基于深度学习的网络入侵检测技术研究

李云亚　张　菲　邵君立
金盾检测技术股份有限公司　江苏南京　210000

摘　要：随着网络攻击手段的不断升级，传统的入侵检测技术已难以应对复杂多变的安全威胁。基于深度学习的网络入侵检测技术，凭借其强大的特征学习能力和自适应能力，逐渐成为解决网络安全问题的有效手段。深度学习能够通过大量数据的训练，自动提取入侵行为的复杂特征，并进行高效的检测与分类。本文将深入探讨深度学习在网络入侵检测中的应用，重点分析深度神经网络（DNN）、卷积神经网络（CNN）和循环神经网络（RNN）等深度学习模型在该领域的技术实现。

关键词：网络入侵检测；深度学习；深度神经网络；卷积神经网络；循环神经网络

引言

随着信息技术的快速发展，网络安全问题愈加严重。传统的网络入侵检测方法依赖于规则匹配和特征库更新，虽然在某些环境下取得了一定的成效，但面对日益复杂的攻击方式和大规模数据流，这些方法的适应性和准确性已显得不足。深度学习技术作为人工智能的重要分支，凭借其在数据分析中的卓越能力，逐渐成为网络入侵检测的新兴技术。深度学习通过自动化特征学习，能够发现数据中潜藏的深层次模式，特别是在处理高维数据和复杂关系方面表现出色。本文将探讨深度学习在网络入侵检测中的应用现状，分析其技术优势，并着重介绍常见的深度学习模型如何在提升入侵检测系统的效率和准确性方面发挥作用。

一、网络入侵检测技术概述

（一）网络入侵检测的基本概念

网络入侵检测是通过对计算机网络流量和系统活动的监控，识别是否存在恶意攻击或不正常行为的一种技术手段。它主要通过分析网络中的数据包、协议、流量模式等信息，及时发现入侵活动。入侵检测系统（IDS）通常被部署在网络边界或者主机上，负责监测潜在的安全威胁。其基本目标是检测未经授权的访问、恶意代码的传播、数据泄露等。随着互联网的发展，入侵方式不断演变，传统的安全防护手段已难以适应复杂多变的网络攻击，因此，网络入侵检测技术的重要性不断上升。网络入侵检测技术通常分为基于签名的检测和基于异常的检测两大类。前者通过已有的攻击模式进行匹配，后者则通过分析正常行为的偏离来识别入侵。有效的入侵检测不仅能够及时发现安全漏洞，还能防止数据泄露和损失[1]。

（二）传统网络入侵检测技术概述

传统的网络入侵检测技术主要包括基于签名的检测方法和基于异常的检测方法。基于签名的检测方法依赖于已知攻击的特征，利用预定义的攻击库进行匹配来识别入侵。此类方法对已知攻击有效，但对于新型或变种攻击的检测能力较弱。另一类传统方法是基于异常检测的技术，该方法通过建立正常行为模型，检测与正常模式偏离的活动，以此识别潜在的入侵。虽然基于异常的检测可以识别未知攻击，但往往面临高误报率的问题。传统技术通常依赖人工定义的规则和特征，这使得系统在处理大规模数据时效率较低，并且难以应对复杂多变的攻击手段。随着网络攻击手段的多样化和智能化，传统的入侵检测技术逐渐显现出不足之处，因此需要引入更加先进的技术手段来提升检测能力和准确性。

二、深度学习模型概述

（一）深度神经网络（DNN）

深度神经网络（DNN）是一种多层的前馈神经网络，它通过多个神经元层的逐层计算，能够自动提取输入数据的高级特征。DNN能够通过反向传播算法进行训练，使得网络的参数得到优化。与传统神经网络相比，DNN具有更多的隐藏层，这使得其能够处理更加复杂的数据模式。在网络入侵检测中，DNN的优势在于能够自动学习到数据中潜在的入侵模式，而不需要人工干预。通过大规模的训练数据，DNN能够逐渐优化网络权重，提升

模型的分类能力。DNN可以处理非线性关系，并且具有强大的自适应能力，对于应对多变的网络攻击行为具有较强的优势。其在入侵检测中的应用，可以提高检测系统对各种已知与未知攻击的识别能力，尤其适合复杂攻击模式的识别。

