杨驰, 俞贵琪, 张建军,等. 基于加权联邦学习和神经网络的工控系统入侵检测[J]. 智能计算机与应用,2025,15(3):79-86.
DOI:10.20169/j.issn.2095-2163.250311

基于加权联邦学习和神经网络的工控系统入侵检测

杨 驰[1], 俞贵琪[2], 张建军[2], 彭 博[2], 贾徽徽[1,3]
(1 公安部第三研究所, 上海 200031; 2 上海市公安局网络安全保卫总队, 上海 201799;
3 上海网络与信息安全测评工程技术研究中心(公安部第三研究所), 上海 200031)

摘 要: 入侵攻击对正常的工业生产流程造成阻塞和破坏,机器学习可实现对入侵检测的分类识别进而加以干预,但工控系统中各数据持有者之间存在的隐私安全壁垒无法将数据整合利用。为了打破数据壁垒,并获得更好的入侵检测分类效果,提出面向工业控制系统的联邦学习和神经网络入侵检测方法,利用多层神经网络实现入侵检测分类,通过联邦学习将各工业数据持有者的数据安全保留在本地,只传输模型和参数信息从而打破数据壁垒。针对各数据持有者的样本量不平衡问题,在联邦学习初始阶段引入权重因子来减少不平衡数据的影响,且为了弥补多维数据下单一分类任务解释性不强问题,在二分类及三分类的电力系统数据集上验证本文方法的有效性,实验结果表明本文方法具有更好的入侵检测识别效果。

关键词: 入侵检测; 工业控制系统; 机器学习; 联邦学习; 数据不平衡
中图分类号: TP181 **文献标志码**: A **文章编号**: 2095-2163(2025)03-0079-08

Intrusion detection in industrial control system based on weighted federated learning and neural network

YANG Chi[1], YU Guiqi[2], ZHANG Jianjun[2], PENG Bo[2], JIA Huihui[1,3]
(1 The Third Research Institute of Ministry of Public Security, Shanghai 200031, China;
2 Shanghai Municipal Public Security Bureau Network Security Protection Corps, Shanghai 201799, China;
3 Shanghai Engineering Research Center of Cyber and Information Security Evaluation (The Third Research Institute of Ministry of Public Security), Shanghai 200031, China)

Abstract: Intrusion attacks cause blockage and disruption to normal industrial production processes, machine learning can realize the classification and recognition of intrusion detection and then intervene, but the privacy and security barriers between various data holders in the industrial control system can not integrate and utilize the data. In order to break the barrier and obtain a better classification effect of intrusion detection, a federated learning and neural network intrusion detection method for industrial control systems is proposed. It uses multi-layer neural networks to achieve intrusion detection classification. Through federated learning, the data of each holder is kept locally, only model and parameter information are transmitted to break the data barrier. Aiming at the unbalanced sample size of each data holder, a weight factor is introduced in the initial stage of federated learning to reduce the impact of unbalanced data. In order to compensate for the weak interpretation of a single classification task under multi-dimensional data, the effectiveness of this method is verified on binary and three classification power system data sets. The experimental results show that this method has better intrusion detection and recognition effect.

Key words: intrusion detection; industrial control system; machine learning; federated learning; data imbalance

0 引 言

工业控制系统[1]在架构和设计上变得越来越复杂,所使用的监控和数据采集系统跨越多个通信协议和物理接口,使得系统之间更加互联。而从远程位置收集数据的方法,会导致硬件和软件存在潜

基金项目: 科技创新 2030—"新一代人工智能"重大项目(2020AAA0109300)。
作者简介: 杨 驰(1997—),男,硕士,主要研究方向:网络与信息安全,联邦学习; 张建军(1985—),男,硕士,主要研究方向:网络与信息安全; 彭 博(1973—),男,学士,主要研究方向:网络与信息安全; 贾徽徽(1986—),男,硕士,助理研究员,主要研究方向:网络与信息安全,Web 安全,安全测评和认证。
通信作者: 俞贵琪(1988—),男,学士,主要研究方向:网络与信息安全。Email:yuguiqi@163.com。
收稿日期: 2023-10-28

在缺陷,使其更容易遭受攻击者入侵。目前,工业环境下电力系统的工作人员严重依赖于专家经验来采取相关行动,但当电力系统在面临网络攻击的情况下,且被伪装和自然事件相仿时,仅依靠人工判断存在着相当大的难度。

机器学习或者深度学习方法作为一种鉴别电力系统干扰的手段目前已经取得可观进展[2],Hink 等学者[3]用 RandomForest 等机器学习算法对鉴别电力系统干扰做了研究和评估,利用电力系统测量数据之间复杂关系,能够准确区分恶意、非恶意和自然干扰等事件。有研究人员针对网络入侵检测采用了 XGBoost[4]、逻辑回归[5]和神经网络算法[6]等方法在入侵检测方面进行了研究验证,另外 Gao 等学者[7]和 Alhowaide 等学者[8]证明集成模型在入侵检测领域的可行性。对于深度学习在入侵检测方面的探索,Muna 等学者[9]用深度学习方法进行无监督学习,从而降低了入侵检测数据维度。贾凡等学者[10]利用卷积神经网络提高了入侵检测识别的分类准确率。

