DOI:10.13878/j.cnki.jnuist.20220919001

王博[1]　万良[1]　刘明盛[1]　孙菡迪[1]

DWT 和 AKD 自动编码器的 DDoS 攻击检测方法研究

摘要

针对 DDoS 网络流量攻击检测效率低及误报率高的问题,本文提出一种基于离散小波变换(Discrete Wavelet Transform,DWT)和自适应知识蒸馏(Adaptive Knowledge Distillation,AKD)自动编码器神经网络的 DDoS 攻击检测方法.该方法利用离散小波变换提取频率特征,由自动编码器神经网络进行特征编码并实现分类,通过自适应知识蒸馏压缩模型,以实现高效检测 DDoS 攻击流量.研究结果表明,该方法对代理服务器攻击、数据库漏洞和 TCP 洪水攻击、UDP 洪水攻击具有较高的检测效率,并且具有较低的误报率.

关键词

DDoS 攻击;离散小波变换;自适应;知识蒸馏;自动编码器

中图分类号 TP393

文献标志码 A

0　引言

在当前网络环境下,DDoS(Distributed Denial of Service)攻击仍然是极具威胁的网络攻击[1].DDoS 攻击者可以利用网络中可用的多台受损设备,向受害者服务器发送大量随机源 IP 地址的虚假报文,以中断用户服务[2].

DDoS 攻击检测一般基于统计学习和机器学习方法.统计学习[3]方法通过定义阈值来区分 DDoS 攻击流量和正常流量.虽然统计学习方法简单实用,但考虑到网络流量的动态特性[4],定义阈值具有一定的挑战性,并且统计学习方法通常比基于机器学习的方法有着更高的误报率.基于机器学习[5]的技术是 DDoS 攻击检测最有效的方法之一,这些技术通常被分为有监督学习和无监督学习.有监督学习需要一些带标签的数据来训练模型,而无监督学习[6]在训练过程中不需要任何带标签的数据.自动编码器[7]神经网络就是利用了无监督学习技术,这类神经网络能够在模型执行结束时重新生成输入数据.

在 DDoS 攻击检测过程中为模型训练提取特征是与机器学习有关的一个重要问题.网络流量具有各种周期性的组成部分[8],例如 TCP 协议的拥塞控制机制以及往返延迟时间.此外,很多研究证明 DDoS 攻击流量具有周期性特征[9],这种周期性特征可以用来区分 DDoS 攻击流量和正常流量.分析周期性最佳的方法是频域分析[10],虽然当前 DDoS 攻击检测有着不同的研究,但所采用的特征大多基于时域分析[11].然而,在利用时域分析对突增的正常流量与 DDoS 攻击流量进行分析时,结果是相似的,正常的突增流量作为一种合法的网络流量,是由成千上万的合法用户同时访问一个 Web 服务器构成的,因为 DDoS 攻击流量与正常流量之间相似度很高,可能会产生很高的误报率.此外,DDoS 攻击流量生成工具可以使 DDoS 攻击流量尽可能与正常流量相似,从而降低检测精度.但是,这些因素并不影响流量的周期性,因为突增的正常流量随着客户端数量动态变化,产生的流量会不断波动,而 DDoS 攻击流量则相对稳定.因此,分析流量的周期性,利用频域分析得到的特征,可以考虑为解决误报率过高的问题提供依据.

从网络流量表中收集统计特征来训练模型已经是一种通用的方法,文献[12]通过从交换机中应用 ANOVA(Analysis of Variance)方法提取统计特征,并使用不同分类器对选择的特征进行性能评估,最终确定了检测 DDoS 攻击的最佳特征集,得到很好的效果.传统的利用

收稿日期 2022-09-19

资助项目 国家自然科学基金(62062020);贵州省教育厅自然科学研究项目(黔教科(2007)015 号)

作者简介

王博,男,硕士生,研究方向为网络空间安全.gs.bowang20@gzu.edu.cn

万良(通信作者),男,教授,研究方向为计算机理论、网络空间安全.lwan@gzu.edu.cn

1　贵州大学 计算机科学与技术学院/公共大数据国家重点实验室,贵阳,550025

统计学习的方法,大多还会用到熵.结合熵的使用,令大多数基于统计学习的方法检测实现高准确率.因此,利用熵的 DDoS 攻击检测方法也被扩展到当前的网络环境当中.文献[13]提出一种基于 DDoS 攻击和正常流量之间熵值变化的防御机制,并且通过该机制可以有效减少特征维度和计算开销,实现了高效的防御.

虽然通过统计特征和信息熵的使用,已经提出了很多成功的方法,但是,在某些情况下可能会出现误报率过高的情况[14].因此,通过将信息熵和机器学习结合的方法,可以很好地缓解误报率过高的问题.文献[15]提出 PSO-BP 神经网络结合熵的 DDoS 攻击检测方法,通过在交换机上对流量计算熵值进行预检测,并提取异常交换机的统计特征,检测 DDoS 攻击是否发生.该方法不仅提高了检测效率,而且减少了误报率.

现有的 DDoS 攻击检测模型大多能实现高准确率[16-19],但是在实际检测过程中,检测速度往往无法与攻击速度持平甚至小于攻击速度,导致无效检测.因此,将模型压缩的方法被研究并使用[20].现在研究者大多从特征降维方面减少模型的参数来增加检测效率,例如文献[21]提出一种正方形草图的特征构造方法,通过设计多维网络流草图结构和网络流测量方法,结合一种新的低速率 DDoS 攻击检测方法(LDDM)对 DDoS 攻击进行检测.该方法不需要提取大量的统计特征,通过 K-L 散度来对比异常草图和正常草图之间的区别,并取得良好的检测效率以及较低的误报率,但是检测精度不高.

