ISSN 1009-3044

*Computer Knowledge and Technology* 电脑知识与技术 Vol.20, No.23, August 2024

E-mail：jslt@dnzs.net.cn http：//www.dnzs.net.cn

Tel：+86-551-65690963 65690964

DOI:10.14004/j.cnki.ckt.2024.1155

基于深度学习的网络流量入侵检测方法研究

李 浩

（江西工业职业技术学院（图文信息中心），江西 南昌330095)

摘要：在电子信息技术快速发展的过程中，网络入侵模式、攻击方式也在不断升级和变化，网络安全问题日益严峻，网络 流量入侵检测面临更大挑战 。为了提高入侵检测的速率、准确率，文章提出一种基于深度学习的网络流量入侵检测新 方法，该方法通过提取流量数据的关键特征，优化粒子群算法 。 同时引入惯性权重模块，并结合极限学习机、AE 自编码 器，有效消除数据噪音的问题，实现对网络流量入侵的有效检测 。实验结果表明，该方法能够精确、高效地对网络流量

入侵进行检测，可以提高网络安全防护效果。

关键词：深度学习；网络流量；入侵检测；检测方法 中图分类号：TP311 文献标识码：A

文章编号：1009-3044(2024)23-0096-04



开放科学（资源服务）标识码(OSID) ：

伴随着网络信息技术的迅猛发展，网络流量入侵 载体更加多样化，网络安全受到严重威胁，亟须加强 网络安全风险管理，提高入侵攻击侧识别速率，保障 资产防护效果 。对此，众多学者围绕网络流量入侵检 测方法开展了大量的研究，取得了一定的研究成果。 宗学军和刘欢欢等[1]针对当下入侵检测系统存在的不 足，将数据平面开发套件技术同 Suricate IDS 进行整 合，利用高效规则匹配算法 NEW-WM，构建高速网络 流量下入侵检测系统，有效解决了以往入侵检测系统 无法准确、实时检测高速工业网络流量的问题，大大 提升了系统数据包捕获处理能力和入侵检测效率；陈 雪倩和步兵[2]提出一种基于网络流量和数据包的入侵 检测系统，基于 CBTC 系统构建 IDS 模型，整合 AR 算 法的网络流量检测模块和数据包检测模块，实现了对 网络流量 、数据包的特征提取和检测 ，有效保障了 CBTC 系统的安全风险，避免了数据篡改、杜绝服务等 问题的发生；边金良[3]提出基于数据挖掘的网络流量 异常入侵检测法，借助数据挖掘的关联分析功能，明 确异常流量特征联系，对异常流量特征开展联合计算 熵值处理，具有良好的异常网络流量入侵检测功能。 上述方法虽然对于网络流量拥有良好的数据采集和 处理能力，提升了网络流量入侵检测效率，缩短了检 测时间，使入侵检测系统消耗大大降低，但在检测精 确度、实时性检测和告警方面仍存在一定的不足，对 入侵行为类型的检测有一定的局限性，适用范围有 限，无法实现大规模广泛应用，对未知或新型攻击的 检测方面仍有待提升，还需进一步提升实际应用性 能 。为了解决上述问题，本研究提出基于深度学习的 网络流量入侵检测方法，联合深度学习、极限学习机 的优势，并通过实验验证该方法的入侵检测性能。

**1** 基于深度学习的网络流量入侵检测方法设计

1.1 关键词维度统计特征

协议、流量包是网络流量入侵检测工作的关键识 别信息 。在对协议这一流量行为信息进行识别时，涉 及 TCP、IP、SLIP、PPP 等几大网络协议，应通过 One- Hot 中单个类型单对应的形式，实现协议同相应数值 的映射处理；而在流量数据包的识别方面，对于特定 平台，提前要在该平台服务器中部署脚本监听端口， 抓取用户访问平台下的用户访问请求、响应等流量信 息[4] 。本研究以字符串的形式存储访问者请求数据。 网络攻击人员采用 GetShell、XSS 注入等方式开展渗 透操作时，会将固定框架语句写入目标码、URL 内 。 本文提出的基于深度学习的网络流量入侵检测方法 的原理为针对不同特征维度选取对应的关键词，并对 关键词出现频率进行记录，从而抓取框架中数据特 征 。通过特征转化，能够将字符串转变为矩阵形式， 便于机器理解和操作，为构建网络流量入侵检测模型 提供了保障 。深度学习下网络流量入侵检测方法的 关键词特征详见如表 1 所示。

