基于深度学习的恶意流量分类检测方法

1024040804 陈果

摘要

目前社会越来越离不开网络，但是使用网络就要面对越来越具有威胁性的恶意流量，这些流量的分类检测问题也越来越受到关注。需要研究有效的方法，并运用方法筛选出恶意流量，从而提高网络资源的安全性。

传统的恶意流量分类采用基于特征学习的机器学习方法，相对来说性能不足以应对日渐复杂的恶意流量威胁，但是以机器学习为基础的深度学习能够更好的解决这个问题。本文主要从模型的角度分析课题，通过研究深度学习的几种经典模型，并比对流量分类特点，选取了卷积神经网络模型 CNNs 作为主要研究模型。经过深入研究，CNNs 模型理论上能够实现流量分类，此外，因为 LSTM 结构也可用于流量分类，因此本文尝试了一种以 CNNs 为主体、加入一些 LSTM 思想的结构设计，并通过实验证明该结构的可行性，能够用于恶意流量分类。此外，通过对比试验，分析了该组合模型在不同的参数设置下的学习效果，可以根据具体的网络安全要求修改对应的数据，在精确度等参数与计算机资源占用之间做出取舍。

关键词: 深度学习; 恶意流量; 卷积神经网络; 流量分类

ABSTRACT

Currently, society is becoming increasingly dependent on the Internet.However, using the Internet also means facing increasingly threatening malicious traffic. The issues surrounding the classification and detection of this traffic have become a major concern in the field. Effective methods need to be researched and utilized to filter out malicious traffic, in order to improve the security of network resources.

Traditional malicious traffic classification adopts machine learning methods based on feature learning, which are relatively insufficient in dealing with increasingly complex malicious traffic threats. However, deep learning, based on machine learning, can better solve this problem. This article mainly analyzes the topic from the perspective of models, and selects convolutional neural network model(CNNs) as the main research model by comparing traffic classification characteristics after studying several classical deep learning models. Through in-depth research, CNNs model can theoretically achieve traffic classification. In addition, because LSTM structure can also be used for traffic classification, this paper attempts a structure design that takes CNNs as the main body and incorporates some LSTM ideas. The feasibility of this structure is proved through experiments,which can be used for malicious traffic classification. In addition, through comparative experiments, the learning effect of this combined model under different parameter settings is analyzed. According to specific network security requirements, corresponding data can be modified, and trade-offs can be made between accuracy, other parameters, and computer resource consumption.

Key words: deep learning; malicious traffic;convolutional neural networks traffic classification

1 绪论

在互联网快速普及的时代，互联网带给我们巨大便利的同时，其所带来的威胁也日益增多，网络安全的问题也渐渐走到大众的面前。其中，恶意流量对用户和互联网本身造成的威胁如何解决，已经成为社会重点关注的问题。但是随着科技进步，恶意流量的形式逐渐多样化，传统的单一的检测方式 (包括简单的规则匹配) 已经不能满足与日俱增的网络安全需求，因此，人们尝试将深度学习的理念融入该领域。深度学习代表的是一种自动学习的模式，因为其能够自主完成学习和工作，因此在各类领域得到广泛的运用，并且能够通过学习，做到与时俱进，应对各领域新的挑战。显然，从大方向看，深度学习能在很大程度上解决当下恶意流量对人们造成的困扰，但是如何完成、能完成到什么程度，还要进行深入的研究。

1.1 1024040804 陈果-课题的研究背景与研究意义

随着互联网和物联网技术的普及和发展，恶意流量对网络安全的威胁也越来越大。传统方法大多关注已经发现的问题，并基于已知的部分去设计检测方式，很难检测到越来越灵活的网络攻击，例如新型的网络攻击方式、无规则数据包、通过加密手段隐藏的攻击等，即便发现了新的攻击方式，在设计检测方法上仍要花费大量时间，这期间恶意流量造成的损失可能是无法估量的。因此，需要开发更加高级的方法来监测和识别恶意行为。

1.2 国内外研究现状

目前国内外有很多对深度学习 [1] 的研究，并不局限于对流量检测的应用。对深度学习的研究可以追溯到上个世纪中期，本来用于人工智能的研究，虽然受限于计算机对数据较弱的处理能力 [2]，但是学者们一直在不断的提出模型，并尝试实现这些模型。到了本世纪，深度学习的研究走上了正轨。

2 基于深度学习的恶意流量分类方法理论

2.1 深度学习概况

深度学习 (DeepLearning) 是一种基于神经网络架构的机器学习方法, 可以从大量的数据中获取抽象的特征和模式，并利用这些特征和模式对未知的数据进行分类、回归、聚类等任务。其发展历程可以追溯到上世纪 60 年代，当时提出了称为感知机的浅层神经网络。但是由于感知机只能解决线性可分问题，无法处理复杂的非线性问题，因此深度学习的发展在很长时间内受到了限制。