针对当前 DDoS 攻击模型结合统计特征的方法导致误报率高,以及模型检测效率低的问题,本文提出一种基于离散小波变换(Discrete Wavelet Transform, DWT)和自适应知识蒸馏(Adaptive Knowledge Distillation,AKD)自动编码器的 DDoS 攻击检测方法.按照时间序列对网络流量应用离散小波变换获得频域特征,有效区分攻击流量和正常流量,达到降低误报率的目的.结合自动编码器神经网络模型,在保证对 DDoS 攻击检测准确性的同时,有效实现非线性表达降维特征.通过自适应知识蒸馏将模型压缩,提高模型的拟合速度,并实现增加检测效率的目的.

1 研究方法

1.1 频域分析和离散小波变换

在频域分析中,傅里叶变换(FT)[22]是分析信号最常用的变换技术之一,但当信号具有非平稳频率时,傅里叶变换不适合进行频域分析.为了解决这个问题,对信号进行短期傅里叶变换(STFT),它将原始信号划分为一些平稳的子窗口,然后对每一子窗口进行傅里叶变换.短期傅里叶变换的主要困难是窗口大小的选择,它需要在时间和频率之间进行权衡.当选择较窄的窗宽时,在时间段上获得了更多的频率,但失去了频率与时间的对应关系.离散小波变换(DWT)[23]被认为是短期傅里叶变换的替代方法,因为这种方法提供了让信号和时间、频率相互对应的方法.此外,由 DWT 在时间和频率上以不同的分辨率分析序列时,可以通过增加粒度来获得攻击流量和正常流量的近似分量和细节分量.因此本文采用 DWT 作为特征提取技术.离散小波为机器学习提供了更合适的特征.

DWT 分为滤波和采样两个步骤.滤波是通过对输入信号应用低通滤波器和高通滤波器来实现的.采样是通过对两个滤波器的输出进行降采样来实现的.当滤波器改变信号的长度时,下采样过程会控制信号的尺度.当采用低通滤波器和高通滤波器对低频和高频信号进行分析时,该过程首先将信号 X 引入一个脉冲响应为 $M=\{m_1,m_2,\cdots,m_q\}$ 的低通滤波器,在保持信号规模不变的情况下去除信号中最大频率一半(即最大频率的 1/2)以上的所有频率.同时,信号 X 通过一个脉冲响应为 $G=\{g_1,g_2,\cdots,g_q\}$ 的高通滤波器,输出的一半被丢弃.上述过程可以表示为

$$y_j^{\text{low}}=\sum_{i=-\infty}^{\infty}m_{2j-1}\times x_i, \tag{1}$$

$$y_j^{\text{high}}=\sum_{i=-\infty}^{\infty}g_{2j-1}\times x_i, \tag{2}$$

其中:y^{low} 为通过低通滤波器的滤波结果,y^{high} 为通过高通滤波器的滤波结果,y^{low} 和 y^{high} 的长度之和与信号 X 总长度相同;m 和 g 分别是低通滤波器和高通滤波器的脉冲响应信号;x 为信号 X 中的离散子信号.

上述过程也被称为子带编码,可以通过将低通滤波器的输出引入到下一层滤波器中来重复进行进一步分解.子带级别的最大数目取决于信号中的数据点数目和滤波器的长度.直到信号比给定小波的滤波器长度 q 短,说明 DWT 计算结束.最大长度定义为

$$L_{\max}=\left\lfloor \log_2\left(\frac{N}{q-1}\right)\right\rfloor, \tag{3}$$

其中,q 为滤波器长度,N 为信号 X 长度.

由于 DWT 得到的高通滤波(详细分量 c_D)和低通滤波(近似分量 c_A)所表达的含义不同,而低通滤波得到的近似分量更能表达信号的整体,所以本文使用近似分量作为统计特征提取的离散信号.

1.2 统计特征

对于离散信号 X,为了区分攻击样本和正常样本,可以分析两者之间的概率分布差异.首先,通过 DWT 得到信号的近似分量 c_A,通过统计 c_A 中信号的平均数、方差、百分位数、偏度、峰度和熵来决定最后的统计特征.因此,本文考虑了均值,方差,10%、20%、30%、45% 和 80% 位数,偏度,峰度和熵作为统计特征,并利用这些特征生成区别于正常样本的判别模型.对于一个给定的离散信号 $X=\{x_1,x_2,\cdots,x_N\}$,式(4)—(7)分别定义均值(μ)、方差(σ)、偏度(S)和峰度(K).百分位数表示位于某一特定百分比位数的值.例如,20% 位数是指在当前信号分布中位于 20% 位置的最小值.文献[24]证明了该方法的有效性.熵是一种计算某一属性对于当前类别的不确定性的方法,定义如式(8)所示.分布离散的属性熵值高,而集中分布的熵值低.

$$\mu=\frac{1}{N}\sum_{i=1}^{N}x_i, \tag{4}$$

$$\sigma=\frac{1}{N-1}\sum_{i=1}^{N}(x_i-\mu)^2, \tag{5}$$

$$S=\frac{1}{N\sigma^3}\sum_{i=1}^{N}(x_i-\mu)^3, \tag{6}$$

$$K=\frac{1}{N\sigma^4}\sum_{i=1}^{N}(x_i-\mu)^4, \tag{7}$$

$$H=\sum_{i=1}^{h}p_i\log_2\left(\frac{1}{p_i}\right), \tag{8}$$

其中,p 为不同属性的概率分布.

1.3 自动编码器神经网络

自动编码器是一种无监督的神经网络,它重构出一个和输入向量维数相同的输出向量,结构如图 1 所示.

自动编码器的基本操作如下:首先,将统计特征整合成输入向量输入到编码器,其次,从中间层重构出与输入相同维数的输出.自动编码器的主要目标是通过应用三个子模块,即编码器、中间层、解码器,来构造一个尽可能接近输入向量的输出向量.编码器将数据编码成低维的特征表达,最中间层表示特征的最低维度,解码器将低维特征重构成高维表达.

