基于联邦强化学习的面向边缘网络的入侵检测方法研究

丁凯[1,2]，黄宜都[1]，陶铭[1]，谢仁平[1]
（1. 东莞理工学院计算机科学与技术学院，广东 东莞 523808；
2. 人工智能与数字经济广东省实验室（深圳），广东 深圳 518107）

摘　要：随着物联网（IoT, Internet of things）设备的迅速普及，针对 IoT 设备的攻击频率和强度不断上升，因而持续更新安全机制以保障物联网设备的安全显得尤为重要。然而，随着公众隐私意识的增强，越来越多的数据集不再对外共享，形成数据“孤岛”现象，阻碍了物联网安全防护能力的提升。为了解决这一问题，提出了一种基于联邦强化学习的入侵检测方法，并通过医疗物联网（IoMT, Internet of medical things）和车联网（IoV, Internet of vehicles）场景下的两个数据集进行实验验证。为模拟真实环境，在每个边缘代理中设计了不平衡的流量样本分布，进而评估全局模型的检测精度和鲁棒性。采用双深度 Q 网络（DDQN, double deep Q-network）为边缘代理的强化学习框架，并通过准确率、精确率、召回率和 F1 分数对实验结果进行评估。实验结果表明，提出的方法具有良好的鲁棒性和检测精度。

关键词：联邦强化学习；入侵检测；物联网安全；物联网
中图分类号：TP301.6
文献标志码：A
doi：10.11959/j.issn.2096−3750.2024.00442

Research on intrusion detection method for edge networks based on federated reinforcement learning

DING Kai[1,2], HUANG Yidu[1], TAO Ming[1], XIE Renping[1]
1. School of Computer Science and Technology, Dongguan University of Technology, Dongguan 523808, China
2. Guangdong Laboratory of Artificial Intelligence and Digital Economy (Shenzhen), Shenzhen 518107, China

Abstract: With the rapid proliferation of Internet of things (IoT) devices, the frequency and intensity of attacks targeting these devices are constantly increasing. Therefore, it’s quite important that security mechanisms are continuously updated to ensure the safety of IoT devices. However, as public awareness of privacy grows, many datasets are no longer shared, leading to the emergence of data silos, which hinders the improvement of IoT security. To address this issue, a federated reinforcement learning-based intrusion detection method was proposed, and experiments were conducted using two datasets from the Internet of medical things (IoMT) and Internet of vehicles (IoV) scenarios. Imbalanced traffic sample distributions were designed for each edge agent to simulate a real-world environment, allowing for the evaluation of the detection accuracy and robustness of the global model. Double deep Q-network (DDQN) was employed as the reinforcement learning framework for the edge agents, and the experimental results were evaluated using accuracy, precision, recall, and F1-score. The results demonstrate that the proposed method exhibits strong robustness and detection accuracy.

Key words: federated reinforcement learning, intrusion detection, IoT security, IoT

收稿日期：2024−10−14；**修回日期：**2024−11−25
通信作者：陶铭，taom@dgut.edu.cn
基金项目：国家自然科学基金资助项目（No. 62001113）；广东省基础与应用基础研究基金资助项目（No. 2021A1515010656）；广东省人工智能与数字经济实验室（深圳）开放研究基金资助项目（No. GML-KF-22-02）；东莞市社会发展科技项目（No. 20221800902472）
Foundation Items: The National Natural Science Foundation of China (No. 62001113), The Basic and Applied Basic Research Funding Program of Guangdong Province (No. 2021A1515010656), The Open Research Fund from Guangdong Laboratory of Artificial Intelligence and Digital Economy (No. GML-KF-22-02), The Social Development Technology Program of Dongguan City (No. 20221800902472)

0　引言

入侵检测系统（IDS, intrusion detection system）与物联网（IoT, Internet of things）设备密切相关，在保障物联网安全方面起着至关重要的作用[1-3]。目前物联网设备广泛应用于各个领域，如智能家居、工业控制、智慧城市和健康监测。物联网设备的计算能力有限且安全措施不足，容易受各种网络攻击的威胁，包括拒绝服务（DoS, denial of service）攻击、分布式拒绝服务（DDoS, distributed denial of service）攻击、恶意软件感染以及数据窃取等。IDS可以通过监控和分析网络流量、系统日志和用户行为来识别和应对潜在的安全威胁和攻击[4-5]。具体来说，IDS能够检测异常的网络活动和未经授权的访问请求，识别已知的攻击模式和未知的威胁，并生成警报通知系统管理员采取必要的措施应对安全事件，从而降低物联网设备被攻陷的风险，但是在实际环境中部署和应用IDS面临独特的挑战和要求。例如，由于物联网设备资源受限，IDS需要在低资源消耗的情况下高效运行。此外，物联网设备的异构性和复杂性要求IDS具备灵活的适应性，能够处理各种类型的设备和通信协议。随着物联网设备数量的增加，IDS还需要具有可扩展性，能够处理大量的数据流和安全事件[6-7]。

当前部署的物联网设备在协议和类型上存在显著差异，针对所有物联网设备更新统一的安全机制或相关硬件和软件的成本高昂[8]。在物联网设备通过无线或有线网络进行日常通信的过程中，其通信行为可能由于工作环境、任务或工作时间的不同而显著变化。这些差异可以作为特征被安全系统利用来判断设备的状态，并提前进行更换或升级。目前许多创新的理念已经付诸实践，包括集中式学习方法，如机器学习和深度学习。

在机器学习方法中，遗传算法（GA, genetic algorithm）和K最近邻（KNN, K-nearest neighbor）算法是用于提取特征的常用方法[9-12]。传统的机器学习分类器，如支持向量机（SVM, support vector machine）、随机森林（RF, random forest）、决策树（DT, decision tree）和高斯朴素贝叶斯（GNB, Gaussian naive Bayes）等，通常需要专家投入大量的时间和精力进行特征选择。为了简化这一过程，深度学习方法被应用于IDS，以减少对复杂特征选择的依赖[13-15]。

在深度学习方法中，长短期记忆（LSTM, long short-term memory）网络和卷积神经网络（CNN, convolutional neural network）是当前研究中常用的网络架构。CNN的空间特征提取能力在IDS的研究中被充分验证[16-18]。NIDS-CNNLSTM[19]结合LSTM和CNN的优点，实现了更高效的特征学习和分类，其中，CNN负责从网络流量数据中学习特征，而LSTM处理时间序列数据。考虑流量特征的维度过高导致计算量太大并影响模型的性能，自编码器（AE, autoencoder）被运用到流量数据中进行特征选择或者降维。Laghrissi等[20]设计了一种基于LSTM的AE神经网络模型，实验结果表明该模型在二分类和多分类任务中均取得了较好的训练和测试准确率。LSTM-AE[21]和DS-SIoDL[22]分别利用LSTM搭建AE架构以实现特征降维和入侵检测。除了机器学习和深度学习方法，基于强化学习的方法也被应用于IDS任务中。DRL-IDS[23]方法采用LightGBM进行特征选择并使用PPO2作为强化学习算法。TA-NIDS [24]是一种基于KNN和PPO的IDS方法，能够在样本数量较少的情况下完成高精度的IDS任务。然而，无论是深度学习还是强化学习，这些工作都是在数据集中的条件下完成的，随着社会对数据隐私的日益关注，越来越多的数据集不再共享，仅在本地进行训练。虽然这种方式避免了传输过程中数据泄露的风险，但也将边缘服务器变成了数据“孤岛”，使其成为一个相对封闭的环境[25-26]。

联邦学习是解决数据“孤岛”问题的有效解决方案之一[27]。目前已有许多基于联邦学习的IDS方法。Popoola等[28]引入了一种基于联邦学习的零日攻击识别技术。He等[29]提出在联邦学习中使用CNN和LSTM作为客户端模型以执行IDS任务。HFedDI[30]在样本不平衡的条件下表现出比FedAvg更高的识别准确率。Al-Naday等[31]提出了结合联邦学习和深度Q网络（DQN, deep Q-network）的IDS方法，使用了全连接搭建神经网络模型。作为结合经验回放的学习方式，DQN可以在一定程度上减少对大量标签数据的依赖。

在上述背景下，本文选择了双深度Q网络（DDQN, double deep Q-network）作为联邦强化学习的学习方法。DDQN是强化学习领域中DQN算法的一种改进版本，它解决了传统Q学习算法中存在的过估计问题[32]。本文还引入了Transformer[33]的

网络结构和组归一化[34]到IDS任务中。前者使得边缘代理模型可以自动提取特征，无须手动进行特征提取，这避免了行为建模中特征信息的丢失；后者能够提高模型在非独立同分布数据上的表现[35]。此外，一种融合历史全局模型的方法被引入到联邦聚合算法中，在FedGKD[36]的工作中，融合了历史模型用于获取全面特征，实验表明该方法可以提高模型的收敛速度。为了全面测试本文所提方法在不同应用环境中的适应性，使用了加拿大网络安全研究所（CIC, Canadian Institute for Cybersecurity）于2024年发布的两个最新数据集进行实验。CICIoV2024[37]数据集用于模拟车联网（IoV, Internet of vehicles）环境，CICIoMT2024[38]数据集用于模拟医疗物联网（IoMT, Internet of medical things）环境，恶意程序攻击物联网设备示意图如图1所示。

本文的主要贡献总结如下。

1) 提出了一种基于联邦强化学习的IDS方法，该方法结合了联邦学习和DDQN。通过在多个边缘代理（EA, edge agent）和全局代理（GA, global agent）实现通信，使得EA能够有效学习全局知识，提高对攻击流量的识别能力，确保了数据隐私。

2) 为了传递历史GA的有效知识从而提高当前模型的性能和学习效率，提出了一种遗传联邦聚合算法，使上一轮GA的知识能够在新一轮GA中得到保留。

3) 提出了一种基于Transformer层和组归一化的模型结构，并用于构建代理的DDQN架构。该架构避免了传统机器学习方法中繁琐的特征选择过程，组归一化通过使用可训练参数来缩放和平移归一化特征，从而恢复模型的表示能力并减轻样本不平衡的影响。

