基于动态类权重的卷积神经网络攻击检测模型

樊荣
（驻马店职业技术学院，河南 驻马店 463000）

摘　要：入侵检测系统（intrusion detection system，IDS）作为物联网安全防御的核心组件，其性能直接影响网络的安全性。然而，入侵检测数据集中类样本的不平衡分布降低了入侵检测系统对少数类样本的检测性能。为解决这一问题，提出一种基于动态类权重的卷积神经网络的入侵检测（dynamical class-weighted-based convolutional neural network intrusion detection，DCID）模型。DCID模型采用一维卷积神经网络（1-D CNN）结构，并引入基于动态类权重的损失函数，使得DCID模型不仅能保持对多数类样本的高检测性能，也能显著提升对少数类样本的检测能力。为验证DCID模型的有效性，使用数据集CICIDS 2017进行实验。实验结果表明，与典型的机器学习模型相比，DCID模型在精确率、召回率和F1值方面表现出明显的优势。此外，还对比了不同损失函数下DCID模型的检测性能，结果表明基于动态类权重的损失函数能够有效提升少数类样本的检测性能。

关键词：入侵检测系统；类分布不平衡；卷积神经网络；损失函数；类权重
中图分类号：TN929.5
文献标志码：A
doi: 10.11959/j.issn.1000−0801.2025113

Dynamical class-weighted-based convolutional neural networks attack detection model

FAN Rong
Zhumadian Vocational and Technical College, Zhumadian 463000, China

Abstract: The intrusion detection system (IDS) as a core component of IoT security defense, was directly impacted in its performance, which in turn affected the overall security of the network. However, the imbalanced distribution of class samples in intrusion detection datasets is found to reduce the detection performance of IDS for minority class samples. To address this issue, a dynamical class-weighted-based convolutional neural network intrusion detection (DCID) model was proposed. The DCID model utilized a one-dimensional convolutional neural network (1-D CNN) structure and introduced a dynamical class-weighted loss function, enabling the DCID model to not only maintain high detection performance for majority class samples but also significantly enhance the detection capability

收稿日期：2024−11−06；**修回日期：**2025−03−21
基金项目：2021年度河南省高等学校重点科研项目（No.21B880052）
Foundation Item: The Key Scientific Research Project of Higher Education Institutions in Henan Province in 2021 (No. 21B880052)

for minority class samples. To validate the effectiveness of the DCID model, experiments were conducted using the CICIDS 2017 dataset. The experimental results demonstrate that, compared to typical machine learning models, the DCID model exhibites significant advantages in terms of precision, recall, and F1-score. Additionally, the detection performance of the DCID model under different loss functions was compared, and the results indicated that the dynamical class-weighted loss function effectively improved the detection performance for minority class samples.

Key words: intrusion detection system, imbalanced class distribution, convolution neural network, loss function, class-weighted

0　引言

随着终端设备功能的提升，物联网（Internet of things，IoT）的终端设备会产生海量数据，形成大规模数据集[1]。然而，大规模数据集往往存在类分布不平衡问题，即部分检测的样本数较多（以下简称大样本），而少数类的样本数很小（以下简称小样本）[2]。若对不平衡的数据集进行训练和预测，则很容易形成有利于大样本的预测结果，降低了模型的整体检测性能。因此，针对高度不平衡的数据集训练无偏模型，仍是一项挑战工作。

尽管小样本本身是小概率事件，但是往往学习小样本的数据能获取更有价值的信息。例如，相比于良性程序，恶意程序是小概率事件，但是学习研究恶意程序进而检测恶意程序比研究正常程序更有意义。原因在于，将恶意程序错误地归为良性程序所产生的后果比将良性程序误判为恶意程序可能更严重。因此，准确地检测小概率事件，有效处理数据集中的不平衡问题也非常重要[3-4]。

传统的攻击检测模型过分依赖样本数分布平衡的数据集，这就导致其在类别分布不平衡数据中容易对少数类别进行错误的检测。因此，良好的攻击检测模型不仅能对多类别数据样本进行准确的检测，还能准确地检测小样本。

深度学习（deep learning，DL）被认为是机器学习领域最关键的突破之一[5-6]，已在各领域广泛使用。例如，研究人员已利用 DL 处理不平衡数据的类别检测问题。深度神经网络（deep neural network，DNN）由一系列的非线性模型组成[7]，并通过这些非线性模型将模型输入与模型输出训练成一个非线性函数关系。神经网络通过损失函数的反馈，优化模型参数，进而优化检测模型。

损失函数的选择决定了所反馈信息的价值，直接影响训练检测模型的效率[8]。目前常采用基于交叉熵（cross entropy，CE）的损失函数[9]。经典的基于 CE 的损失函数不考虑数据实例的占比，而是给每个数据实例赋予同等重要性（同等的权重），但这容易忽视对少数类别实例的观察。因此，基于 CE 的损失函数并不太适应于类分布不平衡的检测任务。