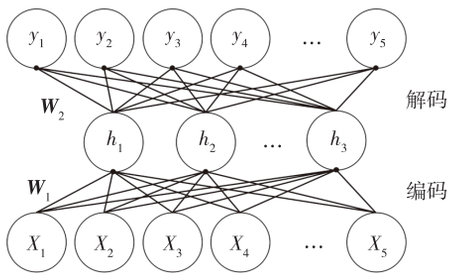


图 1 自动编码器网络结构

Fig. 1 Auto-encoder network architecture

自动编码器的结构可以分为两部分:1)编码器 $f=\mathbf{R}^D\rightarrow\mathbf{R}^M$;2)解码器 $g=\mathbf{R}^M\rightarrow\mathbf{R}^D$,其中 D 为输入数据与重构数据的维度,M 为中间层的维度.对于信号样本 $x\in X$,自动编码器的中间层可表示为

$$h=f(\boldsymbol{W}_1x+b_1). \tag{9}$$

自动编码器的输出为通过解码器重构的数据,可以表示为

$$y=g(\boldsymbol{W}_2h+b_2), \tag{10}$$

其中 $\boldsymbol{W}_1,\boldsymbol{W}_2,b_1,b_2$ 为网络参数,f 和 g 为激活函数,其中 $\boldsymbol{W}_2=\boldsymbol{W}_1^{\mathrm{T}}$.

重构数据与输入数据的重构误差为

$$L=\sum_{n=1}^{N}\|x^n-y^n\|^2+\lambda\|\boldsymbol{W}\|^2, \tag{11}$$

其中,λ 为正则化系数,通过最小化重构误差来训练模型并学习网络参数.

1.4 自适应知识蒸馏(AKD)

知识蒸馏技术是一种模型压缩技术,它使用了一种"教师-学生"的训练方法.采用 AKD 的原因在于:线下训练的网络模型需要大量数据去拟合复杂的模型,会出现模型规模大、模型推理速度慢的问题,而实际线上部署的模型对模型推理速度和部署资源有着严格的要求限制.随着 DDoS 攻击的发生,攻击流量会迅速增加,大模型会出现推理不及时的情况,导致网络拥堵.

知识蒸馏基本方法:通过训练出最优的自动编码器模型作为教师模型 T-Net,再通过减少模型规模设计学生模型 S-Net,利用 T-Net 模型训练 S-Net 模型,可以让 S-Net 模型学习到 T-Net 模型的泛化能力.

知识蒸馏基本步骤:第一步是训练 T-Net 模型,第二步是在设置温度 t 条件下,蒸馏 T-Net 的知识到 S-Net.

蒸馏过程的 Loss 如式(12)所示:

$$L = \alpha L_{\text{soft}} + \beta L_{\text{hard}}, \tag{12}$$

其中,L_{soft}是 T-Net 模型的损失和 S-Net 模型损失的交叉熵:

$$L_{\text{soft}} = -\sum_{j}^{N} p_j^t \log q_j^t, \tag{13}$$

其中,q 为 S-Net 的 Loss 如式(14)所示,p 为 T-Net 的 Loss 如式(15)所示:

$$q_i = \frac{\exp(z_i/t)}{\sum_j \exp(z_j/t)}, \tag{14}$$

$$p_i = \frac{\exp(v_i/t)}{\sum_j \exp(v_j/t)}, \tag{15}$$

其中:z 为学生模型对于类别 i 的逻辑概率;v 为教师模型对于类别 i 的逻辑概率;t 为温度,作用在于平滑 softmax 损失函数的概率分布,t 越高,概率分布越趋向于平滑,负标签的信息会相应放大,模型会更加关注负标签.

L_{hard} 为 $t=1$ 时 p 和 q 的交叉熵,如式(16)所示:

$$L_{\text{hard}} = -\sum_{j}^{N} p_j^1 \log q_j^1. \tag{16}$$

L_{hard}表示非蒸馏条件下的 S-Net 损失,α 和 β 作为超参数,其取值是一个重要问题,选择不合适的数值会增加模型的训练时间,所以本文提出一种自适应的变化方法.先随机初始化 α 和 β,当前迭代训练结束时,利用梯度法改变 α 和 β 的值,如式(17)和(18)所示:

$$\alpha = \alpha \pm a \frac{\partial}{\partial \alpha} L(L_{\text{soft}}, L_{\text{hard}}), \tag{17}$$

$$\beta = \beta \pm a \frac{\partial}{\partial \beta} L(L_{\text{soft}}, L_{\text{hard}}), \tag{18}$$

其中,a 表示学习率,且 α 和 β 会随着损失值 L 的变化而变化,规则如下:

步骤 1,变化 α 的值;

步骤 2,当 L 增加,反向变化 α 的值,并变化 β 的值;

步骤 3,当 L 增加,反向变化 β 的值,然后反向变化 α 的值;

步骤 4,当 L 减少,返回步骤 1;

步骤 5,当所有情况都会导致 L 增加时,停止.

2 基于 DWT 和 AKD 的 DDoS 攻击检测

本节介绍基于离散小波和自适应知识蒸馏的 DDoS 攻击检测方案,具体步骤如图 2 所示.

2.1 特征提取

特征提取模块用于从网络流量数据包中提取统计特征,如图 3 所示.

