戴睿敏, John M Acken. 基于 XGBoost 的入侵检测系统评估研究[J]. 智能计算机与应用,2025,15(9):26-32. DOI:10.20169/j.issn.2095-2163.25032601

基于 XGBoost 的入侵检测系统评估研究

戴睿敏[1,2], John M Acken[2]
(1 南京邮电大学 通信与信息工程学院, 南京 210003; 2 波特兰州立大学 电子与计算机工程学院, 波特兰 97201, 美国)

摘 要: 入侵检测系统(Intrusion Detection System,IDS)在保障网络安全中扮演关键角色,但传统方法在面对复杂多变的攻击场景时仍存在检测精度不高、响应滞后等问题。为提升入侵检测系统的实用性和实时性,本文提出一种基于 XGBoost 的双阶段检测框架。该方法首先通过子段分类实现粗粒度筛查,随后对可疑段落进行数据点级别的细粒度分析,全面提升检测效率与精度。在 ROSPaCe 数据集上开展的实验表明:该模型在等错误率(EER)阈值为 0.7 条件下,准确率可达 95.99%,假阳率显著降低,平均检测延迟控制在可接受范围内。此外,针对零日攻击的模拟试验揭示:训练数据的多样性对于模型的泛化能力具有重要影响,数据覆盖不足将显著削弱检测效果。本文研究结果表明,准确率、检测延迟与误报控制之间存在复杂的非线性关系,三者应作为入侵检测系统评估中的关键指标予以综合考量,从而为物联网与工业控制系统等场景下的 IDS 部署提供理论支持和优化依据。

关键词: 入侵检测系统; 评估指标; 攻击检测延迟; XGBoost; 机器学习
中图分类号: TP393.0 **文献标志码**: A **文章编号**: 2095-2163(2025)09-0026-07

Evaluation of Intrusion Detection System based on XGBoost

DAI Ruimin[1,2], John M Acken[2]
(1 School of Communications and Information Engineering, Nanjing University of Posts and Telecommunications, Nanjing 210003, China; 2 Department of Electrical and Computer Engineering, Portland State University, Portland 97201, USA)

Abstract: Intrusion Detection System(IDS) play a vital role in ensuring network security, but traditional approaches often struggle with detecting sophisticated and evolving threats, resulting in low accuracy and high response latency. To enhance the practicality and real-time performance of IDS, this paper proposes a dual-stage detection framework based on XGBoost. The system firstly performs coarse-grained filtering through subsegment classification, followed by fine-grained analysis at the data point level to improve detection precision. Experiments conducted on the ROSPaCe dataset demonstrate that the proposed model achieves an accuracy of 95.99% under the Equal Error Rate (EER) threshold of 0.7, while significantly reducing false positive rates and maintaining acceptable detection latency. Furthermore, zero-day attack simulations reveal that the diversity of training data is critical to the generalization capability of the model. The study highlights the non-linear trade-offs between accuracy, detection latency, and false alarm control, suggesting that all three metrics should be jointly optimized to improve the deployment effectiveness of IDS in scenarios such as IoT and industrial control systems.

Key words: Intrusion Detection System; evaluation metric; attack detection latency; XGBoost; machine learning

0 引 言

近年来,物联网(Internet of Things, IoT)快速发展,深刻改变了人机交互模式。为应对其固有的延迟与计算负载问题,边缘计算成为一种前沿范式。然而,其分布式特性带来了新的网络安全挑战,边缘设备部署分散、资源受限、易受物理接触,成为攻击目标[1-3]。攻击常通过通信或固件漏洞造成信息泄露与服务中断,且设备异构性加剧了安全防护难度。

入侵检测系统(Intrusion Detection System, IDS)通过监控网络状态,识别潜在攻击,最早提出于 1980 年[4]。尽管已有众多成熟产品,但当前 IDS 普遍存在假阳率高的问题[5],误报频繁,促使研究者致力于提升检测率与降低误报率[6-7]。现有 IDS 的

基金项目: 国家自然科学基金(62071249)。
作者简介: 戴睿敏(1999—),男,硕士研究生,主要研究方向:机器学习。
通信作者: John M Acken(1951—),男,博士,教授,博士生导师,主要研究方向:信息安全。Email: acken@pdx.edu。
收稿日期: 2025-03-26

