物联网场景下基于蜜场的分布式网络入侵检测系统研究

吴昊[1,2]，郝佳佳[1,2]，卢云龙[1,2]
（1. 先进轨道交通自主运行全国重点实验室，北京 100044；2. 北京交通大学电子信息工程学院，北京 100044）

摘 要：为了解决物联网网络入侵检测系统无法识别新型攻击、灵活性有限等问题，基于蜜场提出了一种能有效识别异常流量和具备持续学习能力的网络入侵检测系统。首先，结合卷积块注意力模块的特点，构建专注于通道和空间双维度的异常流量检测模型，从而提高模型的识别能力。其次，利用联邦学习下的模型训练方案，提高模型的泛化能力。最后，基于蜜场对边缘节点的异常流量检测模型进行更新迭代，从而提高系统对新型攻击流量的识别准确度。实验结果表明，所提系统不仅能有效检测出网络流量中的异常行为，还可以持续提高对异常流量的检测性能。

关键词：网络入侵检测系统；联邦学习；蜜场；卷积块注意力模块；物联网
中图分类号：TN929.5
文献标志码：A
DOI: 10.11959/j.issn.1000−436x.2024020

Research on distributed network intrusion detection system for IoT based on honeyfarm

WU Hao[1,2], HAO Jiajia[1,2], LU Yunlong[1,2]
1. State Key Laboratory of Advanced Rail Autonomous Operation, Beijing 100044, China
2. School of Electronic and Information Engineering, Beijing Jiaotong University, Beijing 100044, China

Abstract: To solve the problems that the network intrusion detection system in the Internet of things couldn't identify new attacks and has limited flexibility, a network intrusion detection system based on honeyfarm was proposed, which could effectively identify abnormal traffic and have continuous learning ability. Firstly, considering the characteristics of the convolutional block attention module, an abnormal traffic detection model was developed, focusing on both channel and spatial dimensions, to enhance the model's recognition abilities. Secondly, a model training scheme utilizing federated learning was employed to enhance the model's generalization capabilities. Finally, the abnormal traffic detection model at the edge nodes was continuously updated and iterated based on the honeyfarm, so as to improve the system's accuracy in recognizing new attack traffic. The experimental results demonstrate that the proposed system not only effectively detects abnormal behavior in network traffic, but also continually enhances performance in detecting abnormal traffic.

Keywords: NIDS, federated learning, honeyfarm, convolutional block attention module, IoT

收稿日期：2023−09−18；**修回日期**：2023−12−13
通信作者：卢云龙，yllu@bjtu.edu.cn
基金项目：中央高校基本科研业务费专项资金资助项目（No.2022JBQY004）；基础科研基金资助项目（No.JCKY2022XXXX145）；国家自然科学基金资助项目（No.62221001）；中国国家铁路集团有限公司科技研究开发计划基金资助项目（No.K2022G018）；北京市自然科学基金资助项目（No.L211013）；中国博士后科学基金资助项目（No.2021TQ0028）
Foundation Items: The Fundamental Research Funds for the Central Universities (No.2022JBQY004), The Basic Research Program (No.JCKY2022XXXX145), The National Natural Science Foundation of China (No.62221001), The Science and Technology Research and Development Plan of China Railway Co., Ltd. (No.K2022G018), Beijing Natural Science Foundation (No.L211013), China Postdoctoral Science Foundation (No.2021TQ0028)

0　引言

随着信息技术的迅猛发展，物联网（IoT, Internet of things）已经成为现代社会中不可或缺的一部分，在诸多领域有着广泛应用，如工业物联网[1]、医疗物联网[2]、智慧城市[3]、智能家居[4]等。然而，网络安全是物联网亟须解决的问题。据估计，到 2025 年，物联网设备数量将增加到 750 亿台[5]，这些设备具有有限的计算资源和异构硬件，但是许多物联网设备生产商希望从先发优势中获益，从而发布缺乏安全性的用户友好型产品，因而这些设备安全防御能力受限，更难以抵抗不断出现的新型未知攻击。

为了降低物联网系统的安全风险，通常在网络的边缘网关上部署网络入侵检测系统（NIDS, network intrusion detection system）来保护网络[6]。NIDS 是一种监测和分析网络流量以及检测潜在攻击行为的技术，其研究自 20 世纪 80 年代开始在全球范围内开展，现已发展成为网络安全体系结构中不可或缺的一部分[7]。然而随着数字技术的不断发展和网络攻击手段的日益多样化，网络流量的数据特征愈加复杂，传统 NIDS 已经难以满足物联网用户对准确识别异常和未知攻击的需求。

针对准确识别复杂流量数据中已知异常流量的需求，NIDS 领域开始使用深度学习对网络流量中的复杂特征进行学习，实现在物联网中的异常流量识别。自编码器（AE, auto encoder）是一种无监督神经网络，可以学习如何有效地压缩和加密数据，常用于入侵检测中的特征提取过程[8-10]。文献[11]提出了一种基于堆叠自编码器的两阶段模型用于构建 NIDS，该模型在异常检测过程中会对网络流量进行打分，利用打分值作为附加特征进行最终决策。文献[12]提出了一种基于并行深度自编码器的新型轻量级架构，关心特征向量中各个值周围的局部和信息，大大减少了参数数量、内存占用和对处理能力的需求。文献[13]提出了一种非对称深度自编码器，使用随机森林的特征提取技术，该方法显示出比深度信念网络更高的效率并减少了训练时间。然而单个检测节点的本地数据总是有限的，为了提高深度学习模型的质量，通常需要将数据集中在一个中心服务器进行集中式深度学习，这可能涉及用户隐私泄露和数据传输安全的问题。联邦学习（FL, federated learning）为解决这一问题提供了一种创新的解决方案[14]。FL 是一种分布式机器学习框架，模型训练和参数更新在本地设备上完成，仅将加密处理后的梯度信息发送至中央服务器进行聚合。此方式不仅保护了物联网用户的隐私，还充分利用了分布式设备上的数据资源，提高了模型的性能和泛化能力。文献[15]提出一种 FL 框架下用于工业物联网的 NIDS，为了检测时间序列数据中的异常，使用由差分自编码器和 LSTM 组成的模型架构，结果表明，该方法获得了 97.9%的 F1 分数，并且与集中式学习架构相比带宽减少了 35%。文献[16]提出一种基于 FL 的具有特征选择技术的入侵检测系统，使用贪心算法来选择针对不同攻击类别实现更好入侵检测准确性的特征，结果表明，FL 极大地提高了准确性。文献[17]提出一种使用堆叠自动编码器和 FedAVG 算法[18]进行模型训练的 NIDS。此外，基于 AE 的 FL 入侵检测系统可以扩展到边缘计算的其他领域，例如，文献[19]利用 FL 来处理医疗信息网络中的异常。

另一方面，针对未知和新型的攻击方式，现有 NIDS 的检测能力也相对有限，需要引入新的技术和方法来提升物联网中 NIDS 的灵活性。蜜场作为一种通过代理扩展诱饵节点部署范围的蜜罐系统，是一种可用于物联网的主动防御系统，利用 NIDS 和重定向装置将异常流量从各个子网节点转发到一个集中部署的蜜罐网络内，实现对真实网络的保护，也可以进一步分析其中的异常行为[20-22]，调整应对攻击的防御方案。蜜场分析新型攻击行为的主要方式是对蜜场的日志数据进行提取分析，例如，文献[23]提出一种基于蜜罐的新型威胁情报技术，使用蜜罐收集网络事件日志数据，然后使用 Elasticsearch 技术，对日志数据进行分析评估，进而识别攻击行为并发现攻击趋势。文献[24]提出一种基于软件蜜罐的入侵检测防御系统，系统由蜜罐服务器、监控和入侵检测系统组成，利用监控系统记录访问行为，进而分析检测新型攻击行为，保护系统安全。

