基于人工智能的网络异常行为检测探析

裴旭光

昆仑数智科技有限责任公司，北京，100010；

摘要：随着数字化进程加速，网络已深度融入社会各领域，但网络攻击手段的多样化与复杂化，给网络安全带来严峻挑战。传统基于规则的防护方法难以应对新型攻击，基于人工智能的网络异常行为检测技术应运而生。本文深入剖析该技术，系统阐述网络异常行为类型及特征，详细解析机器学习、深度学习等人工智能技术在网络异常行为检测中的应用。研究表明，人工智能技术为网络安全防护提供新方向，通过技术创新有望进一步提升网络安全防护水平。

关键词：人工智能；网络异常行为；机器学习；深度学习

DOI：10.69979/3041-0673.25.12.002

引言

在数字时代下，网络已成为经济、社会和生活运转的核心基础设施。从个人日常社交娱乐到企业核心业务运营，再到国家关键信息系统保障，网络的重要性愈发凸显。然而，网络空间的安全形势却不容乐观，分布式拒绝服务攻击、恶意软件传播、网络钓鱼等异常行为频发，给个人隐私、企业利益和国家安全带来严重威胁。传统基于规则匹配和特征库的检测方法，在面对快速演变的攻击手段时，暴露出检测滞后、误报率高、难以处理海量数据等问题。人工智能技术凭借强大的学习和分析能力，为网络异常行为检测提供了创新途径。本文旨在深入研究人工智能在网络异常行为检测中的应用，探索其技术原理、实践效果、面临挑战及未来发展方向，助力提升网络安全防护能力。

1 网络异常行为相关阐述

网络异常行为是指在网络环境中，违反正常网络活动模式、规则以及安全策略的网络行为。这些行为除了因系统故障、配置错误等因素，还可能是恶意攻击行为。网络异常行为会对网络正常运行、数据完整性、保密性造成极大威胁。常见的网络异常行为包括以下几点：

（1）DDoS 攻击。该攻击方式是通过控制大量的僵尸网络，向目标服务器大量发送请求，将目标服务器资源耗尽，使其无法正常提供相关服务，如 2018 年 GitHub 遭受的 DDoS 攻击，峰值流量达到 1.35Tbps。

（2）恶意软件入侵。恶意程序包括病毒程序、木马程序、蠕虫程序等，这些程序可通过网络传播，将目标计算机系统感染（可分为有目的感染和无目的感染），切取目标计算机的敏感信息、破坏系统文件、控制被感染设备。

（3）网络钓鱼。攻击者伪装成合法结构或个人，发送欺诈性电子邮件、消息，诱使用户提供账号、密码、手机号、银行卡密码、身份证号等敏感信息。

（4）内部威胁。内部威胁主要是因企业内部人员误操作或有意攻击行为造成的网络异常情况。

目前，网络异常主要表现为流量异常、行为模式异常、协议异常。其中，流量异常表现为流量短时间内激增或激减，或者产生异常的流量模式。如网络在正常使用时非常稳定，一旦突然产生持续的高流量峰值，且与日常业务流量模式不匹配，则可能受到了 DDoS 攻击。行为模式异常为用户或系统行为不符合常规模式，如账号在短时间内在多个地区登录，或系统在非工作期间自动进行大量信息传输工作等。协议异常为网络协议使用错误或违反协议范围，如 TCP 协议中，正常三次握手过程被破坏，很可能是受到攻击。

2 人工智能技术在网络异常行为检测中的应用

2.1 基于机器学习的异常行为检测

机器学习算法可对大量网络历史数学进行学习，搭建正常网络行为的模型。以支持向量机（SVM）算法为例，该技术可将网络数据映射到高维空间，从中找出最优的分类超平面，划分正常行为数据、异常行为数据。在训练期间，算法通过不断调整分类超平面参数，降低分类误差值。其基本原理采用以下求解方法：

$$\min_{w,b}\frac{1}{2}\|w\|^2+C\sum_{i=1}^{n}\xi_i$$

$$s.t.y_i(w^T\phi(x_i)+b)\geq 1-\xi_i,\xi_i\geq 0,i=1,...,n$$

式中，w 是分类超平面的权重向量，b 是偏置项，C 是惩罚参数，ξ_i 是松弛变量，y_i 是样本的类别标签（+1 表示正常，-1 表示异常），$\phi(x_i)$ 是将样本 x_i 映