1　理论基础

1.1　联邦学习

实现联邦学习涉及如下关键步骤：1) 在每个客户端上初始化模型；2) 每个客户端使用本地数据训练本地模型w_i；3) 完成本地模型训练后，从所有客户端收集最新的w_i，并将其聚合以更新全局模型w_g；4) 更新后的全局模型w_g被分发回客户端进行下一轮训练。上述过程会重复多轮$n=\{1, 2, 3, \cdots\}$，直到模型收敛或满足预定的停止标准。在更新全局模型时，使用加权平均方法进行更新，全局模型的更新式为

$$w_g = \frac{\sum_{i=1}^{I} \eta_i w_i}{\eta} \tag{1}$$

其中，η_i表示本地数据的大小，η表示所有本地数据的总量。在聚合过程中，每个EA被平等对待。进一步将聚合方法简化为

$$w_g = \sum_{i=1}^{I} \frac{w_i}{I} \tag{2}$$

本文所提联邦聚合示意图如图2所示，具体聚合算法加式(3)所示。历史遗传知识在GA中的影响由预设参数P决定，$P \in (1, 0)$。

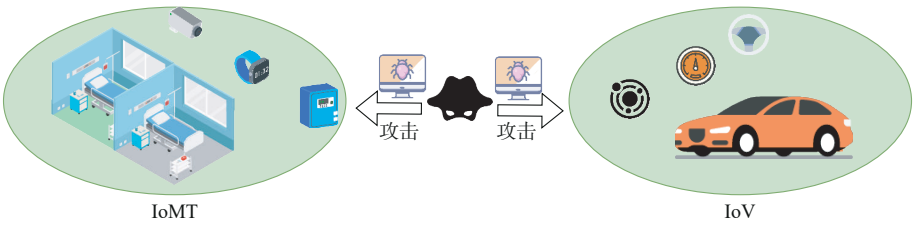


图1　恶意程序攻击物联网设备示意图

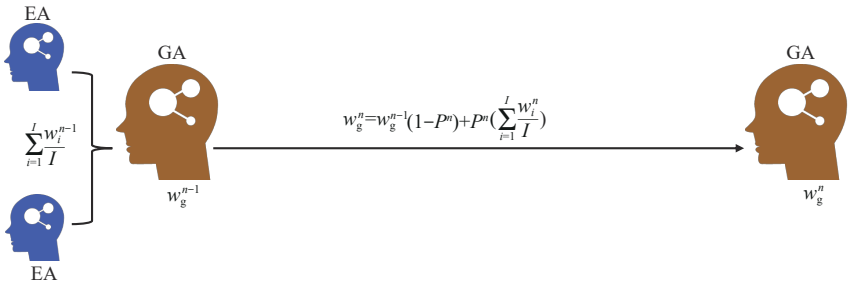


图2　本文所提联邦聚合示意图

$$w_{\mathrm{g}}^{n}=w_{\mathrm{g}}^{n-1}\left(1-P^{n}\right)+P^{n}\left(\sum_{i=1}^{I}\frac{w_{i}^{n}}{I}\right) \tag{3}$$

1.2　深度强化学习

DDQN是一种流行的强化学习方法，它使用两个神经网络Q_{current}和Q_{target}作为决策核心，DDQN如图3所示。在IDS任务中，动作是离散的，每一次动作都代表当前的流量所属的类别。

1) 状态S：数据集中的每个数据点代表代理与环境交互中的状态。设x_t为数据集中第t个样本，该样本对应在某一回合中的时间步t的状态s_t。状态序列随着每次训练回合的更新而打乱。

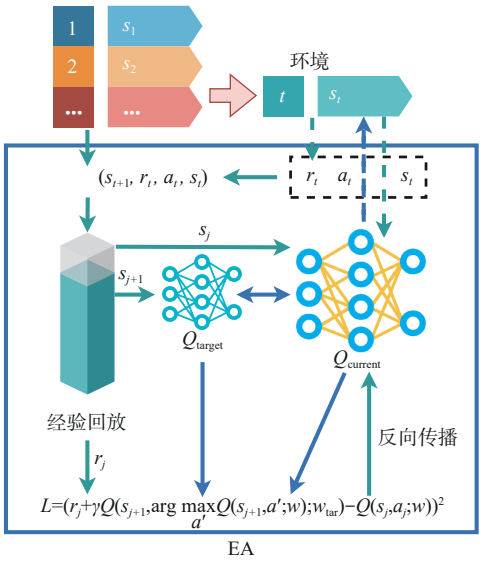


图3　DDQN

2) 动作A：代理所采取的动作a_t对应离散的分类动作，这些动作匹配的真实标签值为y_t。

3) 奖励R：当代理的动作与标签匹配时，给予正反馈作为奖励。

$$R(s_t,a_t,y_t)=\begin{cases}1, & a_t=y_t\\ -1, & a_t\neq y_t\end{cases} \tag{4}$$

4) 策略π：在环境与EA交互的过程中，每一回合会重洗样本序列以构建马尔可夫链$\{s_1, a_1, r_1, s_2, a_2, r_2, \cdots, s_t, a_t, r_t\}$，策略函数$\pi$选择具有最高$Q$值的动作$a'$。

$$\pi(s_t)=\arg\max_{a'}Q(s_t,a') \tag{5}$$

Q函数根据Bellman方程进行更新，其期望表达为

$$Q(s_t,a_t)=\mathbb{E}_{\pi}[r_t+\gamma\max_{a'}Q(s_{t+1},a')] \tag{6}$$

其中，γ是折扣因子，用于考虑未来奖励的重要性。DQN的损失函数定义为预测的Q值与目标Q值之间的均方误差

$$L(w)=\mathbb{E}_{\pi}\left[\left(r_t+\gamma\max_{a'}Q(s_{t+1},a';w)-Q(s_t,a_t;w)\right)^2\right] \tag{7}$$

在DDQN中，Q_{current}每步都会更新，而Q_{target}更新频率比Q_{current}低，是其滞后版本。这有助于解决过估计问题。DDQN的损失函数为

$$L(w)=\mathbb{E}\left[\begin{pmatrix}r_t+\gamma Q(s_{t+1},\arg\max_{a'}Q(s_{t+1},a';w);w_{\mathrm{tar}})-\\ Q(s_t,a_t;w)\end{pmatrix}^2\right] \tag{8}$$

其中，w表示Q_{current}的参数，w_{tar}表示Q_{target}的参数。

损失函数相对于参数w的梯度为

$$y_t'=r_t+\gamma Q(s_{t+1},\arg\max_{a'}Q(s_{t+1},a';w);w_{\mathrm{tar}}) \tag{9}$$

$$\nabla_w L(w)=-2\left(y_t'-Q(s_t,a_t;w)\right)\cdot\nabla_w Q(s_t,a_t;w) \tag{10}$$

其中，

$$\nabla_w Q(s_t,a_t;w)=\frac{\partial Q(s_t,a_t;w)}{\partial w} \tag{11}$$

本文引入了经验回放用于打破连续数据样本之间的相关性并减少过拟合风险。在训练过程中，经验被存储在固定长度为l的回放缓冲区中。每次更新网络时，从该缓冲区中采样长度为batch-size<l的经验批次。当前批次表示为j，下一个批次为j+1。

1.3　联邦强化学习

联邦DDQN-IDS训练流程如图4所示，DDQN和联邦学习的结合可以提高全局模型对动态环境的鲁棒性。理论上讲，全局模型仅关注模型架构，而不关心这些模型是来自于强化学习还是深度学习。在联邦学习中，全局模型通过对多个客户端模型进行加权平均来获得，在联邦强化学习中也是如此。GA模型通过聚合多个EA模型来获得。因此，在本文所提的方法中，EA充当客户端，GA模型通过聚合来自多个EA的Q_{current}来获得。每个EA都拥有自己的独立IoT流量数据集，因此每个EA被视为处于不同的环境中。

1.4　模型架构

Transformer模型由于其自注意力机制和捕捉长期依赖的能力，已经在自然语言处理及其他任务中取得了革命性的进展。基于Transformer层和组归一化的模型架构如图5所示。该架构由嵌入层、Transformer层、组归一化层和输出层组成，在此之后进行动作选择步骤。

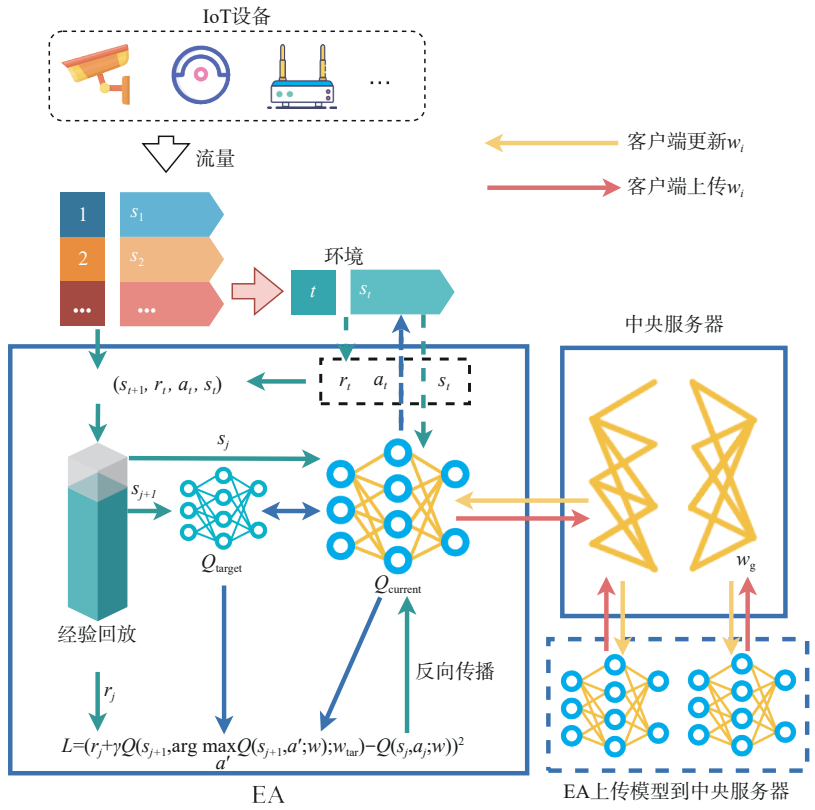


图4　联邦DDQN-IDS训练流程

首先，在训练之前，流量特征会经过预处理步骤对流量信息进行归一化，以统一量纲和范围，得到模型的输入$\boldsymbol{X}$，特征维度为d_t。本文模型中，第一层是线性层，它既作为嵌入层，也将输入$\boldsymbol{X}$转换为Transformer层所需的向量$\boldsymbol{H}$。

$$\boldsymbol{H}=\boldsymbol{X}\boldsymbol{W}_t+\boldsymbol{b}_t \tag{12}$$

其中，$\boldsymbol{W}_t$和$\boldsymbol{b}_t$是可学习的参数，$\boldsymbol{H}$的特征维度为d_{m}。

多头自注意力机制为每个头计算注意力分数。为了计算注意力分数，输入向量$\boldsymbol{H}$被映射为3个不同的向量，即$\boldsymbol{Q}$、$\boldsymbol{K}$和$\boldsymbol{V}$，分别代表查询向量、键向量和值向量。

$$\boldsymbol{Q}=\boldsymbol{H}\boldsymbol{W}_{\boldsymbol{Q}},\boldsymbol{K}=\boldsymbol{H}\boldsymbol{W}_{\boldsymbol{K}},\boldsymbol{V}=\boldsymbol{H}\boldsymbol{W}_{\boldsymbol{V}} \tag{13}$$

