引用格式:赵英, 王丽宝, 陈骏君, 等. 基于联邦学习的网络异常检测[J]. 北京化工大学学报(自然科学版),2021,48(2):92-99.

ZHAO Ying, WANG LiBao, CHEN JunJun, et al. Network anomaly detection based on federated learning[J]. Journal of Beijing University of Chemical Technology (Natural Science), 2021,48(2):92-99.

基于联邦学习的网络异常检测

赵 英 王丽宝 陈骏君 滕 建

(北京化工大学 信息科学与技术学院, 北京 100029)

摘 要: 作为一类网络安全的基础研究,网络异常检测技术目前还存在检测准确率低、误报率高以及缺乏标签数据等问题。为此提出一种融合联邦学习和卷积神经网络的网络入侵检测分类模型(CNN-FL),可有效解决多个参与者在不共享隐私数据的情况下进行一个全局模型的协作训练时所带来的问题。该模型无需汇集模型训练所需要的数据进行集中计算,只是传递加密的梯度相关数据,即可利用多源数据协同训练同一模型,并解决缺乏标签数据的问题。随后将该模型应用于二分类和多分类方法中,并在同一基准数据集 NSL-KDD 上进行了实验比较与分析,实验结果表明,与其他研究方法相比,所提 CNN-FL 分类模型在二分类以及多分类中具有较高的识别性能和分类精度。

关键词: 联邦学习; 网络异常检测; 深度学习; 卷积神经网络(CNN)

中图分类号: TP393 **DOI**: 10.13543/j.bhxbzr.2021.02.012

引 言

随着网络信息技术的迅猛发展,互联网已成为人们日常工作和生活中必不可少的一部分,为人们带来了极大的便利,但同时也时刻威胁着人们的财产与信息安全,因此,进行网络信息安全研究具有重要意义。目前,作为网络安全研究的一个重要方向,网络入侵检测方法[1]已经成为网络安全技术领域研究的热点。

近年来,各种机器学习技术被应用于网络异常检测领域[2-4],但由于传统机器学习需要人工选择特征,存在特征选择困难的问题,需要进行多次测试才能获取分类效果最佳的数据特征组合。深度学习技术是目前解决这一问题的最有效的一种途径。该技术能够自动学习原始数据中的特征,不需要进行人工选择,在自然语言处理、图像识别以及语音识别等领域都显示出较为优秀的识别分类性能[5]。因此,这一技术被越来越多的研究者应用到网络异常检测模型中,并获得了较好的效果[6-8]。

收稿日期: 2020-04-29

第一作者: 男,1966 年生,教授,博士生导师

E-mail: zhaoy@mail.buct.edu.cn

目前,缺乏标记数据是网络异常检测面临的重大挑战之一。如何利用不同数据源的网络流量数据来共同训练网络异常流量检测模型并保护数据隐私是一个亟待解决的问题。联邦学习(federated learning)[9]是解决多源数据共同训练模型的一种有效途径,这一概念是由 Bernd 等[10]最先提出的。联邦学习的宗旨是在不共享隐私数据的情况下进行协同训练,其不用汇聚模型训练所需要的数据进行集中计算,只是传递加密的梯度相关数据,利用多源数据协同训练同一模型[11]。鉴于传统网络异常检测模型存在的检测准确率低、误报率高以及缺乏标签数据等问题,本文提出一种融合联邦学习和卷积神经网络的网络入侵检测分类模型(CNN-FL)来检测网络异常。该模型能够利用多源数据协同训练同一模型,解决数据孤岛以及标签数据缺乏的问题;并在 NSL-KDD 数据集上进行实验验证,结果表明,本文模型在二分类以及多分类实验中具有较高的分类精度,较传统的机器学习方法具有更好的分类性能。

