Software Engineering and Applications 软件工程与应用, 2025, 14(6), 1258-1269
Published Online December 2025 in Hans. https://www.hanspub.org/journal/sea
https://doi.org/10.12677/sea.2025.146111

Hans汉斯

融合大语言模型的可信网络异常入侵检测方法

王　付

佛山市殡仪馆，广东　佛山

收稿日期：2025年10月13日；录用日期：2025年12月11日；发布日期：2025年12月19日

摘　要

当前网络异常入侵检测多依赖黑箱模型，缺乏可解释性，难以支持工程师快速理解告警原因并及时响应，从而严重制约了系统的安全防护效率。为此，本文提出一种融合大语言模型的可信网络异常入侵检测方法。该方法首先使用XGBoost对网络流量进行异常识别，并利用SHAP分析特征对判定结果的贡献，揭示模型决策依据；随后，借助大模型将检测结果与SHAP解释整合为自然语言报告，提供直观易读的告警说明与处置建议。实验结果表明，该方法在保持较高检测精度的同时显著提升了模型的可解释性与实际可操作性。

关键词

网络异常入侵检测，大语言模型，安全可信

Trusted Network Anomaly Intrusion Detection Method Integrating Large Language Model

Fu Wang

Foshan Funeral Home, Foshan Guangdong

Received: October 13, 2025; accepted: December 11, 2025; published: December 19, 2025

Abstract

Current network anomaly intrusion detection often relies on black-box models, which lack interpretability and hinder engineers' ability to quickly understand alert causes and respond promptly, thereby severely limiting the efficiency of security protection systems. To address this issue, this paper proposes a trustworthy network anomaly intrusion detection method that integrates large language models. The approach first employs XGBoost to identify anomalies in network traffic and utilizes

文章引用: 王付. 融合大语言模型的可信网络异常入侵检测方法[J]. 软件工程与应用, 2025, 14(6): 1258-1269.
DOI: 10.12677/sea.2025.146111

SHAP to analyze the contribution of features to the detection results, thereby revealing the decision-making basis of the model. Subsequently, a large language model is leveraged to integrate the detection results and SHAP explanations into a natural language report, providing intuitive and easily understandable alert descriptions along with actionable recommendations. Experimental results demonstrate that, while maintaining high detection accuracy, the proposed method significantly enhances model interpretability and practical usability.

Keywords

Network Anomaly and Intrusion Detection, Large Language Model, Security and Trustworthiness

1. 引言

随着网络攻击手段的日益复杂，传统的入侵检测系统(Intrusion Detection System, IDS)正面临愈发严峻的挑战[1]。尽管基于深度学习的检测模型在检测精度方面表现突出，但其内部决策过程往往缺乏可解释性，使得安全运营人员难以理解模型的判定依据。作为典型的黑箱模型，这一特性不仅增加了误报处理的成本，也削弱了告警响应的及时性，从而影响整体防护效能[2]。因此，提升入侵检测系统的可解释性与可信度，已成为亟待解决的关键问题。

近年来，深度学习方法，尤其是卷积神经网络(Convolutional Neural Network, CNN)和循环神经网络(Recurrent Neural Network, RNN) [3]-[5]，在网络异常入侵检测中取得了显著进展。这类方法能够自动提取高维网络流量特征，并在分类准确性上表现优异[6]。然而，深度模型的黑箱特性导致其决策过程对安全运营中心工程师而言往往不可见，降低了模型在实际应用中的可解释性与可操作性。模型决策缺乏透明度不仅会造成误报与漏报的增加，也使得系统难以获得安全专家的信任，从而制约了其在关键场景下的落地应用[7]。

针对上述可解释性缺失的问题，学界逐渐从“高精度检测”转向“可信智能检测”的研究方向。可解释人工智能(Explainable Artificial Intelligence, XAI)的出现为解决这一瓶颈提供了新的思路，相关技术近年来受到广泛关注。其中，基于特征贡献度分析的 SHAP 方法能够有效揭示模型决策的内部逻辑，通过量化各特征对预测结果的影响，从而提升模型的透明度与可理解性。进一步地，将大语言模型引入检测系统，可以基于模型解释结果生成自然语言报告，将复杂的检测逻辑转化为人类易于理解的语义描述，显著提升系统的可操作性与人机交互体验。

本文提出了一种融合大语言模型的可信网络异常入侵检测方法。首先，利用使用 XGBoost 对网络流量进行异常识别，并利用 SHAP 分析特征对判定结果的贡献，揭示模型决策依据。随后，结合大语言模型生成自然语言报告，使检测结果及解释内容以可理解、可追溯的形式呈现。通过在公开数据集上的实验验证，结果表明该方法在保持高检测精度的同时，显著提升了入侵检测系统的可解释性与实用性，为构建智能化、可信赖的网络异常入侵检测体系提供了有效解决方案。

2. 可信网络异常入侵检测模型

本章旨在从系统层面设计一套可信网络异常入侵检测模型，其处理流程遵循“输入–检测–解释–报告”的完整范式。为系统性地展开论述，本章主要内容划分为四个部分，依次对应上述流程环节：2.1.