然而工业控制系统对数据隐私的保护使机器学习或者深度学习方法面临数据孤岛难题。Mcmahan 等学者[11]提出联邦学习方法使设备能够以一种分布式方法进行学习,从而不再需要与集中式方法一样共享数据,减少了数据在网络传输当中被泄露的可能。Nguyen 等学者[12]提出结合联邦学习的一种异常检测分布式系统,将联邦学习应用到了异常检测领域,证明了联邦学习在检测领域的可行性。最近,王蓉等学者[13]不仅利用卷积神经网络提升了入侵检测的准确率,而且还利用联邦学习区别于一般入侵检测模型的特点保证了数据隐私安全,但缺少对不平衡数据和客户端的选择处理。因此,本文提出了基于联邦学习和神经网络的工控入侵检测方法和客户端数据不平衡处理算法。使用 *MSE* 损失函数解决数据类别不平衡,针对工业环境下各数据拥有者因隐私安全而导致的数据孤岛问题,提出加权联邦学习算法实现不平衡数据终端的安全共享。最后面对工业环境下复杂多维的数据,二分类任务不能涵盖所有工程问题,本文利用所提出的算法在二分类和三分类数据集上进行实验,更加有效地证明了所提算法的实用性。

1 联邦学习基本思想

联邦学习[11]是一个分布式机器学习框架,可使客户端用户能够在本地数据的虚拟集合上协作训练出一个共享的全局模型,而无需将数据从本地环境中移出。中央服务器负责协调由多个训练轮回组成的联邦学习过程,服务器在每轮训练开始时都将当前的全局模型传递给参与用户,每个客户端用户利用本地数据集训练本地模型,并仅将训练后的模型信息发送给中央服务器。服务器在接收到所有客户端发送的更新后,就会更新全局模型。联邦学习相比集中化学习克服了隐私和通信挑战,不再需要将所有数据集进行聚合,只利用本地的机器学习或者深度学习模型就能学习到分散的数据。

一个联邦学习框架结构如图 1 所示。图 1 中,在服务器进行模型初始化后,向客户端分发初始化模型,多个客户端用户通过本地数据集训练本地模型,且经由服务器聚合模型后再更新各客户端用户的本地模型,从而获得一个更优的全局模型。

图 1 联邦学习框架

Fig. 1 Federated learning framework

2 提出的入侵检测算法

本文基于工业控制系统设计的联邦学习入侵检测方法,可实现分布式环境下的联邦学习入侵检测,其方法流程如图 2 所示。图 2 中,从多数据来源的工业系统中获得数据,并将数据分别进行数据预处理,然后把预处理后的数据输入进各数据来源地的本地模型中,本地模型训练通过联邦学习方式,将训练的模型参数进行上传和下载,最后可根据训练好的中央服务器聚合模型来对新的工业入侵数据进行分类检测。

2.1 数据采集

实验用到的电力系统攻击数据[3]由密西西比州立大学提供,该数据集由模拟复杂电子设备和监控系统交互的电力系统产生。基于网络的智能电力系统结构如图 3 所示。4 个智能电子继电器 $IED_1 \sim IED_4$ 分别控制断路器 $R_1 \sim R_4$,变电站网络负责控制继电器和传输信息,监控系统不仅记录了 4 个继电器的各工作物理信息,并记录了相应继电器的面板、警报和继电器日志信息。对于该电力系统除了正常的工作

断电事件外，由于变电站通过网络连接，攻击者可依靠网络来注入攻击命令并伪装使系统断电，以此扰乱工作人员进而无法辨别断电事件被攻击状态，这扰乱了系统的正常运行，且造成工业数据记录的不真实性。

