基于同态加密和群签名的可验证联邦学习方案

李亚红*①② 李一婧① 杨小东③ 张 源② 牛淑芬③

①(兰州交通大学电子与信息工程学院 兰州 730000)

②(电子科技大学计算机与工程学院 成都 610054)

③(西北师范大学计算机工程学院 兰州 730000)

摘 要：在车载网络(VANETs)中，联邦学习(FL)通过协同训练机器学习模型，实现了车辆间的数据隐私保护，并提高了整体模型的性能。然而，FL在VANETs中的应用仍面临诸多挑战，如模型泄露风险、训练结果验证困难以及高计算和通信成本等问题。针对这些问题，该文提出一种面向联邦学习的可验证隐私保护批量聚合方案。首先，该方案基于Boneh-Lynn-Shacham (BLS)动态短群聚合签名技术，保护了客户端与路边单元(RSU)交互过程中的数据完整性，确保全局梯度模型更新与共享过程的不可篡改性。当出现异常结果时，方案利用群签名的特性实现车辆的可追溯性。其次，结合改进的Cheon-Kim-Kim-Song (CKKS)线性同态哈希算法，对梯度聚合结果进行验证，确保在联邦学习的聚合过程中保持客户端梯度的机密性，并验证聚合结果的准确性，防止服务器篡改数据导致模型训练无效的问题。此外，该方案还支持车辆在部分掉线的情况下继续更新模型，保障系统的稳定性。实验结果表明，与现有方案相比，该方案在提升数据隐私安全性和结果的可验证性的同时，保证了较高效率。

关键词：隐私保护；联邦学习；车载自组网；可验证聚合；群签名

中图分类号：TN918; TP309.7 **文献标识码**：A **文章编号**：1009-5896(2025)03-0758-11

DOI: 10.11999/JEIT240796

1 引言

随着智能交通系统的发展，车载网络(Vehicle Ad-hoc NETworks, VANETs)成为研究热点。特别是在全球模型合作训练的背景下，联邦学习(Federated Learning, FL)作为一种先进的分布式深度学习框架[1]，在VANETs中的应用备受关注。FL保障数据本地化的同时，充分利用车辆中的数据资源，支持复杂的机器学习任务[2]，提升了数据处理的精度与效率，增强了数据的安全性与隐私保护。

尽管FL在隐私保护方面具有优势，但在模型更新过程中，仍面临梯度推测敏感数据、恶意客户端重建数据或注入恶意更新等安全威胁。为应对这些挑战，研究者提出了差分隐私、安全多方计算和同态加密等技术，以加强模型更新阶段的隐私和安全保护。差分隐私通过向数据或模型中添加噪声来保护隐私，但会影响模型精度[3,4]。安全多方计算能够有效保护隐私，但其设计复杂且通信开销较大[5]。相比之下，同态加密技术无需解密即可直接进行计算，既能保障数据隐私，又减少了频繁的加解密操作和交互需求，从而显著降低了计算和通信开销，特别适合高频交互的车载网络环境。通过应用同态加密技术，Tamilarasi等人[6]为VANETs提供了一种改进的隐私和安全保护方案，有效解决了用户消息隐私与安全的关键问题。且该方案优化了密钥生成过程，显著提升加密计算的整体性能。Wibawa等人[7]提出的方案结合BFV(Brakerski-Fan-Vercauteren)同态加密，有效降低了VANETs中数据泄露的风险，同时提升了系统的计算效率。Zhang等人[8]提出了一种结合分布式选择随机梯度下降与Paillier同态加密的轻量级隐私保护联邦学习方案，该方案在降低复杂密码系统计算成本的同时，防止了本地梯度泄露。

群签名技术在车载网络联邦学习中的隐私保护中具有显著优势。它不仅确保签名者匿名，防止外部攻击者识别身份，还具备可追溯性，授权机构可在争议或恶意行为发生时揭示签名者身份，从而增强系统安全性。此外，群签名保证了数据完整性，验证签名是否来自合法成员，有助于在保护隐私的同时确保数据的可信性和准确性。例如，Wang等人[9]结合了假名和群签名技术，避免用户直接暴露

收稿日期：2024-09-14；改回日期：2025-02-17；网络出版：2025-02-21

*通信作者： 李亚红 liyahong@lzjtu.edu.cn

基金项目：国家自然科学基金(62461032)，甘肃省科技计划(22JR5RA158, 22JR5RA350)，甘肃省高校教师创新基金(2023A-041, 2023-ZD-234)，兰州交通大学-天津大学联合创新基金(LH2024003)

Foundation Items: The National Natural Science Foundation of China (62461032), Gansu Science and Technology Plan (22JR5RA158, 22JR5RA350), Gansu Province University Teachers Innovation Fund Project (2023A-041, 2023-ZD-234), Lanzhou Jiaotong University-Tianjin University Joint Innovation Fund Project (LH2024003)

真实身份的问题。An等人[10]采用动态群签名方案，允许成员匿名签署消息，实现了身份隐私保护。Zhang等人[11]提出了一种基于可溯源的群签名方案，确保在信息共享时数据的完整性，并能够追溯恶意车辆的真实身份。

现有的群签名方案主要用于身份验证，但在大规模数据聚合的真实性验证方面仍存在局限性，尤其在分布式环境中，恶意服务器或客户端可能伪造更新模型，影响全局模型的准确性，威胁系统安全。此外，VANETs中频繁的车辆掉线或网络中断问题，进一步挑战了数据聚合过程的稳定性，进而影响模型训练的可靠性。因此，在VANETs环境中，构建稳健且持续的FL方案以确保数据聚合和模型训练的稳定性，变得至关重要。为此，批量聚合验证成为有效解决方案，它不仅能在保护隐私的同时验证聚合结果的真实性，防止恶意节点干扰，还能优化计算和通信开销，提高系统效率。文献[12]提出了一种基于短群签名的批量聚合验证方案，解决了因设备数量增加而导致的计算延迟问题。文献[13]通过非交互式零知识证明和伪随机生成器生成的随机矩阵，利用两阶段的聚合验证方案以及边缘设备，有效减少了计算和通信开销。文献[14]通过成对掩码增强隐私保护，并引入批量聚合验证机制，确保在不泄露敏感数据的前提下，提高模型训练的效率。

综上所述，VANETs中FL模型训练面临多重挑战。首先，现有方案通常需多轮通信，导致聚合过程复杂，且消耗大量计算资源和带宽。其次，隐私保护和聚合结果验证机制不完善，可能导致数据泄露或聚合结果不准确，威胁系统安全。最后，车辆频繁掉线或网络中断影响聚合稳定性，降低训练可靠性。针对这些问题，本文提出了一种高效且安全的可验证批量聚合方案，旨在提高数据隐私保护并提升系统效率。主要贡献如下：

(1) 高效可验证的批量聚合方案：提出基于双线性对的BLS动态短群签名方案，简化了群签名生成过程，利用线性同态哈希验证聚合结果，通过批量验证聚合梯度更新，显著减少计算负载和通信开销，提升系统整体效率。

(2) 模型隐私保护：引入改进的CKKS加密技术，确保车辆数据和训练模型的隐私性，满足选择明文攻击下的密文不可区分性要求，有效保护数据安全和训练过程隐私。

(3) 非交互特性：方案允许车辆提交加密信息后离线，即使部分车辆掉线，依然保障模型更新的正常进行，确保系统稳定性。

2　预备知识

2.1　改进的CKKS同态加密

CKKS同态加密方案适用于实时数据计算场景。能够在多次同态运算后保持较高的计算精度，支持对密文进行加法和乘法运算，确保在加密环境中完成复杂计算任务。该方案广泛应用于需要保护数据隐私的场景，满足数据隐私和计算效率的双重需求[15]。改进后的CKKS方案结合线性同态哈希函数，实现了对聚合结果的可验证性以及数据隐私的保护。具体过程如下：

(1) 密钥生成：密钥生成器生成两对公私钥$(p_{\mathrm{a}}, s_{\mathrm{a}})$和$(p_{\mathrm{b}}, s_{\mathrm{b}})$。公钥$p_{\mathrm{a}}$和$p_{\mathrm{b}}$用于加密，私钥$s_{\mathrm{a}}$和$s_{\mathrm{b}}$用于解密。此外，通过私钥和随机参数，生成用于支持同态加法的密钥。