1.2 数据初处理

在提取字符串特征结构时，可以应用关键词维度 统计(KDS) 处理法 。KDS 法对 One-Hot 离散文本语义 特征提取法、词袋文本特征提取法进行了融合，将两 种提取法的核心技术充分发挥出来，可以对专项特征 进行针对性的提取，创新了入侵检测方向的特征提取 思路和方式 。KDS 法可以在无序长字符数据中对关 键特征进行提取，形成满足学习模型的数据信息，具 有降低冗余数据、改善算法模型运行效率的优势 。本

收稿日期：2024-04-29

作者简介：李浩(1977—) ，男，江西靖安人，副教授，硕士，主要研究方向为计算机网络与信息安全。

96 网络通信与安全  本栏目责任编辑：代 影

中国知网 https:Www.cnki . net

第20 卷第23 期 (2024 年8 月）

表 1 关键词特征列表



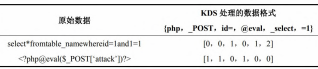
入侵检测模型可以利用 KDS 处理法，实现无规则字符 串数据向结构化数据的转变，有效统计各维度的词 量，充分体现不同流量包的结构差别，以便带入算法 模型中进行运算，促进网络流量入侵检测模型的构建 和学习 。在 KDS 处理法的保障下，可以有效整合模型 各项功能，例如：统计分析、匹配及关键词对比等，进 一步优化数据处理方式，以监督学习的形式提取数据 特征，获得较优的数据特征匹配结果，有效消除冗余 数据，保障算法模型的运行速率[5] 。具体流程如下：

1) 在特征层中将各关键词看作独立特征，通过向 量 格 式 编 排 特 征 层 数 据 ，表 达 式 为 ：*features* = [ *select union null* ...... *root* ]。

2) 将上述表达式进行映射处理，设定特征层第 n 个 关 键 词 为 *Xn*，得 到 映 射 层 表 达 式 ：*X* = [ *X*1 *X*2 *X*3 ...... *Xn* ]。若第 n 个特征出现的标记 为 *Cn*，出现频率计数为 *Cn*，借助公式对各条网络流量 中关键词的出现频率进行统计，得到 *Cn* = 最 终特征向量表达式为：*Hn* = [ *C* 1 *C*2 *C*3 ...... *Cn* ]。

3) 完成数据实例转化对比，建立包括未经处理数 据、KDS 法处理后数据的转化表，调整数据格式如表 2 所示。

表2 KDS 转化事例



1.3 数据压缩与数据降维处理

本研究方法结合自编码器的稳定特征学习性能， 对网络流量数据主成分进行提取，完成数据的压缩、 降维处理，有效避免了噪声数据和冗余数据的影响， 增强了算法的监测效率 。 自编码器神经网络模型建 立在无监督算法的基础上，由输入、隐藏和输出三层 结构组成，其中输入和输出层的节点数相同 。若输入 向量为 *x* = {*x*1, *x*2 *x*3 ... *xn* }，隐藏层编码处理公式为： *h* = *σ* (*Wx* + *b*)。

*Computer Knowledge and Technology* 电脑知识与技术

数据压缩处理结果为 *h* = {*h*1, *h*2, *h*3 ... *hn* }，W 为输 入层 、隐藏层间的权重值 ，*σ* (·) 为激活函数 ，b 为偏 置值。

通过输出层进行数据解码，得到公式： = *σ* (*W' h* + *b'*) 。其中，偏重、偏置量分别是 W'、b'，输出结果 是  。针对隐藏层输出结果进行还原处理，还原为原 始输入数据，并开展降维操作，由此增强入侵检测模 型性能 。采用逐层贪婪训练法，自编码神经网络可以 对数据开展预训练，应用 BP 神经网络调整网络模型。 而鉴于本研究方法将极限学习机、单隐藏层自编码器 进行了整合，所以需要预训练自编码器，确定输入层 到隐藏层间的初始权重，通过随机梯度下降算法和 BP 神经网络，得出隐藏层的最优输出权值和损失函 数的最小值。

