doi:10.14132/j.cnki.1673-5439.2025.05.009

基于自适应噪声与混合注意力的联邦学习异常检测

许 建[1,2],任 义[1],周 浩[1],戴 华[1,2],杨 庚[1,2]

(1. 南京邮电大学 计算机学院,江苏 南京 210023
2. 南京邮电大学 江苏省大数据安全与智能处理重点实验室,江苏 南京 210023)

摘要:现有基于联邦学习架构的分布式异常检测难以兼顾异常检测性能与数据隐私保护。为此,提出了一种基于自适应噪声与混合注意力机制的联邦学习异常检测模型。该模型以卷积神经网络为基础,通过融合空间与多头的混合注意力机制对复杂特征进行多维度、深层次提取,实现高精度的异常检测。其次,基于本地和中心化差分隐私,通过自适应噪声添加与隐私预算分配,进一步提高了模型的隐私性和鲁棒性。基于公开数据集NSL-KDD及UNSW-NB15对模型进行了实验验证。实验结果表明,与现有主流方案相比,该模型能够在保证用户数据隐私性的同时,实现更高质量的异常检测。

关键词:联邦学习;异常检测;隐私保护;自适应噪声;混合注意力机制
中图分类号:TP309 **文献标志码**:A **文章编号**:1673-5439(2025)05-0074-11

Federated learning anomaly detection based on adaptive noise and hybrid attention

XU Jian[1,2], REN Yi[1], ZHOU Hao[1], DAI Hua[1,2], YANG Geng[1,2]

(1. School of Computer Science, Nanjing University of Posts and Telecommunications, Nanjing 210023, China
2. Jiangsu Key Laboratory of Big Data Security and Intelligent Processing, Nanjing University of Posts and Telecommunications, Nanjing 210023, China)

Abstract: Existing distributed anomaly detection models based on federated learning can hardly deal with the balance between anomaly detection performance and data privacy protection. In this regard, a federated learning anomaly detection model is proposed based on adaptive noise and hybrid attention mechanism. First, built on the convolutional neural network, this model integrates spatial and multi-head hybrid attention mechanisms to extract complex features in a multidimensional and deep manner, enabling high-precision anomaly detection. Second, based on both local and centralized differential privacy, this model utilizes the adaptive noise and the privacy budget allocation to further improve the privacy and robustness. Validated experiments are exerted on public datasets NSL-KDD and UNSW-NB15. The results show that compared with the existing mainstream approaches, the proposed model can achieve higher-quality anomaly detection while ensuring user data privacy.

Keywords: federated learning; anomaly detection; privacy protection; adaptive noise; hybrid attention mechanism

收稿日期:2024-10-18;**修回日期**:2025-03-17 **本刊网址**:http://nyzr.njupt.edu.cn
基金项目:国家自然科学基金(62372244)资助项目
作者简介:许建,男,博士,副教授,xuj@njupt.edu.cn
引用本文:许建,任义,周浩,等.基于自适应噪声与混合注意力的联邦学习异常检测[J].南京邮电大学学报(自然科学版),2025,45(5):74-84.
Citation: XU Jian, REN Yi, ZHOU Hao, et al. Federated learning anomaly detection based on adaptive noise and hybrid attention[J]. Journal of Nanjing University of Posts and Telecommunications(Natural Science Edition), 2025, 45(5): 74-84.

近年来,随着机器学习技术的快速发展,其在网络异常检测领域被广泛应用并展示出了卓越的性能[1]。但是,随着网络环境的日益复杂和数据隐私保护需求的不断增加,传统的集中式数据处理方式面临着隐私泄露和高质量数据匮乏的问题[2]。为此,研究者提出了一种分布式机器学习架构联邦学习(Federated Learning,FL)[3]。基于联邦学习架构的异常检测模型,不仅能够提供分布式的数据异常检测[4],而且可以在数据不互通的前提下,实现不同数据源、不同机构间的联合建模,进一步降低了由于数据传输带来的各项风险[5]。尽管基于联邦学习的异常检测实现了跨域的数据联合检测,并在一定程度上降低了隐私泄露风险,但其也面临诸多问题亟需解决[6]。首先,在高维数据中进行有效特征选择和提取的过程较为复杂,特征的冗余性和相关性使得检测模型的复杂性过高[7],且容易导致过拟合等异常情况[8]。其次,在联邦学习架构中,客户端与服务器之间的梯度数据传输,使得模型始终面临着以梯度反推攻击为主的安全威胁[9]。

针对高维数据的特征提取问题,可以通过添加注意力模块以强化数据特征。注意力机制通过获取每张特征图重要性的差异,实现不同特征值的权重分配,并利用计算结果反向指导特征图的权重更新。Chia 等[10]结合 Transformer 及卷积神经网络(Convolutional Neural Network,CNN),首次证明了结合自注意力机制(self-attention)和 CNN 的机制在异常检测领域表现良好。Wang 等[11]经过比较发现,将自注意力机制引入神经网络中能够显著提高神经网络的性能。Woo 等[12]首次发现对于 CNN 中的特征图来说,在通道内部同样存在着大量信息,并提出了卷积注意力模型(Convolutional Block Attention Module Network,CBAM)。该模型通过构建空间注意力模块以及通道注意力模块并进行加权聚合,能够获得更为全面的特征信息。然而,CBAM 中提出的通道注意力机制仅适用于图像检测领域[13],并且没有充分考虑全局特征以及局部特征的融合问题。与通道注意力相比,在异常检测尤其是时间序列数据的异常检测领域,多头注意力不仅能够从数据的多个维度进行长短期特征分析,而且能够通过并行计算来降低时间开销。因此,结合空间注意力与多头注意力的混合注意力机制能够更好地实现时间序列数据的特征强化和提取。

针对联邦学习在梯度传输中面临的安全问题,Abadi 等[14]提出将差分隐私[15]引入联邦学习的参数传递过程中,通过在共享模型参数中引入噪声,有效防止逆向工程和重识别攻击。对于用户级别的差分隐私算法,如 DP-Fedavg[16]、DP-FedSGD、CS-DPFL[17],其中的随机梯度下降法以及固定噪声策略存在着局部最优解的问题,往往导致优化过程容易陷入局部收敛,从而无法充分探索全局最优解,影响模型的整体性能[18]。目前为解决这类问题,主要从噪声添加的角度进行考虑[19]。然而,在差分隐私联邦学习中广泛应用的高斯噪声[20],其策略为固定的噪声大小,僵硬的策略不仅会降低模型灵活度与准确性,同时也会干扰模型对于数据特征的学习[21]。

针对上述问题,本文提出了一种基于自适应噪声与混合注意力的联邦学习异常检测方案,其主要贡献如下:

(1)提出了基于空间注意力以及多头注意力的混合注意力机制,通过对通道以及通道信息的交互进行平衡处理,实现了对数据特征的强化学习,控制了噪声对于模型的干扰,进一步提高了异常检测的性能。