总体架构设计、2.2. 特征提取与预处理、2.3. 异常检测模型实现，以及 2.4.可解释性报告生成。

2.1. 总体架构

可信网络异常入侵检测模型的总体结构如图 1 所示。系统以滑动时间窗接入实时或离线的网络流量，并通过多线程特征管线生成数值化样本，将样本输入 XGBoost 模型完成判定，应用 SHAP 方法计算每个样本的特征贡献分解，最后将“判定 + 解释”的结构化结果交由大模型生成可读的告警与处置建议，实现从黑盒输出到人机协同的可信闭环。

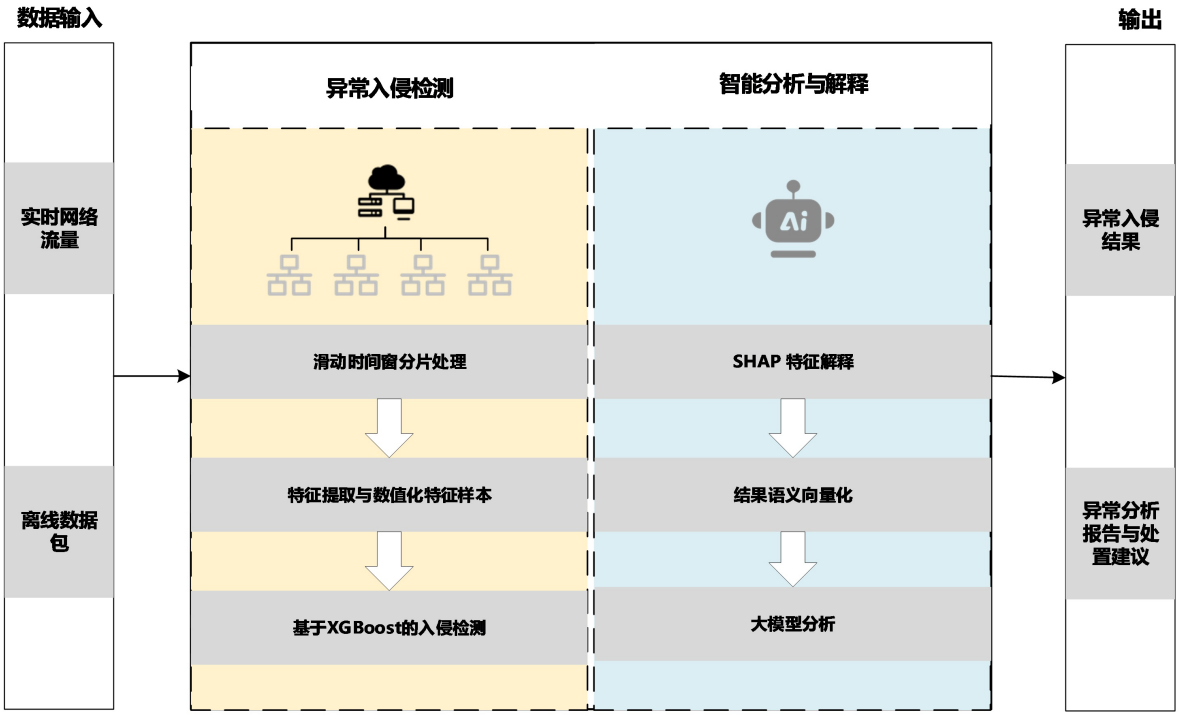

Figure 1. Model architecture diagram
图 1. 模型架构图

2.2. 特征提取与预处理

为了保证入侵检测模型的准确性与稳定性，系统在数据进入分类模型之前进行了系统化的特征提取与预处理。网络流量首先按照五元组(源 IP、目的 IP、源端口、目的端口、协议类型)进行会话聚合，并以滑动时间窗的方式生成样本。每个时间窗 $[t, t+W)$ 内的流量数据被视为一个观测样本，其特征向量记为：

$$\boldsymbol{x} = [x_1, x_2, \cdots, x_d]^{\mathrm{T}} \tag{1}$$

其中 d 表示特征维度，$\boldsymbol{x}$ 是模型的输入。该特征体系旨在全面反映网络流量的结构性与统计性特征，主要涵盖四个维度：基础流特征、统计流特征、TCP 标志特征。