基于类频率的权重策略[10-11]可缓解类别不平衡所产生的损失。基于类频率的权重策略是类出现的频率越高（类样本数多），其权重越小，反之，频率越低（类样本少），权重越高，但是该策略在大规模的真实数据集上的性能较差。

针对类分布不平衡问题，本文提出基于动态类权重的卷积神经网络的入侵检测（dynamical class-weighted-based convolutional neural network intrusion detection，DCID）模型。DCID 模型通过一维卷积神经网络（one-dimensional convolutional neural network，1-D CNN）构建检测模型，并采用类权重策略构建损失函数，使模型不只“偏袒”大样本数据，也能倾斜于小样本数据，提升 DCID 模型在小样本数据上的检测性能。为验证 DCID 模型的性能，选用 CICIDS 2017 数据

集进行仿真实验。结果表明，相比于传统的机器学习算法，DCID模型提高了检测性能，并且基于动态权重的策略在小样本实例上的F1值优于基于CE的策略。

1 系统模型及预备知识

1.1 系统模型

令$\boldsymbol{D}=\{\boldsymbol{x}_i,y_i\}_{i=1}^{n}$表示$n$个类型的样本实例的训练集。$\boldsymbol{x}_i$表示第$i$个样本（实例）的输入矢量，$y_i$是$\boldsymbol{x}_i$对应的样本标签。假定共有$c$类标签，即$y_i\in\{1,2,\cdots,c\}$。用$X$和$Y$分别表示特征空间和标签空间。对于任意一个样本实例$i$，$\boldsymbol{x}_i\subset X$表示该实例的输入特征矢量，而$y_i\subset Y$表示真实的类标签，且$y_i=\{1,2,\cdots,c\}$。

检测器的功能就是将输入的特征空间映射至标签空间$f:X\to Y$，并通过最小化损失函数$L\left(f\left(\boldsymbol{x};\theta\right),y\right)$训练并更新模型参数，进而有效地学习样本数据。对于给定的损失函数L和检测器f，用$R_L(f)=E_D\left[f\left(\boldsymbol{x};\theta\right),y\right]$表示风险函数，其中$E_D$表示风险分布的期望。

1.2 深度神经网络

对于特征矢量$\boldsymbol{x}=(x_1,\cdots,x_d)$，其具有$d$个体特征，可将具有$H$层的深度神经网络表示成一个非线性函数$f_\theta$。

对于任意$\theta_h=\{W_h,b_h\}$，其中$h=1,\cdots,H$，而W_h表示第h层的权重矩阵，b_h表示第h层的偏差矢量。因此，将DNN表示成一个复杂的特征转换过程：

$$a(\boldsymbol{x})=g\left(W_H\cdot g\left(\cdots g\left(W_2\cdot g\left(W_1\cdot\boldsymbol{x}+b_1\right)+b_2\right)\cdots\right)+b_H\right) \tag{1}$$

其中，$g(\cdot)$表示映射函数，由卷积和激活函数构成。

将第j个神经元在第h层的激活函数的输出值$a_j^{[h]}$表示为：

$$a_j^{[h]}=g^{[h]}\left(\sum_k w_{jk}^{[h]}a_k^{[h-1]}+b_j^{[h]}\right) \tag{2}$$

其中，$g^{[h]}$表示第h层的激活函数；$w_{jk}^{[h]}$表示第h层的权重连接，下标j和k表示第$h-1$层的神经元j至第h层的神经元k；$b_j^{[h]}$表示第h层中第j个神经元的偏差项。

最后一个隐藏层的特征矢量被映射至输出空间Y，进而产生模型输出。再利用softmax函数将输出归一化0至1的数。softmax层共有c个神经元，模型检测结果是类别j的概率为：

$$P\left(y=j|\boldsymbol{x}\right)=\frac{\exp\left(a(\boldsymbol{x})^{\mathrm{T}}W_j^s+b_j^s\right)}{\sum_{j=1}^{c}\exp\left(a(\boldsymbol{x})^{\mathrm{T}}W_j^s+b_j^s\right)} \tag{3}$$

其中，$a(\boldsymbol{x})$表示倒数第二层（penultimate层）的输出；W_j^s和b_j^s表示在penultimate层的j个神经元至softmax层s的权重和偏差项。

为了获取最优的模型参数，神经网络需通过优化器对损失函数$L\left(f\left(\boldsymbol{x}_i;\theta\right),y_i\right)$进行迭代更新：

$$\arg\min_{\theta}\ \frac{1}{n}L\left(f\left(\boldsymbol{x}_i;\theta\right),y_i\right) \tag{4}$$

其中，θ表示模型参数；n表示抽样的样本数。

现多类检测问题常利用CE作为损失函数。基于CE的损失函数可表述为：

$$L_{\mathrm{CE}}\left(\hat{y},y\right)=-\sum_{j=1}^{c}y_j\log\left(\hat{y}_j\right) \tag{5}$$