(2)提出了一种基于自适应噪声的联邦学习算法,基于本地和中心化差分隐私,实现隐私预算分配以及自适应噪声添加,并证明了该机制在数据传输以及广播阶段满足差分隐私。

(3)基于 NSLKDD、UNSW-NB15 数据集,对方案性能进行验证与分析。实验结果表明,本文模型在异常检测的准确性、精确性、召回率等方面均优于同类方案。

1　模型架构

本文基于边-端双层联邦学习架构,提出了一种结合混合注意力机制和自适应差分隐私的异常检测模型。该模型的设计目标包括 3 个方面,一是提高联邦学习架构下分布式异常检测的性能,二是增强数据交互过程的隐私保护能力,三是提升模型在复杂场景下的稳定性和泛化能力,其整体框架如图 1 所示。联邦架构中每个客户端不需要与其余客户端进行通信,保证了本地数据集的隐私安全。

本方案选择交叉熵损失函数,通过计算归为异常类与实际异常之间的差值作为评价指标,其计算如式(1)所示。

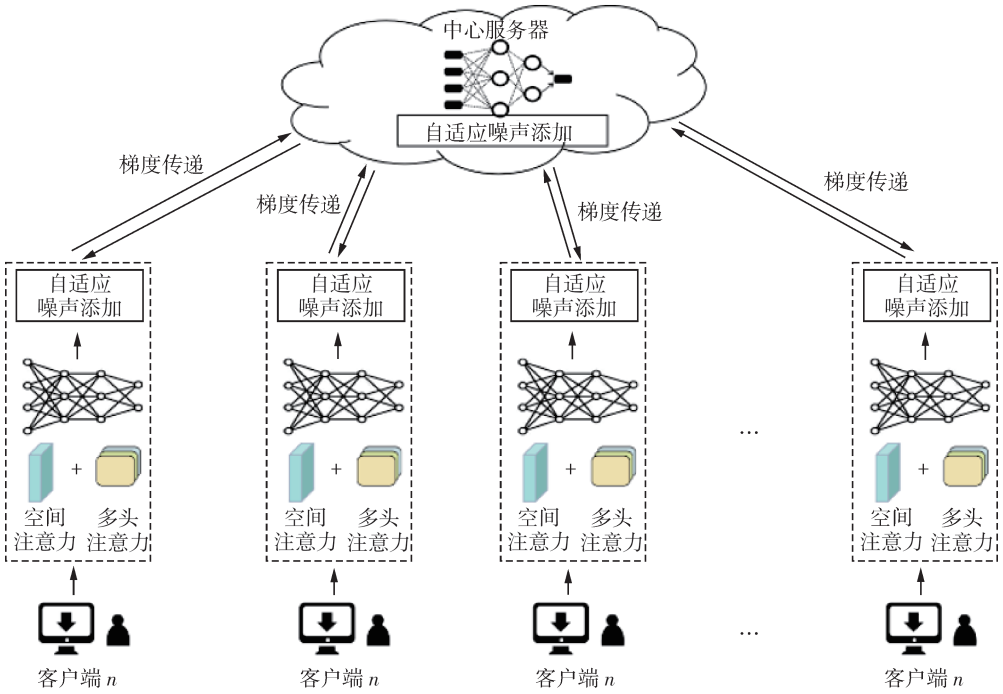

图1 整体模型框架

$$L=\frac{1}{n}\sum_i L_i=\frac{1}{n}\sum_i[-y_i\cdot\log(p_i)+(1-y_i)\cdot\log(1-p_i)] \tag{1}$$

其中,L代表损失函数,y_i及p_i为对应事件及其概率。

在模型训练过程中,每个客户端对于本地数据集$(d_1,d_2,\cdots,d_n)$进行归一化等预处理。然后,经由本地异常检测模型进行训练,获取一个本地模型的相关张量,该张量包括神经网络权重等信息。再经过损失函数对模型迭代优化,进一步获取最优本地模型参数。每个客户端将各自的模型参数进行自适应噪声添加后上传至中心服务器。中心服务器在收到各客户端的参数后进行加权聚合,并基于中心差分隐私完成噪声添加后,将全局模型参数发放至各客户端。客户端接收到服务器返回的全局模型参数后,重组该模型并调用本地数据进一步训练,获得新的本地模型参数。重复迭代上述行为,直到任务结束或达到迭代轮数上限。

客户端和服务器端的具体工作流程分别如算法1和算法2所示。

算法1 客户端工作流程

输入:客户端数据集

输出:局部模型参数

初始化:模型初始化,设置总迭代轮数R、时间T、模型参数以及优化器等各项数据

每个客户端使用差分隐私联邦学习模型训练:

基于混合注意力架构模型进行训练;

计算注意力分数、权重以及求和;

训练模型经过反向传播,获得模型梯度;

计算裁剪系数;

参数裁剪 ;

计算本地自适应噪声;

本地差分隐私:自适应噪声添加 ;

上传模型参数至中心服务器。

算法2 中心服务器工作流程

输入:客户端的局部模型参数

输出:全局模型参数

初始化:设置总迭代轮数R、时间T、模型参数以及优化器等各项数据;

收到所有客户端上传的参数;

参数加权聚合;

添加中心差分隐私;

将全局参数广播到客户端;

2 基于混合注意力的异常检测模型

2.1 异常检测模型框架

本文提出了一种基于混合注意力机制的异常检测模型,该模型以CNN为基础架构,融合空间注意力以及多头注意力形成混合注意力模块,旨在增强模型对于局部空间特征和全局依赖关系的捕捉能力[22],从而提高检测分类的精度和鲁棒性,其整体模型架构如图2所示。

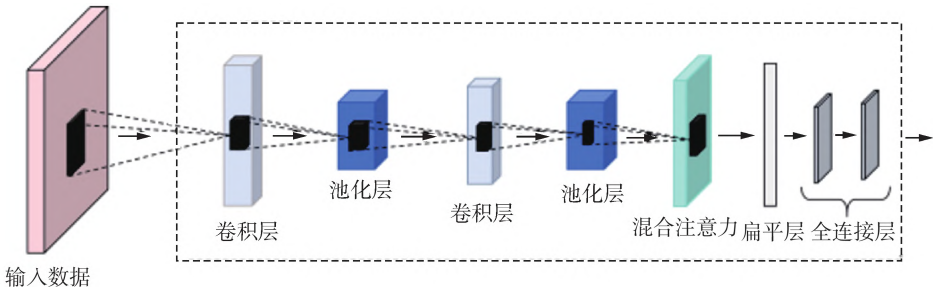

图2　检测模型架构

在该检测模型中,由于高维数据自身具有复杂的特征,因此首先需要对特征进行切割,并选取最后一列特征作为后续分类的标准即目标变量。接着,对切割完成的数据进行归一化处理以提高收敛速度。最后,进一步重塑数据尺寸,使其符合数据序列的要求,若批处理的大小为 N,步长为 TL,切割后特征的维度为196,则预处理后的数据尺寸为(N, TL, 196)。