$\boldsymbol{W}_{\boldsymbol{Q}}$、$\boldsymbol{W}_{\boldsymbol{K}}$和$\boldsymbol{W}_{\boldsymbol{V}}$是用于将输入向量$\boldsymbol{H}$线性转换为$\boldsymbol{Q}$、$\boldsymbol{K}$和$\boldsymbol{V}$向量的权重矩阵。

对于每个头，注意力输出的计算式为

$$A(\boldsymbol{Q},\boldsymbol{K},\boldsymbol{V})=\mathrm{softmax}\left(\frac{\boldsymbol{Q}\boldsymbol{K}^{\mathrm{T}}}{\sqrt{d_{\mathrm{k}}}}\right)\boldsymbol{V} \tag{14}$$

其中，$\boldsymbol{Q}\boldsymbol{K}^{\mathrm{T}}$用于计算两者的相似度，$d_{\mathrm{k}}$表示$\boldsymbol{Q}$、$\boldsymbol{V}$和$\boldsymbol{V}$的维度，其中，$d_{\mathrm{k}}=d_{\mathrm{m}}/h$。为了防止数值的不稳定，尤其是在高维空间中，通常将相似性分数$\boldsymbol{Q}\boldsymbol{K}^{\mathrm{T}}$除以$d_{\mathrm{k}}$的平方根进行缩放。随后，将所有头的输出连接起来并进行线性变换。

$$\mathrm{MA}\left(\boldsymbol{Q}_i,\boldsymbol{K}_i,\boldsymbol{V}_i\right)=\mathrm{Concat}\left(\mathrm{head}_1,\cdots,\mathrm{head}_h\right)\boldsymbol{W}_{\mathrm{o}} \tag{15}$$

其中，$\boldsymbol{W}_{\mathrm{o}}$为变换矩阵，每个头的表示为

$$\mathrm{head}_i=A\left(Q_i,K_i,V_i\right) \tag{16}$$

在多头自注意力机制之后是前馈神经网络，其对自注意力机制之后的输出进行进一步的非线性变换，从而增强特征的表达能力。前馈神经网络由两层线性变换组成，维度为d_{FF}，中间通过激活函数（通常是ReLU）。

$$\mathrm{FFN}\left(M\left(\boldsymbol{H}\right)\right)=\mathrm{RELU}\left(\boldsymbol{H}\boldsymbol{W}_{\mathrm{f1}}+\boldsymbol{b}_{\mathrm{f1}}\right)\boldsymbol{W}_{\mathrm{f2}}+\boldsymbol{b}_{\mathrm{f2}} \tag{17}$$

$\boldsymbol{W}_{\mathrm{f1}}$和$\boldsymbol{W}_{\mathrm{f2}}$、$\boldsymbol{b}_{\mathrm{f1}}$和$\boldsymbol{b}_{\mathrm{f2}}$分别为两层前馈神经网络的权重矩阵和偏置向量。另外，Transformer层中子层的输入和输出使用残差连接，然后进行层归一化（LayerNorm），目的是将注意力机制产生的输出标准化，使得模型能够更稳定地学习从输入到输出的映射。其在模型中的结构为

$$M\left(\boldsymbol{H}\right)=\mathrm{LayerNorm}\left(\boldsymbol{H}+\mathrm{MA}\left(\boldsymbol{Q},\boldsymbol{K},\boldsymbol{V}\right)\right) \tag{18}$$

其中，MA($\boldsymbol{Q}$, $\boldsymbol{K}$,$\boldsymbol{V}$)为多头注意力模块的输出。

$$\boldsymbol{Z}=\mathrm{LayerNorm}\left(M\left(\boldsymbol{H}\right)+\mathrm{FFN}\left(M\left(\boldsymbol{H}\right)\right)\right) \tag{19}$$

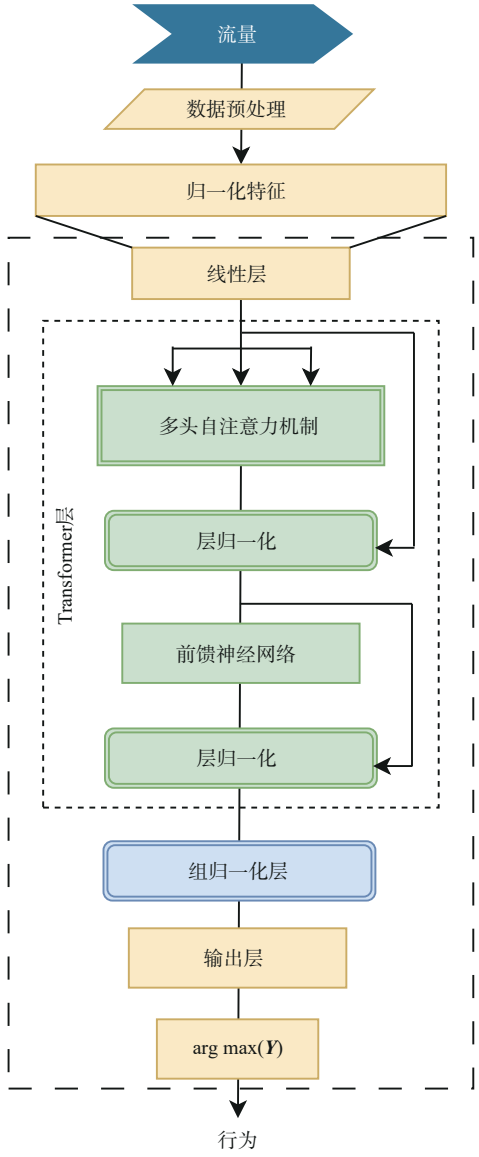


图5　基于Transformer层和组归一化的模型架构

最终得到单层Transformer层的输出是$\boldsymbol{Z}$。如果堆叠多层Transformer层，则$\boldsymbol{Z}^{n-1}$是前一层的输出，$\boldsymbol{Z}^{n}$为

$$\boldsymbol{Z}^{n}=\mathrm{LayerNorm}\left(M\left(\boldsymbol{Z}^{n-1}\right)+\mathrm{FFN}\left(M\left(\boldsymbol{Z}^{n-1}\right)\right)\right) \tag{20}$$

本文所提方法Transformer层只有一层，之后添加一层组归一化层，组归一化可以在保持模型复杂度的同时，通过归一化减少特征值分布的波动。其将输入特征划分为G组，每组有S个通道，其中，$S=C/G$，C表示通道总数。每个组g包含S个通道，每组都独立进行归一化，组归一化不改变输入的形状，它仅在组内部的通道维度上做归一化处理，因此组归一化对批量大小并不敏感，从而确保了模型的稳定，降低了联邦学习中全局模型经过聚合之后权重发散的程度。对于每个组g，归一化特征计算如下

$$\hat{Z}_{g,s}=\frac{Z_{g,s}-\mu_g}{\sqrt{\sigma_g^2+\varepsilon}} \tag{21}$$

其中，μ_g和σ_g^2是组g中特征的均值和方差

$$\mu_g=\frac{1}{S}\sum_{s=1}^{S}Z_{g,s} \tag{22}$$

$$\sigma_g^2=\frac{1}{S}\sum_{s=1}^{S}(Z_{g,s}-\mu_g)^2 \tag{23}$$

其中，S表示每组中的通道数，ε是用于数值稳定的常数，归一化值随后会进行缩放和平移

$$G(\boldsymbol{Z})=\gamma\hat{Z}+\beta \tag{24}$$

其中，γ和β是可学习的缩放和偏移参数。

组归一化层的输出$G(\boldsymbol{Z})$经过最终的线性层，得到最终输出$\boldsymbol{Y}$

$$\boldsymbol{Y}=G(\boldsymbol{Z})\boldsymbol{W}_{\mathrm{out}}+\boldsymbol{b}_{\mathrm{out}} \tag{25}$$

其中，$\boldsymbol{W}_{\mathrm{out}}$是权重矩阵，$\boldsymbol{b}_{\mathrm{out}}$是偏置向量，$\boldsymbol{Y}$表示每个类别的逻辑值。该层将学习到的表示转换为所需的输出维度d_{out}。

最后，使用arg max(·)函数确定估计的动作（Action）

$$\mathrm{Action}=\arg\max(\boldsymbol{Y}) \tag{26}$$

1.5　算法总结

本小节阐述了所提方法的各个部分，运行逻辑为：1) 每一个参与联邦聚合的EA都利用基于Transformer层和组归一化层的神经网络模型搭建DDQN架构。其中，Q_{current}和Q_{target}的网络结构完全一致。Q_{current}用于选择动作，Q_{target}用于计算Q_{current}选择的动作的Q值；2) 每个EA都与各自的环境交互，完成一轮训练之后，EA上传Q_{current}的参数w到GA进行聚合；3) GA不参与训练，只进行全局模型的聚合和测试，其模型结构与EA完全一致。模型聚合方法如式(3)所示；4) EA完成聚合后，将全局模型下发到EA中进行训练。

在全局模型未到达预定标准前，逻辑2)、逻辑3)和逻辑4)重复执行。本文所提方法确保了EA在数据不共享的情况下，能够有效地学习全局知识，并且利用Transformer层的高级特征提取能力和强化学习的奖励机制训练具有优越性能的边缘模型。