1 联邦学习与卷积神经网络模型

1.1 联邦学习

联邦学习作为一项人工智能技术,其设计目的主要是在保障数据交换的同时确保信息安全,保护

个人数据隐私。联邦学习解决了多计算节点在不交换原始数据的情况下，共同训练全局模型的问题。

在联邦学习中，全局模型通过分布式的方式在大量参与者中进行训练。为了避免服务器访问本地数据，参与者只在本地训练模型，并且只与服务器共享模型参数来更新全局模型，这一技术对于许多分布式学习场景来说具有很大的优势。典型的联邦学习体系结构如图 1 所示。假设在联邦学习中有 K 个具有相同目标的参与者联合训练一个模型，在每一次迭代时，服务器将全局模型 M 分发给参与者，参与者通过本地数据单独训练模型。在本地训练完成后，每个参与者将模型参数发送回服务器，服务器通过平均各参与者的模型参数来更新全局模型。全局模型的更新过程如式(1)所示。

$$M_{t+1} = M_t + \alpha \frac{1}{K} \sum_{k=1}^{K} G_t^k \tag{1}$$

式中，α 为更新权重，M_{t+1} 为 $t+1$ 次迭代时的全局模型，G_t^k 为迭代 t 次时第 k 位参与者上传的模型参数。

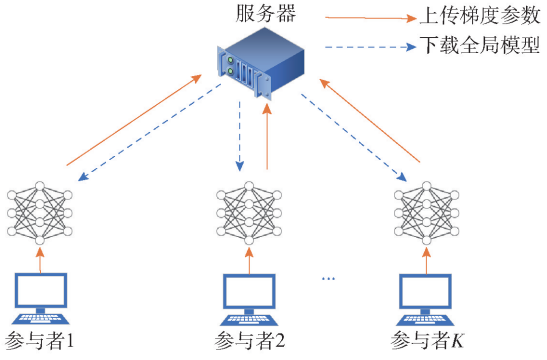

图 1　联邦学习体系结构

Fig. 1　Federated learning architecture

1.2　卷积神经网络

卷积神经网络(CNN)的概念最先是由 LeCun 等[12]提出的，LeCun 首次将反向传播技术应用于神经网络，并将其命名为卷积神经网络。但由于当时计算能力有限，无法满足反向传播技术所需要的巨大计算量，导致对卷积神经网络的研究一直停滞不前。随着半导体技术的飞速发展，计算能力不断提高，基于卷积神经网络的图像识别算法在各类竞赛中均取得了较好的识别效果，由此卷积神经网络技术逐渐被人们所熟知。近年来，越来越多的研究者将卷积神经网络技术应用于图像和语音识别领域，并取得了显著的研究成果[13-14]。同时，这一技术也被越来越多的公司应用到最新的研发产品中，包括 Google 的 Google Net 以及人工智能领域极为热门的 Alpha Go。

图 2 为 CNN 的网络结构图。CNN 模型通过卷积和池化操作对相邻像素点进行处理，为了增强图片信息的连续性，CNN 模型只处理图片中每一块的小像素集，不再单独对每一个像素点进行处理。在图像识别中，CNN 模型能够去除图片中大量无关的参数，保留图片中较为关键的数据特征，以获取较好的识别效果。CNN 方法处理过程如式(2)所示，首先将输入的初始图像与线性滤波器进行卷积运算，然后加上偏置项，最后再经过激活函数来获取特征图。

$$h_{ij}^k = \tanh((W^k * \boldsymbol{x})_{ij} + b_k) \tag{2}$$

式中，h^k 为给定层上第 k 个特征映射，W^k 为滤波权重，$\boldsymbol{x}$ 为给定灰度图中相应区域的像素值矩阵，b_k 为偏置，tanh 为激活函数。

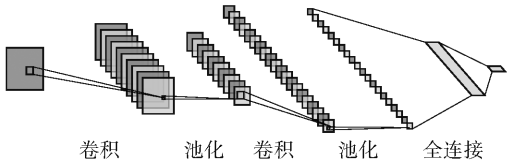

图 2　卷积神经网络结构

Fig. 2　Convolutional neural network structure

1.3　基于 CNN 的网络异常检测方法

在网络异常检测领域，CNN 模型在获取局部特征以及处理具有统计平稳性和局部关联性的数据方面较其他机器学习方法具有更加优良的特性[15]。CNN 模型通常是由输入层、卷积层、池化层、全连接层以及输出层这 5 部分组成。基于 CNN 的网络异常检测模型原理图如图 3 所示。