出于轻量化的目的,本文选用了两层卷积层及最大池化层来进行初步特征提取。输入数据 $\boldsymbol{X}$ 尺寸为(N, TL, 196),经过尺寸为3,输出通道为64的卷积层及池化核大小为2、步长为1的池化层后,数据尺寸变为(N, TL-1, 64)。接着经过尺寸为3,输出通道为128的卷积层及池化核大小为2、步长为1的池化层后,数据尺寸变为(N, TL-2, 128),作为初步特征提取结果 $\boldsymbol{Y}$,其可表示为

$$\boldsymbol{Y} = \mathrm{Conv2D}(\boldsymbol{X}, W_{\mathrm{conv}}) \tag{2}$$

其中,W_{conv} 为卷积核,Conv2D 表示二维卷积层。

对于注意力层的输入 $\boldsymbol{Y}$,本文提出了结合多头注意力以及空间注意力的混合注意力机制,旨在通过对 $\boldsymbol{Y}$ 在不同维度的特征进行更为精细的学习与建模,从而提升模型在捕捉局部与全局特征上的表现,其具体执行过程如图3所示。

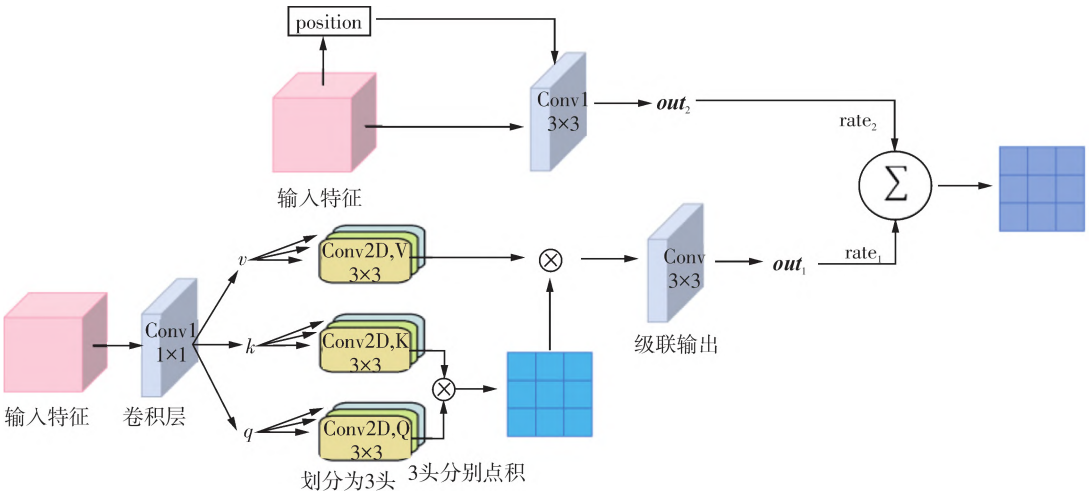

图3　混合注意力示意图

首先,根据输入的高 H 和宽 W,通过线性插值生成一个[-1,1]范围内的位置向量,用于表示相对位置从而捕捉空间信息,其横轴、纵轴分别表示为

$$\boldsymbol{loc}_w = \left[-1, \frac{2}{W-1}, \cdots, 1\right] \tag{3}$$

$$\boldsymbol{loc}_h = \left[-1, \frac{2}{H-1}, \cdots, 1\right] \tag{4}$$

进一步将上述相对位置编码组合成一个二维位置向量,生成一个最终形状为[2,W,H]的位置矩阵 $\boldsymbol{pe}$,从而获取完整二维网络的坐标,为多头注意力提供位置信息。

在通道注意力分支上,对输入数据 $\boldsymbol{X}$ 进行3个独立的线性变化获取用于检索的查询向量—查询 $\boldsymbol{q}$(query)、可能的输入特征—键 $\boldsymbol{k}$(key)以及关联信息—值 $\boldsymbol{v}$(value),该过程在每一个头上进行运算,分别表示为

$$\boldsymbol{q} = \boldsymbol{X} \cdot \boldsymbol{W}_q \tag{5}$$

$$\boldsymbol{k} = \boldsymbol{X} \cdot \boldsymbol{W}_k \tag{6}$$

$$\boldsymbol{v} = \boldsymbol{X} \cdot \boldsymbol{W}_v \tag{7}$$

其中,$\boldsymbol{W}_q$、$\boldsymbol{W}_k$和$\boldsymbol{W}_v$为随机生成的权重参数,并且权

重参数会在随后的反向传播过程中迭代优化。考虑到需要从多个维度对数据进行精细化特征学习,而头数量的增加又会导致更高的计算开销,平衡的选择一般是将维度取值为3[23]。因此,3个头从3个维度处理输入数据$\boldsymbol{X}$的查询、键以及值,来并行地计算输入信息中所选取的多个信息。

接着计算$\boldsymbol{q}$和$\boldsymbol{k}$的点积操作获得注意力分数矩阵,其作用为衡量$\boldsymbol{q}$和$\boldsymbol{k}$的相似度,该值越高则表示query越关注key。将该矩阵进行归一化获得注意力权重矩阵,并将权重与$\boldsymbol{v}$相乘进行加权求和,得到了注意力输出结果$\boldsymbol{out}_1$,公式为

$$\boldsymbol{out}_1 = \mathrm{soft\,max}\left(\frac{\boldsymbol{q}\cdot\boldsymbol{k}^{\mathrm{T}}}{\sqrt{d_k}}\right)\cdot\boldsymbol{v} \tag{8}$$

其中,d_k为键的维度。

由于多头注意力从多个维度对数据特征进行分析,而不会改变输入的特征,因此当最后将3个头的输出级联在一起时,获得的多头输出结果$\boldsymbol{out}_1$维度依旧为(N, TL-2, 128)。

在空间注意力分支上,生成的与输入数据维度相同的位置编码张量($\boldsymbol{pe}$),在通道维度上包含了数据每个位置的x、y坐标信息,与输入特征一起传入空间注意力中定义的卷积层,获得输出$\boldsymbol{out}_2$。相较于引入固定池化层的空间注意力,结合位置编码的动态权重空间注意力能够更好地捕捉特征图的细节操作,输出$\boldsymbol{out}_2$如下

$$\boldsymbol{out}_2 = \sigma(\mathrm{Conv}(\boldsymbol{X}))\odot\boldsymbol{X} \tag{9}$$

同样地,由于空间注意力不会改变输入的维度,仅调整特征图的权重,因此最终输出$\boldsymbol{out}_2$的维度为(N, TL-2, 128)。将多头注意力和空间注意力的输出通道进行维度拼接,并经过全连接层,深度可分离卷积层,利用$rate_1$和$rate_2$将$\boldsymbol{out}_1$和$\boldsymbol{out}_2$进行加权操作,形成最终的综合特征$\boldsymbol{Y}'$,作为模型最终特征输出。

$$\boldsymbol{Y}' = rate_1\cdot\boldsymbol{out}_1 + rate_2\cdot\boldsymbol{out}_2 \tag{10}$$

