在不同的数据分布上进行训练, 而联邦学习的全局模型泛化能力不够强, 使车辆无法很好地适应测试数据中的新分布, 导致训练准确率高于测试准确率. 上述实验结果表明 FL-DSGD 方案在准确率以及模型传输需要的通信开销有较好的平衡性, 更符合车联网场景下建立 V2V 通信链路的稳定性与即时性.

图 3 比较了 FL-DSGD、perfect-FL 以及 DSGD 三种方案在通信链路可靠性程度为 0.7 的迭代训练和测试损失以及迭代训练和测试准确率. 实验结果显示, 在损失方面, perfect-FL 方案的训练损失和测试损失分别为 0.07 和 0.4, 而 FL-DSGD 方案的训练损失和测试损失分别为 0.1 和 0.44. FL-DSGD 与 perfect-FL 方案的训练和测试损失相近, 并且比 DSGD 方案



(a) 三种方案的训练及测试损失对比



(b) 三种方案的训练及测试准确率对比

图 3 三种方案的性能对比

Fig. 3 Performance comparison of three schemes

图 4 表示在通信链路可靠性程度不同的情况下, FL-DSGD 方案的训练准确率对比. 本文分别使用具有链路可靠性不同的矩阵权重  $k=[0.3, 0.5, 0.7, 0.9]$  来评估 FL-DSGD 方案在训练准确率上的表现. 从实验结果可以看出, 随着通信链路可靠性权重降低, FL-DSGD 方案的训练准确率收敛过程都会变慢. 此外, 当  $k$  为 0.7 和 0.9 时, 通信链路的可靠性较高, FL-DSGD 的训练准确率差距只有 0.01. 但当  $k$  为 0.3 和 0.5

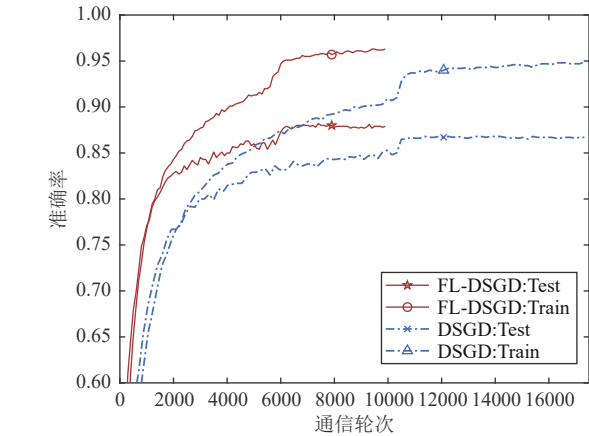


图 2 perfect-FL 与 FL-DSGD 方案通信轮次对比

Fig. 2 Comparison of communication rounds between FL-DSGD and DSGD

降低了约 0.1. 在准确率方面, perfect-FL 方案处于完美通信链路下传输模型, 因此可以获取较优的效果, 其次为 FL-DSGD 方案. 在准确率方面, perfect-FL 方案的收敛速度最快, 且可以达到 97% 的准确率, 而 FL-DSGD 方案的训练准确率比 DSGD 方案提高了 2%. 这是因为 FL-DSGD 方案利用来自所有不可靠通信链路的信息并使用部分接收到的消息进行更新, 而 DSGD 方案仅根据车辆本地数据训练网络模型参数并进行更新.

时, 链路可靠性较低, 可靠性权重之间的性能差距会增加. 这是因为当存在较多传输成功概率很小的链路时, 均匀权重仍然对这些链路分配相同的权重, 从而阻碍了收敛过程, 导致整体的收敛速度降低.