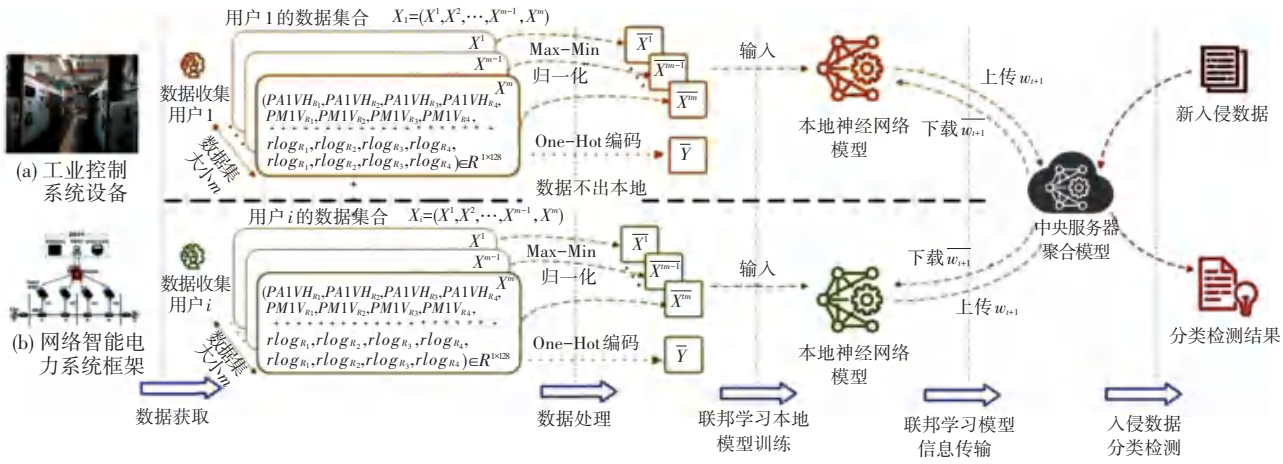


图 2　工业控制系统联邦学习入侵检测流程图

Fig. 2　Flowchart of federated learning intrusion detection for industrial control system

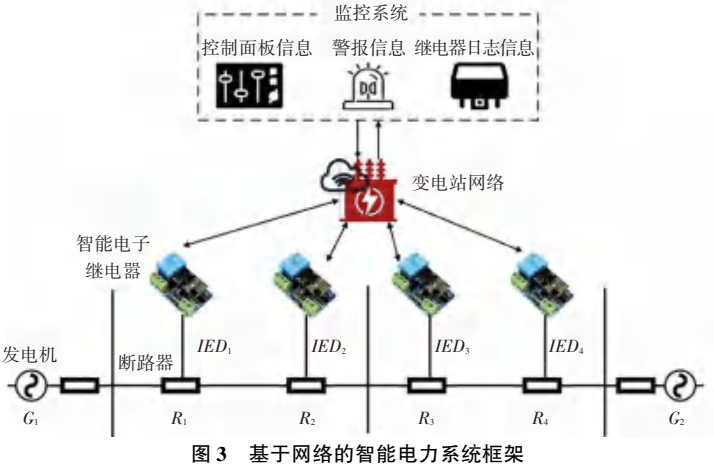


图 3　基于网络的智能电力系统框架

Fig. 3　Network-based smart power system framework

数据集包含 37 种电力系统事件场景，其中二分类数据集是将 37 种事件场景归类为攻击场景（28 种）和正常事件（9 种）；三分类数据集是按照自然事件、无事件和攻击事件的场景进行分类。每条数据样本包含 128 个特征信息，其中 116 个特征信息由每 29 个测量类型的 4 个继电器单元组成，另外 12 个特征信息记录了 4 个断路器单元对应的控制面板信息、警报信息和继电器日志信息。二分类及三分类数据集的样本标签比例见表 1。

表 1　电力系统数据集二分类和三分类的标签分布

Table 1　Distribution of labels for dichotomous and trichotomous classification of power system datasets　%

数据类型	Attack	Normal	
		Natural	NoEvents
二分类	71.0	28.9	28.9
三分类	71.0	23.3	5.6

2.2　问题设定和数据处理

实验中给定一个电力系统中各数据收集用户的本地数据集集合（$i \in N$）。这里，N 为系统中不同本地持有数据集的用户个数。其中，单个用户的本地数据表示为 $X_{i\in N}^{m} = \{X_i^1, X_i^2, \cdots, X_i^{m-1}, X_i^m\}$，这里 m 为样本数量。第 i 个用户的第 j 条数据可表示为 $\boldsymbol{X}_i^{j\in m} = (PA1VH_{R1}^j, PA1VH_{R2}^j, PA1VH_{R3}^j, PA1VH_{R4}^j, PM1V_{R1}^j, PM1V_{R2}^j, PM1V_{R3}^j, PM1V_{R4}^j, \cdots, rlog_{R1}^j, rlog_{R2}^j, rlog_{R3}^j, rlog_{R4}^j, slog_{R1}^j, slog_{R2}^j, slog_{R3}^j, slog_{R4}^j) \in R^{1\times 128}$，研究的目标是检测出当前数据的事件类型 $\hat{Y}$。