基于上述分析，本文旨在研究物联网下的网络入侵检测系统，结合蜜场技术提高网络入侵检测系统对新型攻击流量的识别能力，并在联邦学习框架下进行深度学习模型的训练和优化。本文主要的研究工作如下。

1) 为了解决当前 NIDS 在有效识别新型攻击流量方面的局限性，本文基于蜜场这一主动防御手段，构建了一套具备持续学习能力的异常流量检测系统。通过利用蜜场与新型攻击进行交互，感知新型

攻击行为的流量数据进行分析，进而不断更新和优化模型，实现对新型攻击行为的高效识别和检测。

2) 利用联邦学习框架训练边缘引流节点的网络入侵检测模型，在保护边缘节点流量数据隐私安全的同时，可以有效减轻蜜场与已知攻击流量交互的负担，并充分利用边缘节点的计算能力。

3) 提出一种结合卷积块注意力模块（CBAM, convolutional block attention module）和卷积自编码器的异常流量检测模型 FedCBAM，通过卷积自编码器从流量数据中提取流量特征，引入关注通道关系和空间关系的注意力机制，分配不同的权重给特征参数，从而减少了开销并提高了模型性能。

1 基础知识

1.1 卷积块注意力模块

在入侵检测模型中引入注意力机制的目的主要是提高表征能力，给重要特征赋予更大权重。卷积块注意力模块是 Woo 等[25]在 2018 年提出的轻量级注意力模块，主要由 2 个子模块组成：通道注意力模块（CAM, channel attention module）和空间注意力模块（SAM, spatial attention module），通过引入 CAM 和 SAM 来增强卷积神经网络的表示能力和感受野，从而提高模型性能。

1) CAM 用于对不同通道的特征进行加权，以提取最具有信息量的通道特征，具体结构如图 1(a)所示。

2) SAM 用于捕捉不同区域之间的空间关系和重要性，是 CAM 的补充，具体结构如图 1(b)所示。

3) CBAM 依次使用 CAM 和 SAM，使模型能够分别学习通道和空间特征，具体结构如图 1(c)所示。

1.2 蜜场

单个蜜罐或多个独立部署的蜜罐在分析网络攻击方面存在局限性，难以满足对全局网络攻击监控、关联和趋势预测的综合需求。同时，维护分散地点的多个蜜罐所带来的昂贵成本也不容忽视。为了应对这些问题，蜜场架构应运而生，该架构将多个蜜罐置于同一资源池中，有效地解决了监控和管理方面的难题。具体而言，在被保护网络的各个监控位置均设置了重定向器，将攻击流量引导至资源池中相应的蜜罐。这种集中式的管理方式，不仅只需一名安全人员便可维护所有蜜罐，还为全局网络攻击的监控和关联分析提供了有力的支持。这种集中化的设计极大地提升了蜜场的可管理性和效率，为应对不断演化的网络威胁提供了可靠的基础。

2 系统方案

2.1 系统架构

基于蜜场的物联网分布式异常流量检测系统整体架构如图 2 所示，根据物联网的感知层、网络层、平台层、应用层的架构特点，本文提出的网络入侵检测系统主要部署在网络层和平台层，即在网络层的边缘网关上部署分布式入侵检测模块，在平台层部署云端蜜场，为应用层提供基于云端蜜场的蜜场管理和入侵检测安全防护。本文所提系统主要由以下 5 个部分组成。

1) 访问设备。位于感知层，生成访问真实业务网络的流量数据，从 NIDS 角度可分为异常流量和正常流量，从防火墙判别可分为白名单流量、黑名单流量和灰名单流量。

2) 边缘检测引流节点。位于网络层的边缘网关，

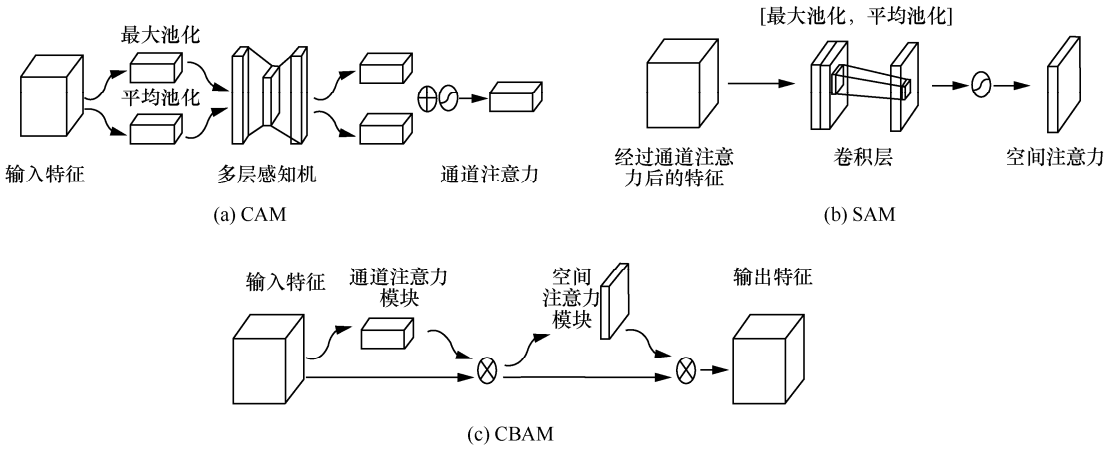


图 1 卷积块注意力模块结构

是具有一定计算能力的边缘设备，执行流量检测和重定向任务，是联邦学习中进行本地模型训练的载体。

3) 云端蜜场。位于平台层，执行对威胁未知流量的分析识别任务，主要是利用蜜场日志信息分析和识别新型异常流量，并基于识别到的新型异常流量数据训练可识别新型异常流量的检测模型，还包含用作联邦学习中心服务器的资源运算平台。

4) 蜜场管理平台。本文所提系统的管理平台，监控云端蜜场和边缘检测引流节点，记录历史流量中的黑名单用户和白名单用户，控制边缘检测引流节点对节点数据重新标记，其中的联邦学习模块管理基于联邦学习异常流量检测模型的训练过程。

5) 真实业务网络。云端蜜场的保护对象，重要保护资产与蜜场中的蜜罐相互映射。经检测的白名单正常流量会直接进入真实业务网络，经蜜场确认为正常的流量也会流入真实业务网络。

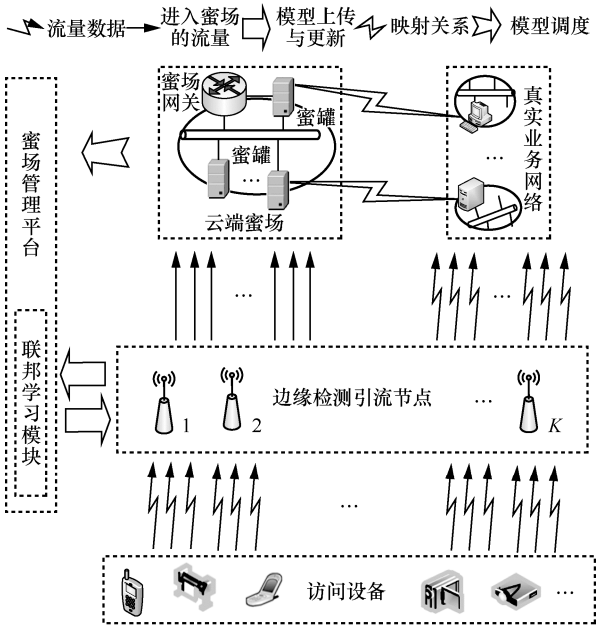


图 2　系统整体架构

本文所提系统设计思路可划分为 2 个主要部分，即联邦学习下的流量检测模型训练策略与基于蜜场的识别新型异常流量的检测模型更新方案。在前者中，本文设计入侵检测模型 FedCBAM，并在联邦学习框架下进行训练。该模型能够有效地辨识与分类异常流量，并将未知异常流量定向到蜜场进行进一步分析。在后者中，本文基于蜜场的构想，实现对新型异常流量的深入分析和识别，并随后更新分布在边缘节点上的入侵检测模型 FedCBAM。2 个方案紧密结合，构建一个在物联网环境下具备持续学习能力的网络入侵检测系统。

