基于机器学习的网络信息安全态势感知系统设计

曾倩倩[1]
ZENG Qianqian

摘 要 面向多源复杂网络环境中的安全威胁识别与响应问题，文章设计了一种基于机器学习的网络信息安全态势感知系统。系统以感知层、分析层与决策层为核心架构，构建了数据采集与预处理、威胁分析与评估、态势展示与联动响应三大功能模块。采用 XGBoost 分类模型与 Isolation Forest 异常检测模型对多维结构化特征向量进行攻击行为判别与风险评分，形成面向动态攻击链的智能识别机制。在仿真网络环境中完成原型系统部署，测试了系统在吞吐能力、模型推理效率、识别准确率与响应延迟等方面的关键指标。实验结果显示，系统可实现 36 800 条 /s 的日志处理能力，攻击识别准确率达 96.4%，联动响应时延低于 1 s，验证了所提方法的可行性与工程适用性。研究结果为构建具备智能决策能力的网络安全防御体系提供了技术支撑与实证依据。

关键词 网络安全；态势感知；机器学习；威胁检测；系统架构

doi：10.3969/j.issn.1672-9528.2025.11.031

0 引言

在信息技术与大数据技术深入发展的背景下，当前大数据网络环境呈现出复杂的变化趋势，数据规模巨大且增长迅速[1]，网络基础设施面临的安全风险呈现出攻击频次高、隐蔽性强、演化速度快等特点。针对态势模糊、攻击路径不透明、威胁响应延迟等突出问题，传统依赖规则匹配和特征库更新的防护方式已难以满足动态环境下的实时识别与快速响应需求。网络安全态势感知作为主动防御的重要技术方向，强调对网络状态、攻击行为和风险演变过程的持续感知与智能决策，正逐步从以展示为主的静态感知系统向融合智能推理与自动联动的综合决策平台演化。

机器学习技术凭借其在数据建模、模式分类和异常检测等方面的优势，已成为提升态势感知系统智能化水平的关键工具。通过构建学习模型挖掘多源异构数据中的潜在攻击特征，可有效实现对未知威胁的识别与行为归因，为安全防护系统提供动态适应与策略调整能力[2]。基于此，构建融合多源数据采集、特征构造、模型推理与策略联动的一体化态势感知系统，成为解决当前网络防御能力瓶颈的关键路径。为此，本文设计并实现了一种基于机器学习的网络信息安全态势感知系统，从系统结构设计、关键功能模块实现及工程性能验证等方面展开研究，探索高动态、高复杂网络环境下智能防御体系的构建方法与实用路径。

1 网络安全态势感知技术

网络安全态势感知技术是一种融合多源异构数据采集、威胁行为建模与动态风险评估的综合性技术体系，旨在实现对网络安全状态的实时认知、趋势预测与响应决策。该技术通过对网络流量、主机日志、用户行为及外部威胁情报等数据进行深度感知与语义解析，构建完整的态势特征向量，并结合威胁传播模型与攻击路径推理机制，对潜在风险进行定量化评估。随着攻击手段的隐蔽化与攻击链的复杂化，传统静态规则或特征匹配方法在时效性和泛化能力方面日益受限，而态势感知则强调动态建模与全局关联分析，支持跨层级、跨域的数据融合与协同判断，具备更高的场景适应性与自动化能力[3]。在实际部署中，态势感知系统通常由感知层、分析层与决策层构成，各层通过信息流闭环联动，形成从数据采集到威胁响应的完整技术链条，为构建主动防御型网络安全体系提供技术支撑。

2 机器学习在态势感知中的应用

2.1 数据特征提取与表示

态势感知系统首先依赖海量数据输入，包括网络流量数据（NetFlow、PCAP）、主机日志（Syslog、Windows 日志）、系统调用序列、IDS/IPS 报警信息、用户操作行为、设备访

1. 山西警察学院 山西太原 030401

[基金项目] 山西警察学院 2025 年度实战领题项目“AI 驱动下多模态社交媒体舆情监测与防控的研究”；山西省社科联 2025 年度重点课题研究项目“以战斗力倍增为核心的公安大练兵新体系构建及实践路径研究”（SSKLZDKT2025297）

问记录等。机器学习模型要求将这些非结构化或半结构化的原始数据转化为可训练的特征向量，因此特征提取过程是建模的第一步[4]。通过基于内容的协议解析（HTTP 字段抽取、DNS 查询域名映射）、时间窗统计（每 10 s 的连接次数、平均包长）以及行为序列编码（系统调用 N-gram 模型），可构建高维度的样本表示。为了消除特征冗余、提升计算效率，进一步引入信息增益、互信息、卡方检验等特征选择方法进行降维。例如，信息增益的计算公式为：

$$\mathrm{IG}(D,A)=H(D)-\sum_{v\in \mathrm{Values}(A)}\frac{|D_v|}{|D|}H(D_v) \tag{1}$$