特征提取步骤如下:

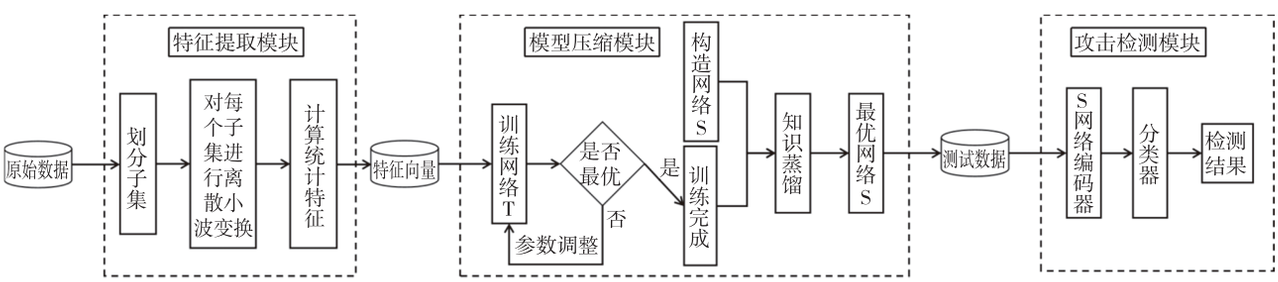


图 2 DDoS 攻击检测总体框架

Fig. 2 General framework for DDoS attack detection

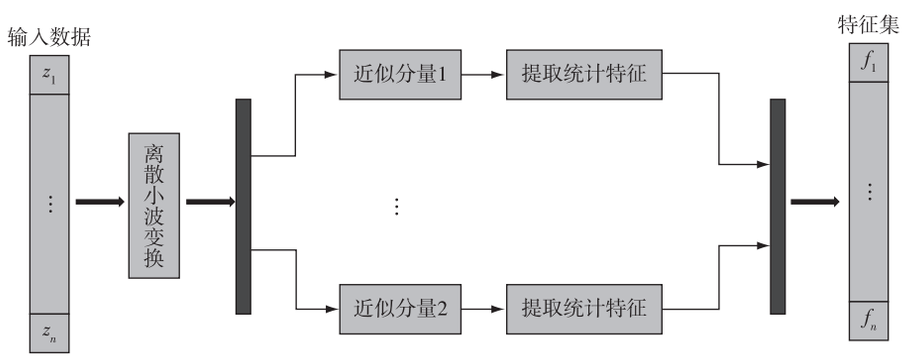


图 3 特征提取

Fig. 3 Feature extraction

步骤 1,将流量数据包统计数量按照相同时间间隔分成 Z 个子集;

步骤 2,对每个子集进行 DWT,得到每一子集的近似分量和详细分量;

步骤 3,计算每个子集近似分量的均值,方差,10%、20%、30%、45%和 80%位数,偏度,峰值和熵的统计特征;

步骤 4,将计算结果和 DWT 计算的近似分量拼接形成特征向量,最终得到特征向量 $\boldsymbol{f}_i \in \mathbf{R}^{1\times n}$.

算法过程如下:

1) 初始化特征列表 F;

2) 初始化当前特征向量 $\boldsymbol{f}$;

3) 初始化特征长度 w;

4) 输入数据 Z;

5) FOR i in Z;

6) 对 i 进行离散小波变换得到近似分量 c_A;

7) 计算 c_A 的均值 μ、方差 σ、百分位数、偏度 K、峰度 S、熵值 H;

8) IF c_A 的长度小于 w;

9) 将长度补充 0 到 w;

10) END IF;

11) $\boldsymbol{f} \| c_A \| \mu$、百分位数、$K$、$S$、$H$,其中 ‖ 为拼接操作;

12) $F \| \boldsymbol{f}$,其中 ‖ 为拼接操作;

13) END FOR.

2.2 模型训练

模型训练模块用于描写 DDoS 攻击检测模型训练框架,如图 4 所示,该 DDoS 攻击模型训练框架具体过程如下:

1) 获取特征集 F 作为自动编码器的输入;

2) 预训练,逐层训练自动编码器神经网络 T;

3) 获得当前数据集训练出的最优模型 T;

4) 构建模型 S;

5) 通过 T 模型蒸馏训练 S 模型;

6) 模型调优.

2.3 检测模型

检测模型模块用于使用学生模型 S 的编码器部分进行 DDoS 攻击检测,当 S 的参数训练确定后,取出编码器部分.通过训练集进行微调后,为测试数据编码.利用分类器对编码结果进行 DDoS 攻击类别鉴定,验证模型 S 的检测能力级泛化能力,如图 5 所示.

检测过程如下:

1) 获得模型 S 的编码器部分;

2) 利用训练集微调 S 模型;

3) 获取测试数据;

4) 通过 S 模型编码器部分将测试数据编码;

5) 通过分类器得到攻击类别.

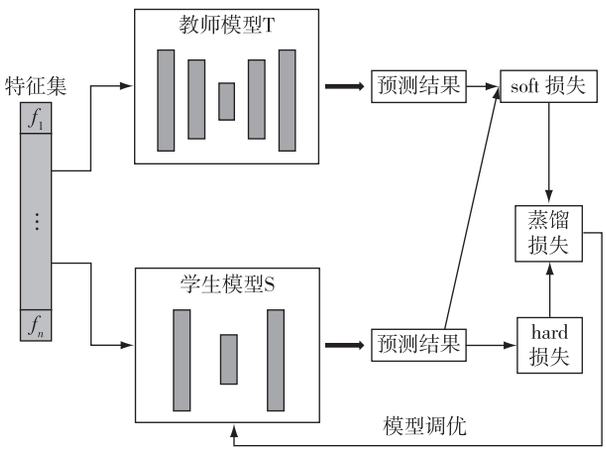


图 4 知识蒸馏

Fig. 4 Knowledge of distillation

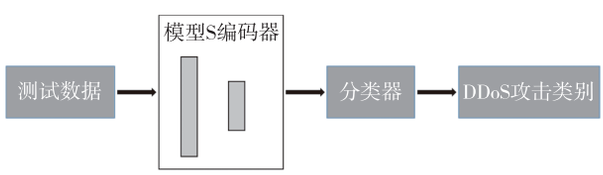


图 5 检测模型

Fig. 5 Detection model

3 实验

3.1 实验环境

本文使用 Keras 框架进行实验模拟,并选择 Python 作为编程语言.硬件:intel(R) Core(TM) i7-10700F 处理器、8 GB 内存、500 GB 固态硬盘、64 位 Windows11 操作系统.