射到高维空间的函数。

网络环境十分复杂，选择合适的特征对异常检测精度有着重要影响。基于机器学习技术的特征选择与提取，可有效识别与网络异常行为有关的信息。常见网络特征包括流量大小、包长度、源/目的 IP 地址、端口号、连接持续时间等。在检测 DDoS 攻击时，流量大小、连接请求速率是至关重要的特征；检测恶意软件入侵时，文件行为特征，如创建、修改、执行频率等也十分关键。采用主成分分析（PCA）等方法，对原始数据降维处理，将冗余信息去除，提升模型训练效率、检测性能。另外，为了适应不断更新的攻击行为，提升检测精度。可采用在线学习算法，如随机梯度下降（SGD），发觉新数据时，模型可实时更新参数，以适应新的网络行为，通过动态更新机器学习模型，及时检测到异常行为。

广泛收集网络流量数据，搭建全面的数据集，用于训练机器学习模型。数据集中包含海量的正常流量数据以及多种类型的异常流量数据，数据越全、越丰富，越能让模型学习到更多类型的异常行为模式特征。除了收集网络服务商提供的网络监测数据外，还可以使用模拟工具人工生成异常流量数据，使收集的数据更加多元化。选择适宜的特征，借助机器学习算法对其优化，提升异常检测精准度。借助信息增益、互信息等方法评估特征重要性并赋予权重，选择信息增益较大的特征作为模型输入。同时，对特征归一化处理，将各类特征值映射到指定范围内，以免某些特征值取值范围过大影响模型训练。对于流量大小、包长度这两个特征，其取值范围相差较大，采用归一化处理，可让其在模型内具备相同权重。

完成模型训练后对其评估，根据评估结果调节模型参数，提升模型泛化能力。常用的评估指标包括准确率、召回率、F1 值等。如若模型在训练集上表现较好，但测试集上性能不足，可能存在拟合问题，此时可通过增加数据集大小、调节模型复杂度、使用正则化方法解决问题。

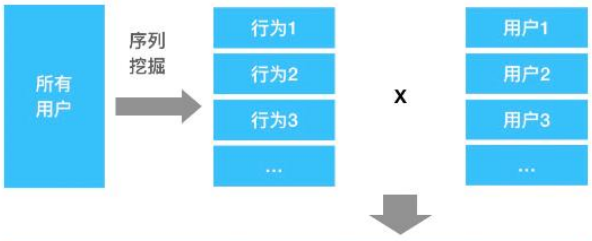
2.2 基于 CNN 与 RNN 的异常检测技术

通过实时收集、分析网络流量数据信息，人工智能可快速识别异常行为。借助网络探针等工具，在网络关键节点采集流量数据，将数据传输给人工智能平台。如 Cisco 的 NetFlow 技术能将收集网络流量源 IP、目的 IP、端口号、流量大小等信息，为人工智能分析提供数据支撑。基于人工智能的流量分析技术可识别各类异常流量模式，包括 DDoS 攻击、僵尸网络等。借助机器学习、深度学习算法学习海量历史数据，搭建正常状态下流量模型，一旦检测到流量数据和正常模式存在差异，即判定为异常流量。模型搭建中，借助聚类算法将相似流量归为一类，异常流量划分到单独类别中。

借助可视化技术，将复杂的流量数据以柱状图、折线图、饼图等图形模式直观展现，有助于安全员快手理解、分析网络状态，呈现流量变化趋势、不同类型流量占比等信息。根据可视化信息，安全员可直观看到网络流量是否存在异常情况。

2.3 基于人工智能的行为序列分析技术

如图 1 所示，基于人工智能的行为分析序列通过搭建用户行为序列模型，该模型可分析、辨别异常操作行为。以马尔科夫模型为例，建设用户的当前行为只与前一个行为有关，通过学习用户历史行为序列，搭建状态转移概率矩阵。到来新的行为序列时，根据状态转移概率矩阵计算其出现概率，概率较低的行为序列可能为异常行为。通过对用户行为序列进行分析，该技术可检测出序列异常行为，如未经授权访问、敏感信息传输等。在企业网络中，员工正常行为序列为先登录内部系统，再访问工作有关文件和数据库。一旦检测到某个严控行为序列突然访问敏感数据目录且没有经过授权流程的情况下，则判定为异常行为。



	行为1	行为2	行为3	...
用户1	1次	2次	5次	...
用户2	1次	3次	1次	...
用户3	2次	3次	2次	...
...	...	...	...	...