最后,将输出$\boldsymbol{Y}'$经过扁平层降维、全连接层映射,传入softmax分类器中获得最终的分类结果。

出于节省通信开销的考虑,本文选择在本地进行多次迭代以获取模型参数的最优解。Adam是一种结合了动量法和RMSProp的自适应优化算法。其核心思想是通过对一阶矩(梯度的移动平均)和二阶矩(梯度平方的移动平均)进行估计,来调整每个参数的学习率,从而加速训练并提高模型的收敛速度,其为目前最广泛使用的优化器。Adam优化器规则如下

$$\boldsymbol{m}_t = \beta_1\boldsymbol{m}_{t-1} + (1-\beta_1)\boldsymbol{g}_t \tag{11}$$

$$\boldsymbol{v}_t = \beta_2\boldsymbol{v}_{t-1} + (1-\beta_2)\boldsymbol{g}_t^{\,2} \tag{12}$$

$$\boldsymbol{m}_t^{\,*} = \frac{\boldsymbol{m}_t}{1-\beta_1^t} \tag{13}$$

$$\boldsymbol{v}_t^{\,*} = \frac{\boldsymbol{v}_t}{1-\beta_2^t} \tag{14}$$

$$\boldsymbol{\theta}_{t+1} = \boldsymbol{\theta}_t - \frac{\eta}{\sqrt{\boldsymbol{v}_t^*+\tau}}\boldsymbol{m}_t^* \tag{15}$$

其中,$\boldsymbol{g}_t$为参数梯度,β_1和β_2为两个加权平均的衰减系数,$\boldsymbol{m}_t^{\,*}$和$\boldsymbol{v}_t^{\,*}$为梯度的偏差纠正后的移动平均值,$\boldsymbol{\theta}_t$包括神经网络的权重矩阵和偏置向量,$\boldsymbol{\theta}_{t+1}$为更新后的参数,$\eta$为学习率,$\tau$为一个极小的常量用于避免分母为0。参数梯度$\boldsymbol{g}_t$为损失函数关于模型参数的偏导数,其主要作用是对每个可训练参数(如卷积层、注意力层、全连接层中的权重和偏置)即$\boldsymbol{\theta}_t$,依据$\boldsymbol{g}_t$初始化一阶矩$\boldsymbol{m}_t$和二阶矩$\boldsymbol{v}_t$。随后在每一轮迭代中,Adam计算当前模型梯度,并根据更新规则更新每个参数,从而达到快速收敛的目的。

3 自适应差分隐私

本文基于差分隐私的安全模型,提出了一种基于自适应噪声的数据隐私保护方案,通过在交互数据中添加自适应噪声,使得攻击者无法分辨出信息来源从而实现模型参数的隐私保护。在本文提出的隐私保护模型中,假设参数数据集D包含某个终端节点T的信息,若在数据集D'去除T的信息,对D及D'进行任意相同的查询操作可得到相同结果,则表明T的信息没有因数据聚合而产生额外的隐私泄露风险。

3.1 自适应噪声机制

在联邦学习框架中,通过结合梯度裁剪以及添加扰动,能够直接影响隐私保护的强度及模型的收敛性能。梯度裁剪可以控制数据样本对于模型的贡献,而添加噪声又能够使得模型的输出具有一定的不确定性,减少对数据特征的依赖,从而降低从模型梯度推断出敏感数据的可能性。文献[24]指出,在联邦学习的训练过程中,客户端的梯度L_2范式会随着迭代次数的增加而减小。因此,在早期迭代中为客户端添加随机噪声给模型梯度带来的影响较小,而在后期添加则影响较大。但由于传统基于固定隐私预算的噪声添加方案不能够随着梯度信息量的变化而变化,会出现噪声冗余或噪声添加量不足的问题[25]。为此,本文提出了基于Adam优

化器的自适应噪声添加和隐私预算分配方案。

首先,在初始化阶段设置全局噪声σ_A、敏感度C_C等参数,通过初始设定的隐私预算ε经过计算获得第i个客户端的噪声为σ_i。进一步根据Adam获得第i个客户端的自适应噪声σ_i',其计算公式如下

$$\sigma_i' = \frac{m_t^*}{\sqrt{v_t^* + \tau}}\sigma_i \tag{16}$$

通过σ_i'可以计算出该客户端消耗的隐私预算ε_i,通过对隐私预算ε_i大小的调整来实现不同环境下的噪声需求。由于在梯度变化量较小时存在着噪声添加量不够的风险,本文通过设置自适应噪声最小值σ_{low}来控制添加噪声的下限,提高全局隐私保护能力,此时添加噪声大小σ_i''表示为

$$\sigma_i'' = \max(\sigma_i', \sigma_{\text{low}}) \tag{17}$$

本地客户端采用自适应噪声添加的方式,为传输的梯度添加扰动,以达到梯度变化量大则添加的噪声大、梯度变化量小则添加的噪声小的目的,本地客户端添加噪声的公式如下

$$\overline{w_t^i} = w_t^i / \max\left(1, \frac{\| w_t^i \|_2}{C}\right) \tag{18}$$

$$\overline{\overline{w_t^i}} = \overline{w_t^i} + N(0, C_{\text{L}}^2 \sigma_i''^2) \tag{19}$$

$$C_{\text{L}} = \text{median}(\| w_1 \|_2, \| w_2 \|_2, \cdots, \| w_i \|_2) \tag{20}$$

其中,在本地化差分隐私联邦学习中w_t^i为客户端i在第t轮通信中的模型参数,$\| w_t^i \|_2$为其L_2范数,C为裁剪阈值,经过裁剪之后模型参数调整为$\overline{w_t^i}$;式(20)中裁剪系数C_{L}为梯度集合的L_2范式的中位数,$N(0, C_{\text{L}}^2\sigma_i''^2)$为高斯分布,均值为0,方差为$C_{\text{L}}^2\sigma_i''^2$;$\overline{\overline{w_t^i}}$为经过裁剪并添加噪声后的模型参数。

在中心服务器上,模型参数加权聚合并添加中心差分隐私的公式如下

$$w_t = \sum_{i=1}^{N}(n_i / \sum_{k=1}^{N} n_k)\overline{\overline{w_t^i}} + N(0, C_C^2\sigma_C^2) \tag{21}$$

$$\sigma_C^2 = \sigma_A^2 - \sigma_l^2 \tag{22}$$

其中,σ_C代表中心差分隐私所添加的噪声大小,σ_l代表客户端已添加的噪声之和,通过对全局噪声以及客户端的自适应噪声的设置,为中心服务器在全局模型参数的发布过程同样添加自适应的噪声σ_C,以实现双向保护隐私。

基于上述概念,利用学习率的自适应的调整作用作为噪声的自适应调整模板,并确定噪声的下限,实现了基于梯度大小的噪声自适应添加。

3.2　隐私性证明

为证明本文提出的自适应差分隐私联邦学习满足差分隐私,需证明包含客户端模型参数的数据集D及相邻数据集D',及其输出o均满足

$$\Pr[M(D) = o] \leqslant \Pr[M(D') = o] \times \text{e}^{\varepsilon} \tag{23}$$

假设查询函数为f,则由高斯公式的概率密度公式可以推得

$$\Pr[M(D) = o] = \frac{1}{\sqrt{2\pi}\,\sigma}\text{e}^{-\frac{(o - f(D))^2}{2\sigma^2}} \tag{24}$$

$$\Pr[M(D') = o] = \frac{1}{\sqrt{2\pi}\,\sigma}\text{e}^{-\frac{(o - f(D'))^2}{2\sigma^2}} \tag{25}$$