其中，y_j为二值变量。

如果实例$\boldsymbol{x}_i$属于类c_j，则y_j=1，否则为0。$\hat{y}_j$表示相应的预测值。对于任意输入的实例$\tilde{\boldsymbol{x}}$，它的预测值为：

$$\hat{y}=\max_j\left[P\left(y=j|\tilde{\boldsymbol{x}}\right)\right] \tag{6}$$

其中，$\hat{y}$表示所有检测中具有最高预测分的类别。

1.3 卷积神经网络

卷积神经网络（convolutional neural network，

CNN）是典型的深度学习方法[12]，其主要由输入层、卷积层、池化层、全连接层和输出层组成。CNN 的典型层次及各层的主要作用如图 1 所示。与传统的全连接网络相比，CNN 通过引入池化层，极大地提升了非线性表达能力。

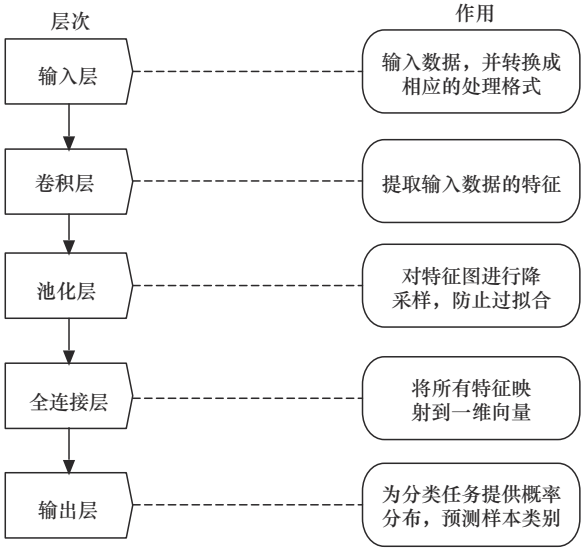


图 1　CNN 的典型层次及各层的主要作用

卷积层是 CNN 的核心层，其负责提供输入样本数据的特征[13]，可形式化表述为：

$$y=F\left(\sum w_{ij}x+b\right) \tag{7}$$

其中，x、y 分别表示卷积层的输入和输出数据；b 为偏置项；$F(\cdot)$ 表示激活函数；w_{ij} 表示二维卷积核。

2　基于 1-D CNN 的入侵检测模型

相比于其他的 DL 模型，基于 1-D CNN 的入侵检测模型具有更好的检测性能。由于 CNN 能够自动对数据进行更好地表示，且无特征选择阶段，更适用于对成本敏感的应用。为此，此次研究利用 1-D CNN 构建入侵检测模型。

本研究采用的 1-D CNN 模型结构如图 2 所示，由 1 个输入层、2 个卷积层、1 个最大池化层、1 个平坦层、1 个密集层和 1 个输出层组成。其中卷积层采用 1-D 滤波内核；最大池化层的下采集率为 2；输出层的层数等于类别个数。采用 ReLU 作为激活函数。输出层采用 softmax 进行多类检测。

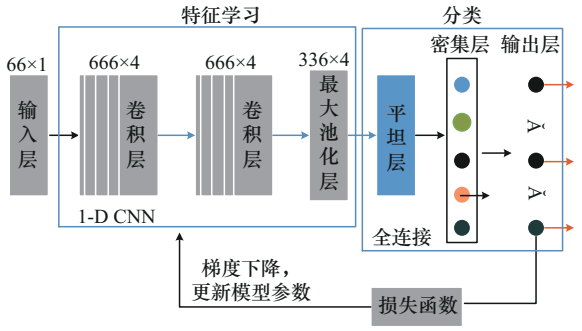


图 2　1-D CNN 模型结构

此外，机器学习和深度学习模型的训练过程是一个优化问题，目标就是最小化损失函数。通过不断调整模型的参数，使损失函数逐渐减小，进而逐渐提升模型的预测能力。

在多类检测问题中，由于不同样本之间的样本数不同，且差异较大，这就形成样本分布不平衡问题，而将权重引入最小化损失函数，能缓解因样本不平衡分布所产生的模型只适用于多样本不适用少样本的问题。为此，DCID 模型采用动态权重的损失函数。由于在仿真中需与其他损失函数进行比较，先阐述基于 CE 的损失函数。

2.1　基于 CE 的损失函数

考虑以 softmax 输出层的 DNN 的模型，将 CE 作为它的损失函数。因此，可通过最小化风险函数，优化 DNN 的模型参数：

$$R_L(f)=E_D\left[f(\boldsymbol{x};\theta),y\right]=-\frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{c}y_{ij}\log\left(f_j(\boldsymbol{x}_i;\theta)\right) \tag{8}$$

