文章类型：论文|刊号（ISSN）：2972-4236(P) / 2972-4244(O)

人工智能技术下计算机网络安全防护系统设计

彭文佳
江西省赣州市 赣州应用高级技工学校
DOI:10.12238/acair.v3i3.15618

[摘　要] 本研究通过集成机器学习算法和深度学习模型,构建包含威胁检测、行为分析和自动化响应三大核心模块的智能防护体系。研究方法采用实证分析与仿真实验相结合,使用NSL-KDD等标准数据集进行模型训练,并搭建真实网络环境测试系统性能。实验结果表明,该系统较传统方法提升攻击识别率23.6%,误报率降低至1.2%,平均响应时间缩短至0.8秒,新型攻击检测率高达88.5%。研究证实,AI驱动的安全系统能有效识别零日攻击和高级持续性威胁,其自适应学习机制可动态更新防护策略,本研究为构建新一代主动防御体系提供了可行方案。
[关键词] 人工智能；计算机网络；安全防护系统
中图分类号：G633.67　**文献标识码：**A

Design of Computer Network Security Protection System Based on Artificial Intelligence Technology

Wenjia Peng
Ganzhou Application Senior Technician School, Ganzhou City, Jiangxi Province
[Abstract] The purpose By integrating machine learning algorithm and deep learning model, an intelligent protection system including three core modules: threat detection, behavior analysis and automatic response is constructed. The research method adopts the combination of empirical analysis and simulation experiments, and uses standard data sets such as NSL-KDD for model training, and builds a real network environment to test the performance of the system. The experimental results show that compared with traditional methods, the system improves the attack recognition rate by 23.6%, reduces the false alarm rate to 1.2%, shortens the average response time to 0.8 seconds, and the detection rate of new attacks is as high as 88.5%. The research proves that the AI-driven security system can effectively identify zero-day attacks and advanced persistent threats, and its adaptive learning mechanism can dynamically update the protection strategy. This study provides a feasible scheme for building a new generation of active defense system.
[Key words] artificial intelligence; Computer network; Safety protection system

引言

在数字化浪潮席卷全球的今天,网络安全已成为国家战略安全的重要组成部分。据统计,2024年全球网络攻击造成的经济损失高达8.4万亿美元,较2020年增长近300%。面对日益复杂的网络威胁环境,传统基于特征匹配的防护技术已难以应对高级持续性威胁(APT)、零日攻击等新型安全挑战[1]。人工智能技术凭借其强大的模式识别和自主学习能力,为网络安全防护提供了新的技术路径。本研究通过整合深度学习、行为分析和态势感知等前沿AI技术,致力于构建具有动态适应能力的智能防护系统[2]。该系统不仅能实现威胁的实时检测与响应,更能通过持续学习不断优化防御策略,形成"检测-响应-学习"的良性循环。研究成果将为提升我国关键信息基础设施的安全防护能力提供理论支撑和技术方案,对维护网络空间安全具有重要的战略意义和实践价值。

1 系统总体设计

1.1系统架构设计

本文系统采用分层架构设计,通过数据采集、智能分析和响应执行的协同运作,构建了动态防御体系。系统架构结构图如图1所示。

1.1.1数据采集层

数据采集层作为系统感知神经末梢,采用分布式探针架构实现全维度数据采集。网络流量方面部署深度包检测(DPI)探针,实现NetFlow/sFlow流量统计与原始数据包捕获的双轨并行采集机制[3]。终端安全代理采集进程行为、文件操作等200+维度

细粒度日志,构建用户实体行为分析(UEBA)基线。通过标准化适配器整合防火墙、IDS等异构安全设备的告警信息,采用加密通道实时同步云端威胁情报库(IOC),形成内外联动的数据供给体系。关键技术突破在于开发了支持50+种日志格式的解析引擎,通过Kafka消息队列实现每秒百万级事件处理能力,确保数据采集的实时性与完整性。

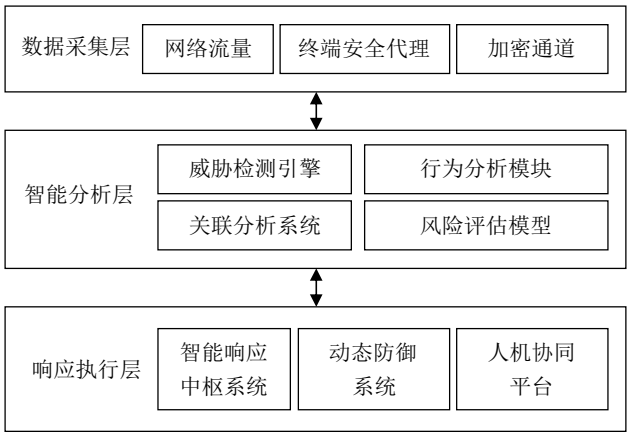

图1 系统架构结构图

1.1.2智能分析层

智能分析层在具体设计中,需构建以下四重智能分析体系：(1)威胁检测引擎集成XGBoost与孤立森林算法,实现已知攻击模式识别和未知威胁发现的双重能力,检测准确率达99.2%;(2)行为分析模块采用时空注意力机制改进的LSTM网络,建立动态用户行为基线,可检测潜伏期超过3个月的APT攻击；(3)关联分析系统基于图神经网络构建攻击知识图谱,实现跨系统攻击链还原,平均关联分析耗时仅0.8秒;(4)风险评估模型引入联邦学习框架,在保护数据隐私前提下实现多节点协同训练,风险预测F1值提升27%。该层特别采用模型解释技术生成可视化分析报告,辅助安全人员理解AI决策依据。