电力系统数据集中各个特征具有不同的取值域，为了使训练效果更可靠，需要将特征数据归一化在一定范围内，这样能够消除不同维度数据之间的异同，减少奇异样本数据导致的影响，从而提升模型的收敛速度。本文使用最小最大值归一化方法，将数据 X_i^m 的特征值转化在 $[-1,1]$ 之间，如下所示：

$$\bar{X}_i^m = 2\frac{X_i^m - \min}{\max - \min} - 1 \tag{1}$$

其中，X 表示原始数据；max 表示样本最大值；min 表示样本最小值；$\bar{X}_i^m$ 表示归一化结果。

为了尽可能保留电力系统数据集类别数值大小的判别信息，本文选择对数据集标签 Y 进行 One-Hot 编码，用到的公式如下：

$$\bar{Y} \overset{\text{One-Hot}}{\leftarrow} Y \tag{2}$$

由式(2)可知，对多个寄存器状态类型加以编码，使得类别变量的每一个类型都对应着一个寄存器的位。得到的电力系统数据集二分类及三分类事件类型 One-Hot 编码结果分别见表 2、表 3。

表 2 二分类标签进行 One-Hot 编码

Table 2 Dichotomous labels for One-Hot coding

类别标签	One-Hot
Natural	(0,1)
Attack	(1,0)

表 3 三分类标签进行 One-Hot 编码

Table 3 Triple categorized labels for One-Hot coding

类别标签	One-Hot
Natural	(0,0,1)
NoEvents	(0,1,0)
Attack	(1,0,0)

2.3 网络结构设计

由于神经网络的输入、输出层的神经元数量由数据集的特征信息和类别数量决定，且输入层与隐含层存在等量关系，所以将神经元数量选择为输入层的 2 倍。当前工业控制系统数据趋于更加复杂多维，单隐层神经网络的特征表达能力有限[14]，其拟合效果不好的问题已不再适用于当前背景下网络结构的选择。为此，本文采用多隐层网络结构，设置 4 个隐含层能更有效处理复杂数据。其中，本文隐藏层和输出层分别采用 ReLU 激活函数和 Softmax 损失函数。

在训练神经网络时，每个类的样本数量具有差异化，类别样本数量少、其准确性将会变得非常低。重采样方法是解决样本不平衡的对策方法之一，但却并不能解决不平衡数据集的根本问题，因此为了保证在有限的样本数量下进行学习的有效性，Kato 等学者[15]指出在类不平衡数据集的学习上，均方误差(Mean Squared Error，*MSE*)损失函数优于在传统分类问题研究上使用的交叉熵(Cross Entropy，*CE*)损失函数，且通过实验证明在类不平衡数据集的学习上，*MSE* 损失优于 *CE* 损失，*MSE* 损失可以使所有类的反向传播的数量相等，并兼顾到类与类之间的关系来学习特征空间，是有效处理类不平衡数据集的方法。Zhou 等学者[16]和 Hui 等学者[17]也同样证明了用均方误差损失函数训练深度神经网络与传统的交叉熵损失函数相比能达到相同、甚至更好的效果。于是本文根据工控系统环境下产生不平衡数据的特点，选择采用 *MSE* 损失函数解决数据不平衡问题。*MSE* 的数学公式具体如下：

$$L_{\text{mse}} = \frac{1}{2}(\hat{Y} - \bar{Y})^2 \tag{3}$$

其中，$\hat{Y}$ 表示模型输出，$\bar{Y}$ 表示编码后的真实标签。

2.4 加权联邦学习

针对工业控制系统，本文采用联邦学习方法实现多用户的分布式学习。首先在每轮联邦学习训练开始前，为了避免了计算资源的浪费，降低隐私泄露的风险，对 N 个训练用户进行随机选择训练，选择出 $K(K \leqslant N)$ 个参与训练的用户集合 K_t。与联邦平均算法(FedAvg)[11]不同的是，本文在此阶段引入权重系数 p_i，计算公式所下所示：

$$p_i = \frac{m_i}{\sum_{n=1}^{N} p_n \times \sum_{n=1}^{N} m_n} \tag{4}$$

每个用户 $i \in N$ 根据持有数据集 m_i 大小得到归一化的概率系数 p_i。加权联邦学习客户端选择方法如图 4 所示，每轮通过概率系数 p_i 来进行 K 次选择，选择出用户集合 K_t 进行训练，这样避免了对于数据量少的客户端在等概率被随机选择时训练出的欠拟合本地模型最终影响到全局模型，通过降低数据量少的客户端的被选择次数能够更好地使数据量多的客户端参与训练，以此获得更鲁棒的模型。

其次，在联邦学习本地训练阶段，根据选择出的训练用户 k，将本地数据 X_k 作为输入，并进行批处理得到批数据 b，输入进设计好的神经网络模型中，结合从服务器接收到的模型参数，计算出批量梯度 g_k^b，最后得到本地更新模型参数 w，具体公式为：