另一方面,根据高斯公式的定义可知

$$\sigma = \frac{\Delta f\sqrt{2\ln(1.25/\delta)}}{\varepsilon} \tag{26}$$

则将式(26)代入式(24,25)可以得到

$$\Pr[M(D) = o] = \frac{\varepsilon}{\sqrt{2\pi}\,\Delta f\sqrt{2\ln(1.25/\delta)}}\text{e}^{-\frac{(o - f(D))^2}{2(\frac{\Delta f\sqrt{2\ln(1.25/\delta)}}{\varepsilon})^2}} \tag{27}$$

$$\Pr[M(D') = o] = \frac{\varepsilon}{\sqrt{2\pi}\,\Delta f\sqrt{2\ln(1.25/\delta)}}\text{e}^{-\frac{(o - f(D'))^2}{2(\frac{\Delta f\sqrt{2\ln(1.25/\delta)}}{\varepsilon})^2}} \tag{28}$$

将式(27,28)代入差分隐私的定义中,即

$$\frac{\Pr[M(D) = o]}{\Pr[M(D') = o]} = \frac{\Pr[M(D) + N = o]}{\Pr[M(D') + N = o]} = \frac{\Pr[N = o - f(D)]}{\Pr[N = o - f(D')]} = \frac{\text{e}^{-\frac{(o - f(D))^2}{2(\frac{\Delta f\sqrt{2\ln(1.25/\delta)}}{\varepsilon})^2}}}{\text{e}^{-\frac{(o - f(D'))^2}{2(\frac{\Delta f\sqrt{2\ln(1.25/\delta)}}{\varepsilon})^2}}} = \frac{\text{e}^{-\frac{(o - f(D))^2}{2(\frac{\Delta f\sqrt{2\ln(1.25/\delta)}}{\varepsilon})^2}}}{\text{e}^{-\frac{(o - f(D') + \Delta f)^2}{2(\frac{\Delta f\sqrt{2\ln(1.25/\delta)}}{\varepsilon})^2}}} = \frac{\text{e}^{-\frac{x^2}{2(\frac{\Delta f\sqrt{2\ln(1.25/\delta)}}{\varepsilon})^2}}}{\text{e}^{-\frac{(x + \Delta f)^2}{2(\frac{\Delta f\sqrt{2\ln(1.25/\delta)}}{\varepsilon})^2}}} \tag{29}$$

由于概率恒为正,则对式(29)取对数可得

$$\left|\ln\left(\frac{\Pr[M(D) = o]}{\Pr[M(D') = o]}\right)\right| = \left|\ln\left(\frac{\text{e}^{-\frac{x^2}{2(\frac{\Delta f\sqrt{2\ln(1.25/\delta)}}{\varepsilon})^2}}}{\text{e}^{-\frac{(x + \Delta f)^2}{2(\frac{\Delta f\sqrt{2\ln(1.25/\delta)}}{\varepsilon})^2}}}\right)\right| = \left|\ln\left(\text{e}^{\frac{(x + \Delta f)^2 - x^2}{2(\frac{\Delta f\sqrt{2\ln(1.25/\delta)}}{\varepsilon})^2}}\right)\right| = \left|\frac{(x + \Delta f)^2 - x^2}{2(\frac{\Delta f\sqrt{2\ln(1.25/\delta)}}{\varepsilon})^2}\right| = \left|\frac{2x\Delta f - \Delta f^2}{2(\frac{\Delta f\sqrt{2\ln(1.25/\delta)}}{\varepsilon})^2}\right| \tag{30}$$

在高斯机制中$\rho \geqslant \sqrt{2\ln(1.25/\delta)}$,结合式(26)取$\sigma = \sqrt{\rho}\,\Delta f/\varepsilon$,则式(30)可变为

$$\left|\frac{2x\Delta f + \Delta f^2}{2\left(\frac{\Delta f\sqrt{2\ln(1.25/\delta)}}{\varepsilon}\right)^2}\right| \leqslant \left|\frac{2x\Delta f + \Delta f^2}{2(\frac{\Delta f\rho}{\varepsilon})^2}\right| = \left|(\frac{2x\Delta f + \Delta f^2}{2\Delta f^2\rho^2})\varepsilon^2\right| \leqslant \varepsilon^2 \tag{31}$$

其中,ε^2为某极小参数,得证。由差分隐私的可组合性,整体联邦学习框架满足差分隐私。

4 实验及结果分析

本实验基于两种公开数据集NSL-KDD和UNSW-NB15,分别模拟大样本以及小样本场景,对方案的异常检测性能进行了验证。以更为复杂的UNSW-NB15数据集为例,本实验拟定了5个客户端并为其分配了大小不等的分集,共有82 332条数据;另分配了一个test数据集用于测试准确性,共有175 341条记录。首先,分别对数据集进行切片,保留0~195列的特征用作计算,第196列的特征作为数据的标签,用于后续的检测。接着,进行归一化,通过切割和填充将原始的一维数据重新组织成适用于时间序列分析的三维数组,最终构成一个符合模型输入要求的三维数组,形状为(-1,*TL*,196)。

4.1 性能指标

本文主要的评价指标包括准确率(Accuracy,Acc)、精确度(Precision,Pre)、召回率(Recall,Rec)以及F1分数(F1 Score,F1),相关介绍及公式如下:

(1)准确率(Accuracy):预测正确的结果占总样本的百分比,计算方法为

$$\mathrm{Acc} = \frac{\mathrm{TP} + \mathrm{TN}}{\mathrm{TP} + \mathrm{TN} + \mathrm{FP} + \mathrm{FN}} \tag{32}$$

(2)精确度(Precision):在所有被预测为正的样本中实际为正样本的概率,计算方法为

$$\mathrm{Pre} = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FP}} \tag{33}$$

(3)召回率(Recall):在实际为正的样本中被预测为正样本的概率,计算方法为

$$\mathrm{Rec} = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FN}} \tag{34}$$

(4)F1分数(F1 Score):综合考虑了Precision和Recall,适用于不平衡类别的情况,计算方法为

$$\mathrm{F1} = 2 \times \frac{\mathrm{Pre} \times \mathrm{Rec}}{\mathrm{Pre} + \mathrm{Rec}} \tag{35}$$

其中,TN表示分到异常类中的异常样本数量,FN表示错分到异常类中的正常样本数量,FP表示错分到正常类的异常样本数量,TP表示分到正常类的正常样本数量,这4类属性构成了总指标混淆矩阵,如表1所示。