在基础流特征中，选取了流持续时间(Duration)、源字节数(SrcBytes)、目的字节数(DstBytes)和协议类型(Protocol)等参数。其中，协议类型为离散变量，通过独热编码(One-Hot Encoding)转化为稀疏二进制向量：

$$Protocol \Rightarrow [p_1, p_2, \cdots, p_C], p_i \in \{0,1\} \tag{2}$$

其中 C 为协议类别总数。此类特征可有效区分不同通信协议的行为差异，为分类模型提供先验结构信息。

统计流特征用于刻画通信流在时间与数据量维度上的动态特征。定义平均流速率(FlowRate)、平均包间到达时间(FlowIATMean)与标准差(FlowIATStd)如下：

$$FlowRate = \frac{SrcBytes + DstBytes}{Duration},\ FlowIATMean = \frac{1}{N}\sum_{i=1}^{N} IAT_i, FlowIATStd = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(IAT_i - FlowIATMean\right)^2} \tag{3}$$

其中 IAT_i 表示第 i 个数据包的到达时间间隔， N 为时间窗内的数据包数量。正常通信的 IAT 通常呈稳定分布，而攻击流量往往表现出突发或周期异常，因此这些特征对异常检测尤为敏感。为增强流量特征的表达能力，还可引入派生指标，如每秒数据包数(FlowPacketsS)、平均包长(MeanPktSize)及字节包比(BytesPerPkt)，以反映通信频度与数据占用率。

TCP 报文头中的控制位标志对识别异常连接行为具有重要作用。提取六种常见标志位：SYN、ACK、FIN、RST、PSH 与 URG，构建二进制特征向量：

$$FlagVector = [SYN, ACK, FIN, RST, PSH, URG], Flag_i \in \{0,1\}. \tag{4}$$

特征提取完成后，需对缺失值、异常值及数据分布进行清理与标准化。对数值型特征，若其分布近似正态，则采用均值插补；若为偏态分布，则使用中位数插补：

$$x_j = \begin{cases} \text{mean}\left(x_j\right), & \text{若} x_j \text{服从近似正态分布,} \\ \text{median}\left(x_j\right), & \text{若} x_j \text{存在偏态或离群现象} \end{cases} \tag{5}$$

异常值通过标准差法检测与修正。若样本特征超出 $\mu \pm k\sigma$ 范围(其中 $k \in [2,3]$)，则将其替换为均值 μ，以防止极端值影响模型训练。

为消除量纲差异并提高模型收敛速度，对连续型特征采用归一化处理。根据实际情况，选择 Min-Max 归一化或 Z-Score 标准化：

$$x_j' = \frac{x_j - \min\left(x_j\right)}{\max\left(x_j\right) - \min\left(x_j\right)}, x_j'' = \frac{x_j - \mu_j}{\sigma_j} \tag{6}$$

其中， μ_j 与 σ_j 分别为第 j 个特征的均值与标准差。离散特征采用独热编码处理，以避免模型将离散值误解为连续量。

由于网络流量中正常样本远多于异常样本，为防止分类器偏向多数类，引入类别加权机制。类别权重定义为：

$$\alpha_c = \frac{1}{\log\left(\beta + q_c\right)} \tag{7}$$

其中 q_c 为类别 c 在训练集中的样本比例， $\beta > 1$ 为平滑系数。该权重在模型训练阶段参与损失函数计算，使少数类样本在梯度更新中获得更高的重要性。

经上述步骤处理后，原始流量数据被转化为归一化、无异常、维度一致的数值向量 $\tilde{\boldsymbol{x}}$。这一特征矩阵不仅保留了网络流量的核心统计特征，还有效消除了噪声与量纲差异，为后续异常检测模型提供了高质量输入数据。此外，通过统一的数据清洗与归一化策略，系统可实现离线批量训练与在线实时检测的

无缝衔接，为 2.3 节的 XGBoost 异常检测模块奠定基础。

2.3. 异常检测

本研究选用 XGBoost 作为入侵异常检测算法，其预测函数可表示为：

$$f(\tilde{\boldsymbol{x}})=\sum_{t=1}^{T} f_t(\tilde{\boldsymbol{x}}), f_t \in \mathcal{F} \tag{8}$$

其中，$\tilde{\boldsymbol{x}}$ 为输入特征向量，$f_t(\cdot)$ 表示第 t 棵树，T 为树的总数，$\mathcal{F}$ 表示所有可能树的集合。模型输出经过 Sigmoid 函数映射为概率值：

$$\hat{p}=\sigma\left(f(\tilde{\boldsymbol{x}})\right)=\frac{1}{1+e^{-f(\tilde{\boldsymbol{x}})}}. \tag{9}$$