(2) 加密算法：给定明文向量$\boldsymbol{m}_i = (\boldsymbol{m}_i[1], \boldsymbol{m}_i[2], \cdots, \boldsymbol{m}_i[d])$，选择随机向量$\boldsymbol{v}_i = (\boldsymbol{v}_i[1], \boldsymbol{v}_i[2], \cdots, \boldsymbol{v}_i[d])$，利用公钥对其进行加密，生成相应的密文$(C_{i1}, C_{i2})$和哈希值。其中$C_{i1} = \mathrm{Enc}_{p_{\mathrm{a}}}(\boldsymbol{m}_i + \boldsymbol{v}_i)$, $C_{i2} = \mathrm{Enc}_{p_{\mathrm{b}}}(\boldsymbol{v}_i)$，哈希值$\mathrm{Hash}(\boldsymbol{m}_i) = \prod_{j=1}^{d} g_j^{\boldsymbol{m}_i[j]}$。

(3) 解密：使用私钥s_{a}和s_{b}和对密文(C_{i1}, C_{i2})进行解密，得到对应的明文$\boldsymbol{m}_i$。

(4) 同态运算：在改进的CKKS同态加密方案中，假设两个明文$\boldsymbol{m}_1$和$\boldsymbol{m}_2$分别对应的密文为(C_{11}, C_{12})和(C_{21}, C_{22})，则两个密文的和，结果仍然是一个密文，$(C_{11}, C_{12}) + (C_{21}, C_{22}) = (C_{11} + C_{21}, C_{12} + C_{22})$，对应的解密结果是两个明文的和$\boldsymbol{m}_1 + \boldsymbol{m}_2$。

3　系统设计

3.1　系统模型

系统模型如图1所示。该系统由3个实体组成：执法机构(Law Enforcement Authority, LEA)、RSU和车辆。在该模型中，LEA和RSU充当聚合服务器。LEA和RSU之间通过有线通道进行通信，而车辆则采用无线传输技术[16]与LEA和RSU进行通信。

LEA：在本模型中，LEA是一个受信任的管理中心，生成系统所需的参数，负责为RSU和车辆提供合法授权和验证，同时验证聚合的掩码模型并更新全局模型。

RSU：RSU位于道路两侧，根据无线传输范围将整个辖区划分为多个区域，组织车辆之间的通信。它配备了具有强计算和存储能力的边缘计算设备。在该模型中，RSU为半诚实实体，可能对车辆的信息感到好奇，但不会与敌手合谋。RSU是有权在发生恶意攻击或检测到恶意车辆时，识别并追踪其身份的机构。

车辆：在本系统模型中，车辆充当FL的客户

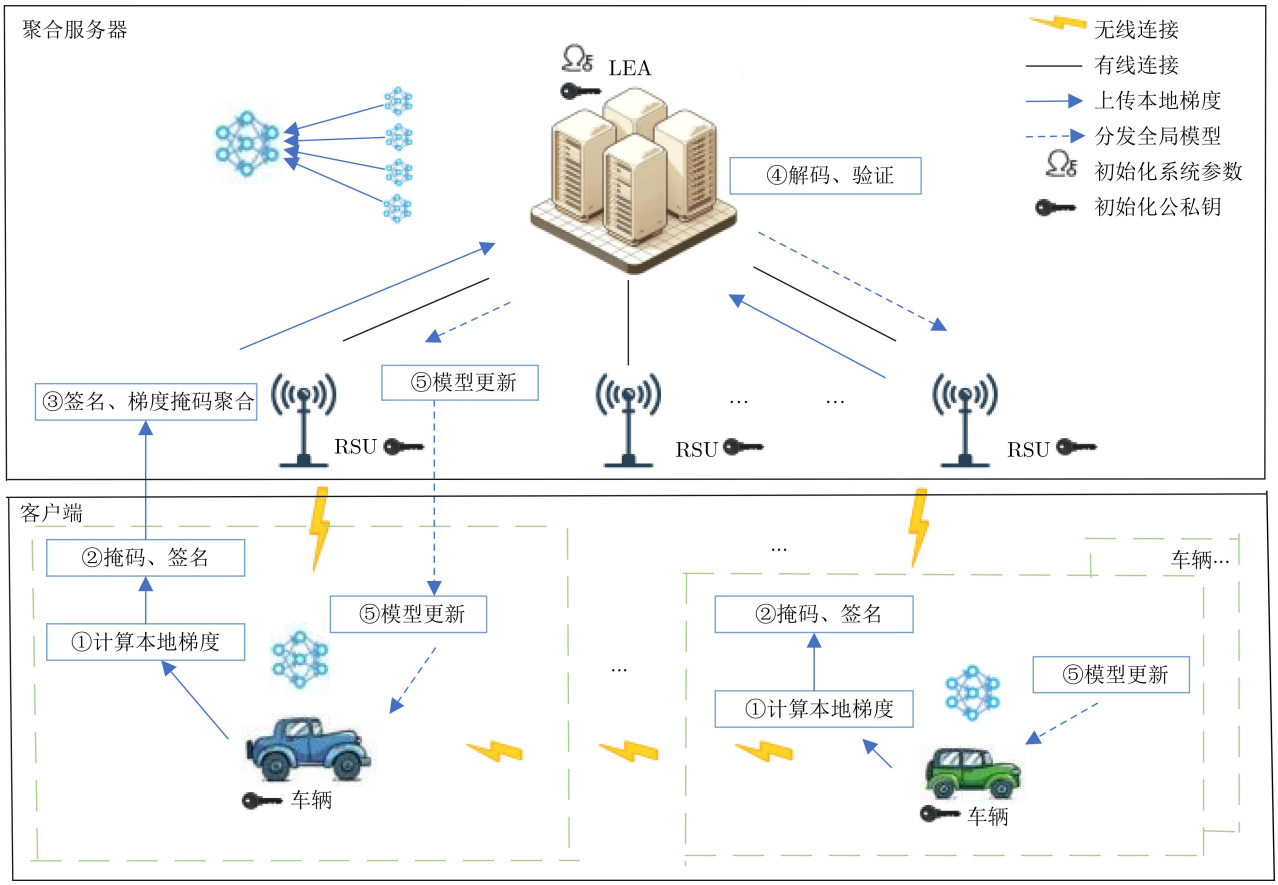


图 1 系统模型

端，负责训练本地模型并更新梯度，此外还要执行对梯度的加密、签名操作。每个客户端将加密后的模型梯度上传至RSU，以保护其隐私。车辆有诚实车辆和恶意车辆。诚实车辆利用本地数据集训练模型，并将本地模型梯度的密文传送给RSU。但是某些恶意车辆在攻击者的控制下，通过上传恶意的模型梯度密文发起攻击[17]，从而破坏整个模型的训练过程。

3.2 工作流程

(1) 初始化：FL训练开始之前，LEA初始化系统参数SP及其公私钥对，并将其发布给车辆客户端和聚合服务器RSU。同时，RSU初始化其公私钥对，车辆则初始化其公私钥及其群成员私钥。

(2) 本地模型训练：FL训练迭代开始时，LEA将(ψ^t, SP, T)转发到RSU。其中ψ^t是T时刻在第t轮的全局模型，SP是用于掩蔽模型梯度的公共参数。随后，RSU向其通信范围内的车辆广播(ψ^t, SP, T)。车辆V_i在每次迭代中自愿加入FL，并基于全局模型进行多次本地训练，以更新局部梯度$\boldsymbol{G}_i^{t+1}$。

(3) 全局模型更新阶段：车辆V_i从RSU下载最新的全局模型ψ^t。随后，车辆更新全局模型并基于其本地数据完成多次训练迭代。在VANETs的FL过程中，每次训练迭代具体过程如下：首先，车辆使用其本地数据集更新本地模型梯度，并对梯度进行掩码处理，生成相应的哈希验证码。之后，车辆与RSU共同完成对梯度掩码的签名。接着，RSU接收并对所有车辆的梯度掩码进行求和，然后对签名进行聚合。最后，LEA负责解密梯度掩码、验证并重建全局模型。为防止半可信RSU欺骗，LEA对聚合结果进行验证，验证通过后，LEA解密梯度掩码以计算最新全局模型，并将其发送回RSU。RSU随后将更新后的全局模型广播给各车辆，车辆根据RSU发送的内容更新其最新的本地模型。

4 所提方案

4.1 系统初始化

给定安全参数μ，LEA选择一个生成元为P的p阶加法循环群G，群G_{T}是同阶的乘法循环群。定义双线性映射$e: G \times G \to G_{\mathrm{T}}$，选择安全的哈希函数$\mathrm{Hash}: \{0,1\}^* \to G$。然后，LEA发布系统参数$\mathrm{SP} = \{G, G_{\mathrm{T}}, e, p, P, \mathrm{Hash}\}$。每个参与者生成如下公私钥：