2.2　联邦学习下的流量检测模型训练方案

联邦学习可以将入侵检测模型 FedCBAM 的训练下放到本地节点，即本文系统的边缘检测引流节点，避免数据离开节点，保护节点数据隐私安全。假设有 K 个本地节点，每个节点都有私有流量数据集 $D_i, i=\{1,2,\cdots,K\}$，节点收到全局模型参数后利用本地数据进行训练，将本地训练的模型参数 w_i 上传到中心服务器，即本文系统的资源运算平台，中心服务器将局部模型参数聚合得到全局模型参数，表示为

$$w=\frac{1}{K}\sum_{i=1}^{K}w_i \tag{1}$$

然后下发到节点进行下一轮训练，最终目标是最小化损失函数，表示为

$$\min F(w)=\frac{1}{|D|}\sum_{i=1}^{K}|D_i|L_i(w) \tag{2}$$

其中，$L_i(w)$ 和 $|D_i|$ 分别表示第 i 个节点的损失函数和数据大小，通过迭代这个过程逐渐提高全局模型的检测能力。

联邦学习下的流量检测模型训练方案的架构如图 3 所示。在此架构中，边缘检测引流节点接收来自各个物联网访问设备的流量，首先复制原始流量包，接着对这一复制的流量进行数据预处理，随后使用已获取的初始检测模型进行本地训练，经过训练后，本地模型的权重参数会被蜜场管理平台的联邦学习模块上传至资源运算平台。上传的多个边缘节点的模型权重将会在资源运算平台上进行聚合。一旦完成聚合，资源运算平台将更新后的模型权重返回给各个边缘节点，使每个边缘节点都得以获得经过联邦学习训练后的模型参数。这种联邦学习的方式使各个边缘节点能够共享并融合彼此经验，同时又不需要将敏感数据集中存储。整个过程旨在提升检测模型的性能和准确度，以实现在分布式物联网环境下的流量检测任务。

引流策略如图 4 所示。在该策略中，边缘检测引流节点利用训练完成的入侵检测模型 FedCBAM 来识别流量。异常流量 F1 经识别后将会在边缘节点被阻止，并且对应的用户会被记录到黑名单中。相比之下，被识别到的正常流量 F2 会经过用户身份检测，之后根据蜜场已有的白名单、黑名单进行流量划分。黑名单下的流量 F21 同样被阻止在边缘节点，而被验证为白名单正常流量的 F22 则可进入

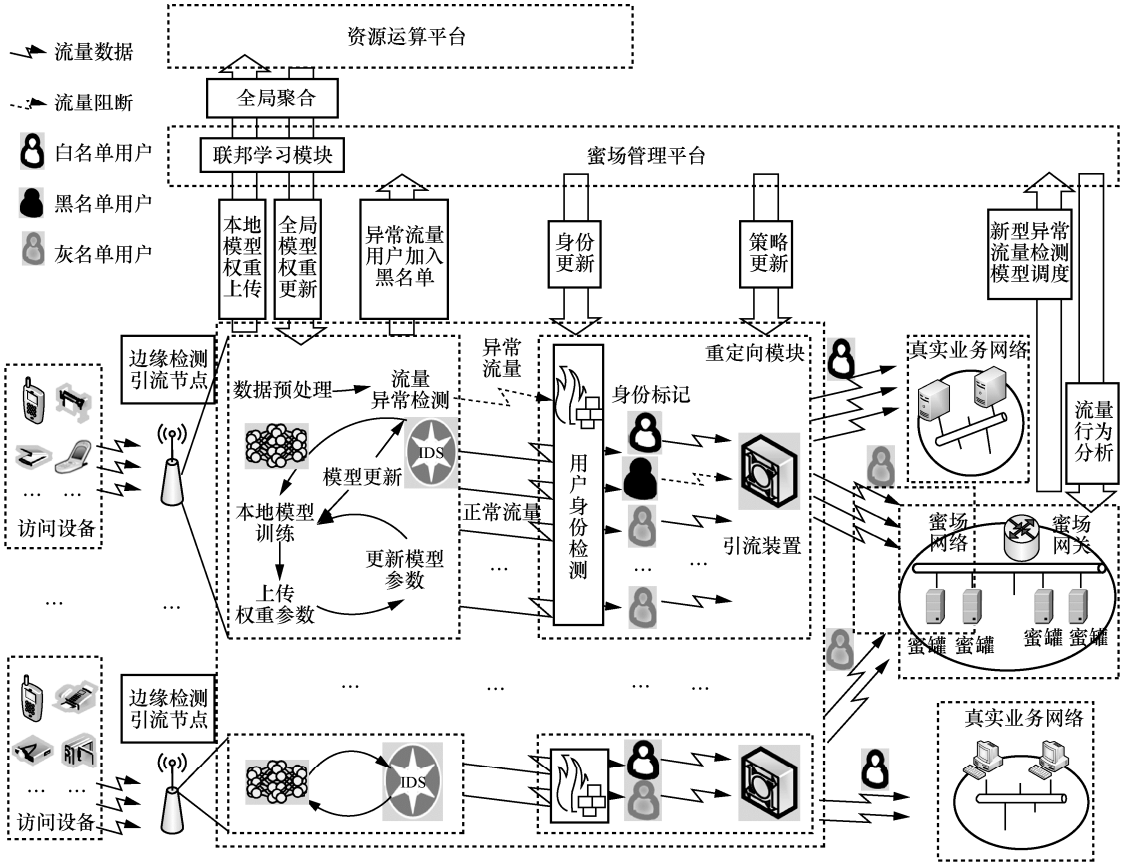


图 3　联邦学习下的流量检测模型训练方案的架构

真实业务网络，在名单之外（灰名单）的威胁未知流量 F23 将经过引流装置重定向到蜜场，以进行对新型异常流量的分析与识别。引流策略的目标是确保在流量检测中既阻止已知的异常流量，也在未知威胁方面保持高度警惕。

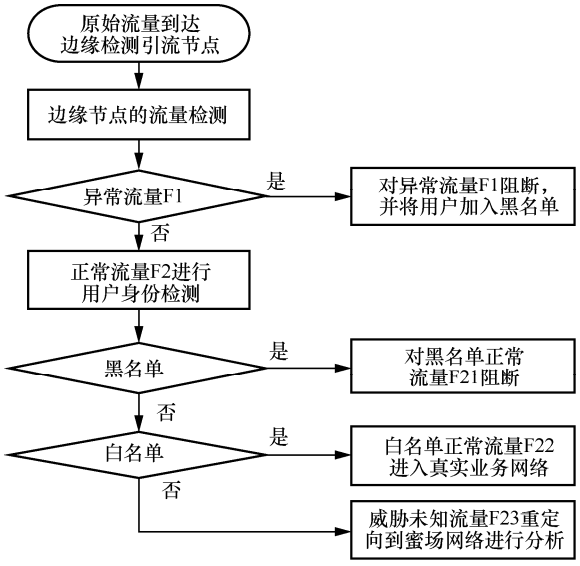


图 4　引流策略

2.3　基于蜜场的识别新型异常流量的检测模型更新方案

如图 5 所示，云端蜜场接收的流量数据都源自当前边缘节点检测模型所认定的正常流量。云端蜜场需要与这些流量数据进行交互，以剖析其中潜藏的新型攻击行为。该过程包括分析流量中的模式、异常行为以及其他潜在的异常指标。通过该分析，云端蜜场能够鉴别出属于正常流量的用户，将其记录在白名单中。与此同时，云端蜜场也会对所发现的新型异常流量进行数据标记，同时还会记录其异常来源的边缘节点。基于这些已标记出的正常与异常流量数据，蜜场进行进一步的 FedCBAM 模型训练，旨在训练一个能够识别新型异常流量的模型。这个训练过程将依赖于已知的正常流量数据和标记的异常流量数据，学习这些流量的特征和模式。

