# 基于机器学习的 DDoS 攻击检测与防护机制研究

刘俊侠

（上海建桥学院，上海 201306）

**摘 要** 分布式拒绝服务（DDoS）攻击已成为威胁网络安全的核心风险，传统检测方法存在误报率高、实时性差、难以应对新型攻击等缺陷。因此，本文提出一种融合特征工程与集成学习的 DDoS 攻击检测与防护机制，以期为 DDoS 攻击的实时检测与主动防护提供技术参考。基于 CSE-CIC-IDS2018 公开数据集的实验结果表明，该机制的检测准确率达 99.73%，误报率仅为 0.21%，检测延迟控制在 50 ms 以内，相较传统机器学习模型与单一深度学习模型展现出较高的工程应用价值与产业化前景优势。

**关键词** 机器学习；DDoS 攻击检测；防护机制

**中图分类号**：TP393.08 **文献标志码**：A **DOI**：10.3969/j.issn.2097-3365.2026.04.008

## 0 引言

随着数字经济的快速发展，当前网络攻击的方式及方法也在不断地升级，其中 DDoS 展示出较强的破坏力以及较高的隐蔽性，通过控制大量僵尸主机向目标服务器发送海量冗余流量的方式，导致服务器资源耗尽、服务不可用，此种攻击的规模与破坏力可不断扩大[1]。以往 DDoS 攻击的检测方法依赖于特征匹配与阈值判断，存在未知攻击的检测能力不足、无法区分正常突发流量与攻击流量、动态适配能力缺乏等问题，而机器学习技术凭借强大的特征学习与模式识别能力，为解决上述问题提供了新路径[2]。因此，本研究基于机器学习的 DDoS 攻击检测与防护一体化机制，对保障关键信息基础设施安全、促进数字经济健康发展具有重要的理论意义与实际应用价值。

## 1 DDoS 攻击类型与特征

### 1.1 常见的 DDoS 攻击类型

按照 DDoS 攻击的攻击目标与技术原理，可将其分为带宽消耗型、资源耗尽型与应用层、新型混合攻击四种类型[3]，如表 1 所示。

### 1.2 DDoS 攻击特征

通过提取其具有强区分度流量特征的方式，精准检测不同类型的 DDoS 攻击[4]；主要包含流量统计、时间序列、协议深层三种特征。

1. 流量统计特征：反映网络流量的静态分布特性。

2. 时间序列特征：捕捉流量的动态变化趋势。

3. 协议深层特征：包括 HTTP 协议特征、TCP 协议特征和加密流量特征等。

表 1 常见的 DDoS 攻击类型

| 类型 | 原理 | 流量特征 | 代表方式 |
|---|---|---|---|
| 带宽消耗型 | 发送海量冗余数据包占用目标网络带宽 | 数据包量大、单包体积小、协议类型单一（UDP/TCP） | UDP Flood、ICMP Flood |
| 资源耗尽型 | 建立大量无效连接消耗目标服务器 CPU/ 内存资源 | 连接数多、会话持续时间短、半开连接占比高 | SYN Flood、ACK Flood |
| 应用层攻击 | 模拟正常用户请求消耗应用服务器资源 | 数据包符合协议规范、请求频率高、payload 合法 | HTTP Flood、Slowloris |
| 新型混合攻击 | 融合多种攻击方式，规避检测规则 | 流量特征复杂、协议类型多样、攻击强度动态变化 | 加密流量攻击、AI 自适应攻击 |

**作者简介**：刘俊侠（1982-），女，硕士研究生，助教，研究方向：模式识别与数据挖掘。

## 2 基于机器学习的 DDoS 攻击检测机制

### 2.1 检测模型整体架构

本文设计的 XGBoost-LSTM 混合检测模型，结合了集成学习的强特征拟合能力与深度学习的时序特征捕捉能力。该模型分为四层：一是输入层，接收经过预处理与特征选择后的网络流量特征数据；二是特征提取层，将原始数据进行标签，去除重复数据、填充缺失值，过滤无效记录，之后进行归一化和时序化；三是训练层，由 XGBoost 模块与注意力机制 LSTM 模块组成，XGBoost 提取特征的非线性关系，LSTM 捕捉流量的时序依赖，注意力机制强化关键时序特征的权重；四是决策层，依据加权融合 XGBoost 与 LSTM 的输出结果，制定合理的防御策略。