其中，θ 表示检测器的参数集；用 $\boldsymbol{y}_i=e_{y_i}\in\{0,1\}^c$ 表示样本实例 $\boldsymbol{x}_i$ 的已经独热编码的标签集，且满足 $\mathbf{1}^T\boldsymbol{y}_i=1,\forall i$。而 y_{ij} 表示 $\boldsymbol{y}_i$ 集中第 j 个元素。

$f_j(\boldsymbol{x}_i;\theta)$为模型的输出，表示对样本实例$\boldsymbol{x}_i$的标签的估计值（概率）。由于输出层为softmax层，则满足$\sum_{j=1}^{n} f_j(\boldsymbol{x}_i;\theta)=1$，$f_j(\boldsymbol{x}_i;\theta)\geqslant 0,\forall j,i,\theta$。

2.2 基于动态权重的损失函数

误差反向传播算法通常用于训练神经网络，它根据训练期间产生的误差成比例地更新模型的权重。然而，这种成比例地更新权重会出现给每个类的数据实例赋予相同重要性的情况。这种策略非常不适用于类分布严重不平衡的情况，所训练的模型也只满足多数类的实例样本，不适用小样本的实例样本，或者说在小样本实例上的检测性能不及在大样本实例上的检测性能。

虽然类分布的不平衡不会影响模型对大样本的检测性能，但是它会影响对小样本的检测性能。事实上，训练模型的目的更倾向于对小样本的检测。原因在于，小样本本质上就是对其观察更少，属于小概率事件，预测它们更难。更希望所训练的模型能够基于仅有的观察对小样本进行学习预测，提高对其的检测精度。此外，正确检测往往比那些错误检测和分布以外实例具有更大的softmax概率[14]。

为此，本研究提出基于动态权重的平衡损失函数（dynamic-weight balanced loss，DWBL），并将其应用至DCID模型，使得那些更难检测的小样本实例被赋予更大的权重。每个类的类权重等于整个数据集中最大类频率与它类频率的比值的对数。

为设置类权重，先令L_{DWB}表示此动态权重的损失函数，其定义如下。

$$L_{\mathrm{DWB}}=-\frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{c}\omega_j^{\left(1-f_j(\boldsymbol{x}_i;\theta)\right)}y_{ij}\log\left(f_j(\boldsymbol{x}_i;\theta)\right)\\ -f_j(\boldsymbol{x}_i;\theta)\left(1-f_j(\boldsymbol{x}_i;\theta)\right)\tag{9}$$

其中，ω_j表示第j个类的权重。

类权重ω_j可视为通过交叉验证从数据集中学习到的超参数，也可将其视为类频率的倒数。考虑ω_j对损失函数的重要性，可采用动态权重策略。因此，可将ω_j设置为最大样本的类频率与类频率的比值的对数，其表达式如下。

$$\omega_j=\mathrm{lb}\left(\frac{\max\left(n_i|i\in c\right)}{n_j}\right)+1\tag{10}$$

其中，n_j表示类别j的样本数（类频率）；$\max\left(n_i|i\in c\right)$表示在$c$个类别中最大类频率。此外，数据集中的每个类的类频率是已知的。

由表达式（10）可知，最大类频率的类权重为1。若类别j不是最大类，则它的权重ω_j大于1，而且类频率越低，其权重将越大。并且，类权重ω_j实质上是对错误检测的惩罚因子。类权重越大，则惩罚越严重。换言之，对类频率低的检测错误惩罚更重，这就逼迫模型更专注对少数类的检测任务。

3 DCID模型的实施流程

入侵检测系统动态监控网络流量，可有效检测正常合法流量中的网络攻击[15]。由于网络入侵仅代表所有网络流量的极少部分，因此良性流量超过恶意流量。绝大多数网络流量将处于“良性”类别中，而罕见的正面情况（恶意网络流量）将处于“攻击”类别中，这一事实造成了极端的类别不平衡问题。因此，入侵检测可视为一个类别分布严重不平衡下的多类检测问题。DCID模型拟通过给不同的类别赋予不同权重，进而处理类别不平衡问题。

基于1-D CNN的DCID模型实施流程如图3所示。该流程主要由数据预处理、模型训练和检测结果输出3部分组成。图3的左半部分是从数据流程角度阐述DCID模型，而右半部分是从CNN模型的层次角度阐述检测的过程。