**2** 实验过程

2.1 基于深度学习的网络流量入侵检测系统模型

在构建网络流量入侵检测系统模型时，将 AE 自 编码器、极限学习机整合起来，针对 AE-ELM 神经网 络中超学习参数对算法模型产生的影响及存在的缺 陷，引入改进后的粒子群优化算法(POS) ，借助优化 PSO 搜索最优化模型学习参数，建立 PSO-AE-ELM 模 型，实现多维数据处理以及对样本数据深层特征的有 效提取，维去噪，缩短训练时间[6] 。针对深度神经网络 的 PSO-AE-ELM 预测模型，可以进行入侵检测识别 的预测，借助粒子群优化算法探寻最优自编码 ELM 的 隐藏层神经节点数量，明确偏差值、权重以及激活函 数，结合最优结果、训练样本，并在模型中导入预测结 果，开展实验验证 。实验验证流程具体如下：

1) 针对自编码器搜集的网络流量数据，开展数据 降维处理操作 ，优化网络流量入侵检测模型运行 效率。

2) 将整合极限学习机和自编码器(AE-ELM) 的神 经网络作为预测模型，在重构模型内输入数据集合， 通过随机梯度下降法(SGD) 对自编码器进行训练，提 取 AE 神经网络特征 。依托 PSO 算法找到最优 AE- ELM 超学习参数，结合优化结果明确隐藏层神经节点 的数目，获取各项参数数值 。其中，AE-ELM 神经网 络下的超学习参数直接影响着算法模型的性能，具体 指的是进行模型训练时在深度学习中设定的参数，包 括：隐藏层数目、梯度下降法循环数、学习率及隐藏层 单元数等。

3) 将实际网络入侵数据导入通过最优参数构建 的 POS-AE-ELM 模型中，结合准确率判定优化后的 模型检测精准度。

2.2 入侵检测模型优化实现过程

POS-AE-ELM 模 型 的 优 化 实 现 过 程 详 见 图 1 所示。

本栏目责任编辑：代 影 i a a a a n u u u  网络通信与安全  97

中国知网 https:Www.cnki . net

*Computer Knowledge and Technology* 电脑知识与技术

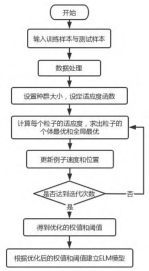


图 1 POS-AE-ELM 模型优化过程

结合图 1，POS-AE-ELM 模型优化过程具体介绍 如下：

1) 针对互联网端的行为流量，利用蜜罐系统进行 搜集和存储，划分和标注行为流量的供给类型，形成 原始数据集，并进行进一步处理获得显性结构数据， 有序进行数据的过滤、类型转变及归一化操作，使实 验样本达到学习模型的标准要求。

2) 按照比例划分实验样本数据，分别形成测试集 和训练集 。结合训练集建立 AE-ELM 网络流量入侵 检测模型，而利用测试集则能够检验模型精确程度， 分析模型检测准确率函数的有效性 。准确率函数指 的是开展机器学习算法检测测试数据时 ，应用 Classification-report 函 数 进 行 测 试 的 准 确 率 运 算 输出。

3) 利用基准粒子群优化算法对 AE-ELM 模型的 学习参数进行优化，并参考准确率函数形成适应度函 数，明确适应度值 。在得到 POS-AE-ELM 模型相关 参数的基础上，确定最优学习参数和优化模型，获得 最优超学习参数组合，即最优超学习参数。

2.3 模型框架结构

基于粒子群优化算法下，种群粒子趋近于最佳粒 子位置时，粒子速度接近 0，会影响种群多样性，致使 种群粒子无法摆脱局部最优，不利于精细局部搜索工 作的开展 。本文算法以标准 PSO 为基础,借鉴遗传算 法的变异操作,引入动态惯性权重,形成改进的粒子群 优化算法 。该算法可有效识别检测目标类别,避免种 群粒子过早收敛[7] 。结合多超学习参数分析结果，也 能够通过下述方式进一步提升模型入侵检测性能 。 一方面，可以将全局搜索视作外部优化层，对改进后 粒子优化算法模型的最优特征粒子进行选择，借助 AE-ELM 模型的适应度函数对特征粒子进行评估；另 一方面也能将局部搜索视作内部优化层，参考遗传算 法的变异操作，在粒子群优化算法中融入动态权重值