(1) LEA随机选择$S_{\mathrm{l}} \in Z_p^*$作为其私钥，计算其公钥$P_{\mathrm{l}} = S_{\mathrm{l}}P$。

(2) RSU随机选择$S_{\mathrm{r}} \in Z_p^*$作为其私钥，计算其公钥$P_{\mathrm{r}} = S_{\mathrm{r}}P$。

(3) 每个车辆V_i随机选择$a_i \in Z_p^*$作为其私钥，计算其公钥$A_i = a_iP$。

假设LEA,RSU和车辆之间的通信是安全的，车辆V_i作为合法的群成员，与LEA执行如下协议：

(1) V_i发送其公钥A_i给LEA，申请成为群成员。

(2) LEA在确定V_i身份后，随机选取$S_{\mathrm{l}i}^{\mathrm{v}} \in Z_p^*$，且当$i \neq j$时，$S_{\mathrm{l}i}^{\mathrm{v}} \neq S_{\mathrm{l}j}^{\mathrm{v}}$，并计算$S_{\mathrm{l}i}^{\mathrm{s}} = (S_{\mathrm{l}} - S_{\mathrm{l}i}^{\mathrm{v}}) \mathrm{mod} p$，分别发送$S_{\mathrm{l}i}^{\mathrm{v}}$和$S_{\mathrm{l}i}^{\mathrm{s}}$给$V_i$和RSU。

(3) 完成该协议后，V_i成为合法群成员，并计算其群成员私钥$s_i = (S_{\mathrm{l}i}^{\mathrm{v}} + a_i) \bmod p$。

4.2 本地模型梯度更新及梯度掩码

假设有n个车辆参与联邦学习训练，定义为$V = \{V_i, i \in [1,n]\}$。所有加入的车辆都使用相同的训练模型，并基于各自的私有数据集进行本地模型训练。在每一轮的训练过程中，LEA都会通过RSU把最新得到的全局模型ψ^t广播给所有车辆。车辆接收到ψ^t后，利用自己的本地数据集进行训练，并更新其本地模型梯度$\boldsymbol{G}_i^{t+1}$。生成梯度的掩码具体包含以下几个步骤：

(1) 车辆V_i随机选择$\boldsymbol{v}_i = (\boldsymbol{v}_i[1], \boldsymbol{v}_i[2], \cdots, \boldsymbol{v}_i[d])$，利用改进的CKKS同态加密算法对梯度$\boldsymbol{G}_i = (\boldsymbol{G}_i[1], \boldsymbol{G}_i[2], \cdots, \boldsymbol{G}_i[d])$进行加密，其掩码为$\mathrm{msk}\{\boldsymbol{G}_i\} = \mathrm{Enc}_{P_{\mathrm{r}}}(\boldsymbol{G}_i + \boldsymbol{v}_i)$和$C_i = \mathrm{Enc}_{P_{\mathrm{l}}}(\boldsymbol{v}_i)$。

(2) 车辆V_i生成验证聚合结果的哈希值，计算梯度的哈希值$\mathrm{Hash}(\boldsymbol{G}_i) = \prod_{i=1}^{d} g_j^{\boldsymbol{G}_i[j]}$，其中$g_j$是群$G$中的元素，$d$为$\boldsymbol{G}_i$的向量维度，$\boldsymbol{G}_i[j]$表示$\boldsymbol{G}_i$中第$j$个元素。

(3) 最后将$(\mathrm{msk}\{\boldsymbol{G}_i\}, C_i)$和$\mathrm{Hash}(\boldsymbol{G}_i)$分别发送给RSU和LEA。

4.3 群签名生成

给定掩码$\mathrm{msk}\{\boldsymbol{G}_i\}$，$V_i$与RSU合作完成对掩码模型的签名。

(1) V_i计算$\sigma_i^{\mathrm{v}} = s_i\mathrm{Hash}(\mathrm{msk}\{\boldsymbol{G}_i\})$，并发送$(\sigma_i^{\mathrm{v}}, A_i)$给RSU。

(2) 当RSU接收到$(\sigma_i^{\mathrm{v}}, A_i)$后，首先根据公钥$A_i$确定$V_i$是否为合法群成员，如果不合法，则拒绝。否则计算$\sigma_i^{\mathrm{s}} = S_{\mathrm{l}i}^{\mathrm{s}}\mathrm{Hash}(\mathrm{msk}\{\boldsymbol{G}_i\})$，然后验证式(1)是否成立。如果不成立，则要求$V_i$重签。

$$e(P, \sigma_i^{\mathrm{s}} + \sigma_i^{\mathrm{v}}) = e(P_{\mathrm{l}} + A_i, \mathrm{Hash}(\mathrm{msk}\{\boldsymbol{G}_i\})) \tag{1}$$

(3) RSU随机选取$b_i \in Z_p^*$，计算$B_i = b_iP$, $S_i = A_i + B_i$, $\alpha_i = b_i\mathrm{Hash}(\mathrm{msk}\{\boldsymbol{G}_i\})$, $\sigma_i = \sigma_i^{\mathrm{s}} + \sigma_i^{\mathrm{v}} + \alpha_i$，群签名为$(\sigma_i, S_i)$。如果签名有问题，需要获取该签名者的真实身份进行追溯，RSU需保存$(A_i, \mathrm{Hash}(\mathrm{msk}\{\boldsymbol{G}_i\}), b_i)$至存储列表。

4.4 掩码聚合和签名聚合

当RSU从n个客户端V_i接收到掩码$\mathrm{msk}\{G_i\}$, C_i和签名(σ_i, S_i)。

(1) 梯度掩码基于同态性质累加计算如式(2)、式(3)的聚合结果

$$\begin{aligned} C_{\mathrm{a}} &= \mathrm{Enc}_{P_{\mathrm{r}}}(\boldsymbol{G}_1 + \boldsymbol{v}_1) + \mathrm{Enc}_{P_{\mathrm{r}}}(\boldsymbol{G}_2 + \boldsymbol{v}_2) + \cdots \\ &\quad + \mathrm{Enc}_{P_{\mathrm{r}}}(\boldsymbol{G}_n + \boldsymbol{v}_n) \\ &= \mathrm{Enc}_{P_{\mathrm{r}}}\left(\sum_{i=1}^{n}(\boldsymbol{G}_i + \boldsymbol{v}_i)\right) \end{aligned} \tag{2}$$

$$\begin{aligned} C_{\mathrm{b}} &= \mathrm{Enc}_{P_{\mathrm{l}}}(\boldsymbol{v}_1) + \mathrm{Enc}_{P_{\mathrm{l}}}(\boldsymbol{v}_2) + \cdots + \mathrm{Enc}_{P_{\mathrm{l}}}(\boldsymbol{v}_n) \\ &= \mathrm{Enc}_{P_{\mathrm{l}}}\left(\sum_{i=1}^{n}\boldsymbol{v}_i\right) \end{aligned} \tag{3}$$

(2) RSU利用其私钥S_{r}对C_{a}解密得到$\sum_{i=1}^{n}(\boldsymbol{G}_i + \boldsymbol{v}_i)$。

(3) 使用LEA的公钥P_{l}对$\sum_{i=1}^{n}(\boldsymbol{G}_i + \boldsymbol{v}_i)$加密，得到新的密文$C_{\mathrm{m}} = \mathrm{Enc}_{P_{\mathrm{l}}}\left(\sum_{i=1}^{n}(\boldsymbol{G}_i + \boldsymbol{v}_i)\right)$，并计算聚合掩码$\mathrm{Enc}_{P_{\mathrm{l}}}\left(\sum_{i=1}^{n}\boldsymbol{G}_i\right) = \sum_{i=1}^{n}\mathrm{Enc}_{P_{\mathrm{l}}}(\boldsymbol{G}_i) = C_{\mathrm{m}} - C_{\mathrm{b}}$。

(4) RSU聚合签名$\prod_{i=1}^{n}\sigma_i = \sigma_1\sigma_2\cdots\sigma_n$。

(5) 返回消息$\mathrm{Msg} = \left(\prod_{i=1}^{n}\sigma_i, \mathrm{Enc}_{P_{\mathrm{l}}}\left(\sum_{i=1}^{n}\boldsymbol{G}_i\right), T, (S_1, S_2, \cdots, S_n)\right)$给LEA，其中$T$为添加的时间戳机制，以确保RSU拒绝任何超过指定最大时间限制的数据。

4.5 梯度验证

LEA收到消息$\mathrm{Msg} = \left(\prod_{i=1}^{n}\sigma_i, \mathrm{Enc}_{P_{\mathrm{l}}}\left(\sum_{i=1}^{n}\boldsymbol{G}_i\right), (S_1, S_2, \cdots, S_n), T\right)$，为了加快梯度验证过程，LEA支持多个梯度消息验证。