另一个问题是缺乏检测未知攻击的能力。由于网络环境变化迅速，攻击变体和新型攻击不断出现。因此，有必要开发能够检测未知攻击的 IDS。传统 IDS 算法主要包括基于签名检测[8]、基于统计分析[9]、基于专家系统[10]和基于流量分析[11]等方法。这些方法通常面临高计算开销、误报与假阴率问题，以及难以检测高级持续性威胁等挑战。为克服传统 IDS 的局限性，近年来研究者广泛引入机器学习技术，以提升入侵检测系统的智能化水平。机器学习方法通过从大量历史数据中学习正常与异常行为之间的差异，能够有效构建数据驱动的分类模型。Enache 等学者[12]提出了一种基于支持向量机（SVM）分类器的入侵检测方法，并利用粒子群优化与人工蜂群算法对 SVM 参数进行优化，显著提升了模型在 NSL-KDD 数据集上的检测率与准确率。Hyderaba 等学者[13]设计了一种基于 K-近邻分类器的入侵检测模型，并在 ISCX 数据集[14]上验证了其性能，最终在不同划分比例和交叉验证下均达到了 99.96% 的准确率，显示出 K-近邻在分类准确性方面的潜力。

尽管研究人员已做了大量工作来提高针对边缘环境的 IDS，但大多数研究主要集中在优化检测或分类准确率[15-17]。评估 IDS 性能时，准确率、精确率、召回率和 $F1$ 分数等指标通常占主导地位，而同样重要的参数、如检测延迟，却往往被忽视。

本研究采用 XGBoost 探讨入侵检测系统在检测延迟与准确率之间的平衡。通过真实网络流量分析和贝叶斯优化（Optuna）参数调整，本文提出了一种优化策略，可提高检测准确率并降低响应时间，从而提供更高效、可扩展的入侵检测解决方案。

1　基于 XGBoost 的入侵检测系统

1.1　XGBoost 简介

XGBoost（eXtreme Gradient Boosting）是一种高效、可扩展的梯度提升决策树（Gradient Boosting Decision Trees，GBDT）实现，由 Chen 等学者[18]在 2016 年提出。相比传统的 GBDT，XGBoost 采用了优化的计算框架，大幅提升了训练速度，同时增强了模型的泛化能力。其本质是通过迭代地构建多棵回归树，不断优化模型预测结果。在第 t 轮迭代中，模型目标是最小化以下正则化目标函数：

$$L^{(t)} = \sum_{i=1}^{n} l(y_i, \hat{y}_i^{t-1} + f_t(x_i)) + \Omega(f_t) \tag{1}$$

其中，l 表示损失函数；f_t 表示第 t 棵树的结构；$\Omega(f_t)$ 表示正则项，用于控制模型的复杂度，具体为：

$$\Omega(f) = \gamma T + \frac{1}{2}\lambda \sum_{j=1}^{T} \omega_j^2 \tag{2}$$

其中，T 表示树的叶子数；ω_j 表示第 j 个叶子的权重；γ，λ 为正则化参数。XGBoost 使用泰勒展开对损失函数进行二阶近似，从而将目标函数简化为关于每个叶子节点的函数：

$$\tilde{L}^{(t)} = \sum_{j=1}^{T} \left[G_j w_j + \frac{1}{2}(H_j + \lambda)\omega_j^2 \right] + \gamma T \tag{3}$$

其中，$G_j = \sum_{i \in I_j} g_i$，$H_j = \sum_{i \in I_j} h_i$。这里，$g_i$，$h_i$ 分别表示样本 i 的一阶和二阶导数，I_j 表示落在叶子 j 上的样本集合。通过对 w_j 求导并令导数为 0，可得到最优化叶子权重：

$$w_j^* = -\frac{G_j}{H_j + \lambda} \tag{4}$$

1.2　基于段的检测框架

传统的入侵检测方法多采用基于数据点的分类机制，即将每一个数据点作为独立样本进行异常识别。这种方法虽然在建模结构上简洁，便于实现，但也存在显著不足。首先，点级检测极易受到噪声干扰，在时间序列数据中表现为标签抖动频繁，导致误报率偏高。其次，该类方法难以捕捉攻击行为的上下文特征，因为每个点的判断仅基于其局部属性，忽略了攻击演化过程中的阶段性特征。此外，数据点级检测缺乏清晰的行为边界，不利于评估检测延迟与响应时间等时间相关性能指标，限制了其在实际部署中的可解释性与控制力。