（二）卷积神经网络（CNN）

卷积神经网络（CNN）是一种特别适用于图像处理的深度学习模型，它通过卷积层、池化层和全连接层的组合，能够有效地提取数据中的空间特征。CNN通过局部感知域和权重共享的方式，减少了模型的复杂度，同时增强了对特征的学习能力。在网络入侵检测中，CNN可以用于处理网络流量数据，通过卷积操作提取出数据包中的重要特征，从而识别潜在的入侵行为。CNN在入侵检测中的优势在于能够对复杂的网络流量进行深层次的特征挖掘，而不依赖于人工定义的特征。通过卷积层的层层过滤，CNN能够自动识别攻击模式，并且具有较强的空间特征学习能力。CNN在处理网络入侵的实时性和准确性方面表现出色，尤其适合大规模网络流量的检测[2]。

（三）循环神经网络（RNN）

循环神经网络（RNN）是一种适用于处理序列数据的深度学习模型。与传统神经网络不同，RNN具有“记忆”功能，能够处理具有时间序列特性的输入数据。在网络入侵检测中，RNN能够对网络流量中的时序信息进行建模，从而识别出基于时间变化的攻击模式。网络流量通常具有时间相关性，攻击行为往往是通过一系列的网络事件逐步展开，RNN能够通过其内在的记忆机制，捕捉到这些时间上的依赖关系。因此，RNN在检测分布式攻击、持久性攻击等复杂网络攻击中具有独特的优势。RNN的一个主要优点是能够处理任意长度的输入序列，对于流量数据中长时间跨度的依赖关系有着良好的建模能力。对于那些通过长时间交互过程实施的攻击，RNN能有效地发现并检测出攻击的迹象。

三、基于深度学习的入侵检测技术

（一）数据预处理与特征提取

数据预处理与特征提取是构建深度学习模型的关键步骤。在网络入侵检测中，原始数据通常是复杂且噪声较多的，直接使用这些数据进行训练容易导致低效的学习结果。因此，需要对原始数据进行清洗、归一化、去噪和填补缺失值等处理，以提高数据质量。比如，对于每个网络数据包，提取其源IP、目标IP、协议类型等特征并进行编码。在特征提取阶段，常用的方法包括基于流量特征的统计分析，例如计算每秒的传输数据量、连接持续时间等。通过这些特征，可以有效减少模型输入的维度，并保留最有意义的信息。对于入侵检测而言，特征提取技术的效果直接影响到后续模型的表现。在一些数据集上，如NSL-KDD，经过特征选择和优化后，检测准确率可提升约5%至10%。数据预处理与特征提取不仅能够提升模型的训练效率，还能显著提高入侵检测的准确率。

（二）基于DNN的入侵检测模型

深度神经网络（DNN）在网络入侵检测中的应用已成为一种主流技术。DNN能够通过多层的神经元结构学习复杂的非线性关系，适应大规模数据集的需求。在入侵检测任务中，DNN通过输入层接收网络流量的多维特征，经过隐藏层的逐层处理后，输出层给出各类行为的预测结果。使用DNN进行入侵检测的优势在于其能够自动学习特征并根据数据的复杂性进行适应。实验结果表明，基于DNN的模型在标准数据集上的分类准确率通常能达到98%以上。例如，使用NSL-KDD数据集进行训练时，DNN模型的检测准确率为98.5%，误报率低于2%。相比传统方法，DNN能够显著减少人工特征提取的工作量，并通过深层次的非线性映射提升对复杂攻击的识别能力。通过适当调整网络结构与训练过程，DNN还能够在面对动态变化的网络环境时具有较强的自适应性。

（三）基于CNN的入侵检测模型

卷积神经网络（CNN）在处理具有局部空间相关性的任务中表现出了强大的能力。对于网络入侵检测，CNN能够通过卷积层提取输入数据的局部特征，如流量的时间序列、协议模式等。CNN采用权重共享机制，减少了模型的参数量，提升了计算效率。在入侵检测中，CNN通过多层卷积与池化操作逐步抽取高级特征，有助于提高检测的精度。使用CNN进行入侵检测时，通常会将网络流量的特征矩阵作为输入，经过卷积和池化后，得到每个数据包或流量段的高层特征表示。实验表明，基于CNN的入侵检测模型在识别攻击类型时，准确率达到99%以上。例如，在CICIDS 2017数据集上，CNN模型的准确率为99.2%，在复杂攻击模式的识别上表现优异。CNN模型的优势不仅体现在自动特征提取上，还在于其能够通过卷积层增强对局部特征的学习能力，从而在多种网络环境下都能保持较高的检测性能[3]。