$$w \leftarrow w - \eta g_k^b \tag{5}$$

其中，η 表示学习率。这样训练好的本地更新模型可以根据如下公式将更新模型 w_{t+1}^k 发送给中央服务器：

$$w_{t+1}^k \leftarrow w_t^k \tag{6}$$

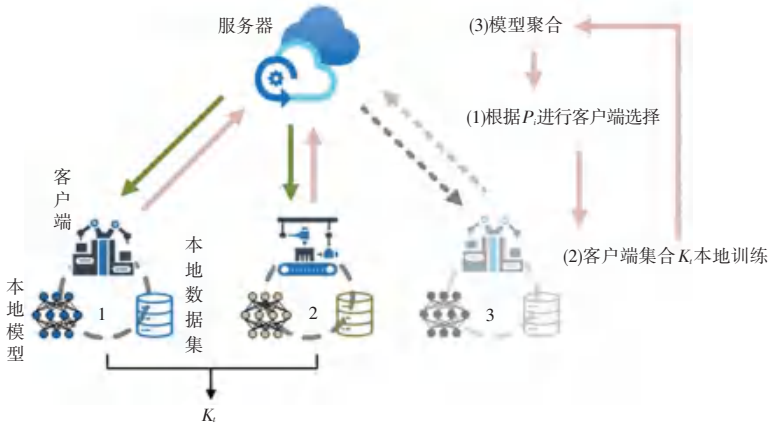


图 4　加权联邦学习客户端选择方法

Fig. 4　Weighted federated learning client selection method

通过这种方式不需要共享各自的数据集，而是通过训练本地模型就可以将模型更新 w_{t+1}^k 发送给中央服务器进行聚合得到全局模型 $\bar{w}_{t+1}$，可由如下公式表示为：

$$\bar{w}_{t+1} \leftarrow \sum_{i=1}^{N} p_i w_{t+1}^n \tag{7}$$

本文提出的加权联邦学习方法的代码描述如下，其中展示了完整的训练流程。

算法 1　加权联邦学习方法

初始参数：K_t 是实际每轮参与训练的客户端集合；E 是本地迭代的次数；T 是全局通信轮次

输入　各用户 i 的本地数据集 $\bar{X}_i^m$

输出　全局模型参数 $\bar{w}_{t+1}$

1. 服务器端：
2. for　全局通信更新轮次 $t = 1,2,\ldots,T$ do
3.　　for　选择的客户端 $j = 1,2,\ldots,K$ do
4.　　生成[0,1)的随机数 r
5.　　初始化累积权重值 $p = 0$
6.　　for 遍历每个客户端 $i = 1,2,\cdots,N$ do
7.　　P +=P_i
8.　　if $r < p$:
9.　　选择当前客户端进入当前训练集合 K_t 中
10.　for 每个参与方 $k \in K_t$ 并行 do
11.　　根据式(6)更新本地模型参数
12.　　发送 w_{t+1}^k 给服务器
13.　　服务器根据式(7)进行加权平均
14. 客户端：
15. for 每个本地迭代 $e = 1,2,\cdots,E$ do
16.　　for 每个批数据 do
17.　　根据式(5)本地更新模型参数

首先，对各客户端的数据量权重使用式(4)进行归一化以符合概率分布，得到 p_i。然后，开始进行 K 次选择，每次选择一个客户端。对于每次选择，生成一个随机数 r，该随机数位于区间 [0,1) 内，根据归一化的权重 p_i 来选择一个选项，在对经过初始化的权重 P 进行累加后，通过判断累加权重值 P 与随机数 r 的大小来选择当前客户端进入当前训练集合 K_t 中，当累积权重值大于等于随机数时，选择该客户端。在获得当前轮次的训练客户端集合后，服务器负责通知该客户端进行并行训练，各被选择客户端根据式(6)及本地数据集 $\bar{X}_i^m$ 更新本地模型参数，并发送 w_{t+1}^k 给服务器，服务器在接收到参数后根据式(7)进行加权平均，直至输出最优的全局模型参数 $\bar{w}_{t+1}$。

3　实验验证

实验环境为移动工作站：配置为 Inter Core i7 11850H 处理器，2.5 GHz 主频，64 GB 内存，Windows 10 操作系统，框架搭建基于 Keras 实现。

3.1　实验评估指标

本文二分类和三分类方法都使用准确率(*Accuracy*)、精确率(*Precision*)、检测率(*Recall*)、综合查全率(*F-Measure*)这 4 个指标对模型进行评估，其计算公式如下：

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN} \tag{8}$$

$$Precision = \frac{TP}{TP + FP} \tag{9}$$

$$Recall = \frac{TP}{TP + FN} \tag{10}$$

$$F\text{-}measure = \frac{2 \times Precision \times Recall}{Precision + Recall} \quad (11)$$

其中，*TP*、*TN*、*FP* 和 *FN* 分别表示真正例、真反例、假正例和假反例。

3.2 实验结果

为了验证本文使用的方法有效性，本文分别使用二分类数据集和三分类数据集进行了广泛实验。

3.2.1 单个用户数据集实验结果

该实验是选取二分类数据集中划分的单个用户数据集，分别选择在电力系统数据上实验的方法 RandomForest[3]，和同样在入侵检测方面获得成功的 XGBoost 方法[4]、LogisticRegression 方法[5]、CNN 方法[10]进行了实验，并用本文设计的单隐层神经网络和多隐层神经网络做了实验对比。实验结果见表 4。由实验结果可知，单隐层的神经网络并不能获得很好的训练效果，多隐层拟合函数的能力更强。且通过和其他方法的对比，本文设计的多隐层神经网络在电力系统数据集的表现上其准确率优于其他方法。