2　实验准备

2.1　参数设置

在实验中，GA 和 EA 进行了 5 轮通信，折扣因子 γ 为 0.1，经验回放缓冲区的最大容量为 10 000。对于 IoMT 实验，d_i=45，$\boldsymbol{H}$ 的维度为 d_{m}，其中，$d_{\mathrm{m}}=d_{\mathrm{FF}}$=64，$h$=2。对于 IoV 实验，$d_i$=153，$\boldsymbol{H}$ 的维度为 128，其中，$d_{\mathrm{m}}=d_{\mathrm{FF}}$=128。遗传知识的影响 P=0.999 5，组归一化的组数设置为 8。

2.2　数据集

1) CICIoMT2024：该数据集主要用于评估医疗物联网设备的安全性。它包含来自 40 台 IoMT 设备（25 台真实设备和 15 台模拟设备）测试平台上的 18 种攻击的网络流量数据。这些攻击被分为 5 类：DDoS、DoS、Recon、MQTT 和 ARP spoofing。实验中，将 DDoS 和 DoS 攻击归为一个类别（TCP），最终样本有 5 类标签：BE（BENIGN）、ARP（ARP spoofing）、MQTT、RE（RECON）和 TCP。随机选择 21 000 条样本，其中一半用于构建不平衡的边缘环境，另一半用于 GA 测试。

2) CICIoV2024：该数据集主要用于研究 IoV 的安全问题。通过对一辆 2019 年福特车辆的完整（ECU, electronic control unit）系统执行 5 种网络攻击采集数据，所有攻击可以归类为 spoofing 和 DoS 攻击，这些攻击通过控制器局域网（CAN，controller area network）协议进行攻击。所有样本可被分类为 6 类：SW（Steering Wheel）、BE（BENIGN）、GAS、SPEED、DoS 和 RPM。除了 BE 和 DoS，其余的攻击都属于 spoofing 攻击。我们随机选择 21 000 条样本，其中一半用于构建不平衡的边缘环境，另一半用于 GA 测试。

构建不平衡数据集之后的各类攻击样本在 EA 中的比例如图 6 所示，每个边缘环境中每种样本类别的比例有所不同，以模拟现实环境中遇到的样本不平衡的问题。

2.3　数据预处理和评价指标

为了提高模型的性能和稳定性，减少梯度消失和爆炸问题，并提高模型的可解释性，本文对数据进行归一化处理，使其具有类似的尺度和分布。

$$X_{\mathrm{norm}}=\frac{X-X_{\min}}{X_{\max}-X_{\min}} \tag{27}$$

本文选择了 4 个指标来评估模型的性能：准确率（Acc, accuray）、精确率（Pre, precision）、召回率（Rec, recall）和 F1 分数（F1, F1-score）。

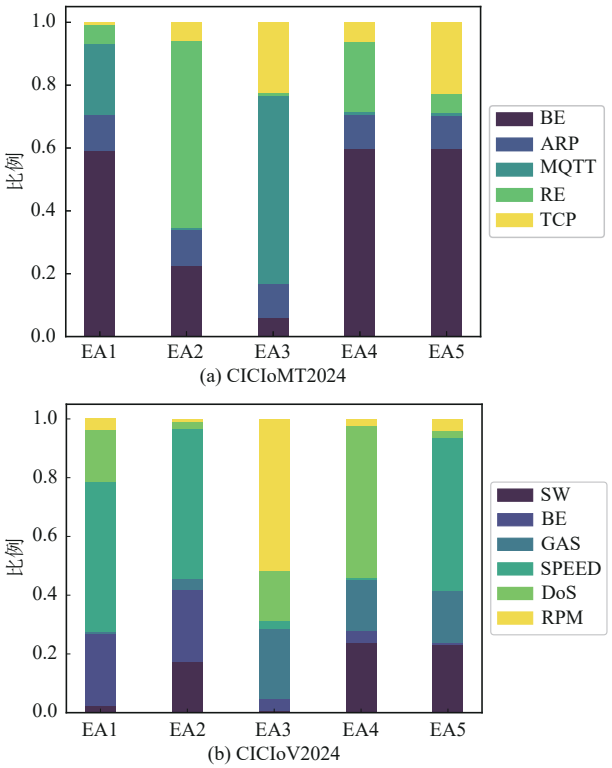


图 6　构建不平衡数据集之后的各类攻击样本在 EA 中的比例

$$\mathrm{Acc}=\frac{\mathrm{TP}+\mathrm{TN}}{\mathrm{TP}+\mathrm{TN}+\mathrm{FP}+\mathrm{TN}} \tag{28}$$

$$\mathrm{Pre}=\frac{\mathrm{TP}}{\mathrm{TP}+\mathrm{FP}} \tag{29}$$

$$\mathrm{Rec}=\frac{\mathrm{TP}}{\mathrm{TP}+\mathrm{TN}} \tag{30}$$

$$\mathrm{F1}=2\times\frac{\mathrm{Pre}\times\mathrm{Rec}}{\mathrm{Pre}+\mathrm{Rec}} \tag{31}$$

其中，TP（true positive）为正确预测的正实例数量，TN（true negative）为正确预测的负实例数量，FP（false positive）为错误分类为正的实例，FN（false negative）为错误分类为负的实例。这些指标对于评估模型的性能以及区分攻击类别的能力至关重要。

2.4　伪代码

联邦强化学习如算法 1 所示，包含两层循环，外层循环决定联邦的通信次数。在内层循环中，EA 获取上一轮 GA 模型用于更新 EA 模型，随后 EA 在上一轮模型的基础上继续进行本地强化学习训练。

算法 1　联邦强化学习

输入　代理的数量 I

输出　全局代理模型 w_{g}

初始化　全局代理模型 w_{g}

for n=1, ⋯, N **do**
  for Agent$_i$ **do**
    w_i=w_g；
    w_i←AGENTTRAIN()；
  end for
  w_g^n←式(3)；
end for

EA 训练算法如算法 2 所示，外层循环负责训练轮数，内层循环决定迭代次数。算法在每一回合开始时，状态 s_t 设定为当前的时间步 x_t，计数器 count 初始化为 0，时间步长 t 初始化为 1，且标志变量 done$_t$ 设置为 FALSE，FALSE 表示当前回合尚未结束。内层循环中，每个时间步 t 根据当前状态 s_t 通过策略 $\pi(s_t)$ 选择动作 a_t。动作执行后，通过函数 STEP(a_t,y_t) 获取下一状态 s_{t+1}、r_t 以及是否结束的标志 done$_t$。随即将当前的经验（s_t，a_t，r_t，s_{t+1}，done$_t$）存储到经验回放缓存器 M 中。随后从缓存器 M 中随机采样一批经验（s_j，a_j，r_j，s_{j+1}，done$_j$），并根据状态 s_{j+1} 选择下一批动作 a'=$\pi(s_{j+1})$。之后利用式(9)计算目标 y_j，通过均方误差（MSE, mean-square error）损失函数计算当前模型参数 w_i 的损失 $L(w_i)$。（TUF, target network update frequency）为 w_{tar} 的更新条件，当计数器 count 等于 TUF 时，则更新目标模型 w_{tar}=w_i，并将计数器重置为 0。

算法 2　EA 训练算法
输入　训练数据 $\boldsymbol{D}$={(x_1, y_1), (x_2, y_2), ⋯, (x_t, y_t)}
输出　边缘代理模型 w_i
初始化　经验回放缓存 M
for e=1, ⋯, E **do**
  s_t=x_t, count=0, t=1, done$_t$=FALSE；
  for t=1, ⋯, T, done$_t$=FALSE **do**
    count=count+1；
    a_t=$\pi(s_t)$；
    s_{t+1}，r_t，done$_t$=STEP(a_t, y_t)；
    保存(s_t, a_t, r_t, s_{t+1}, done$_t$)到 M 中；
    随机从 M 中抽取(s_j, a_j, r_j, s_{j+1}, done$_j$)；
    a'=$\pi(s_{j+1})$；
    y'_j=式(9)；
    $L(w_i)$=MSE(y'_j, $Q(s_j; a_j; w_i)$)；
    if count % TUF=0; **then**
    w_{tar}=w_i, count=0；
  end for
end for

判断当前时间步是否为终点如算法 3 所示，并且返回与环境交互的结果。

算法 3　STEP
输入　a_t，y_t
输出　s_{t+1}, r_t, done$_t$
if a_t=y_t **then**
  r_t=1；
else
  r_t=−1；
if $t \geqslant T$ **then**
  done$_t$=TRUE；
else
  done$_t$=FALSE；
return s_{t+1}, r_t, done$_t$

3　实验与分析

3.1　CICIoMT2024

IoMT 攻击识别结果见表 1，展示了当边缘代理数量（NEA, number of edge agent）为 5 时，GA 在测试集上的结果。结果表明 GA 对 TCP 的识别能力优于其他 3 类攻击，Acc 和 Rec 均达到 0.997 0。MQTT 是仅次于 TCP 被高效识别的攻击。另外，GA 对 ARP 的识别能力相对有限，Acc 为 0.839 5，Pre 为 0.747 0。GA 对 BE 和 RE 的识别 Acc 分别为 0.960 9 和 0.902 8。当 EA 数量为 5 时，IoMT 数据集的训练过程和结果如图 7 所示。图 7(a)、图 7(b)展示了 5 个 EA 的训练损失和累积奖励，不同 EA 之间的损失存在差异，EA1 和 EA2 相对于 EA3 的损失更高，其中 EA3 的损失低于 0.125。图 7(d)是 IoMT 的识别结果混淆矩阵图。样本类别和标签的键值关系为：BE：0，ARP：1，MQTT：2，RE：3，TCP：4。其中 ARP 被错误识别的比例最大，126 个测试个例被识别为 BE。与此同时，BE 也有 81 个测试样本被识别为 ARP，这表明 ARP 和 BE 的行为相对于其他攻击类别更加相似。RE 有 69 和 146 个测试样本被识别为 BE 和 ARP，但 BE 和 ARP 只有 1 个错误的

表 1　**IoMT 攻击识别结果**

类别	Acc	Pre	Rec	F1
BE	0.960 9	0.907 1	0.960 9	0.933 2
ARP	0.839 5	0.747 0	0.839 5	0.790 5
MQTT	0.980 9	0.990 1	0.980 9	0.985 5
RE	0.902 8	0.996 7	0.902 8	0.947 4
TCP	0.997 0	0.996 6	0.997 0	0.996 8