3.2 实验数据

本文大部分实验是基于 CIC-DDoS2019 数据集(https://www.unb.ca/cic/datasets/ddos-2019.html)进行的,该数据集由 Lashkari 在 New Brunswick University(UNB)网络上收集.CICDDoS2019 数据集分 2 天采集,本实验使用的是第 1 天采集的攻击,类型包括:数据库漏洞攻击 MSSQL、UDP 协议洪水攻击、TCP 协议三次握手确认攻击 SYN.

为了验证本文提出方法的通用性,本文还采用 Iot-23 数据集(https://www.stratosphereips.org/datasets-overview)进行了实验.Iot-23 数据集模拟利

用感染的物联网设备实施 DDoS 攻击并获取攻击流量数据.以上数据集均来自于真实的网络环境.

3.3 评价指标

在对比实验中,本文采用了准确率(Accuracy,A)、精确度(P)、F1 评价指标(F1_Score,F_1)、误报率(False Positive Rate,FPR)、召回率(Recall,R)、均方误差(Mean Square Error,MSE)等作为本文的评价优劣的标准.

$$A=\frac{\mathrm{TP}+\mathrm{TN}}{\mathrm{TP}+\mathrm{TN}+\mathrm{FP}+\mathrm{FN}}, \tag{19}$$

$$P=\frac{\mathrm{TP}}{\mathrm{TP}+\mathrm{FP}}, \tag{20}$$

$$F_1=\frac{2\times P\times R}{P+R}, \tag{21}$$

$$\mathrm{FPR}=\frac{\mathrm{FP}}{\mathrm{FP}+\mathrm{TN}}, \tag{22}$$

$$R=\frac{\mathrm{TP}}{\mathrm{TP}+\mathrm{FN}}, \tag{23}$$

$$\mathrm{MSE}=\frac{1}{N}\sum_{i=1}^{N}(y_{\mathrm{true}}-y_{\mathrm{pred}})^2, \tag{24}$$

式(24)中,N 为样本个数,y_{true} 为真实值,y_{pred} 为预测值.式(19)—(23)参数定义如表 1 所示.

表 1 参数定义

Table 1 Parameter definitions

实际值	预测值	
	攻击流量	正常流量
攻击流量	真阳性 TP	假阴性 TN
正常流量	假阳性 FP	真阴性 FN

3.4 参数设置

通过反复调优,并吸取前人的经验进行大量实验,对比各项参数的优劣.通过反复实验,最终确定的参数如表 2 所示.

表 2 参数设置

Table 2 Parameter settings

参数	数值	参数	数值
激活函数	Tanh、Relu	迭代数	100
优化器	Adam	Batch_size	8
分类损失函数	最大值交叉熵	学习率	0.0001
回归损失函数	均值对数方差	S-Net 结构	256-64-256(3 层)
T-Net 编码器结构	256-128-64-32-4(5 层)		

3.5 结构分析

3.5.1 T-Net 层数分析

网络层数以及每层所使用的激活单元的个数对于模型编码来说具有重要作用.随着网络层数和激活单元的增加,编码出的抽象特征能更深层次表达原始数据.因此,为了得到比较优秀的网络模型结构,通过设置不同深度的编码器结构进行对比试验,测试不同结构编码器的性能,对比结果如图 6 所示,横纵坐标分别为维度 1 和维度 2,表示二维坐标点.通过使用不同模型结构的编码器部分对测试集进行分类预测结果如表 3 所示.通过实验结果可以得到,当网络层数达到一定程度时,模型各项指标将不再增加.所以选择 5 层结构作为 T-Net 编码结构.

表 3 T-Net 分类实验

Table 3 T-Net classification experiment

T-Net 编码结构	层数	精确度/%	召回率/%	F1 评价指标/%	准确率/%
256-128-4	3	96.18	96.90	96.53	96.07
256-128-64-4	4	97.56	96.53	97.04	97.31
256-128-64-32-4	5	98.12	98.39	98.25	97.71
256-128-64-32-16-4	6	97.98	96.99	97.48	97.74

3.5.2 T-Net 知识蒸馏 S-Net

通过已确定的 T-Net 结构,S-Net 相对于 T-Net 来说拥有更少的层数和激活单元,找到最优的 S-Net 结构是一项挑战.现设置 3 种不同结构的 S-Net,通过对比试验分析 T-Net 对不同 S-Net 的蒸馏效果,训练过程如图 7 所示.可以发现当中间层越小蒸馏过程越不稳定且准确率最终会总体呈现降低趋势,说明在模型拟合过程中寻找不到最优解.通过对比不同 T-Net 蒸馏 S-Net 前后误差,最终确定 S-Net 结构为 256-64-256 三层结构,其中不同数值代表不同层激活单元个数,如表 4 所示.