2.2 在流量分类中可应用的深度学习模型

深度学习目前拥有各种各样的模型，不同的模型之间拥有不同的适用范围，因此，选出其中适合流量分类的模型十分重要，这就要对流量本身特征做出一定分析。

2.3 CNN

卷积神经网络 (Convolutional Neural Network，CNN) 是深度学习的重要分支，其主要应用于计算机视觉领域，如图像分类、目标检测等。CNN 以其独有的卷积操作和池化操作 [3]，可有效提取图像特征，从而实现图像分类的任务。本小节将详细研究 CNN 的工作原理。

表 1: CNN 典型网络结构

层类型	输入尺寸	输出尺寸	主要特性
卷积层 (Conv)	224×224×3	224×224×64	使用 3×3 卷积核，ReLU 激活函数
池化层 (Pool)	224×224×64	112×112×64	最大池化，2×2 窗口
全连接层 (FC)	4096	1024	包含 Dropout(0.5) 防止过拟合 L2 正则化约束权重

注: 表格展示了 VGG 网络的简化结构

以上是典型的卷积神经网络 (CNN) 结构，如表 1 所示。

3 基于 CNN 的恶意流量检测方法

3.1 CNN 在流量分类检测中的优势

综上所述，CNN 模型在流量检测中具有自动特征提取、可扩展性、处理时序特征、快速训练和推理速度、高鲁棒性等优点，十分适合解决流量分类相关的问题。

3.2 CNN 的具体工作原理

卷积神经网络主要由卷积层、池化层、激活函数和全连接层组成。

1. **局部感受野（Local Receptive Fields）**：
   - 核心思想：模拟生物视觉系统的感受野机制，每个神经元仅处理局部区域的输入信息
   - 实现方式：通过卷积核（Kernel）在输入数据上滑动，每次处理 3×3 或 5×5 的小窗口
   - 技术优势：显著减少参数数量，保持空间局部特征关联性
2. **卷积运算（Convolution Operation）**：
   - 特征抽取：卷积核作为特征检测器，不同核分别提取边缘、纹理、角点等初级特征
   - 多通道处理：对 RGB 图像或流量数据特征图，采用 3D 卷积核同步处理多通道
3. **激活函数（Activation Function）**：
   - 作用：引入非线性转换能力，增强模型表达能力
   - 典型函数：ReLU（Rectified Linear Unit）：$f(x) = \max(0, x)$
   - 优势：解决梯度消失问题，计算复杂度远低于 Sigmoid/Tanh
4. **池化层（Pooling Layer）**：
   - 目的：降维采样，保持特征平移不变性
   - 主要方法：
     - 最大池化（Max Pooling）：提取局部最显著特征
     - 平均池化（Average Pooling）：保留区域均值信息
   - 效果：减少参数数量约 75
5. **层级特征抽象（Hierarchical Feature Abstraction）**：

   卷积网络通过多层堆叠形成特征提取金字塔：
   - 初级层：提取基础特征（如边缘、角点）
   - 中级层：组合基础特征（如纹理、形状组件）
   - 高级层：生成语义特征（如完整对象）
6. **全连接层（Fully Connected Layer）**：
   - 作用：将高维特征映射到类别空间
   - 连接方式：前层所有神经元与本层所有神经元全连接
   - 输出处理：Softmax 函数计算类别概率分布

反向传播优化：通过梯度下降算法（如 Adam 优化器）和反向传播更新卷积核权重：

$$\Delta W = -\eta \frac{\partial \mathcal{L}}{\partial W} + \lambda \cdot W$$

其中 η 为学习率，$\mathcal{L}$ 为损失函数（如交叉熵），λ 为 L2 正则化系数。

在恶意流量检测任务中，CNN 将原始网络流量数据重组为类图像结构：

$$\text{特征图} = \begin{bmatrix} p_1^t & p_2^t & \cdots & p_n^t \\ p_1^{t+1} & \vdots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ p_1^{t+k} & \cdots & \cdots & p_n^{t+k} \end{bmatrix}$$

通过多层特征变换，最终识别出 DDoS、SQL 注入、木马通信等攻击流量的深层模式特征。

3.2.1 1024040804 陈果-卷积层

一种叫空洞卷积的卷积方式，定义 F 为一个离散函数时，卷积运算公式可表示为：

$$(F * k)(p) = \sum_{s+t=p} F(s)k(t)$$

3.2.2 池化层

如果输入数据体尺寸为 W1、H1、D1，有两个超参数：空间大小 FF 和步长 SS；输出数据体的尺寸 W2、H2、D2，则其中公式如下：

$$W_2 = \frac{W_1 - F}{S} + 1$$
$$H_2 = \frac{H_1 - F}{S} + 1$$
$$D_2 = D_1$$

4 实验

4.1 实验软件及其操作

目前深度学习主流的实现方式有 Python 语言、MATLAB 语言等，但是 MATLAB 语言做深度学习测试的过程相对复杂，本文作者水平受限，加上本文作者的指导教师推荐使用 Python 语言，故本章节选取 Python 语言。