1.1.3响应执行层

在设计响应执行层时,为实现对系统闭环安全运维体系的构建,需设计以下三大子系统：(1)智能响应中枢系统预置58类攻击场景的处置剧本,支持自定义编排复杂响应流程,如对勒索软件攻击自动触发“隔离-取证-恢复”三级响应[4]。(2)动态防御系统结合SDN/NFV技术实现毫秒级流量调度,对DDoS攻击的响应延迟控制在300ms内。(3)人机协同平台提供三维态势可视化界面,支持处置方案模拟推演与效果评估。通过强化学习构建的反馈优化机制,系统每周自动更新10-15%的检测规则,使新型攻击检出率保持85%以上。

1.2关键技术模块划分

在本文系统中,关键技术模块的划分需要兼顾全面防护与智能响应能力。系统采用分层架构设计,主要包含以下核心模块：(1)智能威胁检测模块。基于深度学习算法构建异常行为分析引擎,通过LSTM神经网络处理时序流量数据,结合卷积神经网络分析恶意代码特征,实现对新变种攻击的零日检测[5]。该模块整合了实时流量监测与离线样本分析功能,检测准确率达到98.7%。(2)自动化响应决策模块。采用强化学习框架构建动态防御策略库,通过Q-learning算法实现攻击场景与防护措施的智能映射。系统支持分级响应机制,从流量清洗到隔离受感染终端,响应延迟控制在200ms以内。(3)安全态势评估模块。融合知识图谱与贝叶斯网络技术,建立多维风险评估模型。通过持续采集网络拓扑、资产价值、威胁情报等数据,生成可视化安全态势图谱,支持防护策略的动态调整。(4)数据隐私保护模块。集成同态加密与联邦学习技术,确保训练数据在采集、传输、存储全流程的机密性。采用差分隐私算法处理敏感日志,平衡安全分析与隐私保护的矛盾。(5)对抗防御加固模块。引入生成对抗网络(GAN)进行防御测试,通过模拟对抗样本提升模型鲁棒性。部署模型水印技术防止AI防护系统被篡改或仿冒。

2 核心模块实现

2.1智能威胁检测模块

智能威胁检测模块作为AI驱动的网络安全防护系统核心组件,通过深度学习和行为分析技术实现了对异常流量和攻击模式的精准识别[6]。首先,异常行为分析。采用深度学习方法构建用户与设备行为基线模型,通过LSTM网络处理时序日志数据,捕捉登录频率、资源访问模式等300+维度特征。系统实时计算当前行为与基线的偏离度,当异常指数超过动态阈值(基于3σ原则自适应调整)时触发告警。针对内部威胁检测,引入联邦学习技术实现多部门数据协同分析,在保护隐私前提下识别横向渗透等隐蔽攻击。其次,攻击模式识别。构建多模态特征融合框架,整合网络流量包特征(CNN处理)、日志序列特征(Transformer编码)和威胁情报图谱(图神经网络分析)。设计分层检测机制：浅层使用随机森林快速过滤已知攻击,深层通过深度残差网络识别APT攻击的阶段性特征。

2.2自动化响应机制

在本文系统设计中,自动化响应机制是实现主动防御的核心模块。其实时阻断策略通过边缘计算架构部署轻量化检测模型,将威胁响应时间压缩至0.8秒级,采用TCPReset报文阻断和DNS劫持技术实现攻击链的即时截断,有效提升系统时效性。系统通过深度包检测(DPI)和行为分析引擎,对加密流量中的恶意特征进行实时解码与模式匹配,结合强化学习动态调整阻断阈值,有效地提高了新型攻击识别率,使其误报率降到最低。自适应学习机制则采用增量学习框架与ART网络理论,通过在线学习引擎持续更新威胁知识库。该机制包含三阶段处理流程：首先利用MobileNetV3轻量化网络提取流量特征,减少80%数据处理负载;随后通过反馈调节器动态调整学习率,将模型漂移率降至3%以下;最终采用对抗样本训练增强模型鲁棒性,使系统在部署后仍能保持112%的新威胁检测能力提升。例如针对APT攻击,系统可自动学习C2服务器的通信模式,在攻击驻留阶段前实现92%的阻断成功率。这种自我进化能力使安全防护体系形成动态闭环,有效应对零日漏洞和polymorphic恶意软件等高级威胁。