四、深度学习网络入侵检测模型的训练与优化

（一）训练数据的构建与准备

训练数据的构建与准备是深度学习入侵检测模型中的基础环节。有效的训练数据能够直接影响模型的学习效果和最终性能。在网络入侵检测中，训练数据通常来源于历史的网络流量数据集，如KDD Cup 99、NSL-KDD等。在数据准备过程中，需根据不同的攻击类型和正常行为标注数据，以保证训练集的多样性和代表性。此外，数据集的规模也是关键因素，大规模的数据集能够有效提升模型的泛化能力。通过增加数据的多样性，例如使用不同类型的攻击样本以及变化的网络环境，可以显著提高模型的鲁棒性。数据增强技术在数据准备中也起到了重要作用，尤其在数据不平衡时，通过对少数类别样本进行增强，能够避免模型偏向于某一类样本。在某些实验中，经过合理的训练数据准备，模型的准确性提高了5%到10%。

（二）训练方法与参数优化

深度学习模型的训练方法和参数优化对模型性能的影响至关重要。在训练深度学习模型时，选择合适的优化算法至关重要。常用的优化算法如Adam、SGD（随机梯度下降）等，通过调整学习率和其他超参数，能够帮助模型在较短的时间内收敛并达到最优状态。在入侵检测中，模型的学习率通常设置为0.001到0.01之间，而批量大小（batch size）则根据数据集大小和内存限制进行调整。除了选择优化算法，参数优化也是提升模型性能的关键步骤。例如，通过调整隐藏层神经元的数量、使用不同的激活函数，能够显著提高模型的分类能力。在一些实验中，通过调整这些超参数，模型的训练误差降低了20%以上，准确率提升了10%左右。使用网格搜索和随机搜索等方法进行参数调优，也有助于找到最佳的超参数组合。

（三）模型评估与性能测试

在深度学习网络入侵检测模型的训练和优化后，评估模型的性能是不可忽视的步骤。常用的评估指标包括准确率、精确率、召回率、F1分数等，这些指标能够全面反映模型在真实环境中的表现。例如，准确率能够反映模型正确分类的比例，精确率和召回率则分别反映模型对攻击样本的识别能力与对正常样本的识别能力。在性能测试中，常用的测试数据集如NSL-KDD、CICIDS 2017等，能够提供多种类型的攻击数据，帮助评估模型的综合表现[4]。实验表明，基于深度学习的入侵检测模型在这些数据集上的准确率通常超过95%，而F1分数在0.9以上。此外，针对模型的实时性表现，测试系统的响应时间和计算效率也是评估模型的重要指标。在某些实时检测系统中，模型的处理速度要求能够在毫秒级别完成，以确保及时检测并响应入侵事件。

结语

通过本文的研究，可以看出，基于深度学习的网络入侵检测技术在提升入侵检测系统性能方面具有显著优势。深度神经网络（DNN）、卷积神经网络（CNN）和循环神经网络（RNN）等模型，能够自动从大量数据中提取特征，学习到复杂的攻击模式，极大地提升了入侵检测的准确率和实时性。尤其在处理复杂的攻击行为时，深度学习模型表现出了强大的适应性和高效性。随着数据量和计算能力的不断增长，深度学习将在网络安全领域发挥越来越重要的作用。尽管目前仍面临训练数据、模型可解释性等问题，但随着技术的不断发展，这些问题有望得到解决。未来，深度学习技术将继续推动网络入侵检测系统的创新，成为保障网络安全的重要手段。

参考文献

[1]吴祖康.基于深度学习的网络入侵检测系统设计[J].网络安全技术与应用，2025，(10)：64-66.

[2]孟禹.基于深度学习的网络入侵检测技术研究[J].鞍山师范学院学报，2025，27(02)：56-60.

[3]杨博.基于深度学习的车载网络入侵检测系统关键技术研究[D].导师：向艳萍.电子科技大学，2025.

[4]曲彦泽.基于深度学习的网络入侵检测关键技术研究[D].导师：马海龙.战略支援部队信息工程大学，2023.