模型训练完成后，经由蜜场管理平台将模型下发到记录中存在新型异常流量的各个边缘节点。在这些节点使用该模型对本地数据进行检测。在该过程中，所有被认定为新型异常流量的样本被识别出

来后，其错误标签会被相应地修改。接着，借助联邦学习下的流量检测模型训练方案，基于这些修改后的流量数据，对各个边缘节点上的入侵检测模型 FedCBAM 进行新一轮的训练，旨在使这些本地模型能够精确地识别并处理新型攻击行为。通过该方案，本文所提系统能够拥有对新型攻击行为的及时识别和更新能力。

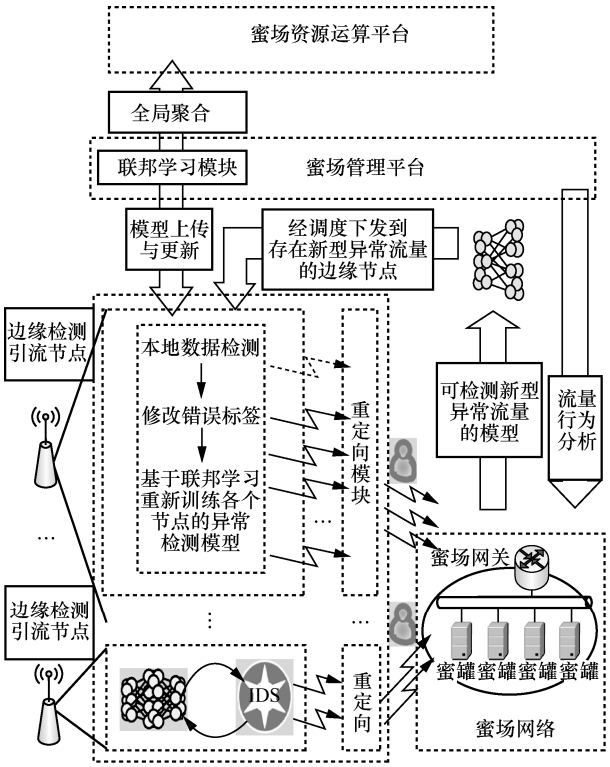


图 5　基于蜜场的识别新型异常流量检测的模型更新方案思路

3　算法设计

本文所提系统主要涵盖 2 个方案下 2 个模型的训练，然而 2 个模型的网络结构相同，主要差异在于其参数权重，因此本节将对联邦学习下的入侵检测算法设计着重论述，同时对基于蜜场的模型训练过程中与联邦学习下的入侵检测算法具有明显差异的部分进行论述。

3.1　联邦学习下的入侵检测算法设计

联邦学习下的入侵检测算法设计过程如图 6 所示，主要分为数据预处理、联邦学习模型训练、流量检测 3 个关键阶段。

3.1.1　数据预处理

进入边缘检测引流节点的数据是原始流量数据，在将这些数据用于模型训练之前，需要经过预处理过程。本文中的数据预处理过程包括以下关键步骤：特征提取、特征数值化以及灰度图转化。首先需要对原始数据进行简单特征提取，从原始数据中提取多维度的特征，这有助于将复杂的原始流量数据转化为更加有意义且易于处理的特征向量。然后需要将数据特征中存在的空值替换，并且为了处理非数值特征，本文使用 One-Hot 编码，将其转化为数值特征，从而使模型能够理解和处理这些特征。此外，为了消除奇异样本数据导致的不良影响，对数值特征进行了 min-max 归一化处理

$$x' = \frac{x - \min(x)}{\max(x) - \min(x)} \tag{3}$$

其中，x 表示特征取值，将其范围映射为 0 到 1 之间，这有助于在训练模型时提高其稳定性和其他性能。最后为了与训练模型的输入格式相匹配，将经过归一化的数据转化为适当的灰度图。

3.1.2　联邦学习模型训练

1) 训练模型结构

本文设计了一个基于 CBAM 卷积自编码器的入侵检测模型 FedCBAM，具体结构如图 7 所示。为了有效识别流量数据中的关键特征，首先，使用卷积自编码器对经预处理后生成的灰度图进行特

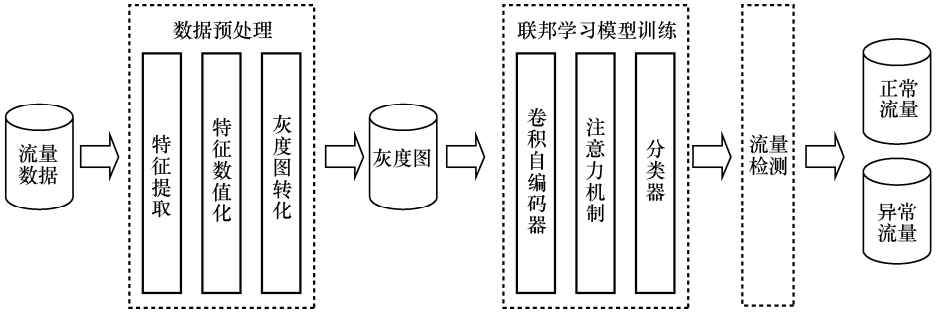


图 6　联邦学习下的入侵检测算法设计过程

征提取；然后，为了更全面地挖掘多维度流量特征，使用 CBAM 机制从通道和空间 2 个维度对关键特征进行加权；最后，使用全连接神经网络完成最终的分类任务。

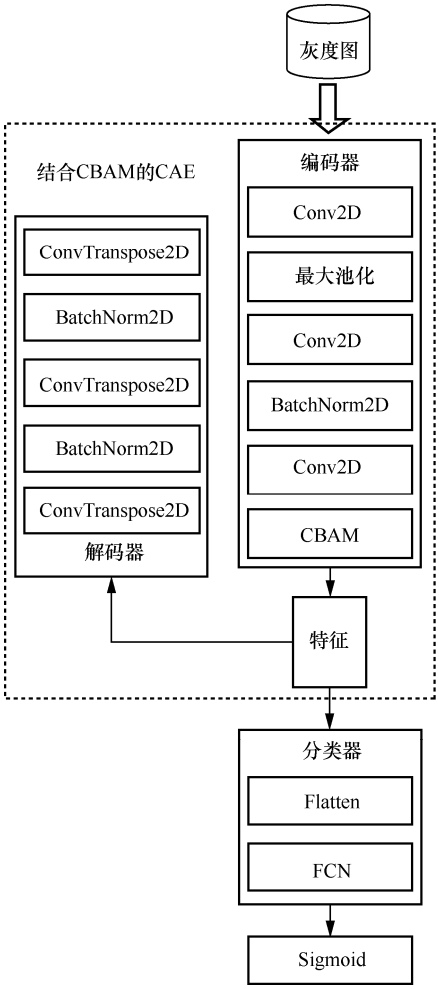


图 7 FedCBAM 结构

2) 联邦学习训练

本文进行异常流量检测模型的联邦学习过程包括 2 个关键步骤：本地训练和全局训练。本地训练算法如算法 1 所示，在每次全局模型聚合之后，得到的第 t 轮的模型参数 w_t 被用于本地训练。在本地训练时基于 w_t 更新本地模型参数，一旦本地训练完成，本地模型的参数就会被上传到资源运算平台。全局训练算法如算法 2 所示，在资源运算平台使用 FedAvg 算法对各个节点的模型参数进行聚合以生成全局模型，随后，这个全局模型会被下发到各个边缘节点上，用作下一轮的本地训练。