当 $\hat{p} \geq \tau$ 时，系统将该样本判定为异常；否则为正常。阈值 τ 可根据验证集的精度—召回率权衡进行自适应调整。

XGBoost 的学习过程通过最小化以下目标函数实现：

$$\mathcal{L}=\sum_{i=1}^{n} \alpha_{y_i} \ell\left(y_i, \hat{p}_i\right)+\sum_{t=1}^{T} \Omega\left(f_t\right) \tag{10}$$

其中，$\ell\left(y_i, \hat{p}_i\right)$ 为样本级损失函数，定义为二元交叉熵：

$$\ell\left(y_i, \hat{p}_i\right)=-\left[y_i \log\left(\hat{p}_i\right)+\left(1-y_i\right) \log\left(1-\hat{p}_i\right)\right] \tag{11}$$

α_{y_i} 为类别加权系数(由 2.2 节定义)，用于缓解类别不平衡问题；$\Omega\left(f_t\right)$ 为模型复杂度正则项，形式为：

$$\Omega\left(f_t\right)=\gamma \cdot T_{leaves}+\frac{\lambda}{2} \sum_j w_j^2 \tag{12}$$

其中，T_{leaves} 表示树的叶节点数，w_j 为叶节点权重，γ 与 λ 分别控制树的结构复杂度和参数惩罚力度。正则项的引入能够限制模型复杂度，防止过度拟合训练样本，从而提升泛化性能。

XGBoost 使用二阶泰勒展开对损失函数进行近似，从而实现高效的贪婪式树结构搜索。在第 t 轮迭代中，目标函数可写为：

$$\mathcal{L}^{(t)} \approx \sum_{i=1}^{n}\left[g_i f_t\left(\tilde{\boldsymbol{x}}_i\right)+\frac{1}{2} h_i f_t^2\left(\tilde{\boldsymbol{x}}_i\right)\right]+\Omega\left(f_t\right) \tag{13}$$

其中，

$$g_i=\frac{\partial \ell\left(y_i, \hat{p}_i\right)}{\partial f^{(t-1)}\left(\tilde{\boldsymbol{x}}_i\right)}, h_i=\frac{\partial^2 \ell\left(y_i, \hat{p}_i\right)}{\partial\left[f^{(t-1)}\left(\tilde{\boldsymbol{x}}_i\right)\right]^2} \tag{14}$$

分别表示一阶与二阶梯度。通过对每个叶节点的样本集合 I_j 进行求和，可得最优权重：

$$w_j^*=-\frac{\sum_{i \in I_j} g_i}{\sum_{i \in I_j} h_i+\lambda} \tag{15}$$

并据此计算分裂增益(Gain)：

$$\text{Gain}=\frac{1}{2}\left(\frac{G_L^2}{H_L+\lambda}+\frac{G_R^2}{H_R+\lambda}-\frac{G^2}{H+\lambda}\right)-\gamma \tag{16}$$

其中 G_L, G_R, H_L, H_R 分别为左右子树的梯度与二阶梯度和。模型在每一步选择使增益最大的特征与阈值进行分裂，从而逐步逼近最优分类边界。

为保证实验的稳定性与可复现性，本研究在模型训练中采用以下主要参数：学习率(Learning Rate)设为 0.1，最大树深(Max Depth)为 6，最小叶节点权重(Min Child Weight)设为 1，子采样率(Subsample)与列采样率(Colsample Bytree)均为 0.8，迭代轮数设为 100。

此外，为防止过拟合，模型在训练过程中采用早停机制(Early Stopping)，当验证集损失连续若干轮无明显下降时停止训练。

在类别极度不平衡的情况下，进一步引入比例权重：

$$scale\_pos\_weight = \frac{N_{neg}}{N_{pos}} \tag{17}$$

以使正负样本在梯度空间中保持平衡。所有实验均基于相同的超参数搜索策略，以确保结果的可比性。

传统的二分类模型通常使用固定阈值 $\tau = 0.5$，但在入侵检测场景中，不同攻击类型的代价不对称，误报与漏报对系统安全的影响差异显著。

因此，本文在验证集上采用基于代价敏感的最优阈值策略：

$$\tau^* = \frac{C_{FP}}{C_{FP} + C_{FN}} \tag{18}$$

其中 C_{FP} 与 C_{FN} 分别表示误报与漏报的代价权重。通过该方法可实现检测率(Recall)与精确率(Precision)的动态平衡。

为了进一步提高概率输出的可信度，引入温度校准(Temperature Scaling)机制，将模型输出的对数几率通过参数 T 进行线性缩放：

$$\hat{p}_T = \sigma\left(\frac{f(\tilde{x})}{T}\right), T^* = \arg\min_{T>0} \sum_i -\log\left(\hat{p}_{T,y_i}\right) \tag{19}$$