表 4 二分类数据集各模型实验训练效果评估

Table 4 Evaluation of the experimental training effect of each model in the dichotomous dataset

模型	*Accuracy*	*Precision*	*Recall*	*F1-Score*
RandomForest[3]	0.913 5	0.947 2	0.939 8	0.943 5
XGBoost[4]	0.847 1	0.849 3	0.973 8	0.907 3
LogisticRegression[5]	0.778 7	0.785 7	0.979 1	0.871 8
CNN[10]	0.885 3	0.843 1	0.828 1	0.835 2
ANN-单隐层	0.869 2	0.811 8	0.842 0	0.824 9
ANN-多隐层	**0.917 5**	0.892 6	0.828 1	0.835 2

对于三分类数据集，也用同样方式与其他方法进行了实验对比。通过实验结果进一步说明本文使用的多层网络模型要优于单层。三分类数据集各模型实验训练效果评估见表 5。由表 5 可知，在检测率上本文的方法要优于其他方法。

表 5 三分类数据集各模型实验训练效果评估

Table 5 Evaluation of the experimental training effect of each model in the triple categorized dataset

模型	*Accuracy*	*Precision*	*Recall*	*F1-Score*
RandomForest[3]	0.927 6	0.891 7	0.885 6	0.888 6
XGBoost[4]	0.885 3	0.892 8	0.782 8	0.824 9
LogisticRegression[5]	0.776 7	0.453 4	0.377 3	0.374 5
CNN[10]	0.859 6	0.870 1	0.658 0	0.721 6
ANN-单隐层	0.871 2	0.879 5	0.871 2	0.721 6
ANN-多隐层	0.905 4	0.883 7	**0.908 4**	0.791 5

3.2.2 联邦学习场景与非联邦学习场景实验对比

该实验选择电力系统数据按照数据不平衡的方式划分成分布式场景下的 4 个参与用户的本地数据集。实验分别用传统方式独立地训练 4 个数据集，在二分类和三分类数据场景分别评估各个数据集的训练效果见表 6，再用联邦学习的方式进行分布式的训练。在联邦平均算法(FedAvg)[11]和本文的改进方法的联邦学习设置中，每轮参与训练的客户端为随机选择的 2 个用户，模型都选择本文设计的神经网络模型。而 FC 方法[13]是设置所有用户都参与训练，且模型为本文设计的卷积神经网络模型。最后评估全局模型的评估结果，其二分类及三分类的实验数据分别见表 6、表 7。本文的方法在二分类数据集上对于 4 项评价指标均高于其他 2 种联邦学习方法及各用户独立训练，其准确率能达到 93.11%。在三分类数据集上，本文方法在 3 项评价指标上高于其他 2 种联邦学习方法及各用户独立训练，其准确率达到 93.96%。

表 6 4 个用户独立训练和联邦场景训练二分类数据集结果

Table 6 Results of four user-independent and federated scenario training dichotomous dataset

模型	*Accuracy*	*Precision*	*Recall*	*F1-Score*
Party 1	0.917 5	0.892 6	0.828 1	0.835 2
Party 2	0.871 2	0.881 2	0.957 0	0.917 5
Party 3	0.925 2	0.933 9	0.964 0	0.948 7
Party 4	0.896 4	0.909 9	0.930 5	0.820 1
FC[13]	0.863 2	0.800 9	0.804 8	0.802 8
FedAvg[11]	0.929 1	0.942 2	0.969 0	0.954 3
本文方法	**0.931 1**	**0.944 5**	**0.969 0**	**0.956 6**