算法 1 FedCBAM 本地训练算法

输入 边缘节点本地训练轮数 N，边缘节点个数 K，下发的全局模型参数 w_t

输出 各个边缘节点训练模型参数 w_{t+1}^k

1) for 每个边缘节点 k do
2) $w_t^k = w_t$；//更新本地模型
3) for $t = 1,2,\cdots,N$ do
4) $w_{t+1}^k = w_t - \eta\nabla L$；
5) end for
6) 将本地模型参数 w_{t+1}^k 上传到资源运算平台；
7) end for

算法 2 FedCBAM 全局训练算法

输入 全局训练轮数 T，边缘节点个数 K，接收的本地模型参数 w_{t+1}^k，$k = 1,2,\cdots,K$

输出 全局最优模型参数 w^*

1) for $t = 1,2,\cdots,T$ do
2) 接收边缘节点上传的本地模型参数；
3) $w_{t+1} = \dfrac{1}{K}\sum_{k=1}^{K} w_{t+1}^k$；
4) 下发更新的全局模型参数 w^*；
5) end for

3.1.3 流量检测

根据算法 1 和算法 2 对 NIDS 模型进行训练。经过训练的 NIDS 模型将能够对流量数据进行异常识别，从而能够及时检测并阻断异常流量。同时，结合身份识别结果和引流策略，系统能够将未知的威胁流量重定向到蜜场，分析其中的未知攻击，进而按照所提方案提升各个边缘节点对新型异常流量的识别能力。

3.2 基于蜜场的识别新型异常流量的检测算法设计

本文中基于蜜场的入侵检测模型训练过程包括 3 个关键步骤：新型异常流量的数据集构建、数据预处理和模型训练。

首先是对新型异常流量的数据集构建。假设存在 K 个边缘检测引流节点，每个节点重定向到蜜场的未知威胁流量包都会添加标记 S_k，对应每个节点。蜜场会持续对部署的蜜罐进行监控，一旦监测到异常行为，例如频繁的扫描活动、异常的访问模式、对漏洞的利用尝试等行为，蜜场就会详细记录与异常流量相关的各种信息，使用 Zeek 等分析工具判断是否有攻击行为，最终建立包含新型攻击流量的数据集。

然后是数据的数据预处理过程，与 3.1.1 节的实现流程一致，需要完成特征提取、特征数值化以

及灰度图转化过程，为模型训练做准备。

最后在模型训练阶段，采用 3.1.2 节所述的基于 CBAM 卷积自编码器的入侵检测模型在资源运算平台进行集中训练。训练完成的模型将根据 S_k 下发到存在新型攻击流量的边缘节点，对其本地数据集的错误标签进行修改。

4　实验分析

4.1　实验设置

4.1.1　数据集设置

本节实验使用 UNSW-NB15 数据集对所提模型进行评估。该数据集是澳大利亚网络安全中心的网络靶场实验室模拟现实网络环境生成的[26-27]，主要由 47 个属性特征和 2 个类别特征组成，并且已经进行了训练集和测试集的分割。本文为了验证所提模型性能，将该数据集进行数据预处理，保留其二分类标签，并划分为异常流量和正常流量两类，标签类别具体分布如表 1 所示。另外，本文将训练数据集分成数量随机的 10 份，作为各个边缘节点的本地数据集，数据大小分布为[8 126, 8 060, 9 171, 8 932, 9 363, 7 486, 7 134, 8 260, 7 278, 7 363]。

表 1　　UNSW-NB15 标签类别具体分布

数据集	标签类别		总和
	异常流量	正常流量	
训练集	19 488	61 685	81 173
测试集	9 625	25 554	35 179
总和	29 113	87 239	116 352

4.1.2　评估指标

由于数据集的数据分布不平衡，本文采用多种评价指标，如准确度（Accuracy）、精确度（Precision）、召回率（Recall）、F1 分数（F1-score）、混淆矩阵（confusion matrix）和 ROC（receiver operating characteristic）曲线。

4.1.3　实验环境

本节实验设置每个边缘节点使用 Intel(R) Core(TM) i7-7700HQ 2.8 GHz 型号的 CPU、NVIDIA GeForce GTX 1050ti 8G 型号的 GPU 作为模型训练和检测的实验环境，所述蜜场依托软极网络的靶场平台进行搭建并完成相关工作，拓扑结构如图 8 所示。经过多次实验，最终优化的相关模型参数设置如表 2 所示。

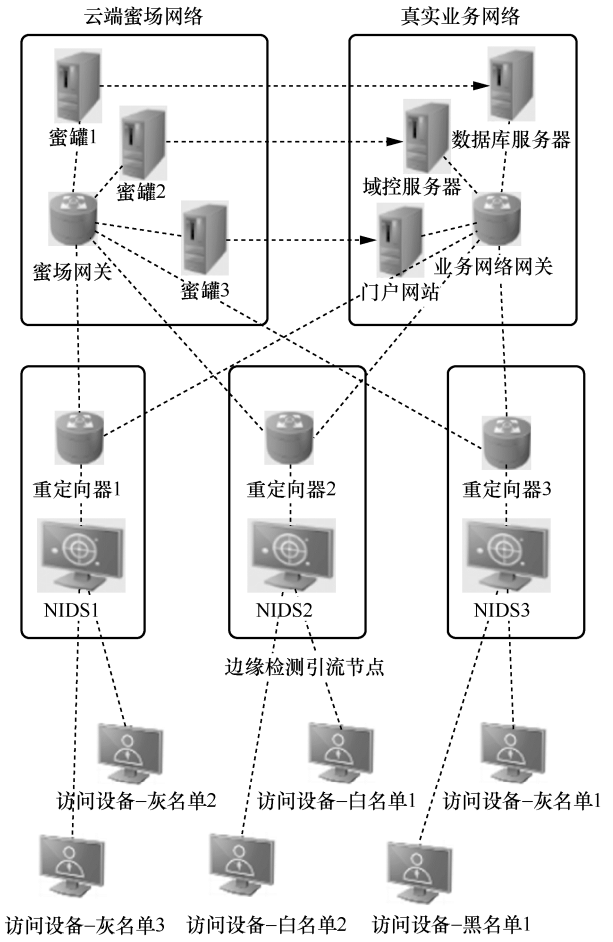


图 8　基于靶场构建的网络拓扑结构

表 2　　模型参数设置

模型参数	参数设置
批次大小	64
损失函数	二分类交叉熵
优化器	随机梯度下降
学习速率	0.005
联邦学习全局训练轮次	100
联邦学习本地训练轮次	1
节点数目	10
本地独立训练轮次	100

4.2　实验结果与分析

本节将设置 3 组对比实验对所提模型进行评估分析。

4.2.1　本地和联邦学习训练模型之间的对比

为了验证联邦学习框架对边缘节点数据的有效利用，在 10 个边缘节点独立训练模型 LocalCBAM，以及基于联邦学习训练所提模型 FedCBAM，2 个模型的各项性能指标如表 3 第 1、4 行所示，表 3 中数据是各节点模型测试得到的平均值。