其中，T 通过验证集最优化获得。当 $T > 1$ 时，模型预测变得更保守；当 $T < 1$ 时，预测置信度增大。温度校准可显著改善模型概率输出的校准误差(Expected Calibration Error, ECE)，使预测结果更符合真实分布。

经过上述训练与校准，异常检测模块最终输出两个结果：

1) 样本的异常判定标签 $\hat{y} \in \{0,1\}$；

2) 经温度校准后的置信度概率 $\hat{p}_T \in [0,1]$。

这两个量将作为输入传递至后续的可解释性分析模块(见第 2.4 节)，在其中通过 SHAP 方法计算各特征的贡献度，并由大语言模型(LLM)生成自然语言告警报告。

2.4. 可解释性报告生成

本研究在异常检测模块之后，引入基于 SHAP 的可解释性分析方法，并结合大模型生成自然语言告警报告，从而实现“模型–解释–运维”的信息闭环。

SHAP 方法通过计算每个特征在模型决策中的边际贡献，量化其对预测结果的重要性。设检测模型的输出为 $f(\tilde{x})$，则特征 i 的 SHAP 值定义为所有特征子集 S 的边际增益的加权平均：

$$\phi_i = \sum_{S \subseteq N \setminus \{i\}} \frac{|S|!(M - |S| - 1)!}{M!} \left(f(S \cup \{i\}) - f(S)\right) \tag{20}$$

其中 N 表示全部特征集合，M 为特征总数，ϕ_i 为特征 i 对最终预测的贡献值，模型输出可以写作特征贡献的加性分解：

$$f(\tilde{\boldsymbol{x}}) = f_0 + \sum_{i=1}^{M} \phi_i \tag{21}$$

其中 f_0 为模型的基线输出(即在无特征条件下的期望预测值)。

在实际计算中，本文采用 TreeSHAP 算法，利用 XGBoost 树模型的结构特性，在多项式时间内准确获得每个样本的 ϕ_i。

正的 SHAP 值($\phi_i > 0$)代表该特征促使模型更倾向于判断为异常，而负值($\phi_i < 0$)则表示该特征抑制了异常预测。通过对 SHAP 值的绝对值排序，可以得到最具解释性的特征集合 $\mathcal{F}_K$：

$$\mathcal{F}_K = \arg\operatorname*{topK}_i \left|\phi_i\right| \tag{22}$$

即贡献度最高的前 K 个特征。

此外，本研究将 SHAP 分析与温度校准概率 $\hat{p}_T$ 结合，形成可信度分层机制。

根据预测概率与不确定度定义综合严重度评分 S：

$$S = \hat{p}_{T,\max} \times \left(1 - \frac{H(\hat{\boldsymbol{p}}_T)}{\log K}\right) \tag{23}$$

其中 $H(\hat{\boldsymbol{p}}_T)$ 为预测分布的熵。系统 S 划分为高(≥0.7)、中(0.4~0.7)和低(<0.4)三个等级，对应不同的响应优先级。

在获得模型判定结果与特征贡献后，系统通过大模型自动生成自然语言告警报告，使复杂的量化分析结果转化为清晰、可操作的文字说明。LLM 以结构化输入集合为基础：

$$\mathcal{U} = \left\{\hat{y}, \hat{p}_T, S, \left(Feature_i, \phi_i\right)_{i\in\mathcal{F}_K}\right\} \tag{24}$$

其中，$\hat{y}$ 表示预测类别标签，$\hat{p}_T$ 表示温度校准后的预测概率，S 表示严重度评分，$Feature_i$ 表示第 i 个输入特征，ϕ_i 表示特征 i 的 SHAP 贡献值，$\mathcal{F}_K$ 表示贡献度排名前 K 的关键特征集合。

综上，通过 SHAP 提供可验证的因果解释，并利用大模型将技术性结果转化为可读报告，实现了从检测到解释再到响应的可信闭环。通过温度校准、特征贡献分析与语义化报告生成三层机制，本研究从输出可靠性、决策透明性与人机可理解性三方面构建了多层次可信保障体系，实现了 AI 安全检测的可验证与可追溯。