表 4 T-Net 蒸馏 S-Net 误差分析

Table 4 T-Net distillation S-Net error analysis

T-Net 蒸馏 S-Net 结构	层数	蒸馏前均方误差	蒸馏后均方误差
256-64-256	3	10.47	1.46
256-32-256	3	16.58	1.52
256-16-256	3	22.67	3.70

3.5.3 压缩效果分析

将 T-Net 和 S-Net 以及知识蒸馏后的 S-Net 的分类效果进行对比,三种网络模型同时对测试数据集

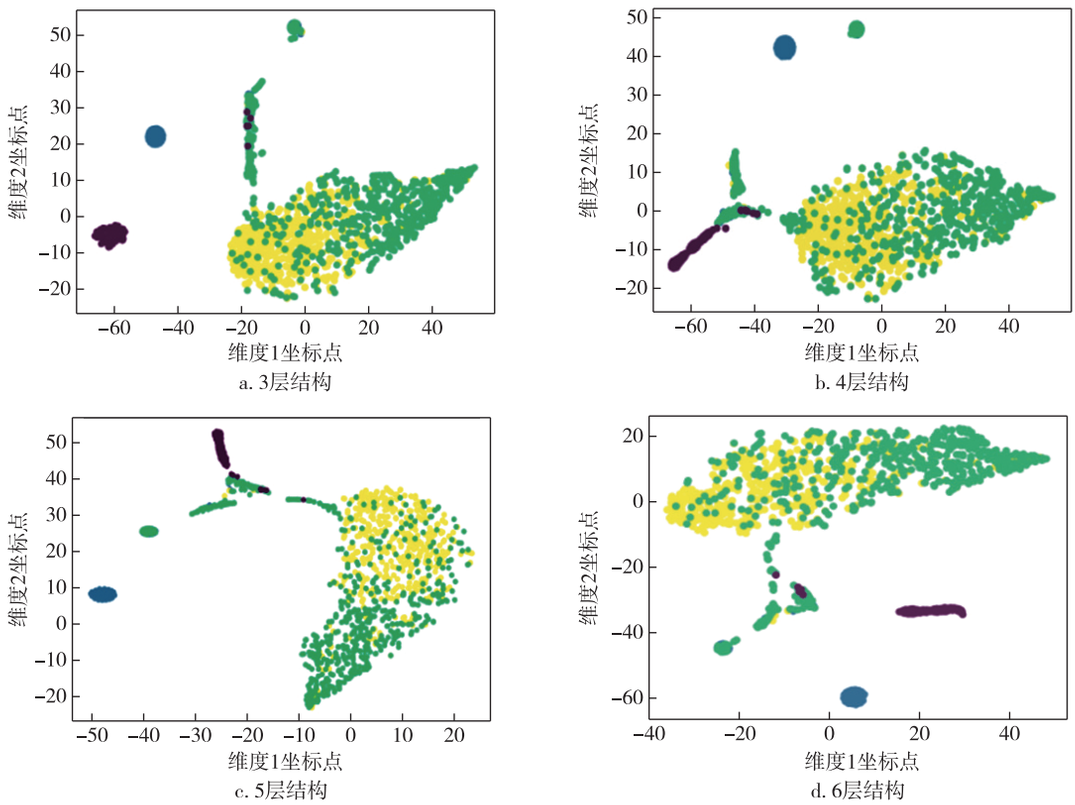


图 6　T-Net 不同模型结构散点图分布

Fig. 6　Scatter plots of different T-Net models

进行数据编码,对比不同结构模型的编码部分对测试集的分类能力、误报率以及检测时间,来验证该方法的有效性,结果对比如图 8 所示.

3.6　DDoS 攻击检测效果分析

通过分析在不同数据集中 S-Net 的效果来验证本文检测方法的有效性,结果如图 9 所示.

为了验证本文方法的有效性,通过对比不同模型在两组数据集的准确率、召回率、F1 评价指标、检测时间、误报率的不同表现来验证,对比结果如表 5 所示,通过实验结果可以得出,本文方法要优于对比方法.

4　结论

1)利用离散小波变换作为频率分析的主要工具,保证了频率和时间的对应关系,很好地区分 DDoS 攻击流量与正常流量的频率变化,并通过离散信号提取平均数、方差、百分位数、偏度、峰度和熵的统计量作为辅助特征,为降低检测误报率提供可能性.

2) 采用自动编码器作为非线性特征降维的神经网络,对高维特征向量进行低维的非线性映射,通过自动编码器将高维特征降维,以及不同层数自动

表 5　不同算法下的计算结果

Table 5　Calculation results under different algorithms

模型	评价指标	CICDDoS2019	Iot-23
CNN[7]	准确率/%	96.07	95.37
	召回率/%	96.90	94.26
	F1 评价指标/%	96.46	94.77
	检测时间/(ms/条)	2.84	2.88
	误报率/%	5.35	5.89
LSTM[17]	准确率/%	97.31	96.52
	召回率/%	96.53	96.33
	F1 评价指标/%	97.01	96.47
	检测时间/(ms/条)	2.99	3.23
	误报率/%	4.19	4.95
DNN[20]	准确率/%	97.74	97.21
	召回率/%	96.99	96.88
	F1 评价指标/%	97.30	97.11
	检测时间/(ms/条)	3.99	4.43
	误报率/%	4.28	4.73
DWT-AKD	准确率/%	97.85	97.36
	召回率/%	98.39	97.79
	F1 评价指标/%	98.24	97.58
	检测时间/(ms/条)	2.01	2.05
	误报率/%	3.53	3.81