表1 混淆矩阵

预测类型	预测为负类	预测为正类
实际为负类	TN	FP
实际为正类	FN	TP

4.2 结果分析

4.2.1 损失函数及收敛性分析

为了更好地确定总迭代轮数,在实现降低模型过拟合可能性与保证轻量化间找到合适的迭代轮数,首先通过损失函数曲线来确定最优值,损失函数变化情况如图4所示。

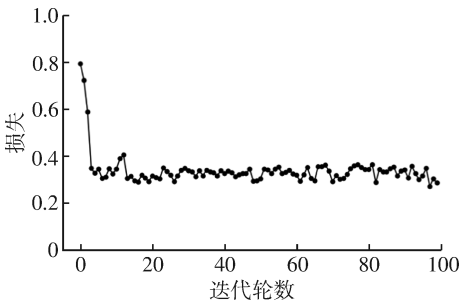

图4 损失函数下降趋势

在联邦学习中,客户端进行本地训练时,每个客户端的损失函数值可能不同,因此全局模型的损失函数趋势是所有客户端本地模型损失函数的加权平均,整体的损失函数的下降趋势可以通过全局聚合后的结果来观察。从图4可以看出,损失函数在训练的初期迅速下降,并随着迭代轮数的增加,趋势逐渐降低、波动逐渐变小,最终趋于一个常数。这是由于,在训练初期,模型的参数远离最优解,损失函数较大,因此随着梯度更新,损失值快速下降,此时,训练数据在参数空间中的探索较为广泛,更新的步伐较大,表明选择了恰当的学习率;随着训练的进行,模型逐渐逼近最优解,损失函数的下降速率逐渐减缓,此时,梯度的幅度逐渐减小,模型更新趋于稳定;在训练的后期,损失函数会趋于平稳,表示模型已经接近最优解,此时的参数更新可能很小,损失函数会达到一个低值并稳定,当损失函数出现停滞时,可以认为模型收敛。在该实验情境下,模型在设定的100轮迭代中,损失函数逐渐趋于0.3,并不再产生较大的波动,表明模型最终收敛。

同时,在整个模型训练过程中,为了减少模型收敛前的总通信开销,在每一轮中,本地模型首先进行多次迭代。在满足损失下降到一定范围内的情况下,即获取了当前轮数下本地模型参数的最优解,上传梯度至中心服务器进行聚合。针对单个客户端,在实验中通过分析损失函数的下降趋势,拟定本地迭代轮数为2,在达到迭代轮数上限时,模型准确度得到显著提高。当一个客户端获取了全局模型参数后,受限于自身条件以及高维数据集等复杂特性,还要考虑网络带宽等实际因素,全局模型初次并不能充分适配本地数据集,因此第一次全局模型在本地的应用效果不佳。基于上述情况,本实验中设定总迭代轮数为40,单次训练中本地模型迭代次数为2,后续工作中将对于单个异常检测模型一轮中迭代次数的自适应优化进行重点研究。

4.2.2　主要参数对模型的性能影响分析

为了探究不同的隐私预算对于模型性能的影响,在各项初始参数不变的情况下,设置学习率为0.01,迭代轮数为40,调整隐私预算ε。在UNSW-NB15数据集中,将每个客户端的隐私预算分别设置为0.5,0.6,0.7,0.8,并且将本文提出的自适应差分隐私联邦算法(FL-ADNP)与不应用差分隐私的FedAvg模型及传统的DP-FL模型进行对比,其中具体实验效果如表2所示。

表2　不同隐私预算下的准确率对比

隐私预算ε	FedAvg	CNN-FL	DP-FL	FL-ADNP
0.5	0.907 3	0.916 5	0.764 1	0.834 6
0.6	0.907 3	0.916 5	0.793 1	0.860 8
0.7	0.907 3	0.916 5	0.812 6	0.896 4
0.8	0.907 3	0.916 5	0.821 3	0.898 1

从表2中可以看出,首先由于FedAvg以及结合CNN的联邦模型(CNN-FL)没有引入差分隐私,因此隐私预算的调整不影响其准确度,也不会受到噪声对准确度的干扰,准确度维持在一个较高的水准,分别为90.73%及91.65%。而传统的DP-FL模型以及本文提出的FL-ADNP模型随着隐私预算的提高,准确率也在不断提高。当隐私预算从0.5提升至0.8时,DP-FL准确度从76.41%提升至82.13%,而FL-ADNP准确率从83.46%提升至89.81%。对数据进行分析,隐私预算提高,模型准确度随之提高,这是由于差分隐私的性质决定的:越高的隐私预算表明着越少的噪声干扰,但是带来的是越弱的隐私保护[26]。另外可以看出当ε从0.7提升至0.8时,准确度提升幅度相较不高,这是因为噪声的非线性效应导致的边界效应:当ε较小时,提升ε可以显著减小噪声,当ε增加到一定程度时,噪声添加量已经相对较小,从而导致ε增加相同的幅度而模型准确度并没有对应地上升。本文提出的FL-ADNP模型牺牲了一定的准确度带来的是隐私保护的加强,而在表现效果上优于传统的DP-FL模型,在相同的隐私预算下具有更高的准确度。

在联邦学习的异常检测模型中,学习速率对于噪声添加和混合注意力机制有着重要影响。学习速率过高时,模型可能在迭代中快速调整参数,难以稳定适应噪声的扰动,从而影响模型的泛化性能。相反,过低的学习速率则会延缓模型收敛,使得噪声对优化过程的影响不显著,导致训练时间过长。因此,适当的学习速率可以在平衡隐私保护和模型性能的基础上,确保噪声添加带来的扰动和混合注意力机制的特征提取效果得到良好协同,进而提升模型的整体表现和准确性。为了探究不同的学习率η对于模型性能的影响,在各项初始参数不变的情况下,在UNSW-NB15数据集中设置η分别为0.000 1,0.001,0.01,0.1,迭代轮数为40,隐私预算ε=0.7,以客户端1为例,其实验结果如图5所示。

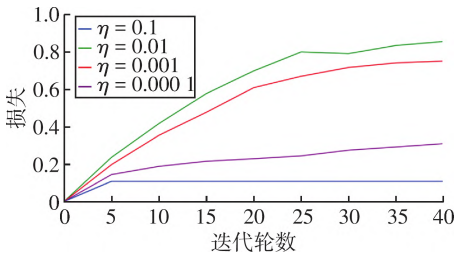

图5　UNSW-NB15数据集不同学习率η对客户端模型准确度的影响

从图5可以明显看出,对于5个客户端来说,尽管迭代过程中数值存在波动,但是整体趋势是一致的。当η为0.1时,自适应噪声会添加得过大导致整个模型失真,最后导致准确率维持在一个极低的数值。当η为0.0001时,可以看出前期收敛速度过慢,模型困于局部最优解,尽管模型准确率呈现持续稳定的上升,但是需要过于庞大的计算次数,最终造成了极大的通信开销[27]。当η为0.01时,相较于0.001的学习率,其为模型提供了更快的收敛速度。在保持稳定性的前提下,较高的学习率能够帮助模型更有效地更新权重,减少训练时间,从而在一定的训练周期内取得更佳的性能表现。可以认为,0.01的学

习率在此实验中为较为理想的选择,既提高了模型的收敛效率,又确保了模型性能的稳步提升。

根据上述实验结果,设定迭代轮数为40,学习率η为0.01,隐私预算ε为0.7,每个客户端在本地运行模型结果的评价指标如表3所示。

表3 客户端评价指标 %

客户端	Acc	Pre	Rec	F1
C1	85.19	80.43	97.80	88.23
C2	95.33	94.98	98.66	96.79
C3	90.87	84.57	99.28	91.34
C4	82.43	81.37	97.53	88.57
C5	88.76	91.06	86.74	88.84