(1) 在检查时间戳T是否最新后，LEA验证式(4)是否成立，若成立，则单个模型梯度$\mathrm{Enc}_{P_{\mathrm{l}}}(\boldsymbol{G}_i)$是有效的，否则拒绝

$$e(P, \sigma_i) = e(P_{\mathrm{l}} + S_i, \mathrm{Hash}(\mathrm{msk}\{\boldsymbol{G}_i\})) \tag{4}$$

(2) 如果式(5)成立，则接受聚合梯度掩码$\mathrm{Enc}_{P_{\mathrm{l}}}\left(\sum_{i=1}^{n}\boldsymbol{G}_i\right)$作为有效聚合掩码，否则拒绝

$$e\left(P, \prod_{i=1}^{n}\sigma_i\right) = \prod_{i=1}^{n}e(P_{\mathrm{l}} + S_i, \mathrm{Hash}(\mathrm{msk}\{\boldsymbol{G}_i\})) \tag{5}$$

4.6 解密并更新全局模型

(1) 给定聚合掩码$\mathrm{Enc}_{P_{\mathrm{l}}}\left(\sum_{i=1}^{n}\boldsymbol{G}_i\right)$，LEA进

行解密计算，使用其私钥$S_{\rm l}$进行解密得到$\sum_{i=1}^{n}\boldsymbol{G}_i$。

(2) 为了防止RSU的欺骗，LEA验证式(6)是否成立

$$\mathrm{Hash}\left(\sum_{i=1}^{n}\boldsymbol{G}_i\right)=\mathrm{Hash}(\boldsymbol{G}_1)\mathrm{Hash}(\boldsymbol{G}_2)\cdots\mathrm{Hash}(\boldsymbol{G}_n) \tag{6}$$

若成立，则RSU在训练期间传递的聚合结果都是正确的，LEA使用式(7)计算在本地更新全局模型

$$\psi^{t+1}=\sum_{i=1}^{|k^t|}\frac{\boldsymbol{G}_i^{t+1}}{|k^t|} \tag{7}$$

其中$|k^t|$为迭代第t次的具有有效梯度的客户端数量。否则，丢弃该聚合值并进入下一轮的迭代训练。

4.7 正确性分析

(1) 车辆签名正确性验证：RSU通过式(8)验证车辆签名的正确性，若成立，则车辆签名是正确的。

$$\begin{aligned}e(P,\sigma_i^{\rm s}+\sigma_i^{\rm v})&=e(P,S_{{\rm l}i}^{\rm s}\mathrm{Hash}(\mathrm{msk}\{\boldsymbol{G}_i\})\\&\quad+s_i\mathrm{Hash}(\mathrm{msk}\{\boldsymbol{G}_i\}))\\&=e(P\mathrm{Hash}(\mathrm{msk}\{\boldsymbol{G}_i\}),S_{{\rm l}i}^{\rm s}+s_i)\\&=e(P\mathrm{Hash}(\mathrm{msk}\{\boldsymbol{G}_i\}),S_{\rm l}-S_{{\rm l}i}^{\rm v}+S_{{\rm l}i}^{\rm v}+a_i)\\&=e(P(S_{\rm l}+a_i),\mathrm{Hash}(\mathrm{msk}\{\boldsymbol{G}_i\}))\\&=e(P_{\rm l}+A_i,\mathrm{Hash}(\mathrm{msk}\{\boldsymbol{G}_i\}))\end{aligned} \tag{8}$$

(2) 群签名正确性验证：假设(σ_i,S_i)是群成员V_i对梯度掩码$\mathrm{msk}\{\boldsymbol{G}_i\}$的一个群签名，根据式(9)

$$\begin{aligned}\sigma_i&=\sigma_i^{\rm s}+\sigma_i^{\rm v}+\alpha_i\\&=S_{{\rm l}i}^{\rm s}\mathrm{Hash}(\mathrm{msk}\{\boldsymbol{G}_i\})+s_i\mathrm{Hash}(\mathrm{msk}\{\boldsymbol{G}_i\})\\&\quad+b_i\mathrm{Hash}(\mathrm{msk}\{\boldsymbol{G}_i\})\\&=S_{{\rm l}i}^{\rm s}\mathrm{Hash}(\mathrm{msk}\{\boldsymbol{G}_i\})+(S_{{\rm l}i}^{\rm v}+a_i)\\&\quad\cdot\mathrm{Hash}(\mathrm{msk}\{\boldsymbol{G}_i\})+b_i\mathrm{Hash}(\mathrm{msk}\{\boldsymbol{G}_i\})\\&=(S_{{\rm l}i}^{\rm s}+S_{{\rm l}i}^{\rm v})\mathrm{Hash}(\mathrm{msk}\{\boldsymbol{G}_i\})+a_i\mathrm{Hash}(\mathrm{msk}\{\boldsymbol{G}_i\})\\&\quad+b_i\mathrm{Hash}(\mathrm{msk}\{\boldsymbol{G}_i\})\\&=S_{\rm l}\mathrm{Hash}(\mathrm{msk}\{\boldsymbol{G}_i\})+a_i\mathrm{Hash}(\mathrm{msk}\{\boldsymbol{G}_i\})\\&\quad+b_i\mathrm{Hash}(\mathrm{msk}\{\boldsymbol{G}_i\})\end{aligned} \tag{9}$$

LEA验证式(10)

$$\begin{aligned}e(P,\sigma_i)&=e(P,S_{\rm l}\mathrm{Hash}(\mathrm{msk}\{\boldsymbol{G}_i\})\\&\quad+a_i\mathrm{Hash}(\mathrm{msk}\{\boldsymbol{G}_i\})+b_i\mathrm{Hash}(\mathrm{msk}\{\boldsymbol{G}_i\}))\\&=e(P,(S_{\rm l}+a_i+b_i)\mathrm{Hash}(\mathrm{msk}\{\boldsymbol{G}_i\}))\\&=e(P_{\rm l}+A_i+B_i,\mathrm{Hash}(\mathrm{msk}\{\boldsymbol{G}_i\}))\\&=e(P_{\rm l}+S_i,\mathrm{Hash}(\mathrm{msk}\{\boldsymbol{G}_i\}))\end{aligned} \tag{10}$$

若成立，则签名是正确的。

(3) 梯度掩码聚合的正确性验证：若LEA验证式(11)成立，则RSU聚合的梯度掩码是正确的。

$$\begin{aligned}e\left(P,\prod_{i=1}^{n}\sigma_i\right)&=e\left(P,\prod_{i=1}^{n}(\sigma_i^{\rm s}+\sigma_i^{\rm v}+\alpha_i)\right)\\&=e(P,\prod_{i=1}^{n}(S_{{\rm l}i}^{\rm s}\mathrm{Hash}(\mathrm{msk}\{\boldsymbol{G}_i\})\\&\quad+s_i\mathrm{Hash}(\mathrm{msk}\{\boldsymbol{G}_i\})\\&\quad+b_i\mathrm{Hash}(\mathrm{msk}\{\boldsymbol{G}_i\})))\\&=e(P,\prod_{i=1}^{n}\mathrm{Hash}(\mathrm{msk}\{\boldsymbol{G}_i\}))\\&=e\Bigg(P\prod_{i=1}^{n}(S_{\rm l}+a_i+b_i),\\&\quad\mathrm{Hash}(\mathrm{msk}\{\boldsymbol{G}_i\})\Bigg)\\&=\prod_{i=1}^{n}e(P_{\rm l}+S_i,\mathrm{Hash}(\mathrm{msk}\{\boldsymbol{G}_i\}))\end{aligned} \tag{11}$$

(4)聚合结果验证：LEA通过式(12)验证聚合结果的正确性，若成立，则RSU在训练期间传递的聚合结果是正确的。

$$\begin{aligned}\mathrm{Hash}\left(\sum_{i=1}^{n}\boldsymbol{G}_i\right)&=\prod_{j=1}^{d}g_j^{\sum_{i=1}^{n}\boldsymbol{G}_i[j]}=\prod_{i=1}^{n}\prod_{j=1}^{d}g_j^{\boldsymbol{G}_i[j]}\\&=\prod_{i=1}^{n}\mathrm{Hash}(\boldsymbol{G}_i)\\&=\mathrm{Hash}(\boldsymbol{G}_1)\mathrm{Hash}(\boldsymbol{G}_2)\cdots\mathrm{Hash}(\boldsymbol{G}_n)\end{aligned} \tag{12}$$