图 5 表示在不同车辆数量和不同链路可靠性程度下, perfect-FL 方案与 FL-DSGD 方案训练准确率的比值. 本文分别使用具有不同车辆数量  $[4, 8, 12, 16]$  和不同链路可靠性矩阵的权重  $k=[0.3, 0.5, 0.7, 0.9]$  来

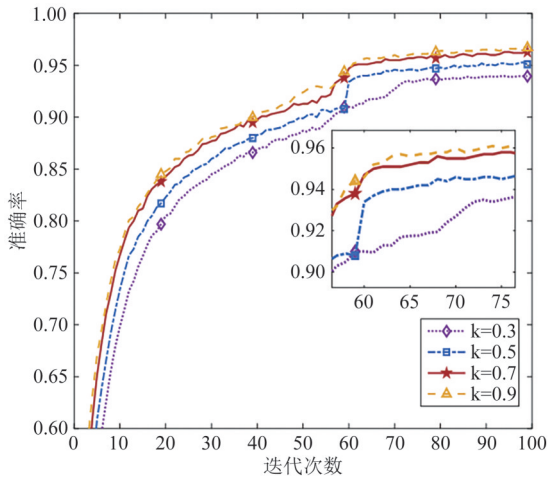


图4 链路可靠性程度不同情况下 FL-DSGD 方案训练准确率对比  
Fig. 4 Comparison of FL-DSGD training accuracy under different link reliability

评估 FL-DSGD 方案在训练准确率的表现. 从实验结果可以看出, 随着链路可靠性程度的降低, perfect-FL 方案与 FL-DSGD 方案训练准确率的比值随之下降. 具体来看, 当车辆数为 4 时,  $k=0.9$  和  $k=0.3$  的准确率比值相差约为 0.02, 但当车辆数为 16 时,  $k=0.9$  和  $k=0.3$  的准确率比值相差约为 0.05. 这是因为随着车辆数量的增加, 模型参数传输过程中的计算和通信开销会增加, 导致训练过程变得更加耗时和复杂, 从而影响了车辆的准确率.

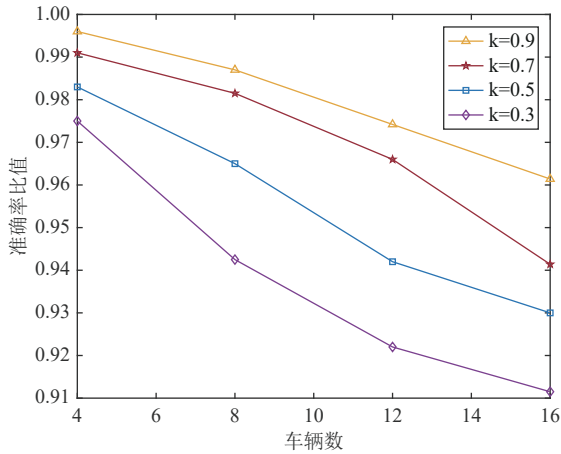


图5 不同车辆数和链路可靠程度下 perfect-FL 与 FL-DSGD 方案准确率比值  
Fig. 5 Accuracy ratio of perfect FL and FL-DSGD schemes under different vehicle numbers and link reliability

图 6 展示了在不同车辆数下 FL-DSGD 方案的训练准确率. 在保证通信链路可靠性分布一致性的前提下, 分别设置了 [4, 8, 12, 16] 辆车, 每个车辆从训练集中随机选择 3000 个样本进行训练. 从图 6 可以看出, 在不同数量的车辆参与联邦学习过程中, 都

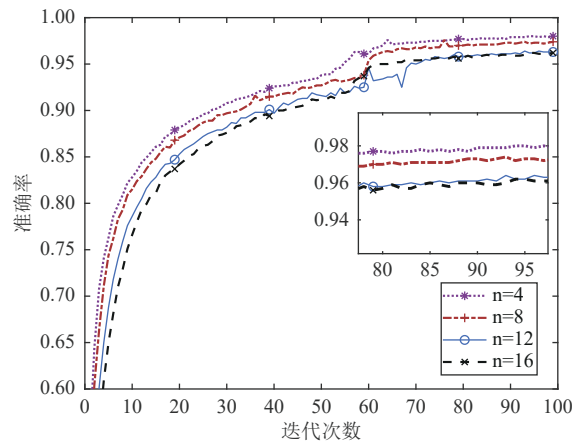


图 6 FL-DSGD 方案在不同车辆数下训练准确率对比  
Fig. 6 Comparison of FL-DSGD training accuracy under different vehicle numbers

能达到收敛的效果, 但随着车辆数量的增加其收敛效果会有一定程度的下降, 其中当车辆数分别为 4 和 16 时, FL-DSGD 方案的训练准确率只相差 0.02, 表明了该方案对车辆数量增多时具有鲁棒性. 虽然更多的车辆参与到联邦学习会导致基站服务器的信息处理压力增大, 但车辆的增加也会导致通信连接更密集, 使整体训练样本增加, 因此训练的模型准确率都保持在 95% 以上.