在训练过程中,观察到尽管所有模型的精确度及其他评估指标随着迭代次数逐渐提升,但部分模型由于对特定数据特征的识别能力较弱,反而对整体训练产生了负面影响。这一现象反映了联邦学习在不同参与方间潜在的不公平性问题[28]:虽然联邦学习可以在全局上促进模型的优化,但个别参与方模型性能的差异性可能对最终结果产生不利影响。针对那些在特定数据特征识别能力上表现较弱的模型,降低其在全局模型更新中的权重,从而减轻其负面影响。相应地,鼓励作出更多贡献的模型,激励机制能够更有效地分配资源,确保整体模型的收敛性和性能不受个别劣质模型的拖累。整个模型最终的训练结果能够达到较为理想的精确度和召回率。

4.2.3 与同类方案的性能对比

为了更好地对比模型的性能,将本方案中的基于混合注意力机制的异常检测模型(FLAC)分别在MNIST、NSL-KDD及UNSW-NB15数据集下模拟不同情况下的检测。本文设计将FLAC、FL-ADNP与循环神经网络、基于二分类的CNN-FL以及M-FedAvg算法进行对比,结果如表4所示。

表4 NSL-KDD数据集模型检测性能对比

算法	Acc	Pre	Rec
FLAC	0.916 7	0.975 2	0.945 6
FL-ADNP	0.915 2	0.964 8	0.942 2
RNN	0.823 9	0.856 7	0.806 7
2-CNN-FL	0.867 2	0.943 5	0.894 1
M-FedAvg	0.895 1	0.967 4	0.935 6

将本文提出的FLAC及FL-ADNP算法与PCA、2DCNN-AE、CNN、FedAvg及结合注意力机制的FedAvg算法进行对比,根据实际应用UNSW-NB15数据集进行环境模拟,指标采用ACC,REC以及F1,部分性能指标如表5所示。

表5 UNSW-NB15数据集模型检测性能对比

算法	Acc	Rec	F1
PCA	0.741 1	0.576 4	0.647 9
CNN	0.859 3	0.910 3	0.883 1
2DCNN-AE	0.908 6	0.972 3	0.921 4
FLAC	0.922 5	0.998 3	0.923 4
FL-ADNP	0.922 4	0.997 8	0.922 7

结合表4~5,在异常检测模型中引入联邦架构,相较于单模型结构能够显著地提高检测的精准度以及隐私性。而本方案提出的联邦学习异常检测模型FL-AC,更是在多个方面表现出优异的效果。尤其是在召回率方面,该模型表现出了卓越的性能,相较于CNN提升约6%,相较于2DCNN-AE提升约1.5%,结果表明本模型在检测实际正例(真正的异常)方面具有很强的能力。高召回率意味着模型能够识别出更多的真实异常,从而减少漏检的情况。考虑到隐私保护,FLAC模型通过引入自适应噪声,构建了自适应差分隐私联邦学习模型FL-ADNP。该模型在精确度、召回率和F1分数上均取得了与原模型相近的效果。这表明,FL-ADNP模型能够在保护用户数据隐私的同时,维持其高效的异常检测能力。通过引入自适应差分隐私方法,模型在确保数据隐私安全的同时,能够最大程度地维持其检测性能。这一策略使得在不同隐私预算下的隐私保护需求与检测精度之间实现了动态平衡,既有效提升了用户数据的隐私保护水平,又确保了模型在检测任务中的表现不受显著影响。

4.2.4 消融实验

为了更好地对比注意力机制对异常检测模型性能的影响,本文设计了消融实验,对多个模型在两个数据集(UNSW-NB15和NSL-KDD)上的准确度进行了对比分析,分别包括传统卷积神经网络(CNN)、联邦学习中的CNN模型(CNN-FL)、基于空间注意力的联邦学习模型(FL-SA)、基于混合注意力的联邦学习模型(FL-AC)以及加入自适应噪声保护机制的联邦学习模型(FL-ADNP)。以准确度为衡量标准,控制其余变量,结果如表6所示。

从表6中的数据可以看出,联邦学习(FL)模型在总体性能上明显优于传统的CNN模型,展示了其在分布式数据环境下的强大处理能力。引入空间

注意力机制的FL-SA模型,相较于CNN和CNN-FL,显著提升了检测效果,这表明空间注意力机制在捕捉特征空间上的表现尤为出色。FLAC模型通过结合空间和通道注意力,进一步提高了对高维和复杂特征的处理能力,显现出在处理复杂数据集上的卓越优势。引入了自适应差分隐私机制的FL-ADNP模型,在增强隐私保护的同时,仍然保持了与FLAC接近的高精度表现,这表明隐私保护策略并未对模型性能产生显著影响。此外,数据样本特征的复杂性以及样本容量的大小,对模型的检测性能具有重要影响。模型对某些特定特征展现了更强的学习能力,因此在后续工作中,针对模型公平性的深入研究将成为进一步提升性能的关键。

表6　消融实验结果

数据集	CNN	CNN-FL	FL-SA	FLAC	FL-ADNP
UNSW-NB15	0.859 3	0.861 7	0.897 1	0.922 5	0.922 4
NSL-KDD	0.834 7	0.837 2	0.889 4	0.916 7	0.915 2

5　结束语

本文围绕网络异常检测问题,针对数据特征冗余以及隐私性保护的问题,从强化特征以及噪声的动态调整两个角度出发,提出了一种结合混合注意力的联邦异常检测模型。其中包括了一种用于特征强化和选取的混合注意力机制,以及一种动态调整客户端与服务器端之间噪声大小的自适应差分隐私机制。通过实验验证了其在异常检测环境中的有效性,相较于其他方法在检测准确率召回率等上都展现了较高的提升,并且与在数据共享情况下训练的模型相比损失极小。

在未来研究工作中,将进一步对提出方法进行改进和完善。首先,本文将继续探索在数据异构及模型异构的复杂场景下,实现高效的模型聚合机制。其次,在优化通信开销方面,如何在客户端自适应选择和多轮本地迭代上做出改进,仍然将是本文后续工作关注的重点。最后,将考虑面对潜在的恶意参与者或好奇参与者,如何进一步提升整体隐私保护效果并确保模型的鲁棒性与安全性,提升方法的实际可用性。

参考文献:

[1] 孙海丽, 龙翔, 韩兰胜, 等. 工业物联网异常检测技术综述[J]. 通信学报, 2022, 43(3): 196-210.
SUN Haili, LONG Xiang, HAN Lansheng, et al. Overview of anomaly detection techniques for industrial Internet of Things[J]. Journal on Communications, 2022, 43(3): 196-210. (in Chinese)

[2] 李杰铃, 张浩. 半监督异常流量检测研究综述[J]. 小型微型计算机系统, 2020, 41(11): 2371-2379.
LI Jieling, ZHANG Hao. Survey on semi-supervised anomaly traffic detection [J]. Journal of Chinese Computer Systems, 2020, 41(11): 2371-2379. (in Chinese)

[3] 杨强. AI与数据隐私保护: 联邦学习的破解之道[J]. 信息安全研究, 2019, 5(11): 961-965.
YANG Qiang. AI and data privacy protection: the way to federated learning [J]. Journal of Information Security Research, 2019, 5(11): 961-965. (in Chinese)

[4] 吴文泰, 吴应良, 林伟伟, 等. 横向联邦学习: 研究现状、系统应用与挑战[J]. 计算机学报, 2025, 48(1): 35-67.
WU Wentai, WU Yingliang, LIN Weiwei, et al. Horizontal federated learning: research status, system applications and open challenges[J]. Chinese Journal of Computers, 2025, 48(1): 35-67. (in Chinese)

[5] TYAGI S, RAJPUT I S, PANDEY R. Federated learning: applications, security hazards and defense measures [C]//International Conference on Device Intelligence, Computing and Communication Technologies (DICCT). 2023: 477-482.