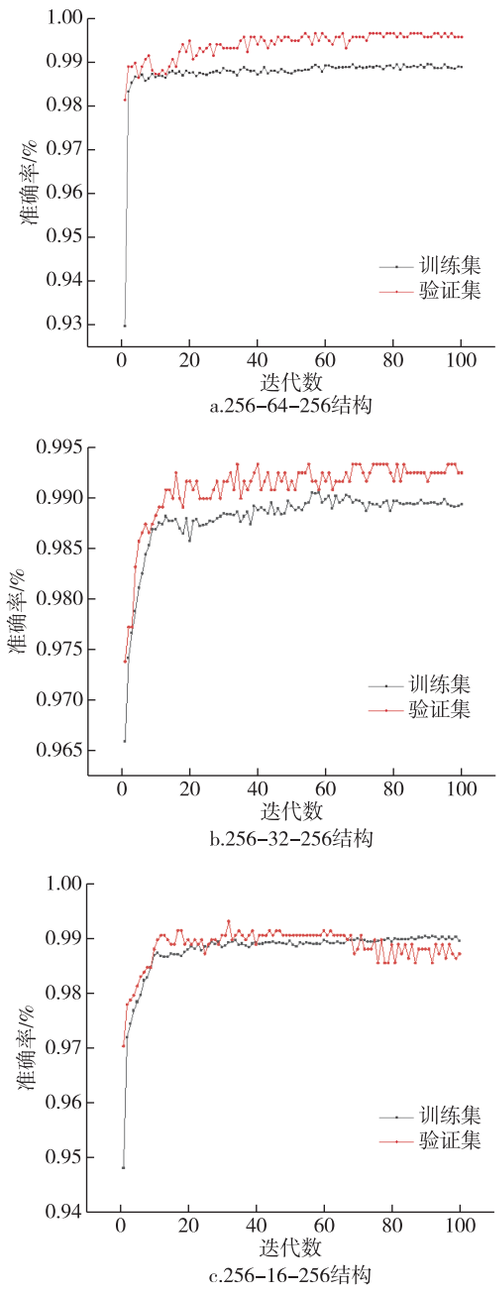


图 7　蒸馏训练准确率曲线

Fig. 7　Distillation training accuracy curves

编码器对数据聚类实验结果对比,验证了该方法的有效性,并利用编码器部分实现有效分类.

3）结合知识蒸馏方法将模型进行压缩处理,并采用自适应的方法使其快速拟合,通过实验证明,压缩后的模型对单条数据的检测时间相对于原模型平均减少了 2.83 ms,并且误报率和检测精度与原模型平均误差均保持在 1%内,证明了该方法的有效性.

4）结合两种数据集验证表明,本文方法在检测准确率及检测效率上具有较大提升,而误报率大幅降低.

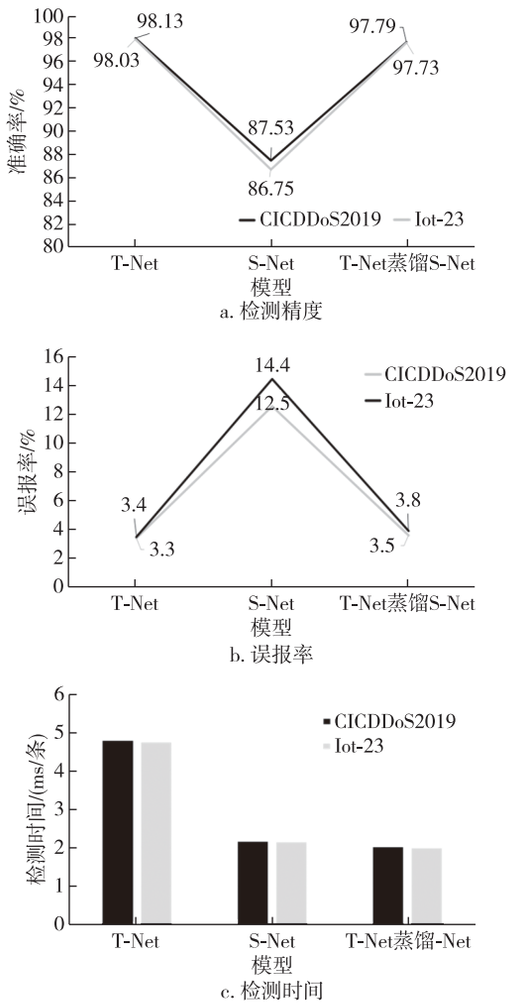


图 8　不同结构模型实验对比

Fig. 8　Experimental comparison of different structural models

参考文献

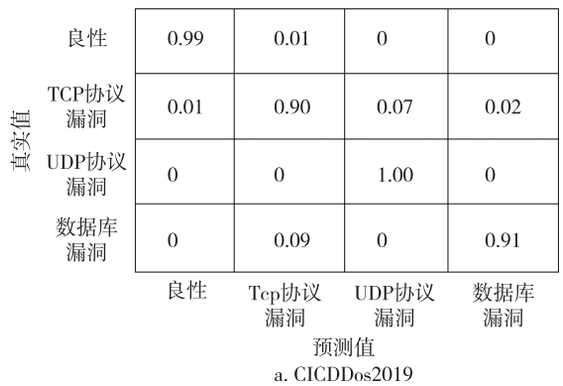
References

[1] 国家互联网应急中心.我国 DDoS 攻击资源分析报告[R].北京:国家互联网应急中心,2022
National Internet Emergency Response Center. China DDoS attack resources analysis report [R]. Beijing: National Internet Emergency Response Center,2022

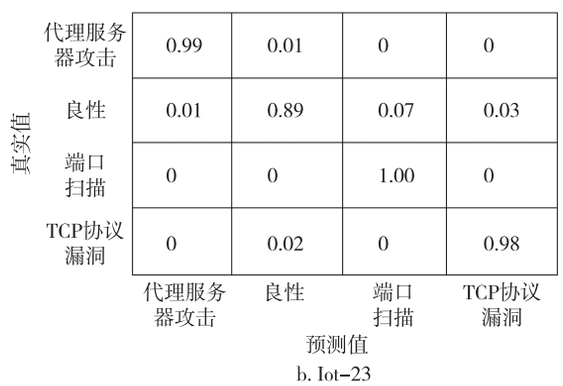
[2] Awan M J, Farooq U, Babar H M A, et al. Real-time DDoS attack detection system using big data approach [J].Sustainability,2021,13(19):10743

[3] Ahuja N,Singal G.DDoS attack detection & prevention in SDN using OpenFlow statistics[C]//2019 IEEE 9th International Conference on Advanced Computing. Tiruchirappalli,India.IEEE,2019:147-152