RE 实例。

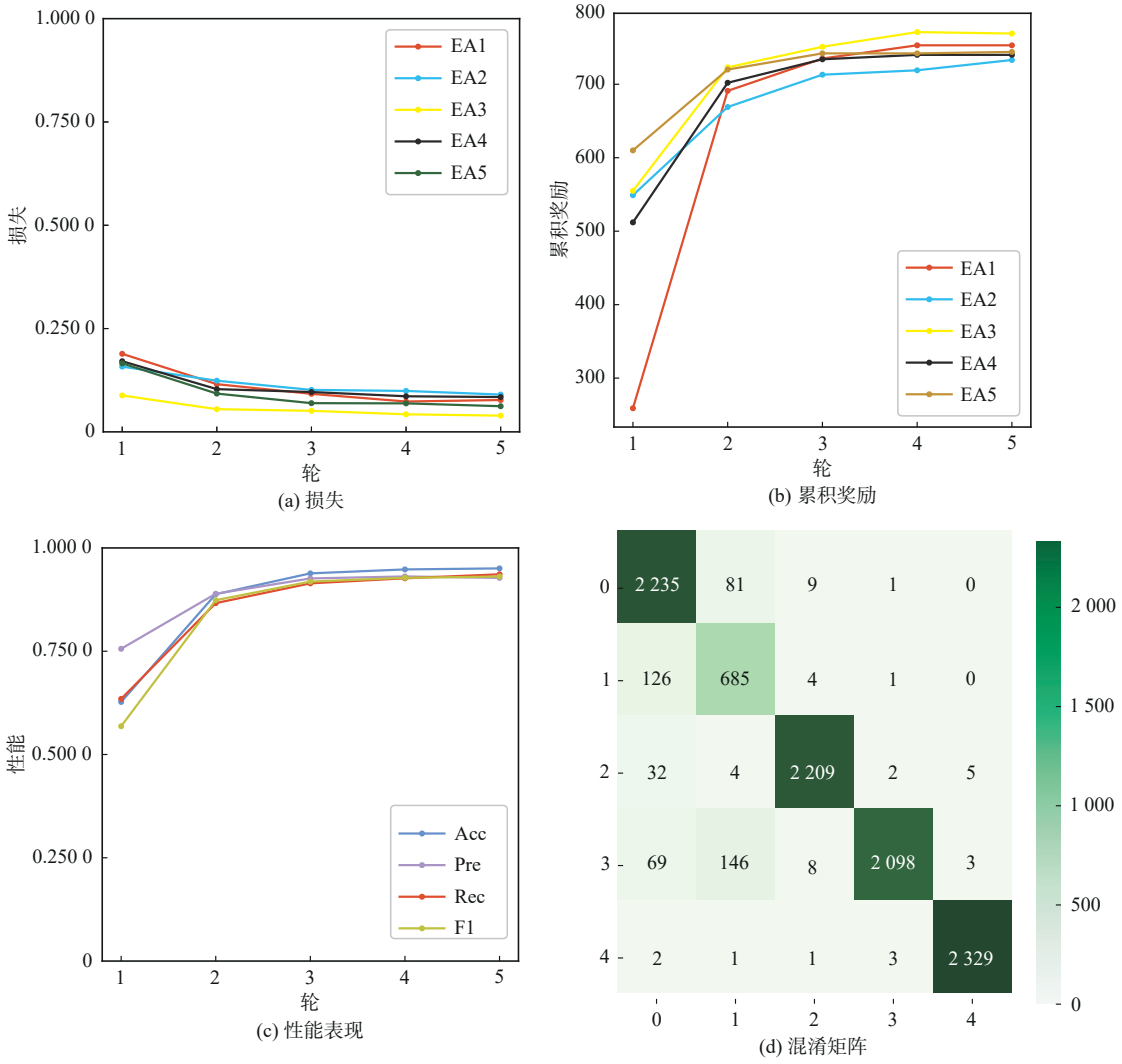


图 7　当 EA 数量为 5 时，IoMT 数据集的训练过程和结果

超参数的影响（IoMT）如图 8 所示，TUF 对 Acc 的影响在图 8(a)中被可视化，当 NEA 为 3 时，TUF 的影响较为显著，当 TUF 等于 50 和 100 时，GA 的测试 Acc 低于 0.92。当 NEA 为 4 和 5 时，TUF 的影响并不明显。

NEA 的影响、批次大小（BS, batch size）的影响和学习率（LR, learning rate）的影响分别见表 2、表 3 和表 4。分别利用 3 个、4 个和 5 个 EA 进行实验，观察 GA 的识别性能。当 NEA 为 3 时，GA 在测试集中的 Acc 表现为 0.943 5，F1 为 0.922 1。随着 NEA 增加到 5，GA 的 Acc 提升至 0.950 5，相较于 3 个 EA 高了 0.7 个百分点。图 8(b)是 NEA 对 GA 性能的影响的折线图，其中 Pre 在 NEA 等于 4 时最高，为 0.933 3，在 NEA 等于 3 时最低，为 0.918 4。

图 8(c)可以观察到学习率的大小对模型性能的影响，LR 设置为 0.001 0 时，在 NEA 为 3、4 和 5 的情况下都达到了最高的 Acc，分别为 0.943 5、0.948 5 和 0.950 5。由表 4 看出，当学习率为 0.005 0 时，GA 的特征学习能力相对于 0.000 5 和 0.001 0 较差。

本文将最优 LR 设定为 0.001 0，对比 BS 对模型的影响，每一批次都是从经验回放队列中随机抽取的样本，实验结果在图 8(d)中展示，当 NEA 为 3 且 BS 为 32 时，GA 的识别 Acc 最低，为 0.930 8。当 NEA 为 4 和 5 时，BS 的影响不显著，但当 NEAs 为 5 时，在 BS 的 3 个参数观察下，GA 的 Acc 表现最

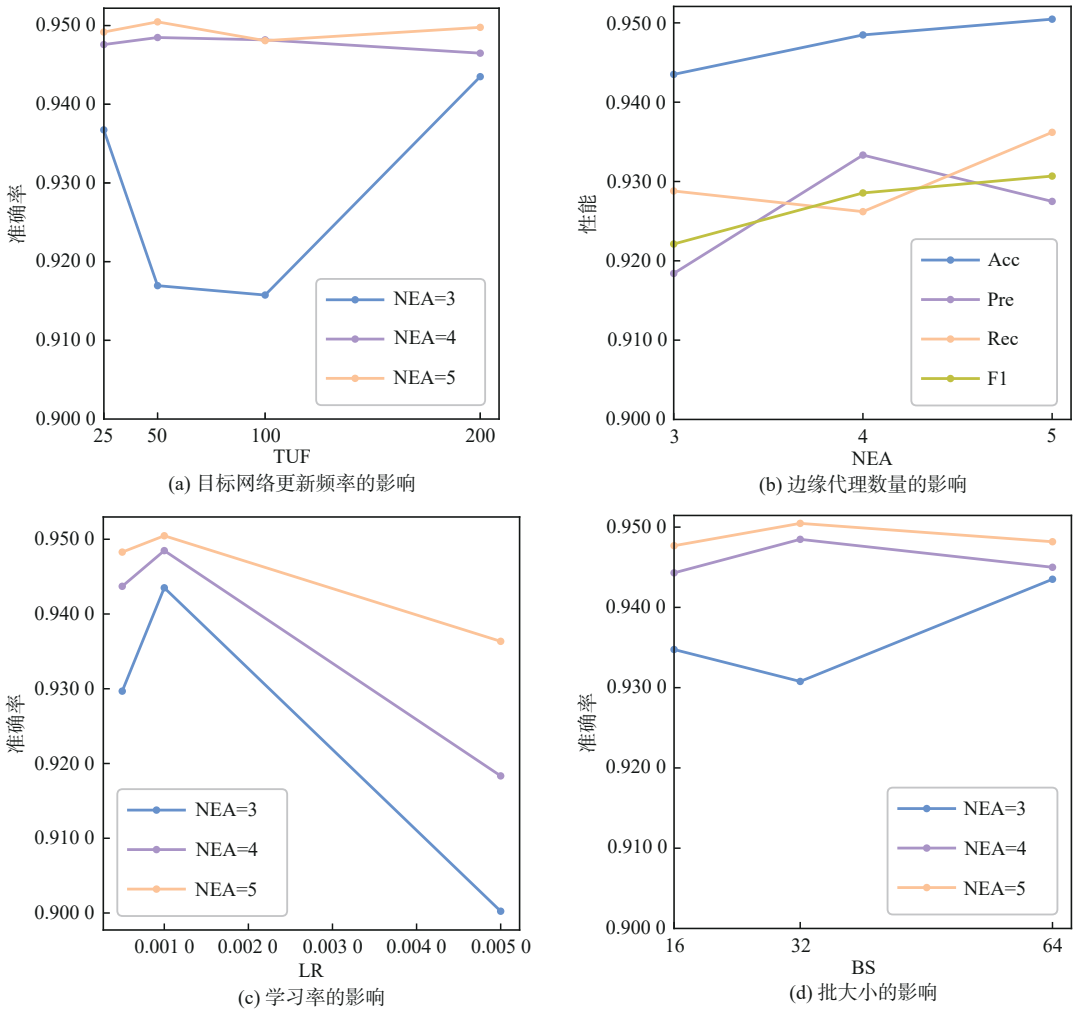


图8　超参数的影响(IoMT)

表2　边缘代理数量的影响(IoMT)

NEA	Acc	Pre	Rec	F1
3	0.943 5	0.918 4	0.928 8	0.922 1
4	0.948 5	0.933 3	0.926 2	0.928 6
5	0.950 5	0.927 5	0.936 2	0.930 7

表3　批大小的影响(IoMT)

BS	NEA=3	NEA=4	NEA=5
16	0.934 8	0.944 3	0.947 7
32	0.930 8	0.948 5	0.950 5
64	0.943 5	0.945 0	0.948 2

表4　学习率的影响(IoMT)