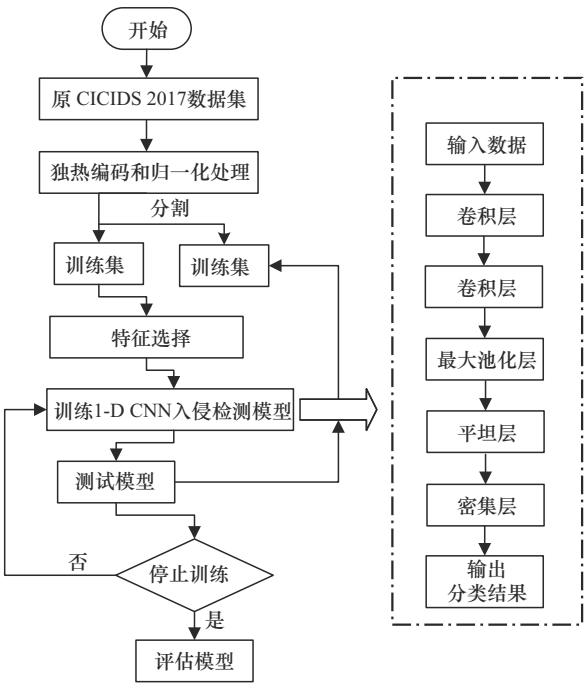


图3　基于1-D CNN的DCID模型实施流程

3.1　数据集描述及预处理

数据集CICIDS 2017中捕获的类似于真实世界的网络流量记录，包括正常和恶意攻击痕迹。这些基于流量的数据是在2017年的5天内捕获的，达到310万条。每个网络流量记录有86个特征表征。在CICIDS 2017数据集中某些攻击类别的样本数极少，占比不到10%。CICIDS 2017数据集存在严重的类别分布不平衡问题。因此，选用数据集CICIDS 2017分析DCID模型的性能。

原数据集CICIDS 2017由每个攻击类的单独攻击文件组成。对这些数据集中的部分攻击进行省略和合并，最终形成11类样本数据。数据集CICIDS 2017样本数据类分布见表1。由表1可知，数据集CICIDS 2017总的样本数为911 421条，其中Benign类样本数达到44 002条，属最多类样本，占比达到48.278 0%。而Bot类样本只有197条，它的样本数最少，占比只有0.216 1%。

为了更好地对数据进行后续处理，对数据进行独热编码和归一化处理。采用最大-最小值方法对样本数据进行归一化处理：

$$x_n=\frac{x-x_{\min}}{x_{\max}-x_{\min}} \tag{11}$$

其中，x表示原始数据；$x_{\min}$和$x_{\max}$分别表示原始数据所在列的最小值和最大值；x_n表示对x归一化后的数据。

3.2　特征选择

考虑实验中选择传统机器学习（machine learning，ML）算法作为基准，对数据进行了特征选择操作，进而选取具有代表性和可区分的特征，提高检测攻击的效率。先进行相关性分析，进而估计特征间的相关性。以0.90为相关系数阈值，去除了32个相关系数大于0.90的特征。然

表1　数据集CICIDS 2017样本数据类分布

类别	样本数/条	占比
正常样本（Benign）	44 002	48.278 0%
基于HTTP洪水的拒绝服务攻击（DoS-Http，DoSH）	23 107	25.352 5%
端口扫描攻击（PortScan，PORT）	15 893	17.437 4%
分布式拒绝服务攻击（distributed denial of service，DDoS）	4 183	4.589 5%
基于黄金眼拒绝服务攻击（DoS-goldeneye，DoSG）	1 029	1.129 0%
针对FTP服务的暴力破解攻击（FTP-patator，FTPP）	794	0.871 2%
针对SSH服务的暴力破解攻击（SSH-patator，SSHP）	590	0.647 3%
耗尽服务器资源的DoS攻击（DoS-slowhttptest，DoSS）	580	0.606 4%
基于缓慢HTTP请求的DoS攻击（DoS-based SLowhttpest，DoSL）	550	0.603 4%
Web攻击（WebA）	218	0.239 2%
僵尸攻击（Bot）	197	0.216 1%

后，利用带有交叉验证的递归特征消除方法选择特征，再利用所选择的特征训练检测模型。

3.3 模型训练及模型输出

将整个数据集的80%作为训练集，剩余的20%作为测试集。通过训练集数据训练图2所示的1-D CNN模型，并采用表达式（9）作为损失函数。通过迭代训练优化模型参数，直至模型收敛。训练结束后，利用已训练的模型对测试集数据进行测试和分析，得到最终检测结果。

4 性能分析

4.1 性能指标及基准算法

采用精确率（Precision）、召回率（Recall）和F1值（F1-score）3个性能指标。F1值融合了精确率和召回率性能，其定义如下。

$$F1 = 2 \cdot \frac{Precision \cdot Recall}{Precision + Recall} \tag{12}$$

其中，Precision和Recall可由式（13）计算：

$$\begin{cases} Precision = \dfrac{TP}{TP + FP} \times 100\% \\ Recall = \dfrac{TP}{TP + FN} \times 100\% \end{cases} \tag{13}$$

其中，TP为发生真阳性（true positive）的次数，FP为发生假阳性（false positive）的次数，FN为发生假阴性（false negative）的次数。