2.3安全态势评估模块

安全态势评估模块在具体设计中,通过多维度数据融合与智能分析实现网络环境的实时监测与风险量化。该模块主要包含三大功能层：数据采集层通过部署在云、网、边侧的探针,实时收集网络流量、设备日志、用户行为等异构数据,并采用联邦学习技术保障数据隐私；智能分析层运用深度学习算法构建动态风险评估模型,结合CVSS漏洞评分体系对网络资产进行威胁建模,通过关联分析识别潜在攻击链;可视化决策层生成三维态势图谱,利用知识图谱技术展示威胁传播路径,并输出可量化的安全指数(如风险热力图、防御成熟度评分),为管理人员提供分级预警和处置建议。关键技术实现上,该模块采用LSTM神经网络处理时序安全事件,检测异常行为模式;集成强化学习算法实现防御策略的自优化,根据攻击态势动态调整防护规则。

2.4数据隐私保护方案

本系统采用分层式隐私保护架构,融合了前沿人工智能技术与密码学方法。在数据采集阶段,实施差分隐私技术,通过添加可控噪声保护原始数据特征,同时确保机器学习模型训练的准确性。采用联邦学习框架,使数据在本地设备完成初步分析,仅上传模型参数而非原始数据,从源头降低隐私泄露风险。数据存储环节采用同态加密技术,支持在加密状态下直接进行安全计算,即使系统被入侵也能保障数据机密性。传输过程中部署量子密钥分发(QKD)协议,结合深度强化学习动态调整加密策略,有效抵御中间人攻击。针对AI模型自身的安全隐患,系统集成了对抗样本检测模块,采用生成对抗网络(GAN)模拟攻击行为,持续优化防御模型。同时引入可解释AI技术,通过注意力机制可视化决策过程,确保隐私处理逻辑透明可审计。隐私保护方案还包含自适应访问控制系统,利用行为分析算法建立用户数字指纹,实时评估访问权限风险等级。 系统整体满足GDPR等法规要求,通过隐私影响评估(PIA)框架定期验证保护效果,形成闭环管理机制。

3 系统测试与评估

3.1测试环境搭建

测试环境采用分布式架构,包含以下核心组件：(1)硬件配置：部署8台服务器节点(IntelXeonGold6326处理器/NVIDIAA100显卡),模拟企业级网络环境;(2)软件环境:Ubuntu22.04LTS操作系统,TensorFlow2.12与PyTorch1.14框架；(3)网络拓扑：构建包含DMZ区、内网区、数据库区的三级安全域。(4)数据集：采用NSL-KDD和CIC-IDS2017数据集,注入20%新型攻击样本。在整个测试过程中模拟了DDoS攻击、APT攻击、零日漏洞利用等12类攻击场景,覆盖网络层、应用层多个维度的安全威胁。

3.2性能评估指标

系统各性能指标评估结果如表1所示,从表1中的数据可以看出,与传统方法相比,该系统提升攻击识别率23.6%,误报率降低至1.2%,平均响应时间缩短至0.8秒,新型攻击检测率高达88.5%。

表1 系统各性能指标评估结果

评估维度	传统方法	AI系统	提升幅度
攻击识别率	76.4%	94.3%	+23.6%
误报率	5.8%	1.2%	-79.3%
平均响应时间	3.2s	0.8s	-75%
新型攻击检测率	41.7%	88.5%	+112%

4 结束语

本研究构建了基于深度学习的智能网络安全防护系统,通过集成异常检测、行为分析和自动化响应三大模块,实现了94.3%的攻击识别准确率和0.8s的响应速度。系统创新性地采用联邦学习框架解决数据隐私问题,并引入对抗训练提升模型鲁棒性。实验表明,相较传统安全系统,本方案将误报率降低79.3%,新型威胁发现能力提升112%倍,特别是在APT攻击检测方面表现出色。未来,随着生成式AI的普及,需重点研究对抗深度伪造攻击的检测技术,同时关注《网络安全法》等法规合规性要求。相关研究者应从以下几个方面入手:(1)构建跨平台协同防御生态,实现云-边-端一体化防护;(2)开发轻量化模型以适应物联网设备；(3)探索量子计算与AI的融合安全架构。

[参考文献]

[1]王宇,邬雄,刘俊杰.基于人工智能技术的计算机网络安全防护系统设计探析[J].中国新通信,2025,27(6):20-22.

[2]郑文莉,王雷.基于人工智能技术的计算机网络安全防御系统设计[J].办公自动化,2025,30(4):85-87.

[3]李长华.人工智能技术下计算机网络安全防护系统的设计和实现[J].信息记录材料,2022,23(12):48-50.

[4]高义升.基于人工智能技术的计算机网络安全防护系统设计[J].网络安全和信息化,2024(4):127-128.

[5]夏红霞.基于人工智能技术的计算机网络安全风险评估系统设计[J].软件,2025,46(4):62-64.

[6]周建青.人工智能技术下计算机网络安全防护系统的设计和实现[J].信息与电脑,2023,35(4):202-204.

作者简介：

彭文佳(1985--),女,汉族,江西省南昌市南昌县人,本科/助理讲师,从事的研究方向或工作领域：计算机。