为克服上述问题，本文提出一种基于段（segment）的检测框架，并在此基础上构建双阶段检测流程。该框架充分依据现实数据集（如 ROSPaCe）中攻击行为的结构特征，即攻击通常并非突发的单点行为，而是通过阶段性动作（如扫描、渗透、破坏）逐步展开。

段、子段与攻击检测延迟定义如图 1 所示，数据按攻击注入时间可以划分为多个段，每个段内部包含正常/观察期（observe）与攻击期（attack）两个子段，分别表示系统在攻击前后的状态。划分点依据数据集中明确标注的攻击开始时间戳进行界定。

检测系统在段结构基础上引入双阶段检测机制。第一阶段以子段为单位提取统计特征（如均值、标准差等）构成固定维度的输入向量，训练 XGBoost 模型进行快速分类，实现对异常子段的粗粒度筛查。该阶段强调效率优先，目标是快速排除大量正常数据，压缩后续处理范围。由于统计聚合过程可能存在信息损失，第一阶段对短时攻击或弱

特征扰动的敏感性有限,因此需进一步通过第二阶段加以补充。

第二阶段在第一阶段判定为异常的子段内部,对所有数据点进行精细化检测,使用原始数值特征输入模型并输出预测概率,结合设定的分类阈值进行最终分类判断。该阶段显著提升了对细粒度攻击行为的识别能力,能够有效弥补第一阶段可能的误判与漏检,同时提供更精准的攻击响应时刻,为检测延迟等时效指标的计算奠定基础。

基于段的检测结构在保持上下文一致性的前提下,通过粗筛与精判协同设计,实现了检测效率、准确率与延迟可控性的统一优化,相较于传统点级检测方法在实用性与扩展性方面具备显著优势。

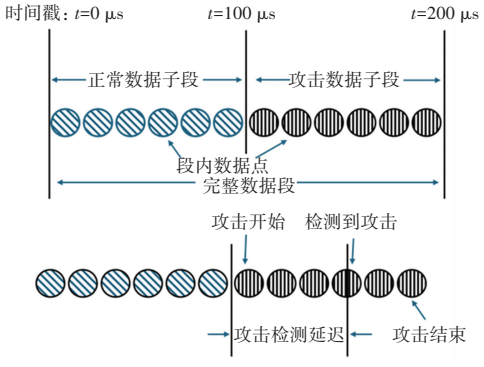

图 1 段、子段与攻击检测延迟定义图

Fig. 1 Definition of segment, subsegment and detection latency

2 数据集与实验方法

本次研究中在配备 i9-12900HX 处理器(16 核)、128 GB RAM 和 RTX 3000 显卡的系统上进行所有实验。为了提高计算效率,使用了 ROSPaCe 数据集的精简版(rospace_reduced,11.9 GB),这个数据集由 Puccetti 等学者[19]在 2024 年发布在 Nature Scientific Data 期刊上。

2.1 ROSPaCe 数据集

ROSPaCe(Robotic Operating System Penetration and Cybersecurity evaluation)专为评估机器人系统和网络物理系统中的安全威胁而构建,数据采集来源涵盖 Linux 操作系统层、网络层与 ROS2 服务层,能够全面反映系统多层级的运行状态与攻击影响[19]。这种跨层结构弥补了传统数据集仅关注网络流量的单一视角,有助于构建更具解释性和泛化能力的检测模型。数据集通过真实系统部署与渗透测试收集而成,避免了传统方法中"模拟正常行为+人工注入攻击"带来的数据分布偏差。ROSPaCe 的数据严格按照"正常行为→攻击行为"顺序排列,形成明显的观察期与攻击期交替结构,可用于精确测量检测延迟,这一关键指标在现有多数数据集中被严重忽视。表 1 列举了 ROSPaCe 数据集涵盖的攻击类型。