式中：D 表示训练集；A 表示待选特征；D_v 表示特征 A 取值为 v 时的样本子集；$H(\cdot)$ 表示集合的信息熵。

该公式衡量特征 A 对训练集 D 的划分贡献，有助于筛选对攻击分类最有效的维度。在具体实现中，特征提取与选择流程通常集成于数据预处理管道中，结合滑动时间窗口、序列向量展开与多维聚合算子构造特征张量输入。对于协议字段嵌套复杂或存在异常编码的数据包，还需采用正则表达式与字符嵌入技术进行结构修复与格式统一，确保输入向量的结构一致性与语义完整性。此外，针对低频特征与缺失特征引入零填充与权重修正策略，以避免模型在训练过程中产生偏置。

2.2 模型选择与适配性分析

完成特征建模后，系统需选择合适的学习模型进行训练与预测。具体模型选择依据目标任务的类别（有监督 / 无监督）、数据标签完整性、攻击行为复杂度等因素动态决策。在已知标签的场景下，常使用支持向量机（SVM）、随机森林（RF）、XGBoost 等模型对攻击行为进行分类，特别适用于 DDoS、扫描、暴力破解等高频攻击的识别[5]。当系统面对未知攻击或标签缺失时，可使用 K-means、DBSCAN 等聚类算法实现无监督识别，或结合孤立森林、自编码器对异常数据进行检测与评分。例如，在识别慢速端口扫描或隐蔽 APT 攻击时，自编码器通过构造重建误差模型，输出潜在风险评分，实现非显式行为的判别。

此外，对于具备时序相关性的安全事件（如多阶段攻击链），机器学习还嵌入长短期记忆网络（LSTM）、Transformer 结构等时间序列建模技术，实现对攻击行为的演进预测。模型训练通常采用批量梯度下降（SGD）、Adam 等优化器，在 GPU 或边缘设备上进行离线训练或在线更新。

2.3 异常检测与威胁识别

机器学习的核心任务之一是提升系统对未知威胁的识别能力。在态势感知场景下，异常检测模块采用无监督或半监督学习构建“正常行为剖面”，并实时对偏离该剖面的事件进行评分与标记。例如孤立森林算法通过构建二叉随机树对样本进行划分，其异常评分公式为：

$$s(x,n)=2^{-\frac{E(h(x))}{c(n)}} \tag{2}$$

式中：$E(h(x))$ 为样本 x 的平均隔离路径长度；$c(n)$ 为样本规模修正因子，评分越高表明越可能为异常。该算法在高维场景下具备良好的扩展性与计算效率，适用于大规模日志与流量检测。

在实际部署中，态势感知系统将上述机器学习模型封装为“威胁检测引擎”，集成在分析层中。系统通过多模型组合（如集成学习框架），并融合规则引擎与知识图谱，实现对复杂攻击链的行为溯源、路径推演与风险分类。检测结果进一步传递至可视化层与响应控制模块，驱动联动封堵、黑名单写入与告警推送等动作，构成闭环控制的智能安全态势感知流程。

3 系统架构与关键模块设计

3.1 态势感知系统总体架构

本系统采用分层式架构设计，划分为感知层、分析层与决策层三大功能域，构成数据驱动、模型驱动与响应驱动相结合的闭环式网络安全态势感知体系。感知层负责多源数据的实时采集与标准化处理，包括网络流量、主机日志和用户行为数据，通过特征提取与归一化生成统一输入格式；分析层作为计算核心，承担机器学习模型的加载与推理任务，实现攻击识别、异常检测与风险评分[6]；决策层依据模型输出进行态势可视化、告警生成与策略联动，实现从威胁发现到响应执行的快速闭环。该架构逻辑清晰、功能解耦，具备良好的扩展性与部署适应性。图 1 为系统的总体架构图。

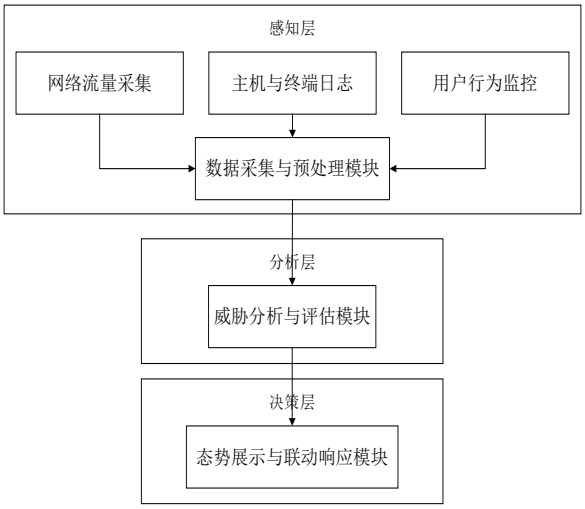

图 1 系统架构图

3.2 数据采集与预处理模块

数据采集与预处理模块位于系统感知层，主要负责多源数据的接入、转换与结构化处理，是保障后续分析模型效率与准确度的关键环节。系统支持对网络流量（NetFlow）、主机日志（Syslog/Windows 事件）、终端行为记录（URL 轨迹、访问频率等）进行并行接入。采集方式采用多线程异步拉取与消息队列缓冲机制，结合 Kafka 与 Flume 实现高吞吐低丢包的数据引入。