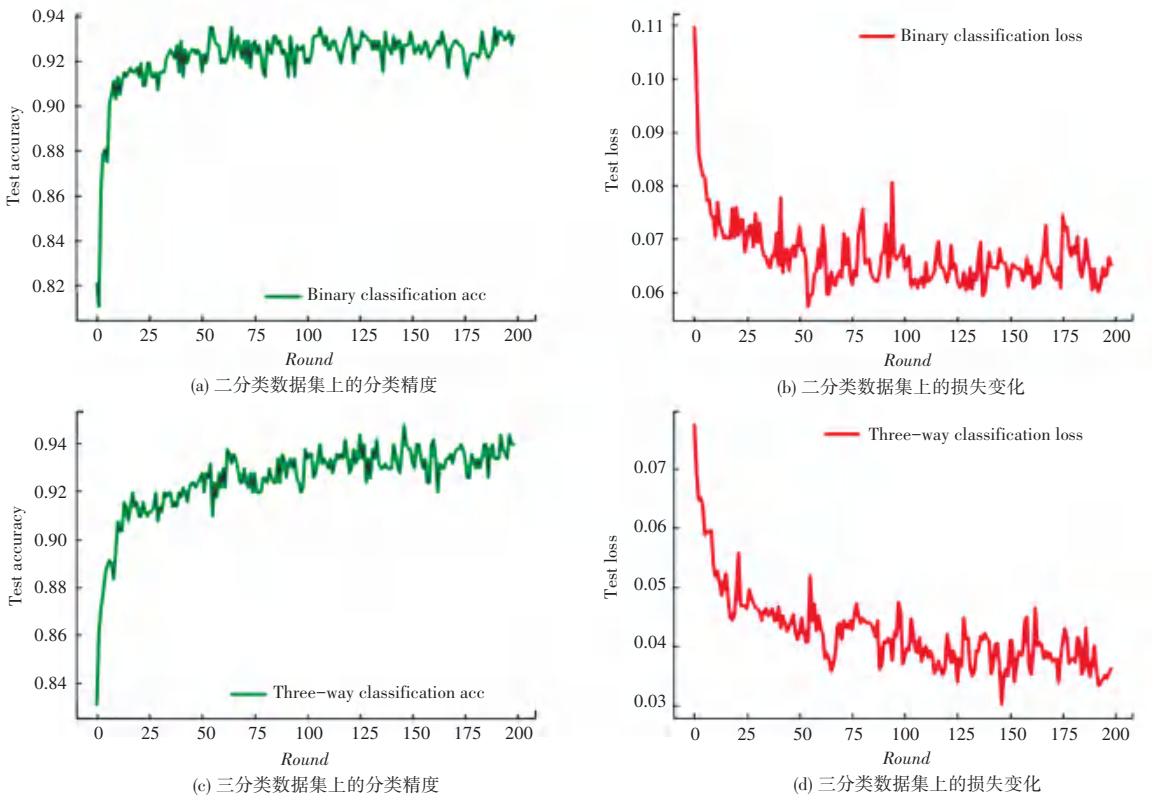
表 7 4 个用户独立训练和联邦场景训练三分类数据集结果

Table 7 Results of four user-independent and federated scenario training triple categorized dataset

模型	*Accuracy*	*Precision*	*Recall*	*F1-Score*
Party 1	0.863 2	0.867 8	0.865 2	0.865 5
Party 2	0.859 1	0.866 6	0.861 2	0.861 2
Party 3	0.916 9	0.903 7	0.921 5	0.890 7
Party 4	0.898 2	0.906 4	0.904 9	0.905 2
FC[13]	0.873 2	0.851 3	0.710 7	0.763 9
FedAvg[11]	0.921 5	0.899 9	0.909 2	0.904 1
本文方法	**0.939 6**	0.896 3	**0.925 1**	**0.908 1**

本文方法在 2 种数据集上训练 200 轮的测试结果如图 5 所示，从 2 种数据集的准确率上升变化中可以看出，初始阶段的准确率较低。但随着训练过

程的进行，准确率逐渐上升直至稳定在一个相对高的水平。这表明模型在训练过程中逐渐学习到了数据的特征，从而提高了预测的准确性。在训练过程后期，数据中的准确率趋于稳定，不再有大幅度的变化，表明模型已经收敛到一个相对最佳的状态。从2种数据集的损失下降变化中可以看出，随着训练的轮次增加，损失值逐渐减小，模型在逐渐适应训练数据，优化了模型的权重和偏差。其训练结果表明通过该方法进行训练验证后是可行的。



(a) 二分类数据集上的分类精度　(b) 二分类数据集上的损失变化

(c) 三分类数据集上的分类精度　(d) 三分类数据集上的损失变化

图 5　加权联邦学习下的入侵检测分类精度和损失变化

Fig. 5　Intrusion detection classification accuracy and loss variation under weighted federated learning

4　结束语

为了提升复杂工业控制系统在网络环境下的入侵检测能力和保证本地数据的安全性，本文提出了基于神经网络的联邦学习入侵检测方法，通过在不同分类问题的数据集上进行的实验证明本文提出的方法在工业控制系统背景下是有效的。本文注意到在工业控制系统背景下数据类别的不平衡性和联邦学习中用户持少量数据而影响全局模型的问题，分别应用 *MSE* 损失函数和在联邦学习训练前引入权重系数来进行改善，实验结果进一步证明所提出方法对模型精度有所提升，并对当今复杂工业控制系统进行入侵检测具有重要和实际意义。

最后，由于本文使用的方法和联邦平均算法都采用同步训练和更新的方式，其优点是模型精度高，收敛速度快。但客户端与聚合服务器之间进行一轮通信的持续时间受到最慢的参与设备的严格限制，在具有异构设备计算能力的工业环境中，对联邦学习过程的完成时间有重大影响。反之，异步联邦学习[18-19]在这方面具有自然优势，但也存在模型质量下降和服务器崩溃的风险。所以对异步联邦学习的进一步优化是今后工作的重点，研究中也同样关注到一种将同步与异步两种方式的优势相结合的半异步联邦学习框架[20-21]。因此未来在异步联邦学习方向的突破也是将联邦学习应用到工业控制系统当中必须解决的问题，及最终适应大范围联邦学习采用同步、异步还是半异步仍需要进一步的研究和探索。

参考文献

[1] 方栋梁，刘圃卓，秦川，等. 工业控制系统协议安全综述[J]. 计算机研究与发展，2022，59(5)：978-993.