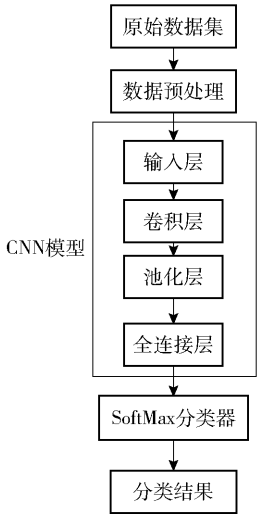

图 3　基于 CNN 的网络异常检测模型原理图

Fig. 3　Block diagram of the CNN-based network intrusion detection system

2 基于 CNN-FL 的网络异常检测方法

为了优化网络异常检测模型,提高网络异常检测模型的准确率,降低误报率,同时为了解决缺乏标签的训练数据的问题,本文提出一种融合联邦学习和卷积神经网络的网络入侵检测模型。使用联邦学习来解决数据稀缺问题并保护用户数据隐私,使多位参与者在不共享隐私数据的情况下协作训练一个全局模型。对于每一位参与者,首先需要对本地数据进行预处理,然后利用 CNN-FL 模型进行特征提取。每一位参与者与服务器只传递加密的梯度相关数据,最后通过 SoftMax 分类器获得分类结果。该算法的原理图见图 4。

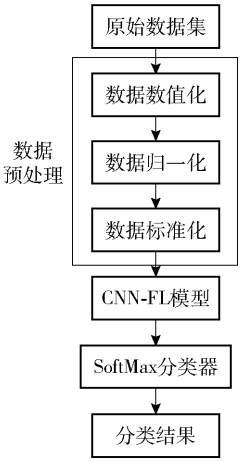

图 4 基于 CNN-FL 的网络入侵检测原理框图

Fig. 4 Block diagram of the CNN-FL-based network intrusion detection principle

如图 4 所示,每一位模型参与者首先对本地原始数据集进行数值化操作,通过 one-hot 编码方式将字符型特征转换为数值型特征,然后采用 min-max 方法对数据集以列为单位进行标准化处理,使数据统一映射到[0,1]区间上,之后将处理后的特征映射至矩阵中并生成灰度图,最后通过 CNN-FL 模型对特征进行提取,并通过 SoftMax 分类器获取分类结果。

对于每位联邦学习的参与者,首先需要对本地原始网络数据进行预处理,并通过数值化、标准化操作将原始数据整理成标准数据格式,然后再通过 CNN-FL 模型进行训练。CNN-FL 模型结构图如图 5 所示。在实际训练中,参与者与服务器只交换加密的梯度相关系数,每位参与者都是独立且平等的个体,并且本文是基于各参与者训练数据规模均等或相差较小的情况进行研究,因此服务器对各参与者上传的模型参数进行算术平均操作。训练过程如算法 1 所示。其中,K 代表 K 位参与者;ω_t 代表迭代 t 次时的全局模型参数;ω_t^k 代表第 k 位参与者在迭代 t 次时的模型参数;η 代表学习率;X^k 代表第 k 位参与者的训练数据集。

算法 1 CNN-FL 模型训练

for Iteration t do

/ * 服务器端: * /

$\omega_t = \frac{1}{K}\sum_{k=1}^{K}\omega_t^k$;

send ω_t to each participant;

/ * 参与者 * /

for Participant k do

$\omega_t^k = \omega_t$;

for Local epoch e do

$\omega_t^k = \omega_t^k - \eta\frac{\partial}{\partial X^k}\mathscr{L}$;

end

end

end

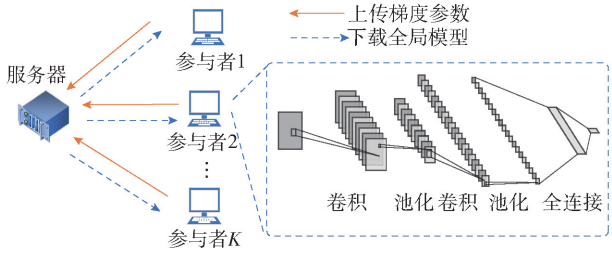

图 5 CNN-FL 模型结构

Fig. 5 CNN-FL model structure

如图 5 和算法 1 所示,每位模型参与者使用本地数据集训练 CNN 模型,在每次迭代过程中,每位模型参与者首先将当前的模型梯度相关系数上传至服务器,服务器通过平均每位参与者最新的梯度相关系数来更新全局模型,每一位参与者在下一次迭代中通过下载最新的全局模型参数,并利用本地数据来训练 CNN 模型。不断循环迭代,直至整体模型达到最优,使得 CNN-FL 模型对每一位模型参与者本地的数据集都具有较好的检测效果。