5 安全性分析

5.1 数据隐私性

定理1 若CKKS同态加密方案是安全的，则本文的方案也同样是安全。

证明 为证明本方案的安全性，假设攻击者A以不可忽略的优势ε在一定时间t内攻破CKKS同态加密算法，则存在一个挑战者B，使其以ε的概率在相同的时间t内攻破CKKS同态加密算法。A和B进行以下交互游戏。

系统建立阶段 给定系统参数μ，产生两对公私钥$(P_{\rm l},S_{\rm l})$和$(P_{\rm r},S_{\rm r})$。使用公钥$P_{\rm l}$对本地模型梯度$\boldsymbol{G}_i$加密，得到梯度掩码$\mathrm{msk}\{\boldsymbol{G}_i\}$，并将LEA的公私钥对$(P_{\rm l},S_{\rm l})$发送给攻击者A和挑战者B。给挑战者B一组密文$(\boldsymbol{G}_0,\boldsymbol{G}_1,C_\omega)$，其中$C_\omega=\mathrm{Enc}_{P_{\rm r}}(\boldsymbol{G}_\omega)$，$\omega\in\{0,1\}$，B的目的是成功区分$C_\omega$是$\boldsymbol{G}_0$的掩码还是$\boldsymbol{G}_1$的掩码。

挑战阶段 挑战者B随机选择$\boldsymbol{v}_i'=(\boldsymbol{v}_i'[1],\boldsymbol{v}_i'[2],\cdots,\boldsymbol{v}_i'[d])$，计算$C_{\omega'1}=\mathrm{Enc}_{P_{\rm r}}(\boldsymbol{G}_\omega)+\mathrm{Enc}_{P_{\rm r}}(\boldsymbol{v}')=$

$C_\omega \mathrm{Enc}_{P_i}(v')$,$C_{\omega'2}=\mathrm{Enc}_{P_1}(v')$ ， 再 将$(\boldsymbol{G}_0,\boldsymbol{G}_1,C_{\omega'})$发送给攻击者A 。

猜测阶段　攻击者A猜测出$\omega'\in\{0,1\}$，挑战者B输出ω'区分C_τ是掩码后的$\boldsymbol{G}_0$还是$\boldsymbol{G}_1$。

假设A能够在给定时间t内以不可忽略的优势ε正确猜测C_ω，则表明B可以以ε的概率来区分改进的CKKS同态算法的密文。因为CKKS同态算法满足选择明文攻击下的密文不可区分性[15]，则本方案也满足同样的安全性质。

定理2　当存在$m<n-1$个恶意车辆与RSU合谋时，也无法猜测出其他车辆的隐私信息。

证明　在本文方案中，最多有$n-1$个恶意车辆上传梯度模型$\boldsymbol{G}_\theta$和随机向量$\boldsymbol{v}_\theta$给RSU，则RSU可获取$\sum_{i=1}^{n}(\boldsymbol{G}_i+\boldsymbol{v}_i)-\sum_{\theta=1}^{n-1}(\boldsymbol{G}_\theta+\boldsymbol{v}_\theta)$，然而仅通过$\sum_{i=1}^{n}(\boldsymbol{G}_i+\boldsymbol{v}_i)-\sum_{\theta=1}^{n-1}(\boldsymbol{G}_\theta+\boldsymbol{v}_\theta)$不能推导出诚实车辆的梯度模型。此外，RSU在掩码聚合算法的计算过程中得到的是加密后的梯度$\mathrm{Enc}_{P_1}\left(\sum_{i=1}^{n}\boldsymbol{G}_i\right)$，因此A不能获取$\sum_{i=1}^{n}\boldsymbol{G}_i$。基于CKKS同态加密算法的安全性，本方案满足隐私保护的要求，可有效抵抗合谋攻击。

5.2 可验证性

本文的验证过程利用哈希函数Hash的同态性。在梯度掩码阶段，车辆将$\boldsymbol{G}_i$掩码之后，计算$\mathrm{Hash}(\boldsymbol{G}_i)$。LEA在执行验证时，若RSU正确执行了聚合计算操作且未对梯度进行任何篡改，那么返回的聚合结果$\sum_{i=1}^{n}\boldsymbol{G}_i$应为准确的。若RSU未诚实地进行聚合，而是对聚合结果进行了篡改，则返回的篡改后的聚合结果$\left(\sum_{i=1}^{n}\boldsymbol{G}_i\right)'$，将导致$\left(\sum_{i=1}^{n}\boldsymbol{G}_i\right)'\neq\sum_{i=1}^{n}\boldsymbol{G}_i$，LEA计算哈希值$\mathrm{Hash}\left(\sum_{i=1}^{n}\boldsymbol{G}_i\right)'=\prod_{j=1}^{d}g_j^{\sum_{i=1}^{n}\boldsymbol{G}_i[j]'}$验证失败。由于哈希函数的抗碰撞特性，式(13)成立

$$\begin{aligned}\prod_{j=1}^{d}g_j^{\sum_{i=1}^{n}\boldsymbol{G}_i[j]}&=\prod_{i=1}^{n}\prod_{j=1}^{d}g_j^{\boldsymbol{G}_i[j]}=\prod_{i=1}^{n}\mathrm{Hash}(\boldsymbol{G}_i)\\&=\mathrm{Hash}(\boldsymbol{G}_1)\mathrm{Hash}(\boldsymbol{G}_2)\cdots\mathrm{Hash}(\boldsymbol{G}_n)\end{aligned}\tag{13}$$

即$\prod_{j=1}^{d}g_j^{\sum_{i=1}^{n}\boldsymbol{G}_i[j]}=\prod_{j=1}^{d}g_j^{\sum_{i=1}^{n}\boldsymbol{G}_i[j]'}$，这意味着解决了群$G$上的离散对数困难问题。然而目前尚无已知的多项式时间算法能解决此难题，因此本方案能够验证本地模型更新时梯度聚合结果的准确性。

5.3 不可伪造性

本方案的签名是由群成员V_i和RSU合作产生。具体来说，群签名由3个部分组成：$S_1\mathrm{Hash}(\mathrm{msk}\{\boldsymbol{G}_i\})$为$V_i$和RSU共同产生的BLS签名；$a_i\mathrm{Hash}(\mathrm{msk}\{\boldsymbol{G}_i\})$为$V_i$对梯度生成的BLS签名；$b_i\mathrm{Hash}(\mathrm{msk}\{\boldsymbol{G}_i\})$为RSU随机生成的BLS签名。因此，缺少任何一方的参与，群签名将不完整，从而本方案满足不可伪造性。

5.4 匿名性

群签名的匿名性要求任意两个不同的群成员V_i和$V_j(i\neq j)$，对同一消息生成的群签名(σ_i,S_i)和(σ_j,S_j)是不可区分的。假设攻击者A具备一定的攻击能力，知道LEA,V_i和V_j的私钥，同时假设RSU是安全的，除目标群签名外，A在RSU的帮助下能够打开任何群签名。通过A提供的消息，V_i和V_j生成群签名(σ_i,S_i)和(σ_j,S_j)。若A能够区分这两个不同成员的签名，则表明A能够计算群$\boldsymbol{G}$上的DH(Diffie-Hellman)问题。假设A知道群签名(σ_i,S_i)来自V_i，并且知道LEA和V_i的私钥，可计算：$\alpha_i=\sigma_i-\beta-\tau_i$，$B_i=S_i-A_i$，其中$\beta=S_{1i}\mathrm{Hash}(\mathrm{msk}\{\boldsymbol{G}_i\})$，$\tau_i=a_i\mathrm{Hash}(\mathrm{msk}\{\boldsymbol{G}_i\})$。因为$B_i$和$\mathrm{Hash}(\mathrm{msk}\{\boldsymbol{G}_i\})$是群$\boldsymbol{G}$的元素，所以存在$a,b\in Z_p^*$，使得$B_i=aP$，$\mathrm{Hash}(\mathrm{msk}\{\boldsymbol{G}_i\})=bP$。根据双线性对映射性质得到式(14)

$$\begin{aligned}e(P,\alpha_i)&=e(P,\sigma_i-\beta-\tau_i)=e(P,\sigma_i)e(P,\beta)^{-1}\\&\quad\cdot e(P,\tau_i)^{-1}\\&=e(P_1+S_i,\mathrm{Hash}(\mathrm{msk}\{\boldsymbol{G}_i\}))\\&\quad\cdot e(P_1,\mathrm{Hash}(\mathrm{msk}\{\boldsymbol{G}_i\}))^{-1}\\&\quad\cdot e(A_i,\mathrm{Hash}(\mathrm{msk}\{\boldsymbol{G}_i\}))^{-1}\\&=e(P_1+S_i-P_1-A_i,\mathrm{Hash}(\mathrm{msk}\{\boldsymbol{G}_i\}))\\&=e(B_i,\mathrm{Hash}(\mathrm{msk}\{\boldsymbol{G}_i\}))\\&=e(aP,bP)=e(P,abP)\end{aligned}\tag{14}$$