表 1 ROSPaCe 数据集中 6 类攻击类型

Table 1 Six types of attacks in the ROSPaCe dataset

序号	攻击类型	简要说明
1	Nmap Discovery	使用 Nmap 工具进行的网络扫描,用于识别系统中的开放端口与主机,属于信息收集阶段的被动攻击
2	Nmap SYN Flood	基于 TCP 协议的 SYN 洪水攻击,向目标发送大量 SYN 请求以耗尽其连接资源,导致拒绝服务(DoS)
3	Metasploit SYN Flood	使用 Metasploit 框架发起的增强版 SYN 洪水攻击,模拟更真实、更复杂的网络攻击流量
4	ROS2 Node Crashing	向 ROS2 节点发送恶意指令或数据包,导致节点崩溃,属于对 ROS2 系统高层逻辑的破坏性攻击
5	ROS2 Reconnaissance	在 ROS2 通信架构下进行的侦察行为,例如监听 DDS 通信或探测节点和话题,目的是收集系统信息
6	ROS2 Reflection Attack	滥用 ROS2 的 publish-subscribe 机制进行反射型攻击,例如造成话题消息风暴,影响系统稳定性或引发拒绝服务

2.2 数据预处理和特征工程

rospace_reduced 数据集[19]包含 3 020 万个数据点和共 60 个特征,见表 2。需要注意的是,时间戳(timestamp)和攻击(attack)不作为模型的特征;首先,对原始特征进行了如下清洗。

(1)恒定特征移除:剔除在所有数据文件中取值不变的特征。

(2)非数值特征剔除:字符串类型数据特征无

法提取其统计特征，因此去除不能直接输入机器学习模型的字符串类型子段（如路径、话题名称等）。

（3）缺失值检查：以 csv0 文件为代表，统计每个特征的缺失率，若超过 1%，则从所有文件中统一移除该特征。

经上述处理后，最终保留了 27 个高质量数值型特征用于后续建模。

表 2　ROSPaCe 数据集中各标签类型的数据分布统计

Table 2　Data distribution statistics for each label type in the ROSPaCe dataset

标签类型	段数量	数据点数量
Normal	1 173	6 645 320
Metasploit SYN Flood	294	14 696 833
Nmap Discovery	362	500 000
Nmap SYN Flood	20	7 993 842
ROS2 Node Crashing	101	5 743
ROS2 Reconnaissance	214	401 494
ROS2 Reflection Attack	182	3 753
总计	2 346	30 246 986

为配合本文提出的两阶段检测策略，在子段分类阶段，需要将每个子段表示为一个固定长度的特征向量。研究采用均值压缩方法：对每个子段内的所有数据点，计算其 27 个数值特征的列均值，从而得到该子段的整体表示。该方法既保留了子段的整体特征趋势，又显著降低了输入维度，提高了分类速度，适用于 XGBoost 等基于树结构的模型。

为保证模型训练的科学性与检测性能评估的可靠性，所有段随机划分为：80%用于模型训练、10%用于验证（超参数调优）、10%用于测试（最终性能评估）。每个子段继承其所属段的划分标签，并将该标签传播到该子段内的所有数据点。因此，如果某个子段被指定为测试集，那么该子段内的所有数据点都将被分配到测试集。这种方法确保了训练集、验证集和测试集的一致性，同时保留数据的时间结构和完整性，避免信息泄漏。

为了验证子段分类的稳健性，在 10 次不同的随机划分下评估了分类性能指标。结果表明，在不同的划分方案下，分类准确率的变化极小（仅 0.01%量级的差异），这表明分类模型在不同的随机划分下具有较强的稳健性。

2.3　实验步骤

为全面评估入侵检测系统的性能，本文采用的多种指标（见表 3），包括准确率、平均检测延迟、假阳率（FPR）、假阴率（FNR）和 $F1$ 分数。综合这些指标可以更全面地衡量入侵检测模型在准确性、稳定性与实时性方面的性能，为不同场景下的部署与优化提供参考依据。

表 3　评估指标及其公式

Table 3　Evaluation metric and their formula

评估指标	公式
准确率	$\frac{TP+TN}{TP+TN+FP+FN}$
平均延迟	$\frac{\sum_{i=1}^{n} T_i - A_i}{n}$（这里，$T$ 表示检测到时间，A_i 表示攻击开始时间）
假阳率（False Positive Rate，FPR）	$\frac{FP}{FP+TN}$
假阴率（False Negative Rate，FNR）	$\frac{FN}{FN+TP}$
$F1$ 分数	$\frac{2TP}{2TP+FP+FN}$