3. 实验及结果分析

3.1. 实验设置

Table 1. Experimental operating environment
表 1. 实验运行环境

名称	配置
操作系统	Ubuntu 22.04 LTS
CPU	Intel Core i7-12700
GPU	NVIDIA RTX 3060 (12GB VRAM)
存储	2 TB SSD

续表

应用大模型	Llama 3 (本地部署)
编程语言	Python 3.10

本实验使用了 CIC-IDS2017 数据集，包含多种网络入侵攻击类型和正常流量，其中 70%的数据用于训练，30%的数据用于测试，确保有效评估模型的泛化能力。在测试集中，类别分布呈现出明显的不平衡现象，PortScan 和 BENIGN 占据了绝大多数样本，而 Bot 和 SSH-Patator 等攻击类别的样本较为稀少。为保证实验的可重复性，实验运行环境的具体配置列于表 1。

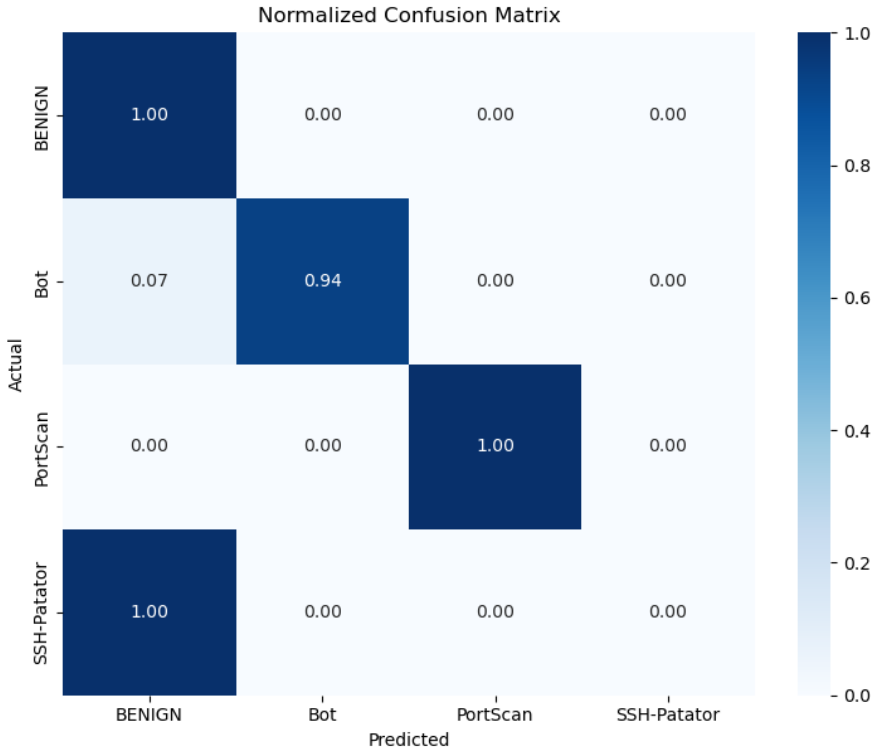

Figure 2. Confusion matrix of intrusion detection results
图 2. 入侵检测结果混淆矩阵

3.2. 实验结果

为全面评估模型性能，首先从分类准确性与泛化能力两方面进行分析。如图 2 所示，通过混淆矩阵分析，BENIGN 和 PortScan 类别的分类准确率接近 100%，表明模型在常见类别上的检测能力较强；然而，针对 Bot 类样本，模型的误判率约为 7%，并且 SSH-Patator 类几乎全部被误判为正常流量，这表明模型在少数类攻击的识别上存在一定局限性。

如图 3 中的 ROC 曲线显示，Micro-average AUC 达到 0.989，表明模型能够在不同决策阈值下维持较高的分类性能。此外如图 4 所示，Precision-Recall 曲线的平均精确度(Average Precision, AP)为 0.973，表明在处理不平衡数据时，模型仍能有效区分攻击流量与正常流量。进一步的置信度校准分析显示，在高置信度区间(0.9~1.0)，模型的预测结果具有较高的准确性；然而，在中低置信度区间，模型存在一定的过度自信现象，如图 5 所示。