由双线性对的非退化性可知：$\alpha_i=abP$，这表明攻击者A计算出了群$\boldsymbol{G}$上的DH问题。然而在群$\boldsymbol{G}$上计算DH问题是困难的，该困难性保证了群签名的匿名性，即攻击者无法通过该签名区分不同的成员。

5.5 可追踪性

在本方案中，群签名的产生过程必须在RSU的帮助下才能完成，因此实现了可追踪性。当群成员V_i生成群签名时，RSU会储存V_i的公钥A_i、掩码的哈希值$\mathrm{Hash}(\mathrm{msk}\{\boldsymbol{G}_i\})$以及随机数$b_i$的信息，这些信息使RSU能在一定情况下揭示签名的具体来源，从而实现群签名的可追踪性。

5.6 鲁棒性

所提改进的CKKS线性同态加密方案无需客户端之间进行交互。当车辆掉线时，其对应的加密信息$(\mathrm{msk}\{\boldsymbol{G}_i\},C_i)$不会被传送至RSU，从而避免了错误聚合结果的生成。该方法不仅降低了对网络稳定性的依赖，也优化了系统的整体通信效率。

6 性能分析

本节从理论和实验两方面对方案的性能进行了评估。首先，分析了客户端和聚合服务器的计算与通信成本，并将其与现有方案进行了对比。随后，对本方案的隐私保护能力进行了量化分析，验证了其在确保数据安全和隐私方面的性能。最后，基于德国交通标志识别基准数据集(German Traffic Sign Recognition Benchmark, GTSRB)，对本方案进行了性能评估，验证了其在实际环境中应用的有效性。

6.1 实验环境

实验环境如下：操作系统为Windows 11(64位架构)，处理器为Intel Core i9-14900HX，内存为16 GB。实验使用Python编程语言，在Visual Studio Code上进行模拟实现。同态加密方案采用改进的CKKS方案，安全参数设置为8 192。各类加密操作执行1 000次的平均时间如表1所示。

6.2 计算开销对比

为分析所提方案的计算性能，本文对文献[12,13]的方案进行分析，表2展示了各方案在客户端和聚合服务器在不同阶段计算开销。

图2基于表2的数据，进一步展示了随着客户端数量增加时计算开销的变化趋势。从图2(a)可以看出，当客户端的数量增加时，所提方案的计算开销相比于文献[12]平均降低了13.5%。相比文献[13]平均降低了53.6%。由此可见，所提聚合方案在计算开销方面相较于文献[12,13]具有明显优势。从图2(b)可进一步观察到，不同客户端数量下聚合服务器的计算开销。随着客户端数量增加，所提方案的聚合服务器运行时间相比文献[12]平均减少了42.4%，相比文献[13]平均减少了33.8%，这表明所提方案在聚合服务器的计算效率上也具有明显优势。这种优势是由于本文采用了BLS动态短群签名，该算法具有较短的签名长度和公钥长度，并且在单次验证过程中仅需计算两个双线性对，显著减少了计算时间。同时，本方案的设计基于同态加法，简化了加密运算的复杂度，进一步降低了整体方案的计算成本。相比文献[12]的群签名方案由于包含幂函数运算，导致客户端计算开销显著增加。而文献[13]中的方案涉及复杂的零知识证明，结构复杂且计算量大，增加了客户端的计算负担。

6.3 通信开销对比

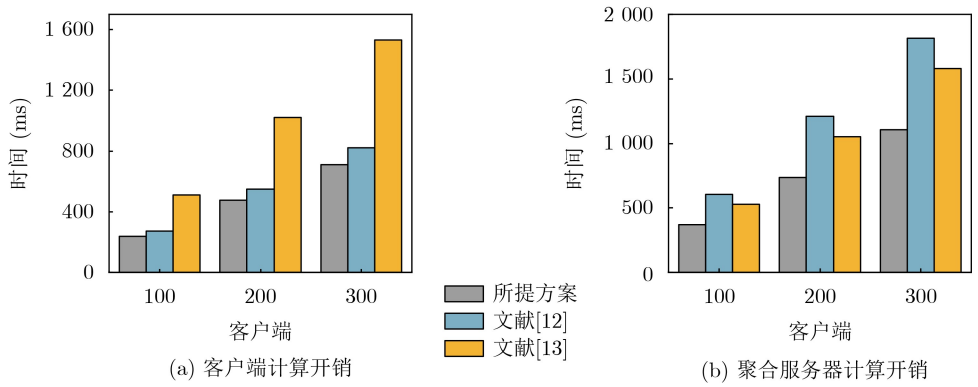
本文对不同方案的客户端与聚合服务器之间的

表 1 密码学操作执行时间

符号	描述	运行时间(ms)
T_{bp}	双线性对操作	1.118 1
T_{h}	映射到G的哈希操作	0.019 3
T_{m}	G下的乘法操作	0.001 1
T_{a}	G下的加法操作	0.000 4
T_{e}	Z_p^*下的指数操作	0.065 0
$T_{\text{o-enc}}$	一次性密码本加密	0.394 0
$T_{\text{o-dec}}$	一次性密码本解密	0.442 0
$T_{\text{dn-enc}}$	DH密钥交换加密	2.761 1
$T_{\text{dh-dec}}$	DH密钥交换解密	0.008 7
$T_{\text{c-enc}}$	CKKS加密	2.350 4
$T_{\text{c-dec}}$	CKKS解密	0.055 8

表 2 计算开销对比

方案	客户端计算开销(ms)	聚合服务器计算开销(ms)
文献[12]	$n(19T_{\text{m}}+13T_{\text{a}}+T_{\text{h}}+2T_{\text{bp}}+T_{\text{e}}+T_{\text{o-enc}})$	$(9n+8)T_{\text{m}}+(5n+2)T_{\text{bp}}+(9n+6)T_{\text{a}}+T_{\text{h}}+2T_{\text{e}}+T_{\text{o-dec}}$
文献[13]	$nT_{\text{o-enc}}+n(19T_{\text{m}}+13T_{\text{a}}+T_{\text{h}}+2T_{\text{bp}}+T_{\text{e}})$	$24nT_{\text{m}}+(4n+2)T_{\text{bp}}+11nT_{\text{e}}+26nT_{\text{a}}+(n+1)T_{\text{h}}$
所提方案	$n(T_{\text{c-enc}}+T_{\text{m}}+T_{\text{h}})$	$(7n-1)T_{\text{m}}+(3n+1)T_{\text{bp}}+10nT_{\text{a}}+(3n+2)T_{\text{h}}+nT_{\text{c-dec}}$



(a) 客户端计算开销　(b) 聚合服务器计算开销

图 2 计算开销对比

通信开销进行了详细的比较和分析，以评估其性能。具体参数定义为：群G元素大小为65字节，Z_p^*元素大小为20字节，时间戳T为4字节，身份ID为10字节。

表3展示了所提方案与文献[12,13]在每轮模型训练中通信开销的对比情况。图3表明，随着客户端数量的增加，所提方案的通信开销明显低于文献[12]。具体来说，所提方案比文献[12]和文献[13]分别节省了70.7%和66.8%。这种优势主要是由于所提方案的开销主要来自发送车辆的掩码模型梯度、哈希值及计算更新全局模型，而文献[12,13]包含了多轮的消息和会话密钥的传输，导致通信开销较高。

6.4 隐私保护性能分析

本节旨在对所提方案的隐私保护性能进行评估与实验分析。在联邦学习中，数据分布式特性要求加密算法在模型训练过程中维护整体数据的隐私性。实验结合深度学习框架和简单加密算术库框架，通过隐私保护强度测算方法，对数据加密的保护能力进行了量化评估。具体隐私保护强度测算的实验分析过程如下：

通过噪声模型模拟了不同加法操作次数下噪声的累积增长情况。噪声累积模型为：$N(k)=N(0)+\alpha_{\text{add}}n_{\text{add}}$，其中$N(0)\approx 10^{-6}$为初始噪声，$n_{\text{add}}$为加法操作的次数，$\alpha_{\text{add}}=10^{-6}$为一次加法引入的噪声增量，当加法操作次数$k\geq 100$时，噪声逐渐接近上限$N_{\text{limit}}=10^{-2}$。