在子段分类任务中，由于数据集仅包含 2 346 个子段，因此参数调优并不必要。然而，为了保证实验的一致性，对数据点分类任务的 XGBoost 模型进行了参数调优，该任务涉及 15 个 CSV 文件中的 30 246 986 个数据点。参数调优使用 Optuna 框架完成，该框架基于贝叶斯优化进行超参数选择。研究中仍需注意的是，reg_alpha、reg_lambda 和 early_stopping_rounds 作为正则化技术被应用于训练过程中，以防止模型过拟合。

数据点分类任务的核心目标是分析攻击检测延迟与分类准确率之间的权衡。Puccetti 等学者[19]通过调整分类阈值研究了假阳率与平均检测延迟之间的关系，本文将在结果分析部分与其研究结果进行对比。然而，本研究更关注准确率与检测延迟之间的关系。

由于 ROSPaCe 数据集中攻击类数据点占比很高，为实现准确率的显著调控，并提升模型在不同检测需求下的适应性，本文采用了二元分类阈值调整策略，通过优化判定标准，有效调节模型输出，在实验和实际部署中实现更合理的性能权衡。

3　实验结果和分析

3.1　子段分类结果

如 2.2 小节所述，测试集由 2 346 个段的 10%

组成,其中包括 117 个攻击子段和 117 个观察子段。分类任务在阈值 0.5 下取得了 95.59%的准确率,假阳率为 0.039 2,假阴率为 0.049。

值得注意的是,由于训练集、验证集和测试集在子段分类和数据点分类任务上的一致性,且 XGBoost 模型参数相同,本次研究预期子段分类的准确率与数据点分类的准确率应保持较高的一致性。后续章节的实验结果验证了这一预期。这一发现进一步支持了子段分类作为快速、粗粒度检测网络攻击的方法,在进行更精细的检测之前提供初步筛查能力。

3.2 数据点分类结果

在数据点分类任务中,测试集并非来自所有 30 246 986 个数据点的 10%,而是根据数据集划分原则进行的。具体而言,子段在划分训练、验证或测试集时,会继承所属段的类别,因此测试集包含 234 个子段中的所有数据点。在不同分类阈值下的分类结果如下。

(1)阈值为 0.5:得到的结果中,准确率为 96.93%;平均攻击检测延迟为 0.86 s;最大攻击检测延迟为 8.59 s;标准差为 1.6 s;假阳率为 0.186 8;假阴率为 0.011 2;*F*1 分数(攻击数据)为 0.98。

(2)阈值为 0.7(等误差率):得到的结果中,准确率为 95.99%;平均攻击检测延迟为 4.72 s;最大攻击检测延迟为 20.7 s;标准差为 4.4 s;假阳率为 0.026 6;假阴率为 0.041 7;*F*1分数(正常数据)为 0.84;*F*1 分数(攻击数据)为 0.98。

结果表明:较低阈值(0.5)减少了检测延迟(0.86 s),但假阳率较高(18.68%);而较高阈值(0.7)降低了假阳率(2.66%),但检测延迟显著增加(平均 4.72 s,最大 20.7 s)。2 种阈值设定下,*F*1 分数变化不大,说明模型在不同条件下均能较稳定地检测攻击,但响应速度和误报控制存在权衡关系。

图 2 展示了平均检测延迟和最大检测延迟随准确率变化的趋势。当从训练集中排除特定攻击类型,而测试集中仍然包含这些攻击时,分类准确率出现波动。准确率与平均检测延迟和最大检测延迟之间的关系是非单调的,而准确率与假阳率(FPR)和假阴率(FNR)之间存在单调关系,见表 4。

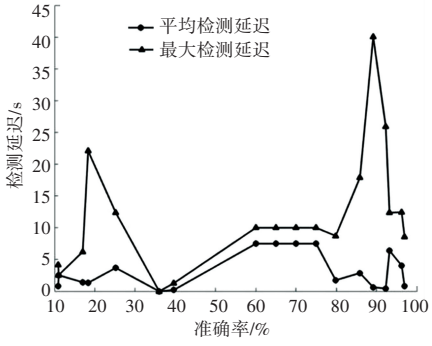

图 2 平均检测延迟与最大检测延迟随准确率变化趋势

Fig. 2 Variation of average and maximum detection latency with accuracy

表 4 不同训练集配置下的分类准确率、检测延迟和假阳率分析

Table 4 Analysis of classification accuracy, detection latency, and false positive rate under different training set configurations