假设当前 CNN-FL 模型有 100 位参与者,迭代 50 次后全局模型参数为 ω_{50},则第 51 次迭代时,每一位参与者初始本地模型参数 $\omega_{51}^k = \omega_{50}$($k$ 的取值区间为[1,100]),第 51 次迭代完成后,全局模型更

新为 $\omega_{51}=\frac{1}{100}\sum_{k=1}^{100}\omega_{51}^{k}$。

3　实验结果与分析

3.1　数据集与环境

目前在网络检测和网络攻击领域，通常采用 KDDcup99 数据集作为算法和模型的测试与评价标准。鉴于 KDDcup99 数据集中存在包含大量重复数据以及未区分训练集与测试集等问题，本文采用 NSL-KDD 数据集作为实验数据集。NSL-KDD 数据集针对 KDDcup99 数据集中存在的问题进行的一系列优化如下。

1）NSL-KDD 数据集针对 KDDcup99 数据集中需要人为划分训练集和测试集的问题，区分了训练集和测试集。

2）NSL-KDD 训练数据集整理并清除了 KDDcup99 数据集中冗余的部分。

3）NSL-KDD 数据集数据量的大小更加合理，训练集中共有 125 973 条数据，测试集中共有 22 544 条数据。

4）NSL-KDD 数据集较其他数据集更能体现出网络异常检测模型的泛化能力。在 NSL-KDD 数据集中，训练集中存在的攻击类型有 22 种，测试集中存在的攻击类型有 39 种，有 17 种网络攻击类型在训练集中是不存在的，因此网络异常检测模型在 NSL-KDD 数据集中识别效果的好坏能更好地体现出模型是否具有较强的泛化能力。

在 NSL-KDD 数据集中，每一条网络流量数据均由 42 维特征组成，其中包括 38 维数值型特征、3 维字符型特征以及 1 维标记特征。NSL-KDD 数据集中的攻击类型分为 Dos、Probe、R2L 和 U2R 这 4 种类型。本文实验中训练集及测试集的数据类别、数量与比例如表 1、2 所示，实验环境配置如表 3 所示。

表 1　训练集的类别、数量与比例

Table 1　Types, quantities and proportions of the training set

类别	数量	比例/%
Normal	67 343	53. 46
Dos	45 927	36. 46
Probe	11 656	9. 25
R2L	995	0. 79
U2R	52	0. 04

表 2　测试集的类别、数量与比例

Table 2　Types, quantities and proportions of the test set

类别	数量	比例/%
Normal	9 711	43. 08
Dos	7 458	33. 08
Probe	2 421	10. 74
R2L	2 754	12. 22
U2R	200	0. 89

表 3　实验环境配置

Table 3　Experimental environment configuration

项目	环境配置
操作系统	Windows10
CPU	Intel Covei7-7700
内存	16 GB
编程语言	Python3. 6
深度学习框架	Pytorch1. 3. 1

3.2　评价指标

本文通过准确率、精确率和召回率等网络异常检测中常用的指标对实验结果进行评价分析，这些指标可以用真阳性(TP)、假阳性(FP)、真阴性(TN)和假阴性(FN)4 个度量标准来表示。

真阳性(TP)：分类结果属于 i 预测的结果也属于 i。

假阳性(FP)：分类结果不属于 i 预测的结果属于 i。

真阴性(TN)：分类结果属于 i 预测的结果不属于 i。

假阴性(FN)：分类结果不属于 i 预测的结果不属于 i。

准确率 A、精确率 P 和召回率 R 的定义如式(3)～(5)所示。

$$A=\frac{n_{\mathrm{TP}}+n_{\mathrm{TN}}}{n_{\mathrm{TP}}+n_{\mathrm{FP}}+n_{\mathrm{TN}}+n_{\mathrm{FN}}} \tag{3}$$

$$P=\frac{n_{\mathrm{TP}}}{n_{\mathrm{TP}}+n_{\mathrm{FP}}} \tag{4}$$

$$R=\frac{n_{\mathrm{TP}}}{n_{\mathrm{TP}}+n_{\mathrm{FN}}} \tag{5}$$

式中，n_{TP}为属于真阳性情况的数据条数，n_{TN}为属于真阴性情况的数据条数，n_{FP}为属于假阳性情况的数据条数，n_{FN}为属于假阴性情况的数据条数。