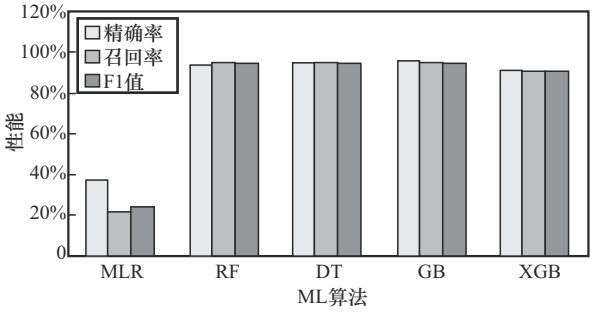
为了更好地分析DCID模型，从2个角度选择基准模型。第一类基准模型：选择传统的ML算法，包括多项式逻辑回归（multinomial logistic regression，MLR）、随机森林（random forest，RF）、决策树（decision tree，DT）、梯度提升（gradient boosting，GB）、分布式梯度增强（XGBoost，XGB）[16]模型。第二类基准模型：选择不同损失函数，包括基于CE的损失函数（cross-entropy loss function，CEL）、带权重的基于CE的损失函数（weighted cross-entropy loss function，WCEL）。其中，WCEL中类的权重等于类频率的倒数。

4.2 DCID模型的平均检测性能

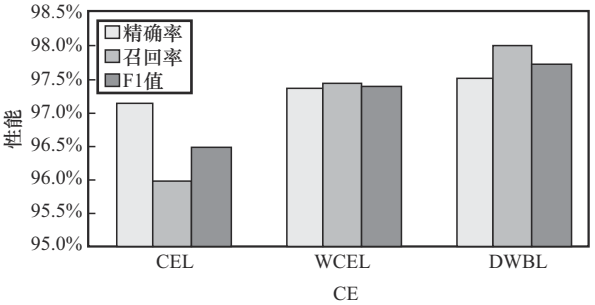
首先，分析各ML算法和基于不同损失函数的1-D CNN模型的性能，性能主要包括精确率、召回率和F1值。入侵检测模型的性能如图4所示。

从图4（a）可知，MLR的检测性能最低。原因在于：MLR算法不能有效地处理类分布不平衡的数据。文献[17]也证实了MLR算法在类别分布不平衡的数据集上的性能不理想的事实。除此之外，RF、DT、GBG和XGB模型的检测性能较理想。RF、DT和GBG模型的F1值均达到95%，而XGB模型的F1值也在90%以上。

相比之下，DCID模型的检测性能优于机器学习模型的检测性能，如图4（b）所示。从图4（b）可知，不论采用哪种损失函数，基于1-D CNN模型的F1值均在96%以上。尽管CEL的F1值最低，但也达到96.50%。相比于CEL和WCEL策略，本文提出的DWBL策略具有最大的F1值，达到97.74%。



（a）基于ML算法的入侵检测模型性能



（b）基于不同CE的1-D CNN模型的性能

图4 入侵检测模型的性能

4.3　DCID 模型对各类样本的检测性能

DCID 模型着重考虑了类分布不平衡对检测性能的影响。由表 1 可知，数据集 CICIDS 2017 样本中的 11 类样本分布不平衡。为此，本小节分析 DCID 模型预测 11 类样本的精确率、召回率和 F1 值。

DCID 模型检测 11 类样本的精确率如图 5 所示。从图 5 可知，相比于 CEL 和 WCEL 策略，DCID 模型所采用的 DWBL 策略提升了在多数样本的精确率。此外，在 11 类样本中，CEL、WCEL 策略和 DWBL 策略均对 Bot 样本的检测准确率是最低的，其次 WebA 样本。原因在于：Bot 样本和 WebA 样本的占比分别位于 11 类样本中的倒数第一和第二位（见表 1）。幸运的是，DWBL 策略预测 Bot 样本的精确率远高于 CEL 和 WCEL 策略。这说明，DCID 模型能够应对类别分布不平衡问题。

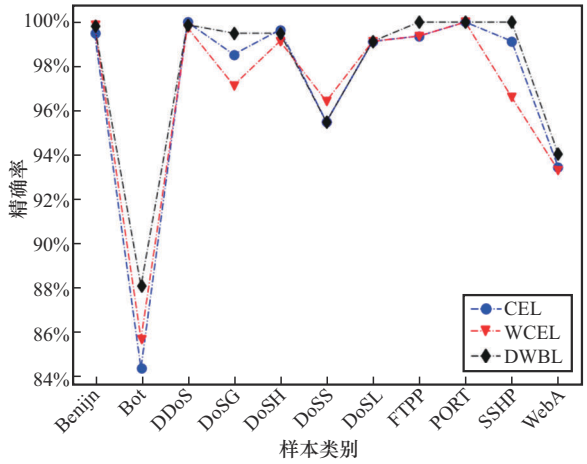


图 5　DCID 模型检测 11 类样本的精确率

DCID 模型检测 11 类样本的召回率如图 6 所示。图 6 中的曲线走势与图 5 中的曲线走势相似。DCID 模型检测小样本的召回率较低。Bot 样本占比最小，DCID 模型对其的召回率较低。基于 CEL 策略下的 DCID 模型的召回率只有约 69.23%，但基于 DWBL 策略的 DCID 模型提高了召回率，将其提升至 95%。这说明，通过设置动态权重能够改善模型对小样本的检测性能。