隐私保护强度通过式(15)计算

$$S=\frac{N_{\text{limit}}-N(k)}{N_{\text{limit}}} \tag{15}$$

通过实验得出表4的数据，实验数据显示，即使在累积噪声逐渐增加的情况下，隐私保护强度略有下降，但仍维持在较高水平。这表明本方案在联邦学习的加密操作中具有强大的隐私保护能力，在保证数据隐私的前提下，实现模型的协同训练，有效抵御隐私泄露的风险。

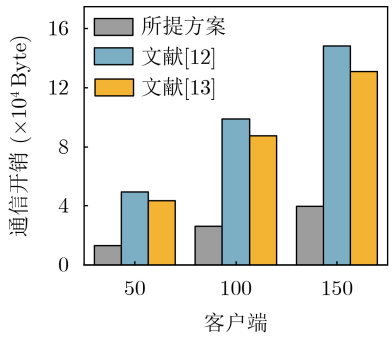


图 3 通信开销对比

表 3 通信开销对比

方案	客户端与聚合服务器间通信	聚合服务器间通信
文献[12]	$7\|G\|+3\|Z_p^*\|+\|T\|$	$7\|G\|+2\|Z_p^*\|+\|\text{ID}\|+\|T\|$
文献[13]	$7\|G\|+2\|Z_p^*\|+\|T\|$	$6\|G\|+2\|Z_p^*\|+\|T\|$
所提方案	$7\|G\|+\|Z_p^*\|$	$7\|G\|+\|Z_p^*\|+\|T\|$

6.5 应用

6.5.1 数据集和模型设置

本文选择了GTSRB数据集进行实验。该数据集由交通标志组成，分为43个类别。利用该数据集执行交通标志识别任务(Traffic Sign Recognition, TSR)。在这个任务中，使用了GTSRB数据集中的超过50 000张交通标志图像，这些图像涵盖了43种常见的交通标志。模型的测试使用了由3个5×5卷积层的基本卷积神经网络架构，训练过程中使用39 000张图像进行训练，12 000张图像用于测试。在每一轮训练中，客户端更新本地模型，学习率为0.01，批次大小为64。

6.5.2 聚合服务器的稳定性

本节通过分析聚合服务器与客户端数量和丢失比例的关系，评估了其稳定性。在该研究中，假设客户端(车辆)退出协议时，会将梯度发送到附近的客户端后退出。图4显示，随着客户端数量从20增加到100，聚合服务器的运行时间客户端数量的增加，从33 ms增加到45 ms，这表明本文方案满足非交互性。客户端在发送梯度掩码后可离线，且即使在运行过程中有部分客户端掉线，也不会影响运行结果。即本文方案可以容忍客户端的丢失，客户端的掉线比例不会显著影响聚合结果。对于TSR任务，RSU每次迭代需要大约40 ms来执行梯度聚合。

6.5.3 训练准确性

本节通过对比文献[8,12–14]中方案，评估了所提方案在联邦学习训练过程中的准确性。如图5所示，所提方案在相同轮数内更快收敛。文献[8,12]依赖整数同态加密，不支持浮点运算，导致梯度更新精度和收敛效率较低。文献[13]虽降低了通信和计算开销，但多轮通信延迟影响了模型收敛速度。文献[14]的线性同态哈希验证机制未优化全局更新方向一致性，导致收敛较慢。而所提方案采用改进的CKKS线性同态加密算法，支持浮点运算，提升

表 4 隐私保护强度数据表

操作次数k	累积噪声$N(k)$	隐私保护强度S
10	1.1×10^{-5}	0.998 9
50	5.1×10^{-5}	0.994 9
100	1.01×10^{-4}	0.989 9
500	5.001×10^{-4}	0.949 9

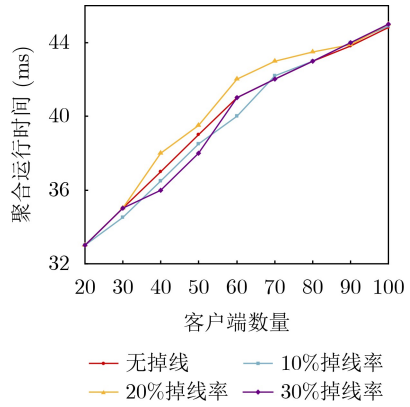


图 4 聚合服务器运行时间

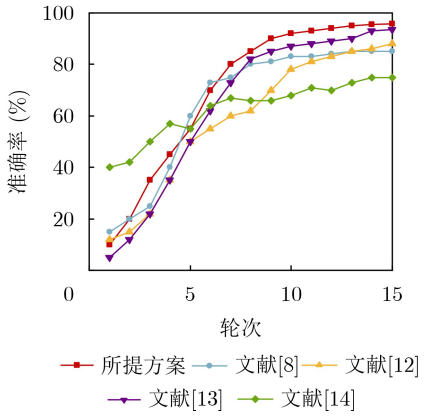


图 5 准确率对比

了梯度更新的精度。并具备非交互式性，加快了训练速度。经过15轮训练，模型基本收敛，准确率超过95%。

7 结束语

本文提出了一种FL的批量聚合方案，以优化其在VANETs中的应用。本方案结合了改进的CKKS线性同态加密算法与基于BLS的动态短群签名技术，在保障客户端非交互性和数据隐私保护的同时，还实现了聚合结果的可验证性以及模型的隐私性。并通过批量处理梯度更新，显著减少了聚合服务器的计算负载。此外，该方案支持车辆在部分掉线的情况下继续更新模型，从而增强了系统的鲁棒性和稳定性。实验结果表明，即使在客户端数量增加的情况下，该方案依然能够保持较低的通信成本和计算时间，为VANETs中的自动驾驶技术提供了可靠的数据安全保障，并支持其在实时交通信息的传播和智能交通系统中的广泛应用。未来的研究将探讨该方案在更大规模实际环境中的应用，并验证其在不同网络条件下的表现。

参 考 文 献

[1] WEN Jie, ZHANG Zhixia, LAN Yang, *et al*. A survey on federated learning: challenges and applications[J]. *International Journal of Machine Learning and Cybernetics*, 2023, 14(2): 513–535. doi: 10.1007/s13042-022-01647-y.

[2] LI Li, FAN Yuxi, TSE M, *et al*. A review of applications in federated learning[J]. *Computers & Industrial Engineering*, 2020, 149: 106854. doi: 10.1117/12.2675351.

[3] 魏立斐, 张无忌, 张蕾, 等. 基于本地差分隐私的异步横向联邦安全梯度聚合方案[J]. 电子与信息学报, 2024, 46(7): 3010–3018. doi: 10.11999/JEIT230923.
WEI Lifei, ZHANG Wuji, ZHANG Lei, *et al*. A Secure Gradient Aggregation Scheme Based on Local Differential Privacy in Asynchronous Horizontal Federated Learning[J]. *Journal of Electronics & Information Technology*, 2024, 46(7): 3010–3018. doi: 10.11999/JEIT230923.

[4] QU Zhiguo, TANG Yang, MUHAMMAD G, *et al*. Privacy protection in intelligent vehicle networking: A novel federated learning algorithm based on information fusion[J]. *Information Fusion*, 2023, 98: 101824. doi: 10.1016/j.inffus.2023.101824.

[5] LI Zhang, XU Jianbo, VIJAYAKUMAR P, *et al*. Homomorphic encryption-based privacy-preserving federated learning in IoT-enabled healthcare system[J]. *IEEE Transactions on Network Science and Engineering*, 2022, 10(5): 2864–2880. doi: 10.1109/TNSE.2022.3185327.

[6] TAMILARASI G, GANDHI K R, and PALANISAMY V. Improved Homomorphic Encryption with Optimal Key Generation Technique for VANETs[J]. *Intelligent Automation & Soft Computing*, 2022, 33(2). doi: 10.32604/iasc.2022.024687.

[7] WIBAWA F, CATAK F O, KUZLU M, *et al*. Homomorphic encryption and federated learning based privacy-preserving CNN training: Covid-19 detection use-case[C]. The 2022 European Interdisciplinary Cybersecurity Conference. Barcelona, Spain, 2022: 85-90. doi: 10.1145/3528580.3532845.

[8] ZHANG Jiale, LIU Yue, WU Di, *et al*. VPFL: A verifiable privacy-preserving federated learning scheme for edge computing systems[J]. *Digital Communications and Networks*, 2023, 9(4): 981–989. doi: 10.1016/j.dcan.2022.05.010.