第20 卷第23 期 (2024 年8 月）

的概念，在增加迭代次数的过程中实现对 PSO 局部化 优化，以此来合理选择局部特征，有效克服标准 PSO 算法容易陷入局部最优的问题,实现全局优化 。其中， 局部搜索、全局搜索是考察算法性能的两大指标，前 者指的是无穷接近最优解的能力，依赖于对解空间进 行按邻域搜索；后者指的是探寻全局最优解位置的能 力 ，可以从最优解中选择最好结果作为最终结果 。 POS-AE-ELM 模型框架结构详见如图 2 所示：

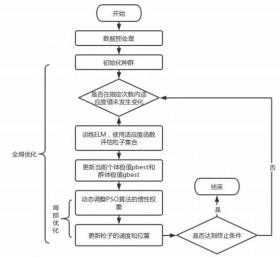


图2 POS-AE-ELM 模型框架示意图

检测网络入侵流量前，首先要初始化改进 PSO 算 法的相关参数，在给定范围内随机生成 20 个初始粒子 形成种群 。然后迭代优化多次，用适应度函数评价各 粒子，更新粒子位置和速度，直到满足迭代停止条件。 最后将全局最优粒子位置对应的参数组合代入 POS- AE-ELM 模型，完成模型训练。

**3** 实验结果分析

3.1 实验环境及数据

搭 建 POS-AE-ELM 模 型 测 试 实 验 环 境 如 表 3 所示：

表3 实验环境

|  |  |  |
| --- | --- | --- |
| 名称 | 型号 | 版本 |
| CPU | 8 核 | AppleM1 |
| 内存 | 16GB | LPDDR4 |
| 操作系统 | MacOs | 7459.101.3 |
| 编程语言 | Python | 3.9.12 |

实验中，PSO-AE-ELM 模型的参数设置如下：种 群粒子数 20，最大迭代次数 200，适应度函数为分类准 确率。

结合实际项目所采集的网络流量数据，将蜜罐系 统部署在服务器或私有云等多节点中，发挥监控系统 的捕捉模块功能来搜集网络流量数据，这些数据均具 有单个标志属性和三个特征属性，属于正常行为流 量[8] 。针对流量数据开展拆分解析，设定相应标签，确 定网络流量入侵检测原始样本数据 。鉴于网络流量

98 网络通信与安全  本栏目责任编辑：代 影

中国知网 https:Www.cnki . net

第20 卷第23 期 (2024 年8 月）

数据无规律且波动变化较大，可以应用 min-max 归一 化法对特征参数进行压缩，使参数处在[0, 1]的范围。 网络流量数据的初始值为 x，xmax 为数据组内最大值， xmin 为数据组内最小值，归一化处理公式表示为：*x'* = *x* - *x*min

。 *x*max - *x*min

3.2 实验结果

迭代次数与最优粒子的适应度成正比关系 ， PSO-AE-ELM 模型准确率与迭代次数关系详见如图 3 所示：

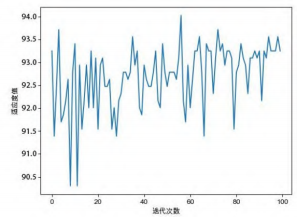


图3 PSO-AE-ELM 模型准确率与迭代次数关系图

结合上图可以看出，随着迭代次数的不断增加， PSO-AE-ELM 模型的准确率逐渐趋于稳定，最优结果 约为 94%，可以对数据进行优化，并大大提升了入侵 见检测识别准确率，表示该算法模型在网络流量入侵 检测领域具有实用性 。为了进一步对比分析该算法 模型同其他模型的性能，在开展识别检测网络流量入 侵行为验证时，应选定准确度、误报率、漏报率及检测 时间四大评价指标，对照分析基于 BP 神经网络检测 法、基准 ELM 模型检测法、本研究检测法的网络流量 入侵检测效果，具体结果详见如表 4 所示。