### 4 结语

针对 C-V2X 通信场景下, 车辆与基站服务器建立 V2I 链路通信进行联邦共识更新, 由于通信链路不可靠导致模型参数在传输过程中丢失的问题. 本文设计了一个联邦抗干扰模型更新方案——FL-DSGD. 该方案考虑到实际 V2X 通信场景中通信链路的不可靠性, 实现了低通信开销的同时对通信链路不可靠具有鲁棒性的效果. 实验结果表明: 在通信链路受到干扰导致不可靠的场景下, FL-DSGD 方案不仅与利用完美通信链路进行联邦聚合的 perfect-FL 方案拥有相似的收敛速率, 且在通信链路可靠性程度发生变化以及参与车辆数量变化的情况下仍有较好的准确率, 证明所提方案鲁棒性较好.

### 参考文献

[1] WANG J, LIU J, KATO N. Networking and communications in autonomous driving: A survey [J]. IEEE Communications Surveys & Tutorials, 2019, 21(2): 1243-1274.  
[2] SUN W, STRÖM E G, BRÄNNSTRÖM F, et al. Radio resource management for D2D-based V2V communication [J].

- IEEE Transactions on Vehicular Technology, 2016, 65(8): 6636-6650.
- [3] LIANG L, LI G Y, XU W. Resource allocation for D2D-enabled vehicular communications [J]. IEEE Transactions on Communications, 2017, 65(7): 3186-3197.
- [4] MEI J, ZHENG K, ZHAO L, et al. A latency and reliability guaranteed resource allocation scheme for LTE V2V communication systems [J]. IEEE Transactions on Wireless Communications, 2018, 17(6): 3850-3860.
- [5] LIU C F, BENNIS M. Ultra-reliable and low-latency vehicular transmission: An extreme value theory approach [J]. IEEE Communications Letters, 2018, 22(6): 1292-1295.
- [6] MCMAHAN H B, MOORE E, RAMAGE D, et al. Communication-efficient learning of deep networks from decentralized data [C]//Artificial Intelligence and Statistics, Fort Lauderdale:PMLR,2017: 1273-1282.
- [7] LI T, SAHU A K, ZAHEER M, et al. Federated optimization in heterogeneous networks [J]. Proceedings of Machine learning and systems, 2020, 2: 429-450.
- [8] FALLAH A, MOKHTARI A, OZDAGLAR A. Personalized federated learning with theoretical guarantees: A model-agnostic meta-learning approach [J]. Advances in Neural Information Processing Systems, 2020, 33: 3557-3568.
- [9] DINH C T, TRAN N H, NGUYEN T D. Personalized federated learning with Moreau envelopes [J]. Advances in Neural Information Processing Systems, 2020, 33: 21394-21405.
- [10] LI T, HU S, BEIRAMI A, et al. Ditto: Fair and robust federated learning through personalization [C]//International Conference on Machine Learning. Virtual:PMLR, 2021: 6357-6368.
- [11] YUAN K, LING Q, YIN W. On the convergence of decentralized gradient descent [J]. SIAM Journal on Optimization, 2016, 26(3): 1835-1854.
- [12] ASSRAN M, LOIZOU N, BALLAS N, et al. Stochastic gradient push for distributed deep learning [C]//International Conference on Machine Learning. Long Beach:PMLR, 2019: 344-353.
- [13] LIAN X, ZHANG W, ZHANG C, et al. Asynchronous decentralized parallel stochastic gradient descent [C]//International Conference on Machine Learning. Stockholm:PMLR, 2018: 3043-3052.
- [14] TANG H, GAN S, ZHANG C, et al. Communication compression for decentralized training [J]. Advances in Neural Information Processing Systems, 2018, 31: 7663-7673.
- [15] KOLOSKOVA A, STICH S U, JAGGI M. Decentralized stochastic optimization and gossip algorithms with compressed communication [C]//International Conference on Machine Learning. Long Beach:PMLR, 2019: 3478-3487.

(责编&校对 雷建云)