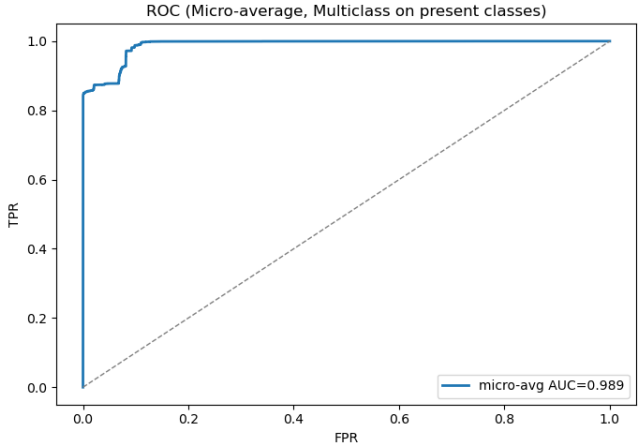

Figure 3. Intrusion detection ROC curve
图 3. 入侵检测 ROC 曲线

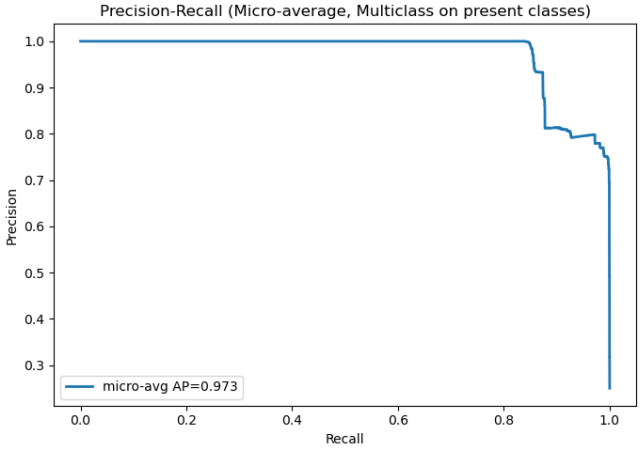

Figure 4. Intrusion detection Precision-Recall curve
图 4. 入侵检测 Precision-Recall 曲线

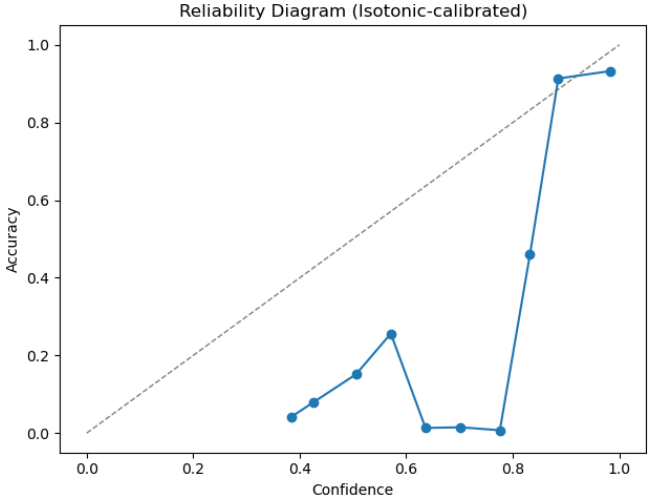

Figure 5. Intrusion detection confidence calibration chart
图 5. 入侵检测置信度校准图

为深入理解模型的预测过程，本实验引入 SHAP 分析对入侵检测结果进行了可解释性分析。如图 6 所示，模型在判断网络流量是否为攻击时，主要依赖于几个关键特征，这些特征对模型输出结果的影响较大。具体而言，数据包数目、数据包长度和目标端口等特征在区分正常流量与攻击流量时起到了重要作用。尤其是在 PortScan 等攻击行为中，数据包的数量和长度显著增加，是模型判定攻击的重要依据。同时 SHAP 分析也揭示了模型在少数类攻击识别中的局限性，尤其是在攻击模式较为隐蔽或与正常流量差异较小的情况下，模型对这些攻击类型的识别能力较弱。