表4 不同方法检测效果对比

|  |  |  |  |
| --- | --- | --- | --- |
| 评价指标 | 检测方法 | | |
| PSO-AE- ELM | ELM | BP |
| 准确率 | 87.01% | 79.68% | 76. 13% |
| 误报率 | 10.39% | 13.79% | 23.87% |
| 漏报率 | 2.6% | 6.51% | 0 |
| 检测时间 | 6.4s | 7.8s | 19.26s |

结合上表可知，本研究所提出的基于深度学习算 法的 PSO-AE-ELM 模型具有良好的性能，通过添加 遗传算法的突变策略，使 PSO 全局优化性能明显改 善，在准确率、漏报率和检测时间方面明显较优，且漏 报率也低 于 ELM 检测 法 。此外 ，后期仅 针对 PSO- ELM 和优化 PSO-AE-ELM 算法模型开展对照分析， 利用数据集进行 100 次迭代训练，进一步检验优化 PSO-AE-ELM 算法模型的有效性，具体结果详见如图

*Computer Knowledge and Technology* 电脑知识与技术

4 所示。

96  94 

92  90  88  86  84  82 

0 20 40 60 80 100

迭代次数

(a) 模型准确率对比图

|  |  |
| --- | --- |
| 001 08 06 04 02 0  数次代迭 | 61  准确率%  41  21  01  8  8  4  2 |

(b) 模型误报率对比图

1.6 

|  |
| --- |
|  |

1.2 

1.0 

0.8 

0.6 

0.4 

0.2 

0.0 



I

0

20 40 60 80 100

迭代次数

(c) 模型漏报率对比图

图4 PSO-ELM 和优化PSO-AE-ELM算法模型性能对比

（虚线表示 PSO-ELM 模型，实线表示优化 PSO-AE-ELM算法

模型）

通 过 对 比 PSO-ELM 和 优 化 PSO-AE-ELM 的 收 敛曲线可以发现，优化后的算法模型在准确率、误报 率、漏报率等指标上均有提升 。本研究所提出的优化 PSO-AE-ELM 算法均得到了提升，其中准确率提升了 8.33%，误报率、漏报率分别降低了 7.6%、0.73%，且运 行时间也缩短了 0.7s，再次验证了优化算法具有良好 的性能，抗噪声能力、降维能力和识别能力较强，且运 行效率、收敛速度和全局寻优能力极具优势，提高网 络流量入侵检测识别的准确性和实时性，在实践应用 中有着很好的可行性。

**4** 结束语

本文算法继承了 ELM 和 AE的优点，通过 PSO 进

行优化，大大提高了检测精度和训练效率 。同时引入 了新的特征提取方法 KDS，可以将字符串特征映射到 （下转第110 页）

准确率%

准确率%

本栏目责任编辑：代 影 i a a a a n u u u  网络通信与安全  99

中国知网 https:Www.cnki . net

*Computer Knowledge and Technology* 电脑知识与技术 第20 卷第23 期 (2024 年8 月）

能力。

此外，案例分析还可以帮助学生更好地理解信息 安全的实际应用 。他们可以学习如何应对真实情境 下的安全事件，了解如何制定安全策略和采取行动， 以保护组织的数据和资产 。这种实际经验是毕业生 进入职场的巨大优势。

根据每年公布的 CVE 漏洞，每年更新校企合作搭 建的信息安全实践平台中的实验项目，以成都信息工 程大学为例，表 1 展示了近期更新的案例库。

表 1 案例库更新

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| CVE 编号 | 漏洞类型 | 是否允许 学生练习 | 难度 水平 | CVE 评分 |
| CVE-2023-46007 | SQL 注入漏洞 | 允许 | 难 | 9.8 |
| CVE-2023-33967 | SQL 注入漏洞 | 允许 | 中 | 8.2 |
| CVE-2023-46854 | XSS 脚本攻击 | 允许 | 中 | 5.4 |
| CVE-2023-46287 | XSS 脚本攻击 | 允许 | 一般 | 6.1 |
| CVE-2023-45798 | 文件包含 | 允许 | 中 | 8.4 |
| CVE-2023-39201 | 文件包含 | 允许 | 一般 | 7.2 |
| CVE-2023-46189 | CSRF | 允许 | 中 | 4.3 |