[9] WANG Peng, and LIU Yining. SEMA: Secure and efficient message authentication protocol for VANETs[J]. *IEEE systems journal*, 2021, 15.(1): 846–855. doi: 10.1109/JSYST.2021.3051435.

[10] AN Haoyang, HE Debiao, BAO Zijian, *et al*. An identity-based dynamic group signature scheme for reputation evaluation systems[J]. *Journal of Systems Architecture*, 2023, 139: 102875. doi:org/10.1016/j.sysarc.2023.102875. doi: 10.1016/j.sysarc.2023.102875.

[11] 张海波, 陈舟, 黄宏武, 等. VANET 系统中基于中国剩余定理

的群内相互认证密钥协商协议[J]. 通信学报, 2022, 43(1): 182–193. doi: 10.11959/j.issn.1000-436x.2022002.
ZHANG Haibo, CHEN Zhou, HUANG Hongwu, *et al*. Intra-group mutual authentication key agreement protocol based on Chinese remainder theorem in VANET system[J]. *Journal on Communications*, 2022, 43(1): 182–193. doi: 10.11959/j.issn.1000-436x.2022002.

[12] XIA Feng, LIU Haiyang, YANG Haowei, *et al*. Batch-Aggregate: Efficient Aggregation for Private Federated Learning in VANETs[J]. *IEEE Transactions on Dependable and Secure Computing*, 2024.1-15. doi: 10.1109/TDSC.2024.3364371.

[13] XIA Feng, WANG Xiaofeng, LIU Haiyang, *et al*. A Privacy-preserving Aggregation Scheme with Continuous Authentication for Federated Learning in VANETs[J]. *IEEE Transactions on Vehicular Technology*, 2024, 73(7): 9465–9477. doi: 10.1109/TVT.2024.3369942.

[14] WANG Ruyan, YUAN Xingmin, YANG Zhigang, *et al*. RFLPV: A robust federated learning scheme with privacy preservation and verifiable aggregation in IoMT[J]. *Information Fusion*, 2024, 102: 102029. doi: 10.1016/j.inffus.2023.102029.

[15] CHEON J H, KIM A, KIM M, *et al*. Homomorphic encryption for arithmetic of approximate numbers[C]. Advances in Cryptology–ASIACRYPT 2017: 23rd International Conference on the Theory and Applications of Cryptology and Information Security, Hong Kong, China, December 3-7, 2017, Proceedings, Part I 23. Springer International Publishing, 2017. doi: 10.1007/978-3-319-70694-8_15.

[16] LEE Y, LEE J W, and KIM Y S. Near-Optimal Polynomial for Modulus Reduction Using L2-Norm for Approximate Homomorphic Encryption[J]. *IEEE Access*, vol. 8, pp. 144321-144330, 2020. doi: 10.1109/ACCESS.2020.3014369.

[17] 王勇. 联邦学习模型安全聚合关键技术研究[D]. [博士论文]. 安徽师范大学, 2024. doi: 10.26920/d.cnki.gansu.2024.000005.
WANG Yong, Research on key technologies of secure model aggregation for federated learning[D]. [Ph. D. dissertation], Anhui Normal University, 2024. doi: 10.26920/d.cnki.gansu.2024.000005.

李亚红：女，博士，副教授，研究方向为密码学与信息安全.
李一婧：女，硕士生，研究方向为联邦学习与密码学.
杨小东：男，博士，教授，研究方向为应用密码学与信息安全.
张　源：男，博士，教授，研究方向为应用密码学与信息安全.
牛淑芬：女，博士，教授，研究方向为云计算和大数据网络的隐私保护.

责任编辑：余　蓉

A Verifiable Federated Learning Scheme Based on Homomorphic Encryption and Group Signature

LI Yahong[①②]　LI Yijing[①]　YANG Xiaodong[③]　ZHANG Yuan[②]　NIU Shufen[③]

[①](*School of Electronic and Information Engineering*, *Lanzhou Jiaotong University*, *Lanzhou* 730000, *China*)

[②](*School of Computer Science and Engineering*, *University of Electronic Science and Technology of China*, *Chendu* 610054, *China*)

[③](*School of Computer Science and Engineering*, *Northwestern Normal University*, *Lanzhou* 730000, *China*)

Abstract:

Objective　In Vehicular Ad-hoc NETworks (VANETs), network instability and frequent vehicle mobility complicate data aggregation and expose it to potential attacks. Traditional Federated Learning (FL) approaches face challenges such as high computational and communication overheads, insufficient privacy protection, and difficulties in verifying aggregation results, which impact model training efficiency and stability. To address these issues, this study proposes a scheme that integrates the Boneh-Lynn-Shacham (BLS) dynamic short group signature with an enhanced Cheon-Kim-Kim-Song (CKKS) homomorphic encryption technique. This approach reduces computational and communication costs, ensures data privacy under chosen-plaintext attacks, and maintains system stability by allowing vehicles to disconnect after submitting encrypted data. The proposed framework enhances privacy, verifiability, anonymity, traceability, and robustness, providing a secure and reliable FL solution for VANETs.

Methods　A batch aggregation scheme is proposed, integrating an improved CKKS linearly homomorphic encryption algorithm with a BLS-based dynamic short group signature technique to address key challenges in

applying FL within VANETs. The improved CKKS linearly homomorphic encryption algorithm mitigates privacy leakage risks in vehicle data and training models. Data security and training privacy are ensured by maintaining ciphertext indistinguishability under chosen-plaintext attacks, preventing attackers from inferring original data from ciphertext and protecting vehicle users' privacy. Linearly homomorphic hashing verifies aggregation result correctness while reducing computational load. This approach also allows vehicles to disconnect after submitting encrypted data, enhancing system robustness and stability. Consequently, model training continuity and reliability are maintained even in dynamic and unstable vehicular network conditions. The BLS-based dynamic short group signature technique simplifies group signature generation, improving aggregation efficiency and reducing computational costs. Combined with batch processing of gradient updates, this method significantly lowers computational and communication overhead on the aggregation server. These techniques collectively enhance system efficiency and ensure adaptability to resource-constrained vehicular environments, providing a practical and effective FL solution for VANETs.

Results and Discussions The proposed scheme significantly enhances computational efficiency, reduces communication overhead, improves privacy protection, and ensures system stability in FL for vehicular networks. In terms of computational overhead, client-side computation is reduced by an average of 13.5% and 53.6%, while the aggregation server's computational cost decreases by 42.4% and 33.8%, respectively (Fig. 2a, Fig. 2b), demonstrating the scheme's ability to efficiently manage large-scale client environments with minimal computational burden. Communication overhead is also significantly minimized as the number of clients increases. By transmitting only masked gradients and hash values, the scheme achieves reductions of 70.7% and 66.8% compared to existing methods, streamlining the aggregation process and eliminating unnecessary data transmission (Fig. 3). This design ensures applicability in resource-constrained vehicular networks. The scheme maintains strong privacy protection, even under increasing noise accumulation. Experimental results confirm that data privacy is safeguarded during training, mitigating the risk of leakage (Table 4). Stability is further demonstrated as the aggregation server's performance remains unaffected by client dropouts, regardless of dropout ratios or the scale of disconnections. Its non-interactive design allows vehicles to go offline after submitting encrypted gradients, enabling the system to function reliably and maintain stable performance in dynamic vehicular environments (Fig. 4). This feature is particularly critical in scenarios involving unstable network conditions or fluctuating client availability. Furthermore, the scheme achieves a convergence rate exceeding 95% within 15 training rounds (Fig. 5). This rapid convergence is facilitated by the improved CKKS homomorphic encryption algorithm, which supports floating-point operations and enhances the precision of gradient updates. By improving gradient accuracy, the scheme enables efficient and stable model training, even in dynamic network conditions. Collectively, these results demonstrate the scheme's ability to address critical challenges in FL for VANETs.

Conclusions The FL batch aggregation scheme proposed in this study addresses data privacy and security challenges in VANETs. By integrating the BLS dynamic short group signature technique with an improved CKKS linearly homomorphic hashing algorithm, data integrity is preserved during interactions between clients and RoadSide Units (RSUs). The confidentiality and accuracy of gradient aggregation results are ensured, effectively preventing model training failures due to potential data tampering on the server side. The scheme also supports model updates despite vehicle disconnections, enhancing system stability. Experimental results demonstrate improvements in data privacy, security, and result verifiability while maintaining high efficiency. Additionally, it achieves low communication costs and reduced computation time as the number of clients increases, demonstrating strong scalability and practicality.

Key words: Privacy preservation; Federated Learning (FL); Vehicular Ad-hoc NETworks (VANETs); Verifiable aggregation; Group signature