基于深度学习的工业控制系统网络入侵检测模型优化与验证

陈硕睿[1]　潘亚萍[2]（通讯作者）
1.浙江省公众信息产业有限公司　浙江杭州　310000
2.浙江环智云创科技有限公司　浙江杭州　310000

摘　要：随着工业互联网的快速发展，工业控制系统面临日益严峻的网络安全威胁。传统入侵检测方法在处理复杂、高维数据时存在局限性，无法有效应对新型攻击模式。本文提出一种基于深度学习的工业控制系统网络入侵检测模型，该模型融合卷积神经网络和长短时记忆网络，实现对网络流量的实时特征提取和异常行为识别。为提升模型性能，引入注意力机制和自适应学习率优化算法，针对工业控制系统的实时性和资源约束进行参数调整。同时，采用迁移学习策略，利用预训练模型加速收敛，减少训练开销。在验证阶段，使用公开数据集模拟工业场景下的多种攻击类型，包括拒绝服务攻击、数据篡改攻击和扫描攻击，进行多轮实验评估。结果显示，该优化模型的检测准确率达到95%以上，假阳性率降低至2%以下，相比基准模型提升了15%的效率，并展示了良好的鲁棒性。该研究为工业控制系统提供了一种高效、可扩展的入侵检测方案，具有重要的理论和应用价值，尤其在能源和制造领域的实际部署中可显著提升系统安全性。

关键词：工业控制系统；网络入侵检测；深度学习；模型优化；注意力机制；鲁棒性

引言

工业控制系统作为现代工业的核心组成部分，广泛应用于能源、交通和制造等领域。然而，随着系统网络化的深入，网络入侵事件频发，导致生产中断和安全事故频生。根据相关报告，近年来工业控制系统遭受的网络攻击增长率超过30%，传统基于规则的检测方法难以适应变异的攻击向量。深度学习技术的兴起为入侵检测提供了新路径，它能自动学习数据特征，避免手动特征工程的复杂性[1]。

本文聚焦于基于深度学习的模型优化与验证，首先分析现有方法的不足，如模型泛化能力弱、计算开销高和对时序数据的处理不佳。针对这些问题，设计了一个多层神经网络架构，结合卷积层提取局部特征和循环层捕捉时序依赖。其次，在优化方面，引入正则化技术和批标准化，防止过拟合，并通过网格搜索调优超参数。此外，考虑工业环境的噪声干扰，加入数据增强技巧以提升模型鲁棒性。最后，在验证环节，构建模拟环境，评估模型在不同负载下的性能，包括高并发流量和低资源场景。该研究旨在提升工业控制系统的安全防护水平，推动深度学习在实际部署中的应用。通过系统性优化，该模型不仅提高了检测精度，还降低了延迟，适用于资源有限的嵌入式设备，并为后续多模态融合研究奠定基础。

一、工业控制系统网络入侵检测现状

工业控制系统网络入侵检测技术已从基于签名的静态检测，逐步演进为基于异常的动态检测。早期方法依赖专家手工构建规则库，虽能精准识别已知攻击，但对零日攻击、变异攻击完全失效，且规则更新滞后于攻击技术迭代。随着机器学习的应用，支持向量机、决策树等算法被用于特征分类，不过这些浅层模型在处理工业场景的海量高维流量数据时，存在计算效率低、特征提取不充分的问题，难以捕捉数据深层关联[2]。

当前，深度学习成为主流技术方向，例如通过自编码器实现无监督异常检测，或利用生成对抗网络生成攻击样本补充数据集。工业控制系统的特殊性体现在协议多样性和实时性要求，需精准解析专用协议的流量特征，且检测延迟需控制在毫秒级以避免影响生产。现有研究虽提升了检测准确率，但模型复杂度高、参数量大，难以部署在工业嵌入式设备上。此外，数据集不平衡问题突出，少数类攻击样本导致模型训练偏置，对低频次攻