[4] Wang M,Lu Y Q,Qin J C.A dynamic MLP-based DDoS attack detection method using feature selection and feedback[J].Computers & Security,2020,88:101645



a. CICDDos2019



b. Iot-23

图 9 S-Net 在不同数据集上的混淆矩阵

Fig. 9 S-Net confusion matrix on different data sets

[5] Doshi R, Apthorpe N, Feamster N. Machine learning DDoS detection for consumer Internet of Things devices [C]//2018 IEEE Security and Privacy Workshops. San Francisco, CA, USA. IEEE, 2018: 29-35

[6] Odumuyiwa V, Alabi R. DDoS detection on Internet of Things using unsupervised algorithms [J]. Journal of Cyber Security and Mobility, 2021: 569-592

[7] Yang K, Zhang J J, Xu Y, et al. DDoS attacks detection with autoencoder[C]//2020 IEEE/IFIP Network Operations and Management Symposium. Budapest, Hungary. IEEE, 2020: 1-9

[8] Huang C, Yi P, Zou F T, et al. CCID: cross-correlation identity distinction method for detecting shrew DDoS[J]. Wireless Communications and Mobile Computing, 2019, 2019: 6705347

[9] Agrawal N, Tapaswi S. Defense mechanisms against DDoS attacks in a cloud computing environment: state-of-the-art and research challenges [J]. IEEE Communications Surveys & Tutorials, 2019, 21(4): 3769-3795

[10] Zhou L Y, Guo H Q, Deng G L. A fog computing based approach to DDoS mitigation in IIoT systems[J]. Computers & Security, 2019, 85: 51-62

[11] Agrawal N, Tapaswi S. Low rate cloud DDoS attack defense method based on power spectral density analysis [J]. Information Processing Letters, 2018, 138: 44-50

[12] Jose A S, Nair L R, Paul V. Towards detecting flooding DDoS attacks over software defined networks using machine learning techniques[J]. Revista Gestão Inovação e Tecnologias, 2021, 11(4): 3837-3865

[13] Mishra A, Gupta N, Gupta B B. Defense mechanisms against DDoS attack based on entropy in SDN-cloud using POX controller [J]. Telecommunication Systems, 2021, 77(1): 47-62

[14] Idhammad M, Afdel K, Belouch M. Semi-supervised machine learning approach for DDoS detection[J]. Applied Intelligence, 2018, 48(10): 3193-3208

[15] Liu Z P, He Y P, Wang W S, et al. DDoS attack detection scheme based on entropy and PSO-BP neural network in SDN[J]. China Communications, 2019, 16(7): 144-155

[16] 周奕涛,张斌,刘自豪.基于多模态深度神经网络的应用层 DDoS 攻击检测模型[J].电子学报,2022,50(2):508-512
ZHOU Yitao, ZHANG Bin, LIU Zihao. Application layer DDoS detection model based on multimodal deep learning neural network [J]. Acta Electronica Sinica, 2022, 50(2): 508-512

[17] Li Y, Lu Y F. LSTM-BA: DDoS detection approach combining LSTM and Bayes[C]//2019 Seventh International Conference on Advanced Cloud and Big Data (CBD). Suzhou, China. IEEE, 2019: 180-185

[18] 张安琳,张启坤,黄道颖,等.基于 CNN 与 BiGRU 融合神经网络的入侵检测模型[J].郑州大学学报(工学版),2022,43(3):37-43
ZHANG Anlin, ZHANG Qikun, HUANG Daoying, et al. Intrusion detection model based on CNN and BiGRU fused neural network[J]. Journal of Zhengzhou University (Engineering Science), 2022, 43(3): 37-43

[19] 赵志强,易秀双,李婕,等.基于 GR-AD-KNN 算法的 IPv6 网络 DoS 入侵检测技术研究[J].计算机科学,2021,48(增刊 1):524-528
ZHAO Zhiqiang, YI Xiushuang, LI Jie, et al. Research on DoS intrusion detection technology of IPv6 network based on GR-AD-KNN algorithm[J]. Computer Science, 2021, 48(sup1): 524-528

[20] Wang Z D, Li Z Y, He D J, et al. A lightweight approach for network intrusion detection in industrial cyber-physical systems based on knowledge distillation and deep metric learning [J]. Expert Systems with Applications, 2022, 206: 117671

[21] Liu X Q, Ren J D, He H T, et al. Low-rate DDoS attacks detection method using data compression and behavior divergence measurement[J]. Computers & Security, 2021, 100: 102107

[22] Liu Z, Hu C Z, Shan C. Riemannian manifold on stream data: Fourier transform and entropy-based DDoS attacks detection method [J]. Computers & Security, 2021, 109: 102392

[23] Fouladi R F, Ermiş O, Anarim E. A DDoS attack detection and countermeasure scheme based on DWT and auto-encoder neural network for SDN [J]. Computer Networks, 2022, 214: 109140

[24] Bornmann L. How to analyse percentile impact data meaningfully in bibliometrics: the statistical analysis of distributions, percentile rank classes and top-cited papers[J]. arXiv e-print, 2012, arXiv: 1206.1741

DDoS attack detection via DWT and AKD auto-encoder

WANG Bo[1] WAN Liang[1] LIU Mingsheng[1] SUN Handi[1]

1 College of Computer Science and Technology/State Key Laboratory of Public Big Data,Guizhou University,Guiyang 550025

Abstract To address the low efficiency and high false alarm rate in detection of DDoS (Distributed Denial of Service) flood attacks,this paper proposes a DWT (Discrete Wavelet Transform) and AKD (Adaptive Knowledge Distillation) self-encoder neural network based approach to detect DDoS attacks.The approach uses the DWT to extract frequency features,the auto-encoder neural network to encode and classify the features,and the AKD to compress the model in order to achieve efficient detection of DDoS attacks.The results show that the approach has high detection efficiency for proxy server attacks,database vulnerabilities & TCP flood attacks,and UDP flood attacks,with low false alarm rate.

Key words DDoS attack;discrete wavelet transform (DWT);adaptive;knowledge distillation;auto-encoder