表 3 各模型的各项性能指标对比

NIDS 模型	Accuracy	异常流量样本			正常流量样本		
		Precision	Recall	F1-score	Precision	Recall	F1-score
LocalCBAM	88.88%	99.93%	59.39%	74.50%	86.73%	99.98%	92.89%
FedAE	91.98%	99.98%	70.70%	82.83%	90.06%	99.99%	94.77%
FedCAE	93.19%	99.82%	75.26%	85.81%	91.47%	99.94%	95.52%
FedCBAM	94.47%	98.76%	80.81%	88.89%	93.24%	99.62%	96.32%

从表 3 可以看出，相比于 LocalCBAM，联邦学习模型 FedCBAM 的准确度有显著提升，提升约 5.6%；针对异常流量样本的检测，FedCBAM 的召回率远高于 LocalCBAM，这是因为联邦学习下的模型训练能够充分利用各个节点的数据来补充攻击流量特征，可以有效提高模型对异常流量的检测能力。值得注意的是，在本文所提基于蜜场的检测系统中，对于正常流量样本的检测，尽管 FedCBAM 的召回率略有下降，但在一定程度上，对于正常流量的误报是可以接受的，系统更关注的是避免异常流量的漏检。此外，F1 分数的提升也进一步验证了联邦学习可以有效提高模型的质量，增强边缘节点的异常流量检测能力。

FedCBAM 与 LocalCBAM 的准确度比较如图 9 所示。可以观察到，LocalCBAM 的准确度存在较大波动，其最大浮动值甚至超过了 7.6%，相比之下 FedCBAM 的波动不超过 0.34%，且整体上高于 LocalCBAM，这是因为本地数据的质量对本地训练具有显著影响，不同节点的本地数据能涵盖的流量特征的完整程度存在差异，引入联邦学习架构可以明显提高各个参与节点的模型准确性，有助于缓解本地数据质量不稳定性所带来的挑战。

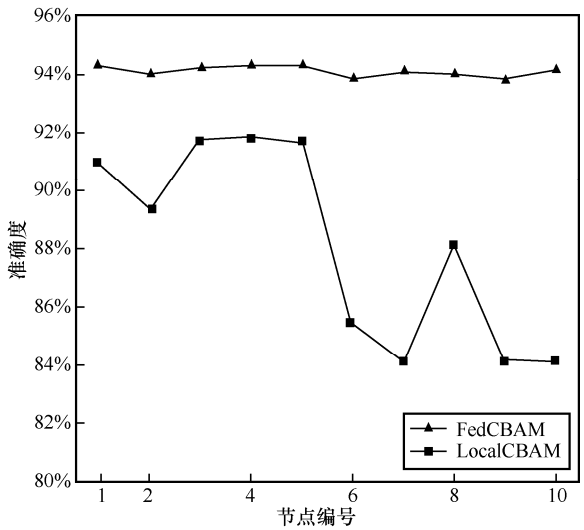


图 9 FedCBAM 与 LocalCBAM 的准确度比较

图 10 展示了在不同参与训练的节点数目下 FedCBAM 的各指标情况。观察可知，参与训练的节点数目对 FedCBAM 模型的质量产生显著影响。随着参与训练的节点数目增加，模型质量呈现逐步提升的趋势，这是因为各个节点所包含的数据呈现多样的攻击流量特征，存在不同的数据分布，而联邦学习框架的引入有助于提升模型的泛化能力，进而增强对攻击流量的识别能力。

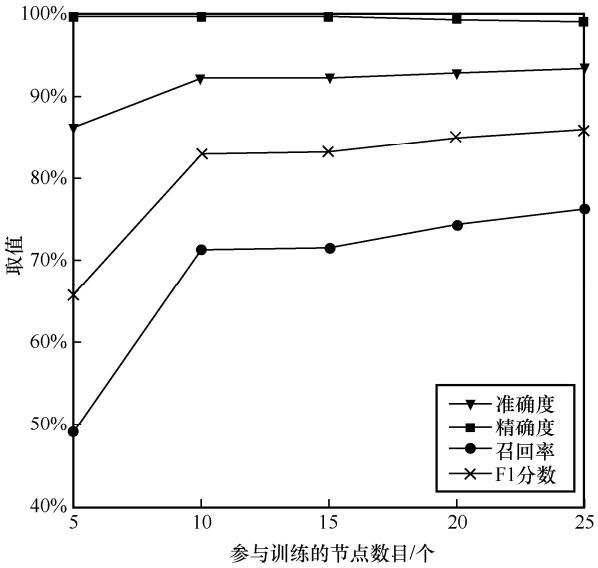


图 10 在不同参与节点数目下 FedCBAM 的各指标情况

4.2.2 FedCBAM 模型和基准模型之间的对比

本文根据文献[28]重现了基于 AE（FedAE）和卷积自编码器（FedCAE）的检测模型，并将其作为基准模型来评估所提模型 FedCBAM 的性能。FedCBAM 模型和 2 种基准模型在训练过程中的准确度变化趋势如图 11 所示，2 种基准模型和 FedCBAM 各项性能指标如表 3 第 2～4 行所示。

从图 11 可以观察到，3 种模型中，FedCBAM 模型在达到较高准确度所需的训练轮次相对较少。这一现象可能归因于 CBAM 机制对关键特征的重点关注。此外，FedCBAM 模型的曲线在达到平稳趋势后表现出最高的准确度，进一步验证了

FedCBAM 模型在性能方面的优越性。从表 3 可以明显看出，FedCBAM 模型在大多数性能指标上优于其他 2 种模型。具体而言，FedCBAM 准确度比 FedAE 高约 2.5%，比 FedCAE 高约 2.3%，达到 94.47%。这种性能提升可以归因于 CBAM 机制的引入，它增强了模型对灰度图中通道和空间特征的提取效果，从而提高了模型的学习能力；针对异常流量样本的检测，虽然 FedCBAM 模型的精确度略低于 FedAE 和 FedCAE，分别降低了 1.22%和 1.06%，但召回率明显高于 2 种基准模型，分别提高了 10.11%和 5.55%，对精确度和召回率综合考量的 F1 分数也分别提高了 6.06%和 3.08%，这表明 FedCBAM 能够更好地完成对异常流量的检测任务，能够更准确地识别出异常行为，证明了 FedCBAM 模型的质量高于基准模型。

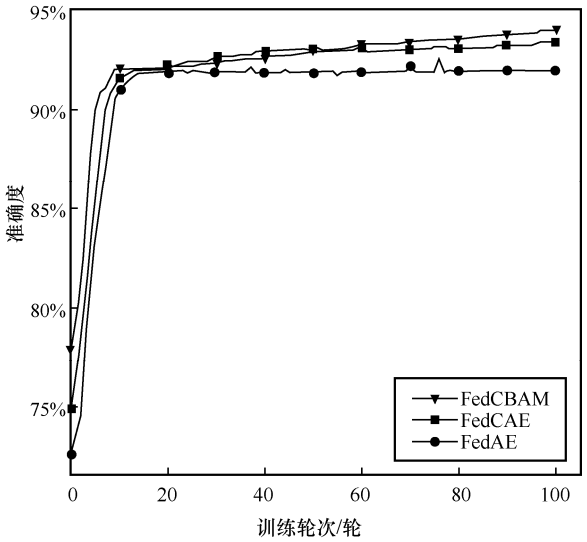


图 11　FedCBAM 模型与 2 种基准模型在训练过程中的准确度变化趋势

仅凭整体指标难以直观地判断各模型的性能差异，如图 12 所示，从 10 个节点中选其一输出其上 3 种模型测试的混淆矩阵。图 12(a)展示了 FedAE 模型的混淆矩阵，可以观察到，FedAE 模型有约 30%的异常流量被错误识别为正常流量。图 12(b)展示了 FedCAE 模型的混淆矩阵，相比于 FedAE，FedCAE 对异常流量的识别稍有改善，但仍然有 25%的异常流量被错误识别。通过对比可以观察到，FedCBAM 模型对异常流量的识别能力表现更佳，尽管对正常流量的识别率略有下降，但这相对于异常流量的检测能力显著提高而言是可以接受的。