4.2 实验结果分析

如下图设置，其中 epochs 表示迭代次数，batch-size 表示批量处理的大小，validation.split 表示使用数据。

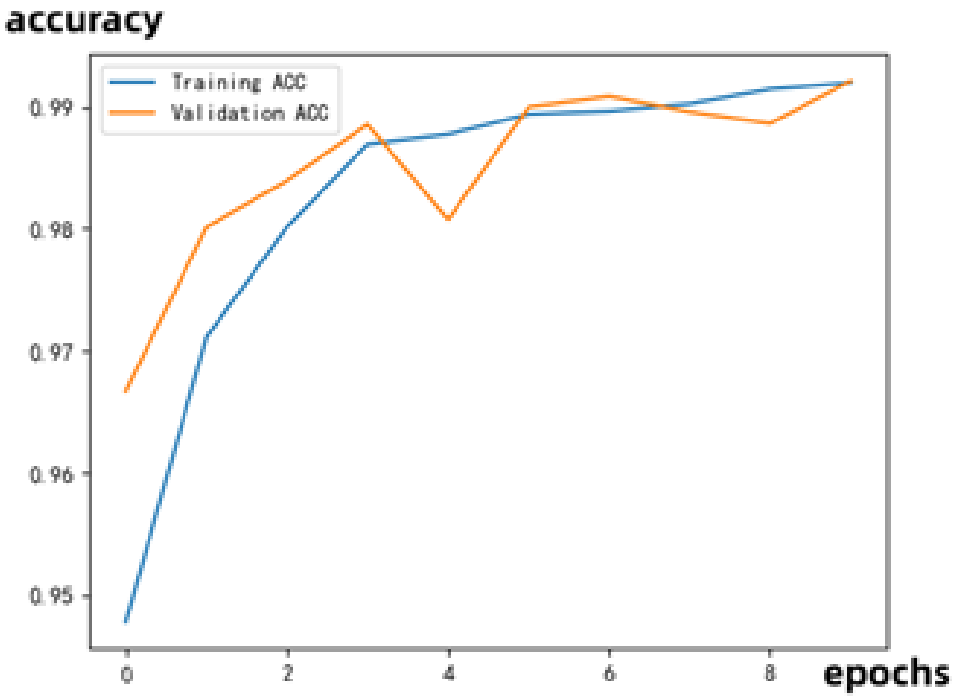

图 1: 改进后的神经网络准确度

5 总结与展望

5.1 总结

本文分析了当今时代恶意流量对网络的危害性，随着日新月异的发展，人们的生活越来越离不开网络，因此，找出一种有效检测恶意流量的方法十分重要。

5.2 1024040804 陈果-展望

本文对基于深度学习的恶意流量分类方法进行研究，得到了一些成果，但是本文也有一些不足。

1. **实时检测能力提升**：未来研究将更关注模型轻量化与推理加速技术，如知识蒸馏、模型剪枝等，使深度学习模型能够满足大规模网络环境下的实时流量监测需求，有望实现微秒级响应速度。
2. **零日攻击防御突破**：通过结合自监督学习和小样本学习方法，新型检测模型将减少对大量标注样本的依赖，能够更快识别从未见过的攻击模式，有效应对新型恶意流量威胁。
3. **多模态联合分析**：融合网络流特征、协议行为模式、终端环境上下文等多维数据源，构建多模态深度学习框架，可显著提高检测精度并降低误报率。
4. **对抗攻防研究深化**：针对攻击者刻意设计的对抗性样本，需要发展具有鲁棒性的防御机制，如对抗训练、特征净化等技术，确保检测模型在对抗环境中的稳定性。
5. **隐私保护计算集成**：结合联邦学习、差分隐私等隐私计算技术，可在不泄露原始数据的前提下实现多方安全协作，解决数据孤岛问题，构建分布式恶意流量检测体系。
6. **可解释性增强**：通过 Attention 机制、特征可视化等技术增强模型透明度，使检测结果具有可解释性，帮助安全分析师理解模型决策依据，建立人机协作防御机制。

随着 AI 芯片算力的持续提升和网络威胁情报共享机制的完善，深度学习驱动的智能威胁检测系统将逐步实现主动防御、预测预警和自动响应的闭环能力，成为新一代网络安全体系的核心防御层。

致谢

毕业即将来临，作为毕业前的最后一段路，毕业设计让我感觉十分充实，在此做出一些感谢。

参考文献

[1] 作者 A. 深度学习发展综述 [J]. 计算机学报, 2010, 33(5): 1020-1030.

[2] 作者 B. 基于深度学习的图像识别方法研究 [D]. 某大学博士学位论文, 2015.

[3] 作者陈果. 基于 1024040804 的 latex 写作方法研究 [D]. 某大学硕士学位论文, 2025.