[2] SADEGHI S，HESAMI N A，MORADI P，et al. Introduction and literature review of the application of machine learning/deep learning to control problems of power systems[M]//NAZARI-

HERIS M, ASADI S, MOHAMMADI - IVATLOO B, et al. Application of Machine Learning and Deep Learning Methods to Power System Problems. Power Systems. Cham:Springer, 2021: 119-135.
[3] HINK R C B, BEAVER J M, BUCKNER M A, et al. Machine learning for power system disturbance and cyber - attack discrimination[C]//2014 7th International Symposium on Resilient Control Systems (ISRCS). Piscataway,NJ:IEEE, 2014: 1-8.
[4] 张阳,姚原岗. 基于 Xgboost 算法的网络入侵检测研究[J]. 信息网络安全,2018(9):102-105.
[5] 金志刚, 苏菲. 基于 FSVM 与多类逻辑回归的两级入侵检测模型[J]. 南开大学学报(自然科学版), 2018, 51(3):1-6.
[6] 罗予东, 陆璐. 基于人工神经网络和遗传算法的网络攻击检测[J]. 计算机工程与设计, 2021, 42(9): 2446-2454.
[7] GAO Xianwei, SHAN Chun, HU Changzhen, et al. An adaptive ensemble machine learning model for intrusion detection[J]. IEEE Access, 2019, 7: 82512-82521.
[8] ALHOWAIDE A, ALSMADI I, TANG Jian. Ensemble detection model for IoT IDS[J]. Internet of Things, 2021, 16: 100435.
[9] MUNA A L H, MOUSTAFA N, SITNIKOVA E. Identification of malicious activities in industrial internet of things based on deep learning models [J]. Journal of Information Security and Applications, 2018,41: 1-11.
[10]贾凡,孔令智. 基于卷积神经网络的入侵检测算法[J]. 北京理工大学学报,2017,37(12):1271-1275.
[11]MCMAHAN B, MOORE E, RAMAGE D, et al. Communication- efficient learning of deep networks from decentralized data[J]. arXiv preprint arXiv,1602.05629v3,2023.
[12]NGUYEN T D, MARCHAL S, MEITTINEN M, et al. DÏoT: A federated self-learning anomaly detect-ion system for IoT[C]// 2019 IEEE 39th International Conference on Distributed Computing Systems (ICDCS). Piscataway,NJ:IEEE,2019: 756-767.
[13]王蓉,马春光,武朋. 基于联邦学习和卷积神经网络的入侵检测方法[J]. 信息网络安全,2020,20(4):47-54.
[14]席磊,何苗,周博奇,等. 基于改进多隐层极限学习机的电网虚假数据注入攻击检测[J]. 自动化学报,2022,49(4): 881-890.
[15] KATO S, HOTTA K. MSE loss with outlying label for imbalanced classification[J]. arXiv preprint arXiv,2107.02393, 2021.
[16]ZHOU Jinxin, LI Xiao, DING Tianyu, et al. On the Optimization landscape of neural collapse under MSE loss: Global optimality with unconstrained features[J]. arXiv preprint arXiv, 2203.01238, 2022.
[17] HUI L, BELKIN M. Evaluation of neural architectures trained with square loss vs cross-entrop-y in classification tasks[J]. arXiv preprint arXiv,2006.07322, 2020.
[18] ZHANG Yu, DUAN Moning, LIU Duo, et al. CSAFL: A clustered semi-asynchronous federated learning framework[J]. arXiv preprint arXiv, 2104.08184,2021.
[19]芦效峰,廖钰盈,LIO P, 等. 一种面向边缘计算的高效异步联邦学习机制[J]. 计算机研究与发展, 2020, 57(12): 2571-2582.
[20] SPRAGUE M R, JALALIRAD A, SCAVUZZO M, et al. Asynchronous federated learning for geospatial applications[C]// Joint European Conference on Machine Learning and Knowledge Discovery in Databases. Cham:Springer, 2018: 21-28.
[21]WU Wentai, HE Ligang, LIN Weiwei, et al. SAFA: A semi-asynchronous protocol for fast federated learning with low overhead [J]. IEEE Transactions on Computers, 2020, 70(5): 655-668.