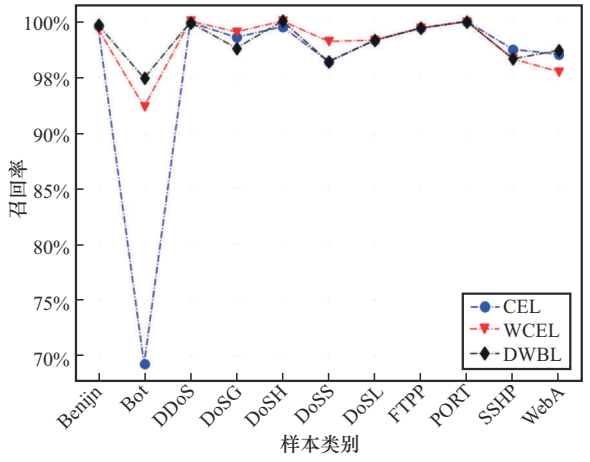


图 6　DCID 模型检测 11 类样本的召回率

DCID 模型检测 11 类样本的 F1 值如图 7 所示。从图 7 可知，对于多数类样本来说，相比于 CEL 和 WCEL 策略，DWBL 策略能获取更高 F1 值，且在小样本上的优势更为明显。Bot 样本和 WebA 样本的占比分别只有 0.216 1% 和 0.239 2%。DWBL 策略在这 2 类样本上的 F1 值高于 CEL 和 WCEL 策略所获取的 F1 值。

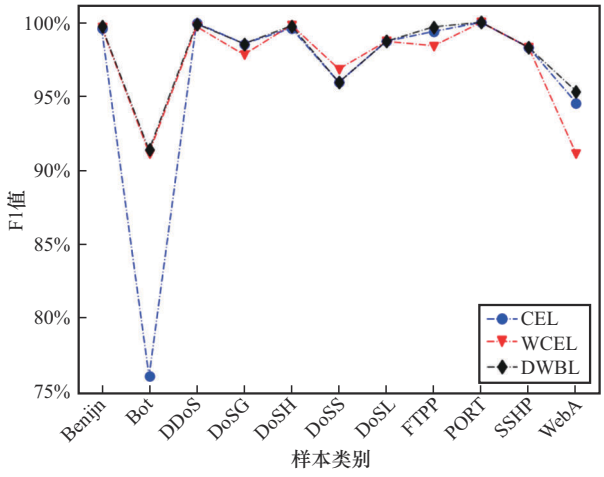


图 7　DCID 模型检测 11 类样本的 F1

此外，WCEL 策略的整个 F1 值低于 DCID 模型，但高于 CEL 策略。原因在于，WCEL 策略也根据类频率设置了类权重。与 DWBL 策略不同的是，DWBL 策略只是简单地将类权重设置为类频率的倒数，而 WCEL 策略是依据表达式（10）设置类权重，这也导致它们在多类样本上的性能相

近，但在小样本上的性能相差较大，如Bot样本和WebA样本。

结合图5、图6和图7的数据可知，DWBL策略通过依据样本占比动态地调整损失函数的权重，能有效地改善模型在小样本上的检测性能不佳问题。

5 结束语

针对入侵检测模型数据集中的类分布不平衡问题，本文提出基于动态类权重的卷积神经网络的入侵检测（DCID）模型。DCID模型采用一维卷积神经网络，并基于类频率动态地设置损失函数的权重，使模型更倾向于小样本的检测。为检测DCID模型的性能，选用数据集CICIDS 2017为样本，对比5种基于机器学习算法的入侵检测模型性能。实验结果表明，本文提出的DCID模型通过采用卷积神经网络，有效地提升了检测样本的F1值、精确率和召回率。同时，选择了基于检测CE的损失函数作为基准，同基于动态权重的损失函数进行对比分析。结果表明，基于动态权重的损失函数可有效提升入侵检测模型在小样本上的检测性能。

参考文献：

[1] 江荣旺, 魏爽, 龙草芳, 等. 基于联邦学习的车联网虚假位置攻击检测研究[J]. 信息安全研究, 2023, 9(8): 754-761.
JIANG R W, WEI S, LONG C F, et al. Research on malicious location attack detection of VANET based on federated learning[J]. Journal of Information Security Research, 2023, 9(8): 754-761.

[2] 林同灿, 葛文翰, 王俊峰. 基于对齐原型网络的小样本异常流量分类[J]. 四川大学学报(自然科学版), 2024, 61(3): 9-20.
LIN T C, GE W H, WANG J F. Aligned prototype network for few-shot anomaly traffic classification[J]. Journal of Sichuan University (Natural Science Edition), 2024, 61(3): 9-20.