3) 定期调研人才需求，使培养方案与时俱进

随着社会和技术的不断发展，企业对信息安全人 才的需求也在不断变化 。为了确保课程培养大纲能 够与企业需求保持一致，定期调研人才需求变得至关 重要。

根据每年企业招聘对于信息安全专业学生的能 力要求，除了进行信息安全基础知识教学外，也引入 了前沿技术讲座，并在大二、大三各阶段，邀请企业导 师来学校给学生授课，与企业建立紧密的合作关系， 以了解他们的需求和期望 。并推荐同学进入到有条 件提供实践机会的企业中实习，以帮助他们应用所学

的知识。

**3** 结束语

网络攻击与防御课程是信息安全及相关专业的 必修课，但课程内容有一定难度，并且知识迭代速度 快 。通过协同育人模式的研究，对网络攻击与防御课 程进行改革和创新，从而激发学生的学习兴趣，提升 学生的学习效果，增强学生的实践动手能力 。从近 2 年的教学改革实践看，学生积极性得到了一定的提 高，实践能力普遍增强，在各级各类比赛获奖人数和 级别也在逐年递增。

在未来的教学中，课程仍需要从实际出发，紧密 结合企业人才需求，拓宽思路，让课堂生动起来，让学 生的主动性调动起来。

参考文献：

[1] 2020-2021 年网络安全人才现状白皮书[OL]. https://wenku. baidu. com/view/2a73a167cb50ad02de80d4d8d15abe23482f03 ae.html.

[2] 2022 年网络安全人才实战能力白皮书攻防实战能力篇[OL]. [https://weibo. com/ttarticle/p/show? id=23094048117827862530](https://weibo.com/ttarticle/p/show?id=23094048117827862530) 45#\_loginLayer\_ 1698761167525.

[3] 2023 网络安全人才市场状况研究报告(baidu.com)[OL].https: //baijiahao. baidu. com/s? id=1774255270483595914&wfr=spi⁃ der&for=pc.

[4] 网络安全人才实战能力白皮书[OL].<http://www.ahjgbzw.gov>. cn/content/detail/6317f2347f8b9aa0218b4567.html.

[5] 王晓燕,董昊 . 基于校企协同育人模式的数据结构课程教学 研究[J].南阳师范学院学报,2022,21(3):60-63.

[6] 王瑾琦,樊卓 . 校企协同育人模式下课程思政建设研究[J].广 西教育学院学报,2023(2):139-143.

【通联编辑：王 力】

（上接第99 页）

模型中，一定程度上解决了数据的高维稀疏问题 。未 来还需要从优化权重初始化、扩充 KDS 词库、丰富评 价指标等方面继续探索，进一步挖掘深度学习在入侵

检测领域的潜力，更好地守卫网络安全。 参考文献：

**[** 1] 宗学军,刘欢欢,何戡,等 . 高速网络流量下实时入侵检测系 统研究与应用[J].网络安全与数据治理,2023(4):56-61,84.

[2] 陈雪倩,步兵 . 基于网络流量和数据包的 CBTC 入侵检测系 统[J]. 中国安全科学学报, 2019,29 (S2): 154-160.

[3] 边金良 . 基于数据挖掘的网络异常流量入侵检测方法[J].信

息与电脑,2022(21):1-3.

[4] 王婷,王娜,崔运鹏,等 . 基于半监督学习的无线网络攻击行 为检测优化方法[J].计算机研究与发展,2020,57(4):791-802.

[5] 张永东 . 基于深度强化学习的网络流量分析与入侵检测[J]. 信息记录材料,2024,25(3):173-175.

[6] 王馨彤 . 基于深度学习的轻量化网络流量异常检测方法研 究与实现[D]. 南京:南京邮电大学, 2023.

[7] 朱平哲 . 基于深度学习的实时网络入侵检测方法[J].安阳工 学院学报,2019, 18(4):48-51.

[8] 何俊鹏,罗蕾, 肖堃,等 . 基于特征值分布和人工智能的网络 入侵检测系统的研究与实现[J].计算机应用研究,2021,38(9): 2746-2751.

【通联编辑：光文玲】

110  u n网络通信与安全  u u u u u u  本栏目责任编辑：代 影

中国知网 https:Www.cnki . net