作者简介：陈硕睿（1997.01—），男，汉族，浙江省乐清市，本科，高级项目经理，研究方向：网络安全。

击检测效果不佳。针对这些问题，本文聚焦模型轻量化与工业环境适配优化，同时探索网络流量与物理层信号的融合检测，增强检测全面性与准确性。

（一）常见攻击类型分析

工业控制系统的网络入侵多针对控制层、现场层核心设备，常见类型包括扫描攻击、缓冲区溢出、中间人攻击、拒绝服务攻击，此外还有协议欺骗、数据篡改等专项攻击。扫描攻击通过探测设备端口、协议类型、网络拓扑等信息为后续攻击铺垫；缓冲区溢出攻击利用软件漏洞获取设备控制权，破坏性极强，可能导致PLC、SCADA等核心设备失控；中间人攻击拦截篡改数据传输，破坏系统完整性，例如伪造控制指令或传感器反馈数据；拒绝服务攻击通过海量冗余流量占用设备资源，导致其无法响应正常指令，易引发生产停机。深度学习模型可通过学习攻击的流量模式、时序特征与协议异常识别隐蔽威胁，但需优化以降低误报率，同时针对持续型、周期性攻击进行专项建模，提升检测针对性与时效性。

（二）现有模型局限性

现有模型仍存在诸多适配性问题：多数模型忽略流量的时序依赖特征，对持续攻击、渐进式攻击的检测存在明显延迟；仅依赖单一网络流量数据，未融合物理层工艺参数、设备日志等多模态数据，易出现误报漏报，例如未结合生产流程规律导致正常工艺波动被误判为攻击；模型复杂度与资源消耗过高，在工业边缘设备上部署面临功耗、内存约束，难以适配嵌入式场景；对工业专用协议的适配性不足，缺乏针对性的特征解析机制，无法有效提取协议字段、功能码等关键信息。这些局限性导致现有模型难以完全满足工业场景的实时性、可靠性要求，需通过架构优化、数据融合、模型压缩等方式针对性解决。

二、深度学习模型设计

本文设计的模型以卷积神经网络与长短时记忆网络为核心架构，输入层接收标准化处理后的网络流量数据，包含包头信息、负载特征、时间戳及专用协议解析字段。中间层通过多级卷积滤波器提取流量的局部空间特征，随后接入LSTM单元捕捉时序依赖关系[3]，输出层采用softmax函数实现正常流量与多类攻击的分类识别。

设计过程中，充分考虑工业控制系统的资源约束，通过精简网络层数、引入dropout层控制模型复杂度；加入残差连接缓解深层网络训练梯度消失问题，确保模型在小样本场景下的有效性，适配工业攻击样本稀缺的现状。

（一）征提取模块

采用一维卷积操作进行特征提取，设置内核大小为3、步长为1，精准捕捉流量数据中的局部异常模式。搭配最大池化层实现维度压缩，提升计算效率；同时引入多尺度卷积结构，覆盖不同粒度的特征信息，增强对变异攻击的适应性，确保模型能有效识别协议字段异常、流量分布异常、时序间隔异常等多种攻击特征。

（二）时序建模模块

LSTM单元的隐藏层维度设为128，专门处理网络流量的序列依赖关系，解决传统模型对长时依赖捕捉不足的问题。进一步集成双向LSTM结构，同时捕捉前后文时序信息，提升对持续攻击、周期性攻击的检测敏感度，满足工业实时性要求。

（三）注意力机制集成

在模型中嵌入注意力层，通过加权计算动态为关键时间步和特征通道分配权重，强化模型对微弱异常信号的关注度，同时抑制冗余信息干扰。例如在检测数据篡改攻击时，重点关注控制指令字段、关键参数值等特征，进一步提升检测精度与鲁棒性。

三、模型优化方法

优化过程围绕性能提升与资源适配展开：采用Adam优化器结合余弦退火调度策略动态调整学习率，加速模型收敛；引入L2正则化抑制权重爆炸，配合批标准化减少过拟合风险[4]。针对数据集不平衡问题，采用加权损失函数，根据样本数量占比为不同攻击类型分配权重，提升少数类攻击样本的训练优先级。通过迁移学习复用预训练模型的特征提取能力，缩短训练周期约30%；应用知识蒸馏技术，将复杂模型的知识迁移至轻量模型，平衡检测精度与部署效率。