LR	NEA=3	NEA=4	NEA=5
0.000 5	0.929 7	0.943 7	0.948 3
0.001 0	0.943 5	0.948 5	0.950 5
0.005 0	0.900 2	0.918 3	0.936 3

优。学习率设定为0.001 0，批次大小分别为64和32时，边缘代理数量为3时的训练结果如图9所示，边缘代理数量为4时的训练结果如图10所示。

3.2　CICIoV2024

当EA数量为5时，IoV数据集的训练过程和结果如图11所示，展示了使用CICIoV2024数据集进行训练的过程和结果，混淆矩阵中样本类别和标签的键值关系为:SW：0，BE：1，GAS：2，SPEED：3，DoS：4，RPM：5。结合图11(a)、图11(b)和图11(c)可以观察到第一轮通信后，损失和识别精度已经接近最优性能，累积奖励亦无明显上升趋势，平均Acc接近于1。IoV攻击识别结果见表5，结果表明，所提方法在所有4个指标上对DoS攻击的检测能力达到了1。另外，对GAS的检测结果与DoS一致。对于SW和RPM，Acc为0.999 4，Pre为1.000 0，Rec为0.999 4，F1为0.999 7。在

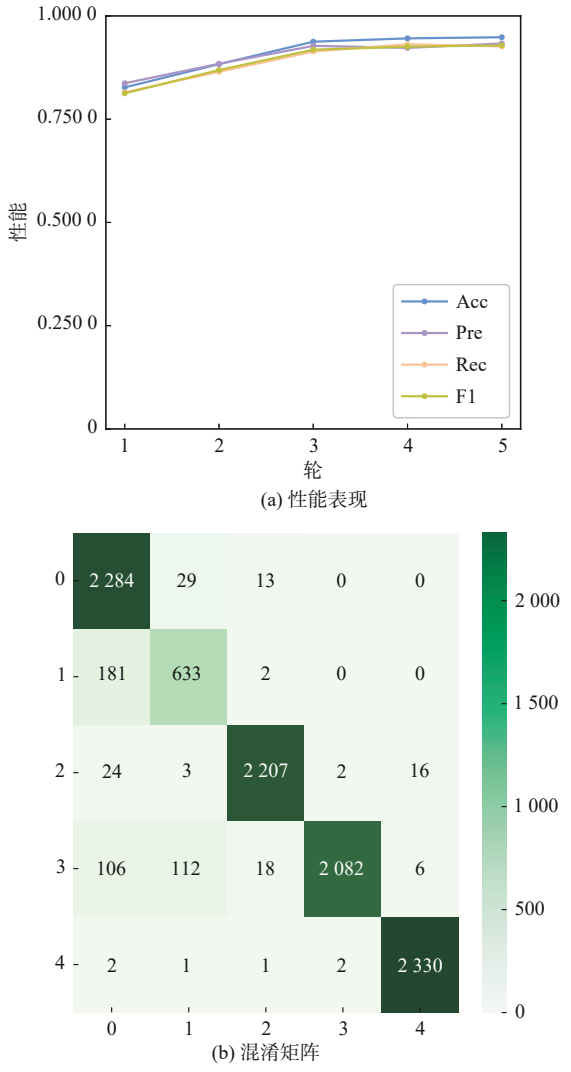


图9　边缘代理数量为3时的训练结果（IoMT）

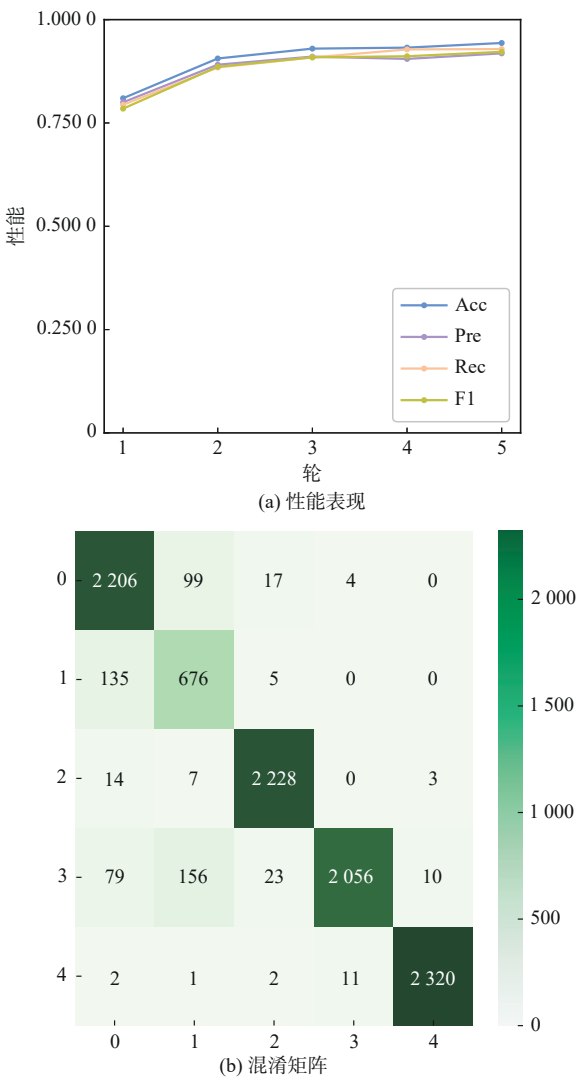


图10　边缘代理数量为4时的训练结果（IoMT）

图 11(d)的混淆矩阵中，被错误识别的实例只有2例，说明GA面对IoV的入侵攻击能够表现优异的辨别能力。

超参数的影响（IoV）如图12所示，边缘代理数量为3时的训练结果（IoV）如图13所示，边缘代理数量为4时的训练结果（IoV）如图14所示，边缘代理的影响、批大小的影响、学习率的影响分别见表6、表7和表8。从图12(a)和图12(b)可以了解TUF和NEA对于入侵的识别精度影响相对LR和BS要小，虽然有所浮动，但精度始终维持在0.99以上。在调整LR的实验中，当LR为0.005 0时，GA的精确度最高仅有0.546 4，而当NEA为3时，Acc仅为0.424 4。设定LR为0.001 0，当NEA为3或4时，结合图13、图14和表6分析可知，GA对于入侵的识别精度几乎一致，差距很小，分别为0.999 8和0.999 7。从图13和图14的混淆矩阵可知，被误判的实例数量分别为3和2，并且都存在RPM被误判至GAS，和BE误判为SPEED的实例。综合表6、表7和表8的实验结果来看，IoV的流量特征相对于IoMT要更加明显，所以识别精确度相对更高。

4　实验结论

4.1　与先进方法的对比

基于联邦学习的IDS方法已在多种环境和领域中得到应用，涵盖了多种数据集、应用场景和任务，每种方法的应用都展示了显著的效果。为了比较目前先进方法和本文方法的性能，我们分别实现了4种先进的IDS方法，这些方法都利用了联邦学习作为IDS的协同训练架构。其中，zero-day和

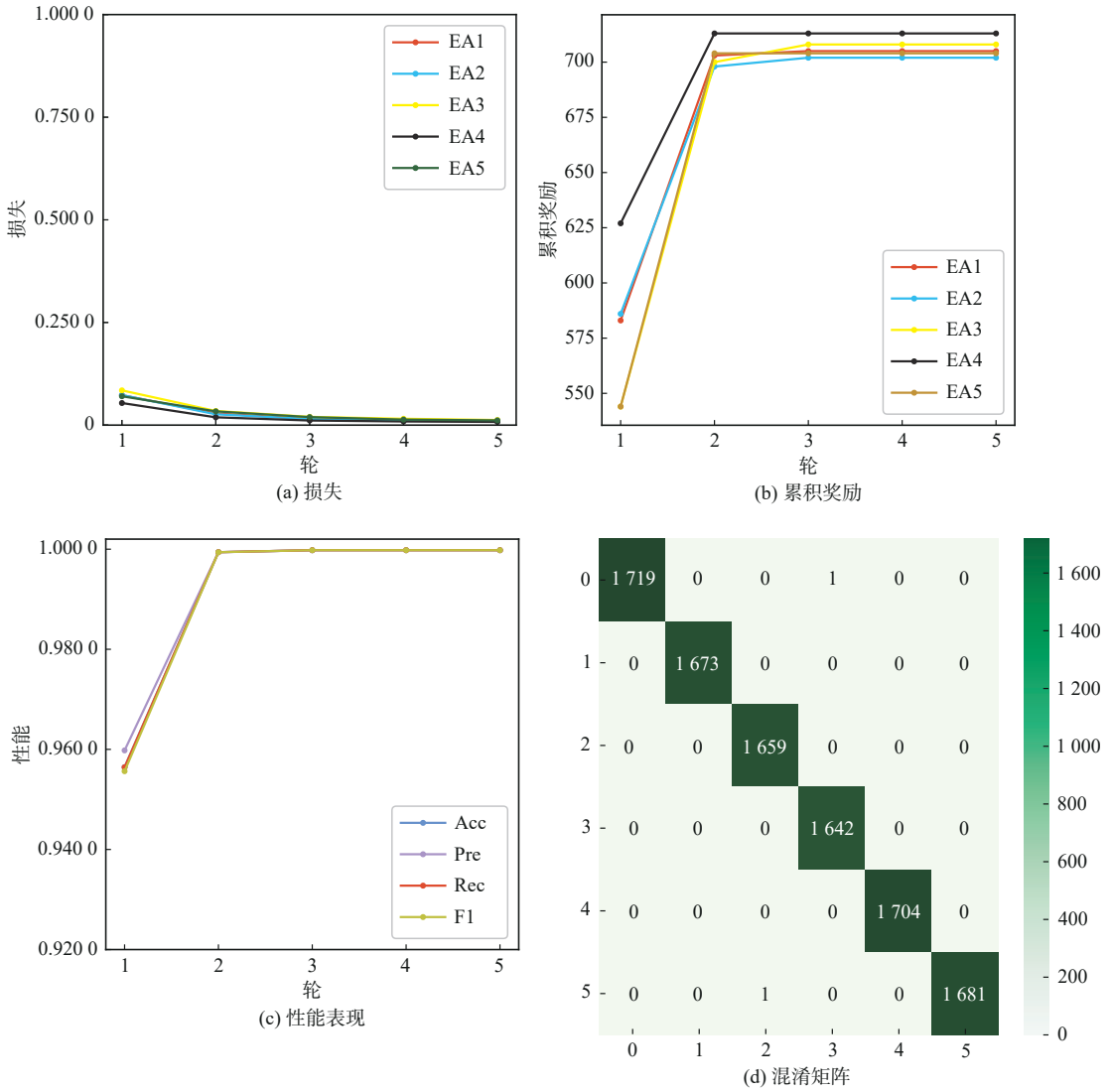


图11　当EA数量为5时，IoV数据集的训练过程和结果

表5　　IoV攻击识别结果

类别	Acc	Pre	Rec	F1
SW	0.999 4	1.000 0	0.999 4	0.999 7
BE	1.000 0	0.999 4	1.000 0	0.999 7
GAS	1.000 0	1.000 0	1.000 0	1.000 0
SPEED	1.000 0	0.999 4	1.000 0	0.999 7
DoS	1.000 0	1.000 0	1.000 0	1.000 0
RPM	0.999 4	1.000 0	0.999 4	0.999 7