### 2.2 数据预处理

首先，数据采集与标注，实验数据采用 CSE-CIC-IDS2018 公开数据集，将流量数据标注为“正常”（标签 0）与“攻击”（标签 1）两类，攻击样本占比 65%，正常样本占比 35%。

其次，数据清洗，对于连续型特征，采用中位数填充；对于离散型特征，采用众数填充；异常值则采用 3σ 准则识别异常值，将超出 [ μ -3 σ , μ +3 σ ] 范围的值替换为 μ ±3 σ（ μ 为特征均值，σ 为标准差）。

最后，数据归一化与时序化，采用 Min-Max 归一化将所有连续型特征映射至 [0,1] 区间，公式表示为：

$$x' = \frac{x - \min(x)}{\max(x) - \min(x)} \tag{1}$$

式（1）中，$x$ 为原始特征值；$x'$ 为归一化后的特征值。

### 2.3 核心模块设计

本文采用 XGBoost 模型提取特征的非线性关系，将 XGBoost 模块的输出为特征的非线性映射向量（维度为 64），作为后续融合层的输入之一。引入注意力机制，增强模型对重要时间步特征的感知能力。注意力机制 LSTM 模块的输出为时序特征向量（维度为 128），作为融合层的另一输入。融合决策层采用加权求和的方式融合 XGBoost 模块与注意力机制 LSTM 模块的输出，公式表示为：

$$o=\omega_1 F_{XGB}+\omega_2 F_{LSTM} \tag{2}$$

式（2）中，$\omega_1$=0.4，$\omega_2$=0.6；$F_{XGB}$ 为 XGBoost 模块的输出向量；$F_{LSTM}$ 为注意力机制 LSTM 模块的输出向量。将融合后的向量输入 Sigmoid 激活函数，得到攻击检测概率，当概率大于 0.5 时判定为攻击，否则为正常。

## 3 SDN-based DDoS 攻击防护机制

为实现检测与防护的闭环，本文提出基于 SDN 的智能防护架构，该架构分为感知层、检测层、决策层、执行层和反馈层五个部分。

1. 感知层：主要包括流量采集、数据预处理、数据传输。

2. 检测层：部署本文提出的 XGBoost-LSTM 混合检测模型，接收感知层传输的特征数据，实时输出检测结果。

3. 决策层：基于检测层的攻击信息，结合网络拓扑、带宽资源、业务优先级等因素，制定个性化防护策略。

4. 执行层：执行层由 SDN 交换机、防火墙、负载均衡器与黑洞路由模块组成，负责执行决策层制定的防护策略。

5. 反馈层：以防护效果为奖励函数，通过 RL 智能体不断学习不同攻击场景下的最优防护策略，实现防护策略的自适应调整。奖励函数公式表示为：

$$R=\alpha P_{normal}+\beta P_{block}-\gamma D \tag{3}$$

式（3）中，$P_{normal}$ 为正常流量通过率；$P_{block}$ 为攻击阻断率；$D$ 为服务延迟，$\alpha$=0.4、$\beta$=0.5、$\gamma$=0.1，均为权重系数。

## 4 实验验证

### 4.1 实验设置

1. 环境：CPU 为 Intel Core i7-12700K（12 核 20 线程），内存 32GB，硬盘 1TB SSD，交换机为 Open vSwitch 2.15.0；操作系统为 Ubuntu 20.04 LTS，深度学习框架为 TensorFlow 2.8，机器学习库为 Scikit-learn 1.0。

2. 数据集：采用 CSE-CIC-IDS2018 公开数据集与自建数据集相结合的方式；其中 CSE-CIC-IDS2018 数据集，10GB，用于模型训练与测试；自建数据集则通过在实验室环境中模拟 UDP Flood、SYN Flood、HTTP Flood 等攻击，采集流量数据 5GB，用于模型泛化能力测试；数据集划分比例为训练集 70%，验证集 15%，测试集 15%。

3. 评价指标 [5]：采用网络安全领域常用的评价指标，包含准确率（Accuracy）、召回率（Recall）、F1 分数（F1-Score）、误报率（FPR）、检测延迟（Detection Latency）。

### 4.2 实验结果

在性能方面，将本文提出的 XGBoost-LSTM 混合模型与传统机器学习模型（SVM、XGBoost）、单一深度学习模型（LSTM）进行对比实验，实验结果如表 2 所示。