3.3　实验方案

为了测试所提出模型，我们模拟了 CNN-FL 模

型训练所需要的环境。实验步骤如下。

1) 数据扩充　NSL-KDD 数据集中有 125 973 条训练数据,为了便于拆分以供 K 位参与者训练模型使用,补充 27 条正常流量数据,使训练数据扩展到 126 000 条。

2)数据数值化　由于 NSL-KDD 数据集中存在字符型特征,需要对其进行数值化操作。数据集每一条网络流量数据均包含 3 维字符型特征,分别为“protocol_type”、“service”和“flag”,需要对这 3 维特征进行 one-hot 编码,将字符型特征转换为数值型特征,以其中的“flag”为例,其对应的 one-hot 编码如表 4 所示。

表 4　Flag 属性 one-hot 编码

Table 4　Flag attribute one-hot encoding

状态	编码
SH	(00000000001)
SF	(00000000010)
S3	(00000000100)
S2	(00000001000)
S1	(00000010000)
S0	(00000100000)
RSTR	(00001000000)
TSTOS0	(00010000000)
RSTO	(00100000000)
REJ	(01000000000)
OTH	(10000000000)

3)数据标准化　由于 NSL-KDD 数据集中数值差异较大,因此对训练集以及测试集均采 min-max 方法,以列为单位进行标准化处理,使数据统一映射到[0,1]区间上。

4)生成图片　为了便于后续使用卷积神经网络处理数据,将经过步骤 3)处理后的 122 维特征映射到 12 × 12 的矩阵中,不足的部分用 0 填充。为了生成 12 × 12 的灰度图,需要将矩阵中的数值乘以 255,从而获得图片中各个点的像素值。图 6 为生成的部分样本图片。

5)数据拆分　将步骤 4)中生成的图片随机均匀分成 K 份,用于 K 位参与者训练模型使用。

6)模型训练　在实际训练中,参与者与服务器只交换加密的梯度相关系数。

本文将从以下两个方面对 CNN-FL 模型在网络异常检测中的可行性进行验证。

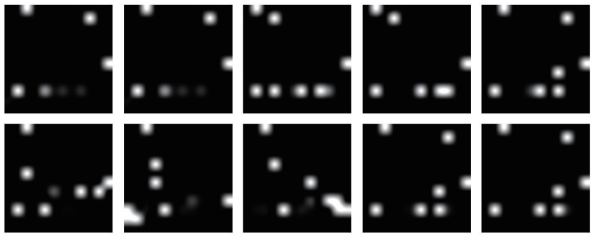

图 6　部分样本图片

Fig. 6　Some sample pictures

(1)从整体角度出发　以网络异常检测二分类为例设置了 6 个不同的场景 K {5,10,20,50,100,1 000}来研究不同参与者模型训练的准确率并与基于卷积神经网络的网络异常检测模型(K = 1)进行对比,详细结果如图 7 所示。

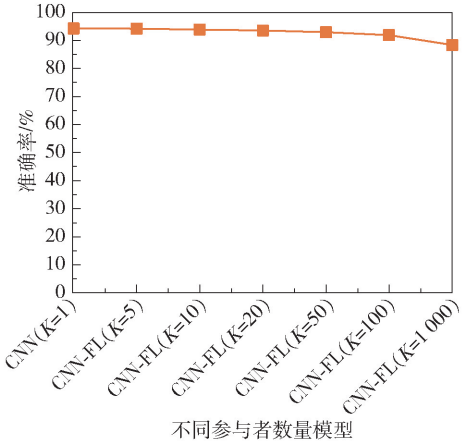

图 7　不同参与者模型准确率

Fig. 7　Model accuracy for different participants

从图 7 可以明显看出,随着参与者数量的不断增加,CNN-FL 模型的准确率逐步下降。与 CNN 模型相比,CNN-FL 模型准确率虽有所下降,但准确率基本相近(迭代 100 次后,CNN 模型的准确率为 94.26%,CNN-FL 模型(K = 100)的准确率为 91.88%)。同时 CNN-FL 模型解决了缺乏标注的训练数据的问题,该模型能够使用多源数据训练同一模型,并保护数据隐私。因此,CNN-FL 模型在网络异常检测中是可行的。