图 13 展示了 FedCBAM 和 FedCAE 的 ROC 曲线。从图 13 可以观察到，这 2 种模型的 ROC 曲线均分布在图像的上半部分。值得注意的是，引入 CBAM 机制后，模型的 ROC 曲线得到进一步改善，其曲线下面积从 0.9887 增加至 0.9951 以上，这表明 CBAM 机制的加入进一步提升了模型的分类性能。

表 4 对本文所提模型和基准模型在模型准确度达到 91%时所用的训练时间进行对比。由表 4 可知，相比于 FedAE、FedCAE，本文所提模型 FedCBAM 的每轮平均训练时间最长，这是因为 FedCBAM 引入了 CBAM 机制，增加了模型本身的时间复杂度；但同时本文所提模型在模型准确度达到 91%时所用的训练总时间最短，相同情况下，与 FedCAE 模型相比训练总时间缩短了 35.7%，与 FedAE 模型相比训练总时间缩短了 49.9%，这是因为 FedCBAM 引入注意力机制，在训练过程中分配不同的权重给特征，提高了模型对通道和空间特征的提取效果，从而减少了达到较高准确度的训练轮次，大大缩短了训练总时间。

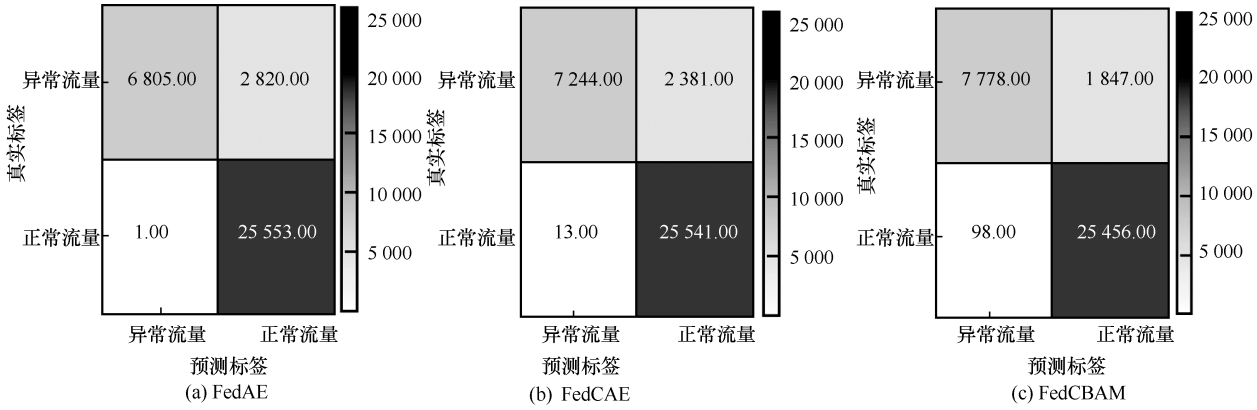


图 12　混淆矩阵

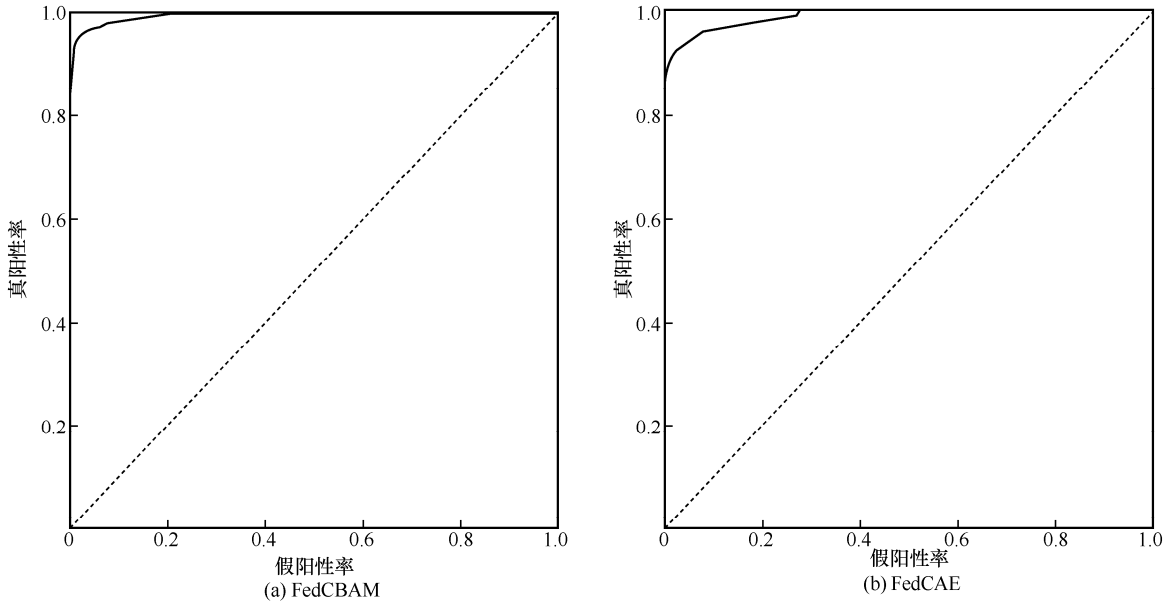


图 13　FedCBAM 和 FedCAE 的 ROC 曲线

表 4　　**各模型的训练时间对比**

模型	每轮平均训练时间/s	训练总时间/s
FedAE 模型	56.97	579
FedCAE 模型	61.58	451
FedCBAM 模型	67.65	290

4.2.3　可识别新型异常流量的现阶段模型与前一阶段模型之间的对比

为验证基于蜜场识别新型异常流量的模型更新方案的可行性，基于联邦学习在边缘节点训练 2 个模型：现阶段模型 FedCBAM-n 和前一阶段模型 FedCBAM-o。

首先，边缘节点在训练 FedCBAM-o 模型时，未能完全识别部分 Exploits 攻击类型的流量，导致这些流量被视为未知威胁流量并重定向到蜜场。在蜜场使用 Zeek 工具对日志数据进行分析的过程中，发现部分流量针对蜜罐中的已知漏洞进行攻击，并且携带特定的攻击载荷，与 Exploits 攻击的行为模式相符，因此训练可识别这些 Exploits 异常的检测模型到各个来源节点进行数据标签的更新，随后基于联邦学习重新训练了各个边缘节点上的 NIDS 模型，即 FedCBAM-n 模型。最终，使用标签更新后的测试集对这 2 个模型的性能进行测试，性能指标如表 5 所示。

通过观察实验结果可以明显看出，在修改错误标签后的训练集上训练的 FedCBAM-n 模型性能相比于更新前的 FedCBAM-o 模型有了显著提升。具体来说，FedCBAM-n 模型的准确率提高约 2.3%，而对异常流量检测的召回率则提高了约 9.2%，这些结果表明，更新后的模型能够识别更多的异常流量，从而更准确地捕捉新型攻击行为，证明了本文所提基于蜜场识别新型异常流量的模型更新方案是可行的。