图 1　用户行为序列模型

为了提升检测精度，该技术需具备实时分析功能，应迅速应对快速变化的网络环境。采用 ApacheFlink 等流计算技术，对实时产生的行为序列数据实时处理，第一时间发现异常行为。

2.4 基于深度学习的异常行为检测

深度神经网络可模拟复杂的网络行为模式，通过多层次抽象提取数据内在特征，强化异常行为检测能力。以卷积神将网络（CNN）为例，其包含了卷积层、池化层、全连接层。先进数据预处理，分为三个步骤，一是

归一化，将数据缩放到相同的范围，例如[0,1]；二是矩阵化，将时间序列数据转换为二维矩阵，便于卷积操作；三是标签化，为正常流量和异常流量分配标签，例如 0 和 1。进而开展模型设计，根据实际需求进行调整，以下是一个典型的 CNN 模型结构：

输入层：接收预处理后的网络流量数据。

卷积层：通过多个卷积核提取局部特征。

激活层：使用 ReLU 激活函数引入非线性。

池化层：对卷积层的输出进行下采样。

全连接层：整合特征并输出分类结果。

最后进行模型训练，其训练步骤如下：

前向传播：计算输入数据的预测结果。

损失计算：使用交叉熵损失函数衡量预测结果与真实标签的差异。

反向传播：通过梯度下降算法更新模型参数。

迭代优化：重复上述步骤，直到模型收敛。

在数据量急速增长、算法持续优化的背景下，深度学习可扩展性使其在大数据环境下的网络异常行为检测中更具优势。如图 2 所示，ensorFlow 和 PyTorch 等分布式学习框架，可借助多台计算机或 GPU 集群并行计算，提高模型训练效率，可实现海量网络数据高效处理，检测出异常行为。

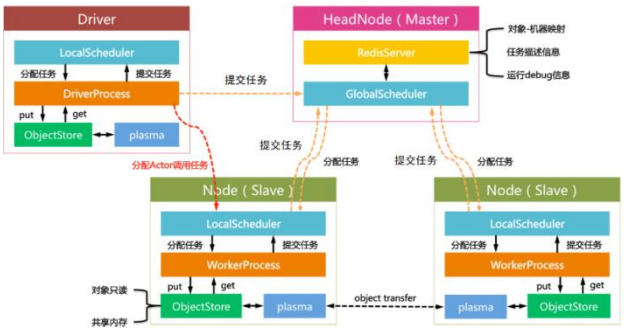


图 2　分布式框架示意图

RNN 是一种具有循环结构的神经网络，能够处理序列数据。其核心思想是通过隐藏状态（hiddenstate）捕捉序列中的时间依赖关系。基本结构包括输入层、隐藏层和输出层。RNN 可以通过时间展开的方式表示为一个深度神经网络。每个时间步的输入、隐藏状态和输出都可以表示为：

输入：(x_t)

隐藏状态：(h_t=f(W_hh_{t-1}+W_xx_t+b_h))

输出：(y_t=g(W_yh_t+b_y))

其中，(W_h)、(W_x)、(W_y)是权重矩阵，(b_h)、(b_y)是偏置向量，(f)和(g)是激活函数。

RNN 中常用的激活函数包括 ReLU、Tanh 和 Sigmoid。在网络异常行为检测中，Tanh 函数常用于隐藏层，因其能够处理正负值并保持梯度稳定。具体模型搭建流程与上述 CNN 几乎一致，在此不做多赘述。将训练好的神经网络模型用于实时检测网络流量，识别异常行为，并发出警报信息。由新的网络流量输入到模型时，模型根据学习到的特征、模式进行判断，如若输出结果表明该流量数据存在异常行为，则出发预警机制，通知安全员及时处理。

3 结束语

综上所述，基于人工智能的网络异常行为检测技术为网络安全领域带来了革命性突破，其通过对机器学习、深度学习等算法的应用，有效提升了异常行为检测的准确性与效率。从理论研究到实际案例应用，均展现出该技术在网络安全防护中的巨大潜力。展望未来，随着多模态数据融合、边缘计算、可解释性人工智能和联邦学习等技术的发展，基于人工智能的网络异常行为检测技术将不断完善。相信通过持续的技术创新与实践探索，该技术将在保障网络安全、维护数字社会稳定运行中发挥更为关键的作用，推动网络安全防护体系迈向更高水平。

参考文献

[1]程铭瑾.基于人工智能的视频监控异常行为检测方法[J].信息记录材料,2024,25(3):136-138.

[2]谭秦红,田应信.基于人工智能的通信网络入侵检测系统设计[J].长江信息通信,2022,35(12):189-191.

[3]王冬梅.人工智能技术在网络安全威胁检测与防御中的应用研究[J].信息与电脑,2024,36(13):123-125.

[4]魏敏.基于人工智能的计算机网络信息安全防护模式研究[J].信息记录材料,2024,25(11):130-132.

[5]李亚辉.基于人工智能算法的大数据网络防御系统研究[J].信息与电脑,2024,36(6):92-94.

[6]童炜华.基于人工智能技术的分布式入侵检测系统设计[J].信息记录材料,2024,25(7):150-152.

作者简介：裴旭光（1978.06-），男，汉族，河北省邢台市人，本科，工程师，研究方向：过程控制自动化与智能化。