准确率/%	平均检测延迟/s	最大检测延迟/s	标准差/s	假阳率	假阴率	训练集中的 CSV 文件
10.91	0.83	4.20	1.02	0.060 1	0.995 1	1,2,3
11.00	2.55	2.55	0.00	0.014 5	0.999 8	1,2
17.00	1.42	6.23	1.75	0.055 5	0.927 3	2(60%)
18.34	1.34	22.08	4.90	0.064 8	0.911 0	1,2,3,4,5
25.20	3.70	12.38	3.39	0.003 6	0.841 4	0,1,2(20%)
36.00	3.00	8.20	0.00	0.483 6	0.659 6	2
39.62	0.21	1.28	0.30	0.086 9	0.668 6	0,1,2
58.88	2.82	8.73	1.86	0.014 0	0.383 2	3
79.85	1.75	8.73	1.54	0.014 0	0.225 0	3
85.88	1.82	7.19	4.55	0.014 2	0.157 1	3,5,12
89.17	0.64	40.05	3.87	0.873 7	0.012 2	13
92.29	3.02	25.92	5.37	0.607 7	0.061 7	0,4(50%)
93.23	4.60	12.38	4.37	0.397 2	0.026 3	7
96.23	4.07	12.45	4.74	0.223 1	0.014 4	6
96.96	0.80	8.59	1.54	0.184 2	0.011 1	全部

3.3　结果分析

实验结果表明，在阈值 0.5 时模型达到最高准确率 96.93%，而在阈值 0.7（EER）下，模型预测更均衡，但平均检测延迟增加至 4.72 s。FPR 在高阈值下显著降低，FNR 略有上升，显示检测策略在误报控制与响应速度之间存在显著权衡。

准确率与检测延迟之间的关系并非单调：部分准确率较低的模型反而具有更长的检测延迟，说明两者不构成线性因果关系。检测延迟的变化更多受训练数据覆盖范围与分类策略的影响。

模拟零日攻击时，移除关键攻击类型会导致准确率骤降至 10.91%，模型几乎失效，显示当前方法对未知攻击的泛化能力较弱。增加训练数据多样性虽可提升检测效果，但可能带来误报率上升，因此需在准确性、延迟与泛化能力之间平衡优化。本研究模型和文献[20]中已发布的结果对比展示在表 5 中。

表 5　ROSPaCe 数据集的数据点分类性能比较

Table 5　Comparison of data point classification performance on the ROSPaCe dataset

模型	性能评估指标				
	准确率/%	回调率（攻击）	*F*1 分数（攻击）	数据点	测试集
Puccetti 等学者[19]	92.70	0.99	0.95	10.08	总数据量的 30%
本研究模型	95.99	0.96	0.98	30.24	117 个攻击子段，117 个观察子段

4　结束语

本研究提出一种基于 XGBoost 的双阶段入侵检测系统，并在 ROSPaCe 数据集上进行了系统评估。实验结果表明，本文模型在准确率、假阳率与检测延迟之间达成良好平衡，能够有效适配复杂的攻击场景。尤其在不同阈值设定下，模型展现出灵活的性能调节能力，支持实际部署中根据需求选择“更低误报”或“更快响应”的策略。

研究还发现，仅优化准确率并不能显著改善检测延迟，说明响应速度应作为独立关键指标纳入系统设计。此外，训练集覆盖范围对检测性能有显著影响，面对未知攻击类型时模型准确率大幅下降，提示当前 IDS 在泛化能力上仍存短板。

未来可进一步引入深度神经网络、自适应学习等机制，提升模型在变异攻击或未知威胁下的鲁棒性。同时建议在更大规模、真实流量环境下开展测试，以验证模型的通用性与可扩展性。

参考文献

[1] XIAO Yinhao, JIA Yizhen, LIU Chunchi, et al. Edge computing security: State of the art and challenges[J]. Proceedings of the IEEE, 2019, 107(8): 1608-1631.

[2] JIN Xin, KATSIS C, SANG Fan, et al. Edge security: Challenges and issues[J]. arXiv preprint arXiv, 2206.07164, 2022.

[3] ACKEN J M, SEHGAL N K, BANSAL D, et al. Security and Trust Metrics for Edge Computing[C]//Proceedings of 2023 IEEE PES Grid Edge Technologies Conference & Exposition (Grid Edge). Piscataway, NJ: IEEE, 2023: 1-6.