5　结束语

本文提出一种能有效识别异常流量和具备持续学习能力的网络入侵检测系统，旨在解决物联网网络入侵检测系统无法识别新型攻击、灵活性有限

表 5　　**更新前后 2 个模型的各项性能指标对比**

NIDS 模型	Accuracy	攻击样本			正常样本		
		Precision	Recall	F1-score	Precision	Recall	F1-score
FedCBAM-o 模型	92.19%	99.79%	71.59%	83.37%	90.33%	99.94%	94.89%
FedCBAM-n 模型	94.47%	98.76%	80.81%	88.89%	93.24%	99.62%	96.32%

等问题。考虑到边缘节点的数据孤岛问题，本文设计了联邦学习下的流量检测模型训练方案和基于蜜场的识别新型异常流量检测的模型更新方案，并且结合卷积块注意力模块的特点，构建专注于通道和空间双维度的异常流量检测模型 FedCBAM。实验结果表明，联邦学习提高了模型的稳健性，相比于 FedAE、FedCAE，本文所提模型对异常流量的识别能力更强，并且可以持续提高对新型异常流量的检测性能。未来研究工作中，将细化入侵检测系统的攻击类型，并对蜜场的识别分析过程进行深入研究。

参考文献：

[1] SISINNI E, SAIFULLAH A, HAN S, et al. Industrial Internet of things: challenges, opportunities, and directions[J]. IEEE Transactions on Industrial Informatics, 2018, 14(11): 4724-4734.

[2] GATOUILLAT A, BADR Y, MASSOT B, et al. Internet of medical things: a review of recent contributions dealing with cyber-physical systems in medicine[J]. IEEE Internet of Things Journal, 2018, 5(5): 3810-3822.

[3] ARASTEH H, HOSSEINNEZHAD V, LOIA V, et al. IoT-based smart cities: a survey[C]//Proceedings of 2016 IEEE 16th International Conference on Environment and Electrical Engineering (EEEIC). Piscataway: IEEE Press, 2016: 1-6.

[4] ALAA M, ZAIDAN A A, ZAIDAN B B, et al. A review of smart home applications based on Internet of things[J]. Journal of Network and Computer Applications, 2017, 97: 48-65.

[5] SCHILLER E, AIDOO A, FUHRER J, et al. Landscape of IoT security[J]. Computer Science Review, 2022, 44: 100467.

[6] NGUYEN X H, NGUYEN X D, HUYNH H H, et al. Realguard: a lightweight network intrusion detection system for IoT gateways[J]. Sensors, 2022, 22(2): 432.

[7] CAO B, LI C, SONG Y, et al. Network intrusion detection model based on CNN and GRU[J]. Applied Sciences, 2022, 12(9): 4184.

[8] MUHAMMAD G, HOSSAIN M S, GARG S. Stacked autoencoder-based intrusion detection system to combat financial fraudulent[J]. IEEE Internet of Things Journal, 2023, 10(3): 2071-2078.

[9] SABIR M, AHMAD J, ALGHAZZAWI D. A lightweight deep autoencoder scheme for cyberattack detection in the Internet of things[J]. Computer Systems Science and Engineering, 2023, 46(1): 57-72.

[10] YAO W, HU L, HOU Y, et al. A lightweight intelligent network intrusion detection system using one-class autoencoder and ensemble learning for IoT[J]. Sensors, 2023, 23(8): 4141.

[11] KHAN F A, GUMAEI A, DERHAB A, et al. A novel two-stage deep learning model for efficient network intrusion detection[J]. IEEE Access, 2019, 7: 30373-30385.

[12] BASATI A, FAGHIH M M. PDAE: efficient network intrusion detection in IoT using parallel deep auto-encoders[J]. Information Sciences, 2022, 598: 57-74.

[13] SHONE N, NGOC T N, PHAI V D, et al. A deep learning approach to network intrusion detection[J]. IEEE Transactions on Emerging Topics in Computational Intelligence, 2018, 2(1): 41-50.

[14] MOTHUKURI V, PARIZI R M, POURIYEH S, et al. A survey on security and privacy of federated learning[J]. Future Generation Computer Systems, 2021, 115: 619-640.

[15] HUONG T T, BAC T P, LONG D M, et al. Detecting cyberattacks using anomaly detection in industrial control systems: a federated learning approach[J]. Computers in Industry, 2021, 132: 103509.

[16] QIN Y, KONDO M. Federated learning-based network intrusion detection with a feature selection approach[C]//Proceedings of 2021 International Conference on Electrical, Communication, and Computer Engineering (ICECCE). Piscataway: IEEE Press, 2021: 1-6.

[17] CETIN B, LAZAR A, KIM J, et al. Federated wireless network intrusion detection[C]//Proceedings of 2019 IEEE International Conference on Big Data (Big Data). Piscataway: IEEE Press, 2019: 6004-6006.

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

[19] CHOLAKOSKA A, PFITZNER B, GJORESKI H, et al. Differentially private federated learning for anomaly detection in e-health networks[C]//Proceedings of the 2021 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2021 ACM International Symposium on Wearable Computers. New York: ACM Press, 2021: 514-518.

[20] AMAL M R, VENKADESH P. H-doctor: honeypot based firewall tuning for attack prevention[J]. Measurement: Sensors, 2023, 25: 100664.

[21] PASHAEI A, AKBARI M E, ZOLFY L M, et al. Early intrusion detection system using honeypot for industrial control networks[J]. Results in Engineering, 2022, 16: 100576.

[22] ELLOUH M, GHALEB M, FELEMBAN M. IoTZeroJar: towards a honeypot architecture for detection of zero-day attacks in IoT[C]// Proceedings of 2022 14th International Conference on Computational Intelligence and Communication Networks (CICN). Piscataway: IEEE Press, 2022: 765-771.

[23] ALMOHANNADI H, AWAN I, AL HAMAR J, et al. Cyber threat intelligence from honeypot data using elasticsearch[C]//Proceedings of

2018 IEEE 32nd International Conference on Advanced Information Networking and Applications (AINA). Piscataway: IEEE Press, 2018: 900-906.

[24] ALBASEER A, ABDALLAH M. Fine-tuned LSTM-based model for efficient honeypot-based network intrusion detection system in smart grid networks[C]//Proceedings of 2022 5th International Conference on Communications, Signal Processing, and their Applications (ICCSPA). Piscataway: IEEE Press, 2022: 1-6.

[25] WOO S, PARK J, LEE J-Y, et al. CBAM: convolutional block attention module[J]. arXiv Preprint, arXiv:1807.06521, 2018.

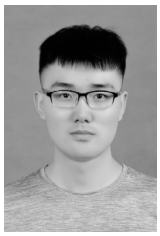
[26] MOUSTAFA N, SLAY J. UNSW-NB15: a comprehensive data set for network intrusion detection systems (UNSW-NB15 network data set)[C]//Proceedings of 2015 Military Communications and Information Systems Conference (MilCIS). Piscataway: IEEE Press, 2015: 1-6.

[27] MOUSTAFA N, SLAY J. The evaluation of network anomaly detection systems: statistical analysis of the UNSW-NB15 data set and the comparison with the KDD99 data set[J]. Information Security Journal: A Global Perspective, 2016, 25(1-3): 18-31.

[28] WANG Z, WANG P, SUN Z. SDN traffic anomaly detection method based on convolutional autoencoder and federated learning[C]// Proceedings of 2022 IEEE Global Communications Conference. Piscataway: IEEE Press, 2022: 4154-4160.

[作者简介]

吴昊（1973− ），女，江苏常熟人，博士，北京交通大学教授、博士生导师，主要研究方向为云边协同智能、异构网络安全等。



郝佳佳（1999− ），男，河南洛阳人，北京交通大学硕士生，主要研究方向为主动防御技术、物联网安全等。

卢云龙（1991− ），男，山东日照人，博士，北京交通大学副教授、硕士生导师，主要研究方向为移动通信、边缘智能、网络安全等。