[6] 陈学斌, 任志强, 张宏扬. 联邦学习中的安全威胁与防御措施综述[J]. 计算机应用, 2024, 44(6): 1663-1672.
CHEN Xuebin, REN Zhiqiang, ZHANG Hongyang. Review on security threats and defense measures in federated learning [J]. Journal of Computer Applications, 2024, 44(6): 1663-1672. (in Chinese)

[7] KHALID S, KHALIL T, NASREEN S. A survey of feature selection and feature extraction techniques in machine learning[C]//Science and Information Conference. 2014: 372-378.

[8] 李开菊, 许强, 王豪. 冗余数据去除的联邦学习高效通信方法[J]. 通信学报, 2023, 44(5): 79-93.
LI Kaiju, XU Qiang, WANG Hao. Communication-efficient federated learning method via redundant data elimination [J]. Journal on Communications, 2023, 44(5): 79-93. (in Chinese)

[9] WEI W Q, LIU L. Gradient leakage attack resilient deep learning[J]. IEEE Transactions on Information Forensics and Security, 2021, 17: 303-316.

[10] CHIA Y K, WITTEVEEN S, ANDREWS M. Transformer to CNN: label-scarce distillation for efficient text classification [EB/OL]. [2024-03-10]. https://arxiv.org/pdf/1909.03508.

[11] WANG Q L, WU B G, ZHU P F, et al. ECA-net: efficient channel attention for deep convolutional neural networks[C]//IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). 2020: 11531-11539.

[12] WOO S, PARK J, LEE J Y, et al. CBAM: convolutional block attention module [C]//European Conference on Computer Vision (ECCV). 2018: 3-19.

[13] 张侣, 周博文, 吴亮红. 基于改进卷积注意力模块与残差结构的SSD网络[J]. 计算机科学, 2022, 49(3): 211-217.
ZHANG Lü, ZHOU Bowen, WU Lianghong. SSD network based on improved convolutional attention module and residual structure[J]. Computer Science, 2022, 49(3): 211-217. (in Chinese)

[14] ABADI M, CHU A, GOODFELLOW I, et al. Deep learning with differential privacy[C]//Proceedings of the ACM SIGSAC Conference on Computer and Communications Security. 2016: 308-318.

[15] KAIROUZ P, BONAWITZ K, RAMAGE D. Discrete distribution estimation under local privacy[C]// Proceedings of the 33rd International Conference on Machine Learning. 2016: 2436-2444.

[16] MCMAHAN H B, RAMAGE D, TALWAR K, et al. Learning differentially private recurrent language models [EB/OL]. [2024-03-10]. https://arxiv.org/pdf/1710.06963.

[17] FU J, CHEN Z L, HAN X. Adap DP-FL: differentially private federated learning with adaptive noise[C]//IEEE International Conference on Trust, Security and Privacy in Computing and Communications (TrustCom). 2022: 656-663.

[18] 叶青青, 孟小峰, 朱敏杰, 等. 本地化差分隐私研究综述[J]. 软件学报, 2018, 29(7): 1981-2005.
YE Qingqing, MENG Xiaofeng, ZHU Minjie, et al. Survey on local differential privacy[J]. Journal of Software, 2018, 29(7): 1981-2005. (in Chinese)

[19] 肖雄, 唐卓, 肖斌, 等. 联邦学习的隐私保护与安全防御研究综述[J]. 计算机学报, 2023, 46(5): 1019-1044.
XIAO Xiong, TANG Zhuo, XIAO Bin, et al. A survey on privacy and security issues in federated learning[J]. Chinese Journal of Computers, 2023, 46(5): 1019-1044. (in Chinese)

[20] WEI K, LI J, DING M, et al. Federated learning with differential privacy: algorithms and performance analysis [J]. IEEE Transactions on Information Forensics and Security, 2020, 15: 3454-3469.

[21] 周传鑫, 孙奕, 汪德刚, 等. 联邦学习研究综述[J]. 网络与信息安全学报, 2021, 7(5): 77-92.
ZHOU Chuanxin, SUN Yi, WANG Degang, et al. Survey of federated learning research[J]. Chinese Journal of Network and Information Security, 2021, 7(5): 77-92. (in Chinese)

[22] 蒋锐, 陈儒娜, 王小明, 等. 基于注意力机制及多分支特征融合的实时语义分割算法[J]. 南京邮电大学学报(自然科学版), 2024, 44(2): 91-100.
JIANG Rui, CHEN Runa, WANG Xiaoming, et al. Real-time semantic segmentation based on attention mechanism and multi-branch feature fusion[J]. Journal of Nanjing University of Posts and Telecommunications (Natural Science Edition), 2024, 44(2): 91-100. (in Chinese)

[23] VASWANI A, SHAZEER N, PARMAR N, et al. Attention is all you need[C]//Advances in Neural Information Processing Systems. 2017: 5998-6008.

[24] JIANG S S, LU M X, HU K, et al. Personalized federated learning based on multi-head attention algorithm [J]. International Journal of Machine Learning and Cybernetics, 2023, 14(11): 3783-3798.

[25] 吴家皋, 蒋宇栋, 刘林峰. 一种自适应的网格化联邦学习客户端调度算法[J]. 南京邮电大学学报(自然科学版), 2025, 45(1): 79-89.
WU Jiagao, JIANG Yudong, LIU Linfeng. Adaptive gridding client schedule algorithm for federated learning[J]. Journal of Nanjing University of Posts and Telecommunications (Natural Science Edition), 2025, 45(1): 79-89. (in Chinese)

[26] 邱鑫源, 叶泽聪, 崔翛龙, 等. 联邦学习通信开销研究综述[J]. 计算机应用, 2022, 42(2): 333-342.
QIU Xinyuan, YE Zecong, CUI Xiaolong, et al. Survey of communication overhead of federated learning [J]. Journal of Computer Applications, 2022, 42(2): 333-342. (in Chinese)

[27] CHEN Z J, LIAO G C, MA Q, et al. Adaptive privacy budget allocation in federated learning: a multi-agent reinforcement learning approach [C]//IEEE International Conference on Communications. 2024: 5166-5171.

[28] ATABEK A, ERALP E, GURSOY M E. Trust, privacy and security aspects of bias and fairness in machine learning [C]//5th IEEE International Conference on Trust, Privacy and Security in Intelligent Systems and Applications (TPS-ISA). 2023: 111-121.

(责任编辑:李小溪)