(2)从个体角度出发　将数据集中的数据随机均匀分成 K 份,分别代表每位用户所拥有的数据集。同样设置 6 个不同的场景 K {5,10,20,50,100,1 000},以二分类为例研究在不同数据规模的情况下,每位用户仅使用本地数据训练的 CNN 网络检测模型的识别准确率,并取均值。随后与相同场景下基于联邦学习的检测模型的识别准确率进行对比,

结果如图 8 所示。

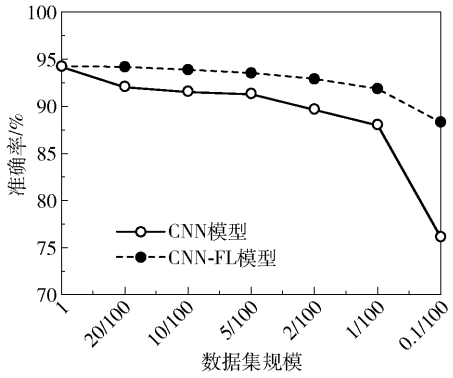

图 8　不同数据规模准确率对比

Fig. 8　Comparison of accuracy of different data scales

K 值越大，代表每位用户拥有的数据集规模越小。图 8 可以明显反映出随着每位用户拥有的数据集规模不断减小，其使用 CNN-FL 模型以及仅使用本地数据训练的 CNN 网络的检测模型准确率均不断下降。但在同等数据集规模的情况下，用户使用 CNN-FL 模型的识别准确率要高于仅使用本地数据训练的 CNN 网络检测模型的识别准确率。在多分类情况下，由于某些攻击类型的数据较少，每位用户本地数据集中该类型数据较少或不存在该类型数据，如果仅使用本地数据集训练模型，会造成模型识别准确率较低甚至无法训练模型的问题，对比效果将会更加明显。因此，在相同数据规模下，用户通过使用 CNN-FL 模型能够获得更好的识别效果，充分验证了该模型在入侵检测领域的可行性。

3.4　绩效评估

本文设计了两个实验来研究 CNN-FL 模型（$K=100$）的二分类（Normal，Anomaly）和五分类（Normal，Dos，Probe，R2L 和 U2R）性能。为了与其他机器学习方法进行比较，同时还设计了对比实验，将 CNN-FL 模型的分类性能与 C4.5 决策树、随机森林、随机树、支持向量机、循环神经网络等机器学习方法进行了对比。

3.4.1　二分类

表 5 显示了二分类实验中测试集上 CNN-FL 模型的分类效果。实验表明，经过 100 次迭代后，CNN-FL 模型具有较高的检测准确率，训练集中准确率为 97.44%，测试集中准确率为 91.88%，如图 9 所示。

与之前研究人员提出的C4.5决策树、随机森林、随机树、支持向量机、循环神经网络等方法在同一基准数据集（NSL-KDD）上进行比较的结果如图 10 所示。很明显，在二分类实验中，CNN-FL 模型的各项性能均优于其他分类算法。

表 5　二分类实验中 CNN-FL 模型的分类效果

Table 5　Classification effect of the CNN-FL model in the binary classification experiment

正确类别	预测结果	
	Anomaly	Normal
Anomaly	11 507	1 326
Normal	505	9 206

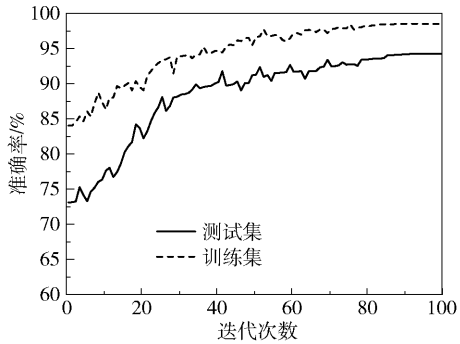

图 9　二分类模型在训练集、测试集上的检测准确率

Fig. 9　Detection accuracy of the binary classification model for the training set and test set