由表 2 可知，本文提出的检测方式能够有效提升

表 2 不同模型在 CSE-CIC-IDS2018 数据集上的性能

| 模型 | 准确率（%） | 召回率（%） | F1 分数（%） | 误报率（%） | 延迟（ms） |
|---|---|---|---|---|---|
| XGBoost-LSTM | 99.73 | 99.58 | 99.65 | 0.21 | 48 |
| SVM | 92.36 | 90.15 | 91.24 | 2.87 | 89 |
| XGBoost | 98.57 | 97.89 | 98.23 | 0.86 | 63 |
| LSTM | 98.12 | 97.34 | 97.73 | 0.98 | 75 |

DDoS 攻击检测的精度与实时性。

在泛化能力方面，采用自建数据集进行测试，结果如图 1 所示，可知模型具有较强的泛化能力。

在防护机制方面，通过在 SDN 实验环境中模拟不同类型、不同强度的 DDoS 攻击，结果如表 3 所示。

由表 3 可知，SDN-based 防护机制能够根据攻击类型与强度动态调整防护策略，实现攻击的精准阻断与正常流量的有效保障。

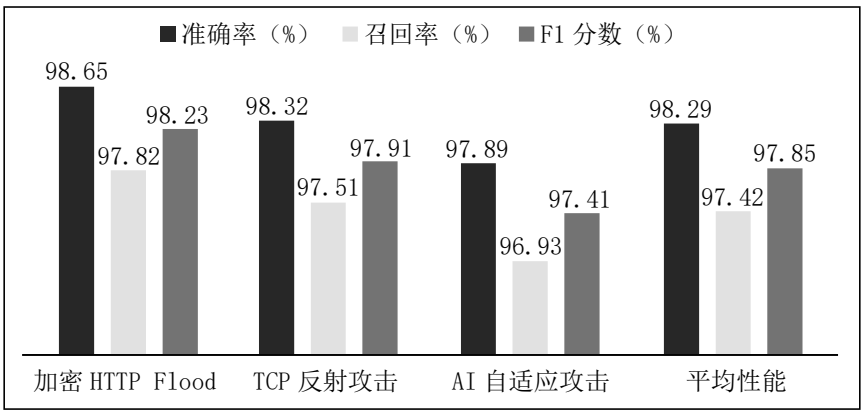


图 1 泛化能力测试结果

表 3 防护机制有效性测试结果

| 攻击 | 强度 | 防护策略 | 阻断率（%） | 流量通过率（%） | 延迟（ms） |
|---|---|---|---|---|---|
| UDP Flood | 低 | 流量限速 | 99.23 | 99.78 | 25 |
| UDP Flood | 高 | 黑洞路由 | 99.91 | 99.56 | 32 |
| SYN Flood | 中 | 流量清洗 + 源 IP 验证 | 99.45 | 99.67 | 41 |
| HTTP Flood | 中 | 流量清洗 + 负载均衡 | 99.38 | 99.72 | 38 |
| 加密 HTTP Flood | 高 | 流量清洗 + 黑洞路由 | 99.85 | 99.43 | 45 |

## 5 结束语

基于机器学习的 DDoS 攻击检测与防护机制，不仅具有强区分度的核心特征集，降低了模型计算复杂度，还集成了机器学习与深度学习的优势，展示出较强的泛化能力；在防护机制方面具有较高的攻击阻断率、正常流量通过率以及可控的服务延迟优势，可有效为网络安全与业务连续性提供基础保障。因此，在未来研究中，可通过提高模型对对抗性 DDoS 攻击的检测能力，进一步优化防护策略的有效性，达到提升混合攻击防护效果的目的。

## 参考文献：

[1] 马立鑫，薛占双，刘海燕 .DDoS 攻击的发展与检测技术研究 [J]. 现代计算机 ,2024,30(20):52-56,62.
[2] 苗浩宇，吴报玉，梁雪婷，等 . 通信网络中 DDoS 攻击动态防御策略优化研究 [J]. 中国宽带 ,2025,21(11):88-90.
[3] 刘会飞，朱雨雷，刘皓天，等 . 基于人工智能的工业物联网安全防护综述 [J]. 物联网技术 ,2025,15(10):118-125.
[4] 杨阳，费晓琛 . 面向物联网通信的人工智能安全机制与防护技术研究 [J]. 微型计算机 ,2025(19):184-186.
[5] 国家市场监督管理总局，国家标准化管理委员会 . 网络安全技术 信息技术安全评估方法 (GB/T 30270-2024)[S]. 北京：中国标准出版社 ,2024.