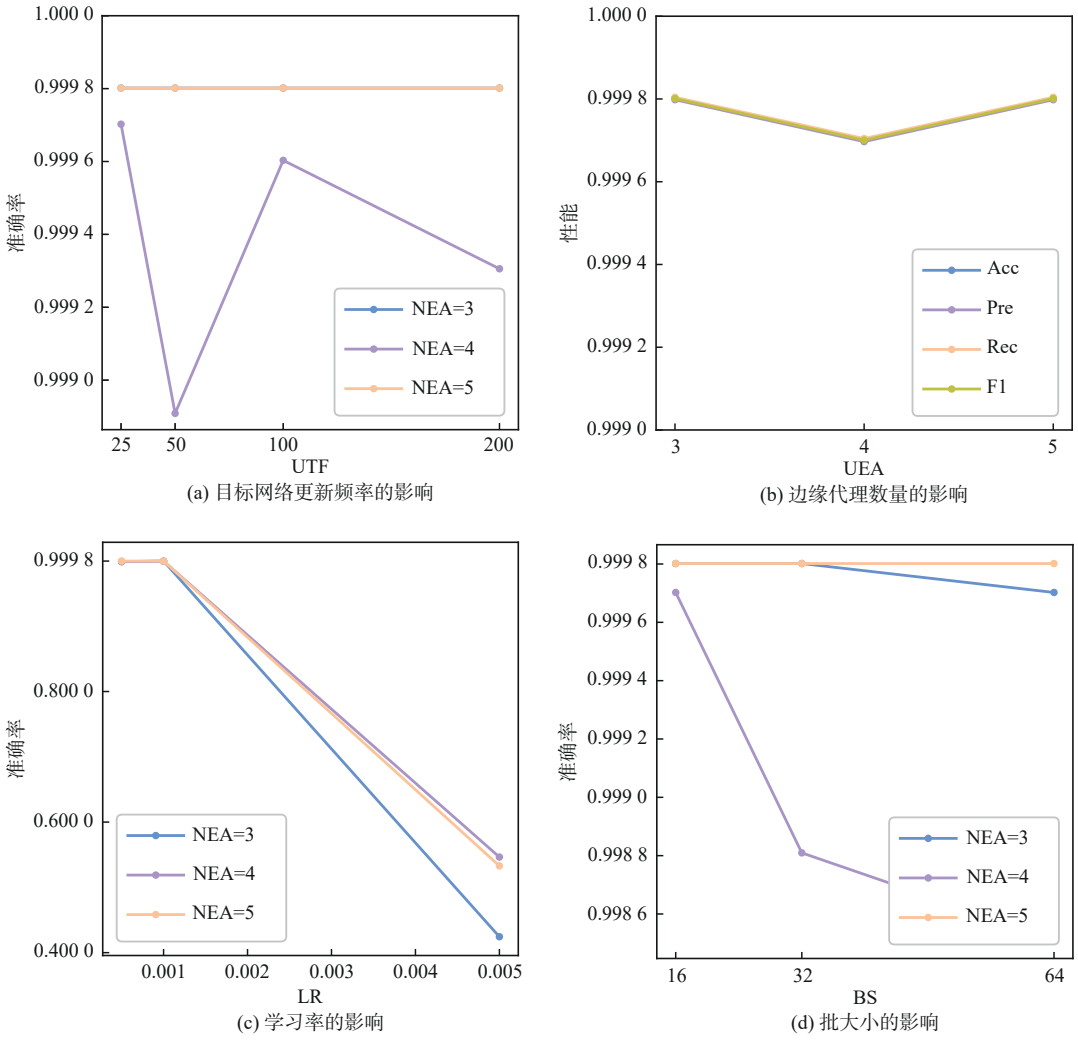
HfedDI考虑了样本分布不平衡的问题，并达到了优异的识别精确度。FedIDI使用LSTM作为客户端模型。FDQN是联邦强化学习在IDS方向上的先进研究成果，被应用于边缘到云生态系统中。为了对比提出的聚合方法的有效性，本文将经典FedAvg的聚合算法作为对比实验之一。IoMT环境中与先进方法的比较结果见表9，IoV环境中与先进方法的比较结果见表10。

经过两个数据集的对比实验，可以观察到本文所提方法在IoMT的环境下领先于目前的先进IDS方法，在IoV的环境下达到与先进方法相当甚至更优的识别精度。综合而言，实验结果证明了将联邦学习和DDQN结合并运用于IDS的可行性，以及在不同环境下的鲁棒性和优异性能。

4.2　缺点和局限性

1) 在实际应用中，面临的挑战通常比实验环境中的问题更加复杂。特别是，EA遇到的攻击类型的差异可能导致GA模型发生权重发散或灾难性遗忘，从而面对不同攻击类别时表现不均衡的识别性能。

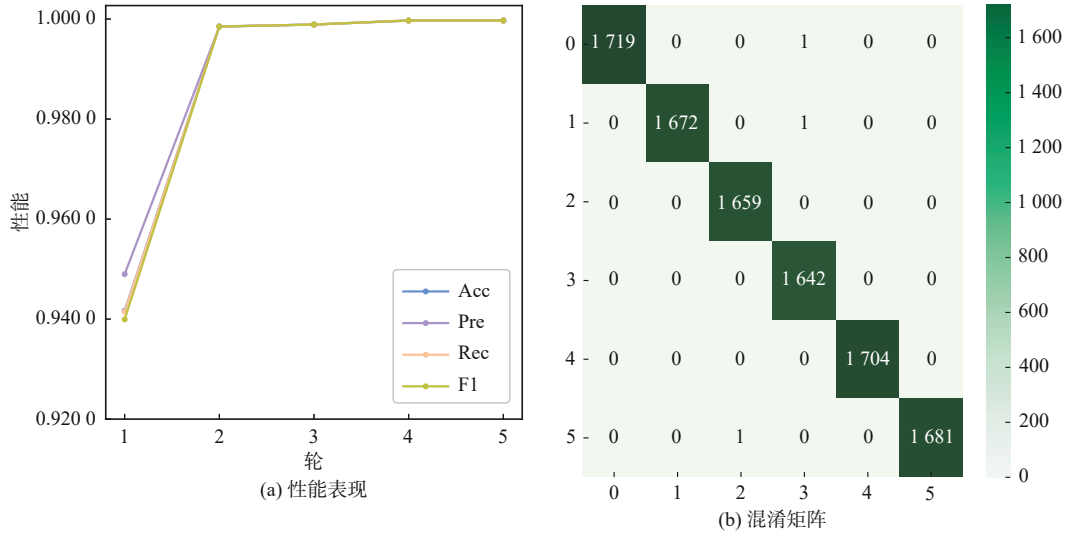
2) 在分布式架构中，入侵检测的效率不仅受数据隐私问题的影响，还受边缘设备的计算能力和通



(a) 目标网络更新频率的影响　(b) 边缘代理数量的影响

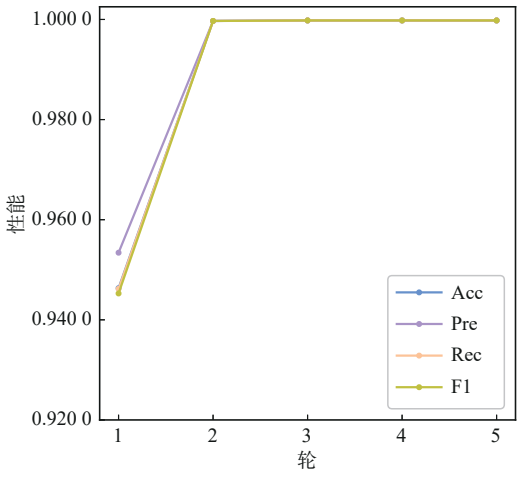
(c) 学习率的影响　(d) 批大小的影响

图 12　超参数的影响（IoV）

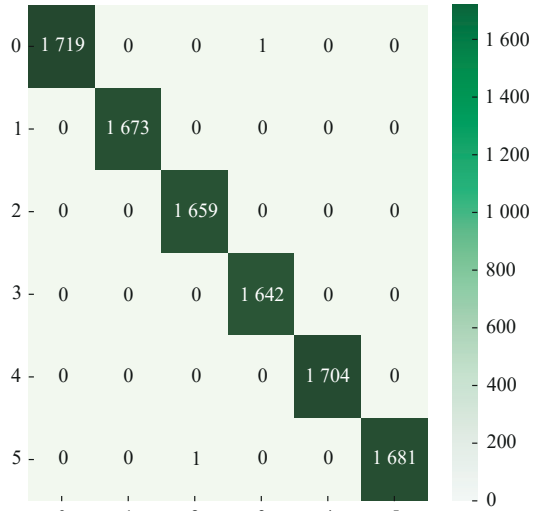


(a) 性能表现　(b) 混淆矩阵

图 13　边缘代理数量为 3 时的训练结果（IoV）



(a) 性能表现



(b) 混淆矩阵

图14　边缘代理数量为4时的训练结果（IoV）

表6　　边缘代理数量的影响（IoV）

NEA	Acc	Pre	Rec	F1
3	0.999 8	0.999 8	0.999 8	0.999 8
4	0.999 7	0.999 7	0.999 7	0.999 7
5	0.999 8	0.999 8	0.999 8	0.999 8

表7　　批大小的影响（IoV）

BS	NEA=3	NEA=4	NEA=5
16	0.999 8	0.999 7	0.999 8
32	0.999 8	0.998 8	0.999 8
64	0.999 7	0.998 5	0.999 8

表8　　学习率的影响（IoV）

LR	NEA=3	NEA=4	NEA=5
0.000 5	0.999 1	0.998 8	0.999 8
0.001 0	0.999 8	0.999 7	0.999 8
0.005 0	0.424 4	0.546 4	0.532 9

表9　　IoMT环境中与先进方法的比较结果

方法	Acc	Pre	Rec	F1
FedAvg[27]	0.946 1	0.927 1	0.926 0	0.925 7
zero-day[28]	0.852 3	0.871 3	0.835 1	0.843 7
FedIDI[29]	0.430 8	0.462 1	0.404 9	0.374 2
HfedDI[30]	0.879 2	0.882 0	0.858 0	0.864 1
FDQN[31]	0.941 5	0.921 4	0.920 1	0.919 6
本文方法	**0.950 5**	**0.927 5**	**0.936 2**	**0.930 7**