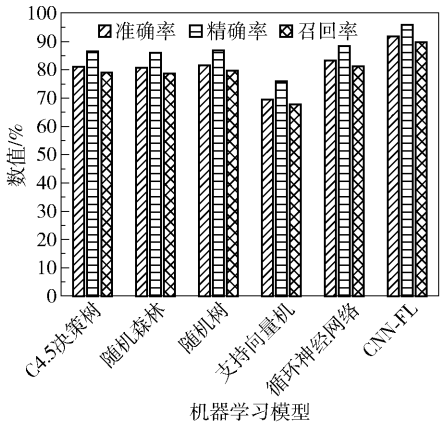

图 10　二分类各模型性能对比

Fig. 10　Performance comparison chart of each model in the binary classification test

3.4.2　五分类

在五分类实验中，CNN-FL 网络检测模型在训练集上的准确率达到 97.47%，在测试集上的准确率达到 82.40%。CNN-FL 模型在测试集上的效果如表 6 所示。表 7 显示了不同攻击类型的检测精确

率和召回率。

表 6 五分类实验中 CNN-FL 模型的分类效果

Table 6 Classification effect of the CNN-FL model in the five-category experiment

正确类别	预测结果				
	Normal	Dos	Probe	R2L	U2R
Normal	9 278	152	267	4	10
Dos	938	6 398	93	8	21
Probe	216	1 07	2 095	3	0
R2L	1 969	10	0	772	3
U2R	152	0	9	5	34

表 7 不同攻击类型的检测精确率与召回率

Table 7 Detection accuracy and recall ratio of different attack types

类型	精确率/%	召回率/%
Dos	95. 97	85. 79
Probe	85. 02	86. 53
R2L	97. 47	38. 95
U2R	50. 00	17. 00

如图 11 所示,CNN-FL 模型的检测准确率较 C4. 5 决策树、随机森林、随机树、支持向量机、循环神经网络等分类算法所获得的准确率要高。由于数据集中样本分布不均匀,与二分类相比,五分类的模型检测准确率有所下降。

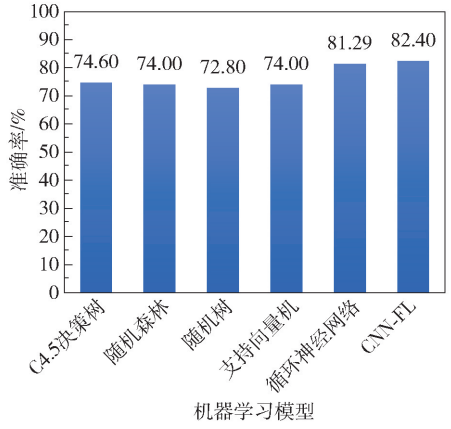

图 11 五分类各模型准确率

Fig. 11 Accuracy of each model in the five classification test

4 结束语

针对目前网络异常检测技术还存在着检测准确率低、误报率高以及缺乏标签数据等问题,本文提出了 CNN-FL 网络异常检测模型。在 CNN-FL 模型中,参与者不会将他们的训练数据共享给第三方,只是传递加密的梯度相关数据,从而保护了数据隐私,同时解决了数据孤岛以及缺乏标签数据的问题。在实验中 CNN-FL 模型具有强大的入侵检测建模能力,在二分类和多分类中均具有较高的准确率、精确率和召回率,优于传统的分类方法。在未来的工作中,将进一步拓展联邦学习的使用领域,并研究长短期记忆网络(LSTM)、双向循环神经网络(双向 RNNs)等深度学习算法与联邦学习模型结合在网络异常检测领域的分类性能。

参考文献:

[1] SANTORO D, ESCUDERO-ANDREU G, KYRIAKOPOULOS K G, et al. A hybrid intrusion detection system for virtual jamming attacks on wireless networks[J]. Measurement, 2017,109: 79 – 87.

[2] ABBES T, BOUHOULA A, RUSINOWITCH M. Protocol analysis in intrusion detection using decision tree[C]// International Conference on Information Technology: Coding & Computing. Las Vegas, 2004:29 – 34.