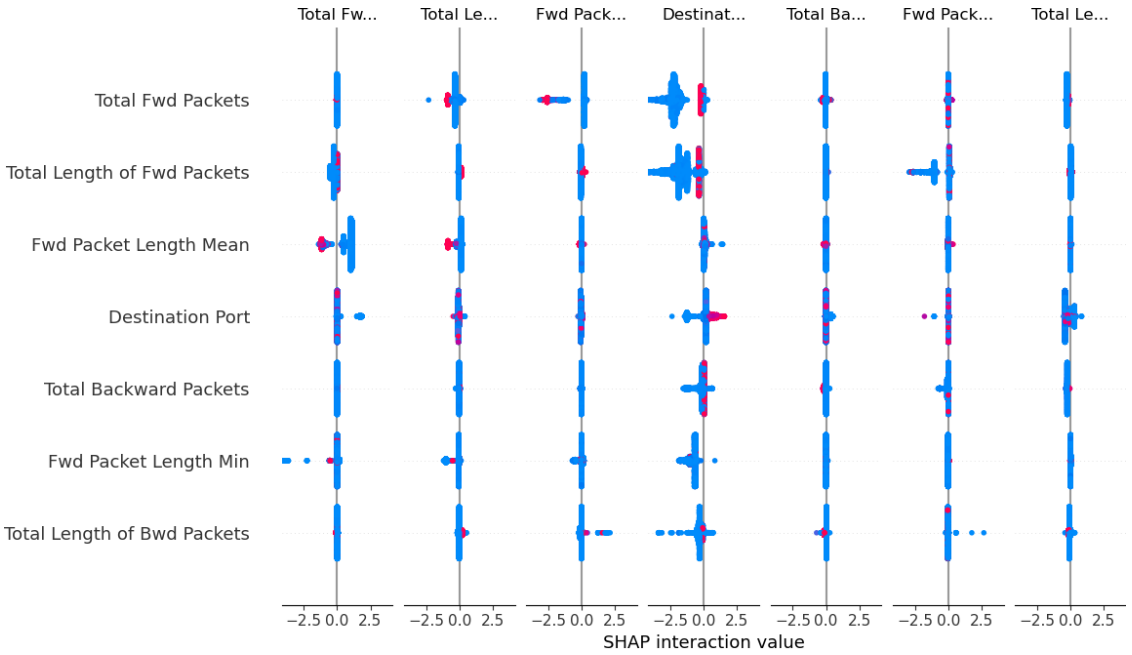

Figure 6. Intrusion detection feature contribution analysis chart
图 6. 入侵检测特征贡献度分析图

为了进一步提升模型的可解释性与可操作性，本实验结合了大语言模型来生成自然语言报告。从结构与内容上看，报告分为三个逻辑递进的模块，形成了一个完整的“特征归因—攻击模式分析—防护响应建议”的研判闭环，生成的报告如图 7 所示。基于 SHAP 分析，模型能够自动生成易于理解的告警报告，这些报告不仅帮助安全工程师理解模型的预测依据，还能提供明确的应对策略。通过大语言模型的结合，复杂的 SHAP 输出得以转化为简洁的自然语言报告，显著提升了告警信息的可读性与操作性。此举不仅增强了安全团队对入侵检测系统的信任，也加快了响应速度，提高了防护效率。

4. 结论

本文针对传统入侵检测系统可解释性不足的问题，提出了一种融合大语言模型的可信网络异常入侵检测方法。该方法在保持高检测精度的同时，通过引入 SHAP 可解释性分析，揭示模型决策过程中各特征的贡献关系，从而使模型判定过程可追溯、可理解。实验结果表明，本方法在多种公开数据集上均取得了较高的检测性能，并显著提升了系统在复杂网络环境下的稳定性与透明度。

此外，通过结合大模型生成自然语言告警报告，系统能够将复杂的模型输出转化为人类易读的安全说明与处置建议，实现了从检测到解释再到响应的智能化闭环。该研究不仅提升了入侵检测系统的可信度与可操作性，也为未来安全运营中心(SOC)的智能化决策支持提供了新的思路与技术方向。

Figure 7. LLM intrusion detection analysis report
图 7. LLM 入侵检测分析报告

参考文献

[1] 蹇诗婕, 卢志刚, 杜丹, 等. 网络入侵检测技术综述[J]. 信息安全学报, 2020, 5(4): 96-122.

[2] 封化民, 李明伟, 侯晓莲, 等. 基于 SMOTE 和 GBDT 的网络入侵检测方法研究[J]. 计算机应用研究, 2017, 34(12): 3745-3748.

[3] 刘衍珩, 田大新, 余雪岗, 等. 基于分布式学习的大规模网络入侵检测算法[J]. 软件学报, 2008(4): 993-1003.

[4] 周杰英, 贺鹏飞, 邱荣发, 等. 融合随机森林和梯度提升树的入侵检测研究[J]. 软件学报, 2021, 32(10): 3254-3265.

[5] 李辉, 管晓宏, 昝鑫, 等. 基于支持向量机的网络入侵检测[J]. 计算机研究与发展, 2003(6): 799-807.

[6] Louk, M.H.L. and Tama, B.A. (2023) Dual-ids: A Bagging-Based Gradient Boosting Decision Tree Model for Network Anomaly Intrusion Detection System. *Expert Systems with Applications*, **213**, Article 119030. https://doi.org/10.1016/j.eswa.2022.119030

[7] Ayad, A.G., Sakr, N.A. and Hikal, N.A. (2024) A Hybrid Approach for Efficient Feature Selection in Anomaly Intrusion Detection for IoT Networks. *The Journal of Supercomputing*, **80**, 26942-26984. https://doi.org/10.1007/s11227-024-06409-x