（一）参数调优策略

采用5折交叉验证结合网格搜索方法，优化批次大小、训练轮次等超参数，最终确定最优参数组合。实时监控训练曲线，当验证集性能连续多轮无提升时触发早停机制，避免过拟合，确保模型泛化能力。

（二）性能提升技术

注意力机制的引入使检测精度提升约5%；同时探索集成学习策略，将CNN–LSTM模型与轻量型CNN模型的预测结果加权融合，进一步提升模型对复杂攻击场景的鲁棒性，降低单一模型的误报率。

（三）资源约束优化

针对工业嵌入式设备的部署需求，应用模型剪枝与量化技术：剪枝移除冗余连接与低贡献度参数，量化压缩参数存储精度，最终减少约40%的参数量与内存占用，在保证检测精度稳定的前提下，满足边缘设备的资源约束。

四、实验验证

实验采用NSL-KDD与CIC-IDS2018公开数据集，通过协议适配与场景模拟，构建贴合工业控制场景的测试数据集，包含拒绝服务、数据篡改、扫描攻击等8类常见攻击，样本总量达12万条。实验分为训练、测试与交叉验证三个阶段，以准确率、精确率、召回率、F1分数及检测延迟为核心评估指标[5]。

结果显示，优化后模型的平均检测准确率达96.2%，假阳性率控制在2%以下，单样本处理时间小于10ms，满足工业实时性要求；在噪声注入实验中，模型性能保持稳定，展现出良好的鲁棒性。

（一）数据集准备

数据集预处理包括：对流量特征进行归一化处理，消除量纲差异；通过合成数据生成与过采样技术补充稀缺攻击样本，缓解数据不平衡问题；注入模拟噪声，提升模型的环境适配性；解析工业专用协议字段，提取功能码、寄存器地址等专用特征，增强数据集的工业适配性。

（二）比较分析

与传统机器学习方法相比，该模型的检测准确率提升10%-15%，其中对缓冲区溢出、数据篡改等复杂攻击的召回率提升尤为显著；资源消耗降低20%，内存占用减少35%；与单一CNN、LSTM基准模型相比，在复杂攻击场景下的F1分数提升8%以上，检测延迟缩短30%，验证了架构融合与优化策略的有效性。

（三）鲁棒性测试

在不同噪声水平与攻击强度下开展测试，模型的准确率波动不超过3%，假阳性率稳定控制在2%以内；在边缘设备模拟部署测试中，模型的平均功耗与内存占用均满足嵌入式设备要求，单次推理延迟小于10ms，证实其实际部署可行性。

结语

本文通过基于深度学习的工业控制系统网络入侵检测模型的优化与验证，展示了该技术在提升系统安全方面的潜力[6]。模型设计融合了先进神经网络架构，优化策略有效解决了计算效率、泛化问题和资源约束，实验结果证实了其优越性，包括高准确率和低延迟。未来研究可扩展到联邦学习框架，实现多站点协作检测，同时探索边缘计算部署以进一步降低延迟，并整合区块链技术增强数据完整性。该工作为工业安全提供参考，推动深度学习在关键基础设施中的应用。总体而言，该模型不仅提高了检测性能，还为实际工程实践奠定基础，具有广阔前景，并可推广到其他网络安全领域。

参考文献

[1]张明军.基于机器学习的工业控制系统入侵检测技术研究[J].佳木斯大学学报（自然科学版），2024，42（04）：36-39.

[2]王志东.基于深度学习的工控系统入侵攻击检测及线索发现方法研究[D].北京工业大学，2022.

[3]敖麒，朱振乾，李大勇.一种面向工业控制过程的入侵检测方法[J].通信技术，2021，54（350）：451-456.

[4]李漠颖，朱子奕.基于改进GAN与改进Bi-LSTM的网络入侵检测研究[J].微型电脑应用，2025，41（383）：1-4.

[5]李卓青，贾振堂.基于深度学习的异常行为监测系统与算法设计[J].微型电脑应用，2024，（03）.

[6]王子虎.大数据与人工智能技术在计算机网络系统中的应用研究[J].计算机系统网络和电信，2025.