[3] DUTTA V, CHORAŚ M, PAWLICKI M, et al. A deep learning ensemble for network anomaly and cyber-attack detection[J]. Sensors, 2020, 20(16): 4583.

[4] KARIMIPOUR H, DEHGHANTANHA A, PARIZI R M, et al. A deep and scalable unsupervised machine learning system for cyber-attack detection in large-scale smart grids[J]. IEEE Access, 2019, 7: 80778 – 80788.

[5] 李硕豪, 张军. 贝叶斯网络结构学习综述[J]. 计算机应用研究, 2015,32(3):641 – 646.
LI S H, ZHANG J. Summary of Bayesian network structure learning [J]. Computer Application Research, 2015,32(3):641 – 646. (in Chinese)

[6] 王明, 李剑. 基于卷积神经网络的网络入侵检测系统[J]. 信息安全研究, 2017, 26(11):32 – 36.
WANG M, LI J. Network intrusion detection system based on convolutional neural network [J]. Information Security Research, 2017, 26(11): 32 – 36. (in Chinese)

[7] YIN C L, ZHU Y F, FEI J L, et al. A deep learning approach for intrusion detection using recurrent neural networks[J]. IEEE Access, 2017, 5: 21954 – 21961.

[8] 朱虎明, 李佩, 焦李成, 等. 深度神经网络并行化研究综述[J]. 计算机学报, 2018, 41(8):1861 – 1881.
ZHU H M, LI P, JIAO L C, et al. A review of parallel-

ization of deep neural networks[J]. Chinese Journal of Computers, 2018, 41(8):1861 – 1881. (in Chinese)

[9] YANG Q, LIU Y, CHEN T J, et al. Federated machine learning: concept and applications[J]. ACM Transactions on Intelligent Systems and Technology, 2019, 10(2):12.

[10] MALLE B, GIULIANI N, KIESEBERG P, et al. The more the merrier-federated learning from local sphere recommendations[M]// HOLZINGER A, KIESEBERG P, TJOA A, et al. Machine learning and knowledge extraction. Charm: Springer, 2017: 367 – 373.

[11] KONEČNÝ J, MCMAHAN B H, YU F X, et al. Federated learning: strategies for improving communication efficiency[EB/OL]. [2017-10-30]. arXiv:1610.05492.

[12] LECUN Y, BOSER B, DENKER J S, et al. Backpropagation applied to handwritten zip code recognition[J]. Neural Computation, 1989, 1(4): 541 – 551.

[13] 李志鹏. 基于神经网络的多源数据攻击检测研究与应用[D]. 成都:电子科技大学,2018.
LI Z P. Research and applications on multi source data attack detection based on neural network [D]. Chengdu: University of Electronic Science and Technology of China, 2018. (in Chinese)

[14] WANG W, ZHU M, WANG J L, et al. End-to-end encrypted traffic classification with one-dimensional convolution neural networks[C]// 2017 IEEE International Conference on Intelligence and Security Informatics (ISI). Beijing, 2017:43 – 48.

[15] LI Z P, QIN Z, HUANG K, et al. Intrusion detection using convolutional neural networks for representation learning[C]// International Conference on Neural Information Processing. Guangzhou, 2017: 858 – 866.

Network anomaly detection based on federated learning

ZHAO Ying　WANG LiBao　CHEN JunJun　TENG Jian

(College of Information Science and Technology, Beijing University of Chemical Technology, Beijing 100029, China)

Abstract: With the rapid development of information technology, network security has become a hot issue in current research. As the basis of network security, current network anomaly detection technology has problems such as low detection accuracy, high false alarm rate and lack of label data. This paper proposes a network intrusion detection classification model (CNN-FL) that combines federated learning and convolutional neural networks, which effectively solves the problems caused by multiple participants training a global model without sharing private data. The model does not need to collect the data required for model training for centralized calculation, since it only transmits encrypted gradient-related data, thereby realizing the use of multi-source data to collaboratively train the same model and solve the problem of lack of label data. We applied the model to both binary classification and multiple methods, and conducted experiments, comparisons and analyses on the same benchmark data set (NSL-KDD). The experimental results show that compared with other methods, our CNN-FL classification model has higher performance and classification accuracy in both binary classifications and multiple classifications.

Key words: federated learning; network anomaly detection; deep learning; convolutional neural network (CNN)