[4] ANDERSON J P. Computer security threat monitoring and surveillance[R]. Washington DC: Anderson Company, 1980.

[5] KHRAISAT A, GONDAL I, VAMPLEW P, et al. Survey of intrusion detection systems: techniques, datasets and challenges[J]. Cybersecurity, 2019, 2(1): 20.

[6] GHARAEE H, HOSSEINVAND H. A new feature selection IDS based on genetic algorithm and SVM[C]//Proceedings of 2016 8th International Symposium on Telecommunications (IST). Piscataway, NJ: IEEE, 2016: 139-144.

[7] TIAN Qiuting. An intrusion detection approach based on improved deep belief network and LightGBM[C]//Proceedings of 2022 6th International Symposium on Computer Science and Intelligent Control (ISCSIC). Piscataway, NJ: IEEE, 2022: 40-44.

[8] JIN Shiyi, CHUNG J G, XU Yinan. Signature-based Intrusion Detection System (IDS) for in-vehicle CAN bus network[C]//Proceedings of 2021 IEEE International Symposium on Circuits and Systems (ISCAS). Piscataway, NJ: IEEE, 2021: 1-5.

[9] WASKITA A A, SUHARTANTO H, PERSADHA P D, et al. A simple statistical analysis approach for Intrusion Detection System[C]//Proceedings of 2013 IEEE Conference on Systems, Process& Control (ICSPC). Piscataway, NJ: IEEE, 2013: 193-197.

[10] BAUER D S, EICHELMAN F R, HERRERA R M, et al. Intrusion detection: an application of expert systems to computer security[C]//Proceedings of International Carnahan Conference on Security Technology. Piscataway, NJ: IEEE, 1989: 97-100.

[11] AJAEIYA G A, ADALIAN N, ELHAJJ I H, et al. Flow-based Intrusion Detection System for SDN[C]//Proceedings of 2017 IEEE Symposium on Computers and Communications (ISCC). Piscataway, NJ: IEEE, 2017: 787-793.

[12] ENACHE A C, PATRICIU V V. Intrusions detection based on Support Vector Machine optimized with swarm intelligence[C]//Proceedings of 2014 9th IEEE International Symposium on Applied Computational Intelligence and Informatics (SACI). Piscataway, NJ: IEEE, 2014: 153-158.

[13] HYDERABA D, NIKHITHA M, JABBAR D M A, et al. Knearest neighbor based model for intrusion detection system[J]. International Journal of Recent Technology and Engineering, 2019, 8(2): 2258-2262.

[14] SOHEILY-KHAH S, MARTEAU P F, BECHET N. Intrusion detection in network systems through hybrid supervised and unsupervised machine learning process: A case study on the ISCX dataset[C]//Proceedings of 2018 1st International Conference on Data Intelligence and Security (ICDIS). Piscataway, NJ: IEEE, 2018: 219-226.

[15] MAROUANE H, DANDOUSH A, AMOUR L, et al. Performance evaluation of machine learning-based misbehavior detection systems in VANETs: A comprehensive study[C]// Proceedings of 2023 International Symposium on Networks, Computers and Communications (ISNCC). Piscataway, NJ: IEEE, 2023: 1-6.

[16] TIAN Zhihong, LUO Chaochao, QIU Jing, et al. A distributed deep learning system for Web attack detection on edge devices[J]. IEEE Transactions on Industrial Informatics, 2020, 16(3): 1963-1971.

[17] ALGARNI A, ACARER T, AHMAD Z. Anedge computing-based preventive framework with machine learning- integration for anomaly detection and risk management in maritime wireless communications[J]. IEEE Access, 2024, 12: 53646-53663.

[18] CHEN Tianqi, GUESTRIN C. XGBoost: A scalable tree boosting system[C]//Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. New York: ACM, 2016: 785-794.

[19] PUCCETTI T, NARDI S, CINQUILLI C, et al. ROSPaCe: Intrusion detection dataset for a ROS2-based cyber-physical system and IoT networks[J]. Scientific Data, 2024, 11(1): 481.

[20] PUCCETTI T, CECCARELLI A. Detection latencies of anomaly detectors: An overlooked perspective? [C]//Proceedings of 2024 IEEE 35th International Symposium on Software Reliability Engineering (ISSRE). Piscataway, NJ: IEEE, 2024: 37-48.