[3] RUWANI M FERNANDO K, TSOKOS C P. Dynamically weighted balanced loss: class imbalanced learning and confidence calibration of deep neural networks[J]. IEEE Transactions on Neural Networks and Learning Systems, 2022, 33(7): 2940-2951.

[4] 窦佳恩, 张瑛瑛, 陈玮. 基于集成学习的物联网攻击检测方法[J]. 兵工自动化, 2024, 43(8): 23-26, 59.
DOU J E, ZHANG Y Y, CHEN W, et al. IoT attack detection method based on ensemble learning[J]. Ordnance Industry Automation, 2024, 43(8): 23-26, 59.

[5] 金志刚, 丁禹, 武晓栋. 融合梯度差分的双边校正联邦入侵检测算法[J]. 信息网络安全, 2024, 24(2): 293-302.
JIN Z G, DING Y, WU X D. Federated intrusion detection algorithm with bilateral correction merging gradient difference[J]. Netinfo Security, 2024, 24(2): 293-302.

[6] 王俊恒, 朱铭铭. 基于多原型指导的小样本家族域名入侵检测算法[J]. 中国电子科学研究院学报, 2023, 18(11): 1049-1057.
WANG J H, ZHU M M. Multi-prototype guided few-shot family domain intrusion detection algorithm[J]. Journal of China Academy of Electronics and Information Technology, 2023, 18(11): 1049-1057.

[7] 刘欢, 肖蔚, 赵长明. 基于融合机器学习算法的网络入侵检测与定位技术[J]. 现代电子技术, 2023, 46(12): 182-186.
LIU H, XIAO W, ZHAO C M. Network intrusion detection and location technology based on fused machine learning algorithm[J]. Modern Electronics Technique, 2023, 46(12): 182-186.

[8] 王赞, 闫明, 刘爽, 等. 深度神经网络测试研究综述[J]. 软件学报, 2020, 31(5): 1255-1275.
WANG Z, YAN M, LIU S, et al. Survey on testing of deep neural networks[J]. Journal of Software, 2020, 31(5): 1255-1275.

[9] 李伟, 黄鹤鸣. 基于双交叉熵的自适应残差卷积图像分类算法[J]. 计算机工程与设计, 2023, 44(12): 3670-3676.
LI W, HUANG H M. Adaptive residual convolution image classification algorithm with dual cross-entropy[J]. Computer Engineering and Design, 2023, 44(12): 3670-3676.

[10] WU Z Y, GUO Y, LIN W F, et al. A weighted deep representation learning model for imbalanced fault diagnosis in cyber-physical systems[J]. Sensors, 2018, 18(4): 1096.

[11] AURELIO Y S, DE ALMEIDA G M, DE CASTRO C L, et al. Learning from imbalanced data sets with weighted cross-entropy function[J]. Neural Processing Letters, 2019, 50(2): 1937-1949.

[12] 陈悦, 杨柳, 李帅, 等. 基于Softmax函数增强卷积神经网络—双向长短期记忆网络框架的交通拥堵预测算法[J]. 科学技术与工程, 2022, 22(29): 12917-12926.
CHEN Y, YANG L, LI S, et al. Traffic congestion prediction algorithm based on CS-BiLSTM framework[J]. Science Technology and Engineering, 2022, 22(29): 12917-12926.

[13] 郭越. 基于改进CNN的工业控制网络入侵检测研究[J]. 机械

设计与制造工程, 2023, 52(6): 103-108.
GUO Y. Research on intrusion detection of industrial control network based on improved CNN[J]. Machine Design and Manufacturing Engineering, 2023, 52(6): 103-108.

[14] 胡佳玲, 施一萍, 谢思雅, 等. 基于轻量级卷积神经网络人脸识别算法的研究与应用[J]. 传感器与微系统, 2022, 41(1): 153-156.
HU J L, SHI Y P, XIE S Y, et al. Research and application of face recognition algorithm based on lightweight CNN[J]. Transducer and Microsystem Technologies, 2022, 41(1): 153-156.

[15] RODDA S, EROTHI U S R. Class imbalance problem in the network intrusion detection systems[C]//Proceedings of the 2016 International Conference on Electrical, Electronics, and Optimization Techniques (ICEEOT). Piscataway: IEEE Press, 2016: 2685-2688.

[16] 徐东方, 徐洪珍, 邓德军. 基于CNN-BLSTM-XGB的入侵检测[J]. 计算机工程与设计, 2024, 45(3): 676-683.
XU D F, XU H Z, DENG D J. Intrusion detection based on CNN-BLSTM-XGB[J]. Computer Engineering and Design, 2024, 45(3): 676-683.

[17] KING G, ZENG L C.Logistic Regression in rare events data[J]. Political Analysis, 2001, 9(2):137-163.

[作者简介]

樊荣（1994– ），男，驻马店职业技术学院科研与信息化处讲师，主要研究方向为计算机应用技术。