表10　　IoV环境中与先进方法的比较结果

方法	Acc	Pre	Rec	F1
FedAvg[27]	0.999 4	0.999 4	0.999 4	0.999 4
zero-day[28]	0.997 8	0.997 8	0.997 8	0.997 8
FedIDI[29]	0.573 1	0.579 7	0.573 4	0.565 7
HfedDI[30]	0.999 4	0.999 4	0.999 4	0.999 4
FDQN[31]	0.999 1	0.999 1	0.999 1	0.999 1
本文方法	**0.999 8**	**0.999 8**	**0.999 8**	**0.999 8**

信开销等因素的影响。这些因素超出本文的研究范围，还需要更多的工作进行探讨。

5　结束语

本文提出了一种基于联邦强化学习的IDS方法，用于保证数据隐私的分布式系统中物联网设备的安全。本文所提方法在IoMT和IoV环境中的IDS任务的实验中证明了有效性。未来的工作可以探讨联邦强化学习在多样化和动态环境中面对攻击的适应性和鲁棒性。此外，优化计算效率可以支持更广泛的应用场景和实时处理需求。这些改进预计将为应对不断变化的IDS任务挑战提供更全面且高效的解决方案。

参考文献：

[1] BACE R G, MELL P. Intrusion detection systems[M]. Special Publication (NIST SP), 2001: 1-51.

[2] LIAO H J, LIN C H R, LIN Y C, et al. Intrusion detection system: a comprehensive review[J]. Journal of Network and Computer Applications, 2013, 36(1): 16-24.

[3] 梁浩然, 伍军, 赵程程, 等. 基于博弈优化边缘学习的物联网入侵检测研究[J]. 物联网学报, 2021, 5(2): 37-47.
LIANG H R, WU J, ZHAO C C, et al. Leveraging edge learning and game theory for intrusion detection in Internet of things[J]. Chinese Journal on Internet of Things, 2021, 5(2): 37-47.

[4] KIZZA J M. System intrusion detection and prevention[M]. Texts in Computer Science. Cham: Springer International Publishing,

2024: 295-323.

[5] YOUNUS Z, ALANEZI M. A survey on network security monitoring: tools and functionalities[J]. Mustansiriyah Journal of Pure and Applied Sciences, 2023, 1(2): 55-86.

[6] MISHRA N, PANDYA S. Internet of things applications, security challenges, attacks, intrusion detection, and future visions: a systematic review[J]. IEEE Access, 2021, 9: 59353-59377.

[7] HEIDARI A, ALI JABRAEIL JAMALI M. Internet of things intrusion detection systems: a comprehensive review and future directions[J]. Cluster Computing, 2023, 26(6): 3753-3780.

[8] LIU Y X, WANG J, LI J Q, et al. Machine learning for the detection and identification of Internet of things devices: a survey[J]. IEEE Internet of Things Journal, 2022, 9(1): 298-320.

[9] AKSOY A, GUNES M H. Automated IoT device identification using network traffic[C]//Proceedings of the ICC 2019 - 2019 IEEE International Conference on Communications (ICC). Piscataway: IEEE Press, 2019: 1-7.

[10] KOSTAS K, JUST M, LONES M A. IoTDevID: a behavior-based device identification method for the IoT[J]. IEEE Internet of Things Journal, 2022, 9(23): 23741-23749.

[11] SALMAN O, ELHAJJ I H, CHEHAB A, et al. A machine learning based framework for IoT device identification and abnormal traffic detection[J]. Transactions on Emerging Telecommunications Technologies, 2022, 33(3): e3743.

[12] VENKATESAN S. Design an intrusion detection system based on feature selection using ML algorithms[J]. Mathematical Statistician and Engineering Applications, 2023, 72(1): 702-10.

[13] LUO Y T, CHEN X, GE N, et al. Transformer-based device-type identification in heterogeneous IoT traffic[J]. IEEE Internet of Things Journal, 2023, 10(6): 5050-5062.

[14] LU Q, YANG Z K, ZHANG H L, et al. MRFE: a deep-learning-based multidimensional radio frequency fingerprinting enhancement approach for IoT device identification[J]. IEEE Internet of Things Journal, 2024, 11(18): 30442-30454.

[15] KALWAR J H, BHATTI S. Deep learning approaches for network traffic classification in the Internet of things (IoT): a survey[J]. arXiv preprint, 2024, arXiv: 240200920.

[16] KILICHEV D, KIM W. Hyperparameter optimization for 1D-CNN-based network intrusion detection using GA and PSO[J]. Mathematics, 2023, 11(17): 3724.

[17] SHARMA B, SHARMA L, LAL C, et al. Explainable artificial intelligence for intrusion detection in IoT networks: a deep learning based approach[J]. Expert Systems with Applications, 2024, 238: 121751.

[18] ZHANG Z, LIU F, GE Y, et al. An intrusion detection method based on depthwise separable convolution and attention mechanism[J]. Chinese Journal on Internet of Things, 2023, 7(1): 49-59.

[19] DU J W, YANG K, HU Y J, et al. NIDS-CNNLSTM: network intrusion detection classification model based on deep learning[J]. IEEE Access, 2023, 11: 24808-24821.

[20] LAGHRISSI F, DOUZI S, DOUZI K, et al. Intrusion detection systems using long short-term memory (LSTM)[J]. Journal of Big Data, 2021, 8(1): 65.

[21] HNAMTE V, NHUNG-NGUYEN H, HUSSAIN J, et al. A novel two-stage deep learning model for network intrusion detection: LSTM-AE[J]. IEEE Access, 2023, 11: 37131-37148.

[22] HAZMAN C, GUEZZAZ A, BENKIRANE S, et al. Enhanced IDS with deep learning for IoT-based smart cities security[J]. Tsinghua Science and Technology, 2024, 29(4): 929-947.

[23] THAREWAL S, ASHFAQUE M W, BANU S S, et al. Intrusion detection system for industrial Internet of things based on deep reinforcement learning[J]. Wireless Communications and Mobile Computing, 2022, 2022(1): 9023719.

[24] HE M S, WANG X J, WEI P, et al. Reinforcement learning meets network intrusion detection: a transferable and adaptable framework for anomaly behavior identification[J]. IEEE Transactions on Network and Service Management, 2024, 21(2): 2477-2492.

[25] LI Q B, DIAO Y Q, CHEN Q, et al. Federated learning on non-IID data silos: an experimental study[C]//Proceedings of the 2022 IEEE 38th International Conference on Data Engineering (ICDE). Piscataway: IEEE Press, 2022: 965-978.

[26] HUANG Y T, CHU L Y, ZHOU Z R, et al. Personalized cross-silo federated learning on Non-IID data[J]. Proceedings of the AAAI Conference on Artificial Intelligence, 2021, 35(9): 7865-7873.

[27] MCMAHAN H B, MOORE E, RAMAGE D, et al. Communication-efficient learning of deep networks from decentralized data[J]. arXiv preprint, 2016, arXiv: 1602.05629.

[28] POPOOLA S I, ANDE R, ADEBISI B, et al. Federated deep learning for zero-day botnet attack detection in IoT-edge devices[J]. IEEE Internet of Things Journal, 2022, 9(5): 3930-3944.

[29] HE Z M, YIN J, WANG Y, et al. Edge device identification based on federated learning and network traffic feature engineering[J]. IEEE Transactions on Cognitive Communications and Networking, 2022, 8(4): 1898-1909.

[30] SUMITRA, SHENOY M V. HFedDI: a novel privacy preserving horizontal federated learning based scheme for IoT device identification[J]. Journal of Network and Computer Applications, 2023, 214: 103616.

[31] AL-NADAY M, DOBRE V, REED M, et al. Federated deep Q-learning networks for service-based anomaly detection and classification in edge-to-cloud ecosystems[J]. Annals of Telecommunications, 2024, 79(3): 165-178.

[32] LI Z, KONG Y B, JIANG C J. A transfer double deep Q network based DDoS detection method for Internet of vehicles[J]. IEEE Transactions on Vehicular Technology, 2023, 72(4): 5317-5331.

[33] VASWANI A. Attention is all you need[C]//Advances in Neural In-

formation Processing Systems. [S.L.: s.n.], 2017: 5998-6008.
[34] WU Y, HE K. Group normalization[C]//Proceedings of the European Conference on Computer Vision (ECCV). [S.L.: s.n.] 2018: 3-19.
[35] CASELLA B, ESPOSITO R, SCIARAPPA A, et al. Experimenting with normalization layers in federated learning on Non-IID scenarios[J]. IEEE Access, 2024, 12: 47961-47971.
[36] YAO D Z, PAN W N, DAI Y T, et al. Local-global knowledge distillation in heterogeneous federated learning with Non-IID data[J]. arXiv preprent, 2021, arXiv: 2107.00051.
[37] NETO E C P, TASLIMASA H, DADKHAH S, et al. CICIoV2024: advancing realistic IDS approaches against DoS and spoofing attack in IoV CAN bus[J]. Internet of Things, 2024, 26: 101209.
[38] DADKHAH S, NETO E C P, FERREIRA R, et al. CICIoMT2024: a benchmark dataset for multi-protocol security assessment in IoMT[J]. Internet of Things, 2024, 28: 101351.

[作者简介]

丁凯（1985– ），男，博士，东莞理工学院计算机科学与技术学院副教授，主要研究方向为物联网、智慧城市、机器人技术和移动互联应用等。

黄宜都（1998– ），男，东莞理工学院计算机科学与技术学院硕士生，主要研究方向为物联网工程、联邦学习和网络流量检测等。

陶铭（1986– ），男，博士，东莞理工学院计算机科学与技术学院教授、副院长，主要研究方向为人工智能、边缘计算和云计算等。

谢仁平（1989– ），男，博士，东莞理工学院计算机科学与技术学院特聘副研究员，主要研究方向为计算机视觉、图像处理、目标检测、图像分割、图像融合、图像拼接和桥梁检测机器人等。