预处理流程包括特征重构、噪声剔除、时间对齐与维度归一。特征向量构建过程中，采用滑动窗口序列建模方法提取上下文相关特征，生成统一格式的样本矩阵。行为数据 $\boldsymbol{X} \in \mathbf{R}^{n \times m}$ 在窗口宽度 w 下，构建特征 [7] 表示为：

$$\boldsymbol{x}_t = \phi([\boldsymbol{x}_{t-w+1}, \boldsymbol{x}_{t-w+2}, ..., \boldsymbol{x}_t]) \tag{3}$$

式中：$\phi(\cdot)$ 表示特征映射函数（如平均值、变化率、突变系数等）；$\boldsymbol{x}_t \in \mathbf{R}^m$ 表示对应时窗下生成的高维特征向量。该处理方式保留了攻击行为的时序动态特性，适用于后续模型进行序列建模与行为关联分析。

最终生成的特征集缓存入内存数据库（如 Redis 或 TimescaleDB）并按时间戳索引供模型调用，实现标准化、高可用的输入数据流。

3.3 威胁分析与评估模块

威胁分析与评估模块作为系统的核心计算单元，部署于分析层，主要完成对结构化数据的分类、异常检测与风险等级判定。该模块以机器学习模型为计算基础，支持多种算法结构的并行加载，包括监督学习模型（如 XGBoost、SVM）与无监督模型（如 Isolation Forest、自编码器）等。模块运行流程分为模型加载、推理执行与风险量化三阶段 [8]。

模型加载阶段，系统根据攻击类型配置加载对应模型权重文件与参数配置，并使用 ONNX 或 TensorRT 加速推理引擎提升执行效率。推理执行阶段，输入特征向量经过模型计算后输出标签（如“扫描”“拒绝服务”）或风险分数（[0, 1] 区间），以 XGBoost 为例，其优化目标函数为：

$$L(\theta) = \sum_{i=1}^{n} l(y_i, \hat{y}_i) + \sum_{k=1}^{K} \left(\gamma T_k + \frac{1}{2} \lambda \sum_{j=1}^{T_k} \boldsymbol{w}_j^2 \right) \tag{4}$$

式中：$L(\theta)$ 为整体损失函数，表示模型的优化目标；n 为训练样本总数；$l(y_i, \hat{y}_i)$为样本 i 的损失函数（如对数损失、平方误差），表示真实标签 y_i 与预测值$\hat{y}_i$之间的差异；K 为树的总棵数；γ 为控制每棵树复杂度的正则化系数；T_k 为第 k 棵树的叶节点数量；λ 为叶子权重的 L_2 正则系数；w_j 为第 j 个叶子节点的权重。

该目标函数由两部分组成：前项用于最小化预测误差，后项作为复杂度惩罚项防止过拟合。训练过程通过梯度提升法迭代优化。

风险评分阶段，系统根据输出结果匹配设定阈值，划分威胁等级，常用三段式划分为：

$$\text{Level}(x) = \begin{cases} \text{高危}, & s(x) \geq 0.85 \\ \text{中危}, & 0.6 \leq s(x) < 0.85 \\ \text{正常}, & s(x) < 0.6 \end{cases} \tag{5}$$

模块还具备模型动态切换与增量学习接口，支持新型攻击行为下的模型热更新能力。

3.4 态势展示与联动响应模块

态势展示与联动响应模块位于系统决策层，负责将模型输出结果实时转化为可视化态势信息，并根据威胁等级执行预定义响应动作 [9]。可视化端采用 WebSocket 机制与前端仪表组件联动，展示主机风险热力图、攻击趋势图及告警日志流。后台响应引擎内置规则驱动逻辑，依据模型输出的风险评分触发联动操作，控制流程如下：

```
def respond(ip, score):
    if score >= 0.85:
        exec(f"iptables -A INPUT -s {ip} -j DROP")
        notify(ip, "HIGH")
    elif score >= 0.6:
        notify(ip, "MEDIUM")


def notify(ip, level):
    msg = f"{level} risk detected: {ip}"
    log(msg)
    send_alert(msg)
```

系统通过接口模块调用防火墙策略写入、发送邮件 / 短信、SOC 平台对接等服务，实现联动闭环。整个响应流程与模型输出实时绑定，事件闭环平均响应时间控制在 1 s 以内，有效提升了系统的威胁处理自动化水平与实战适应性。

4 系统实现与实验验证

4.1 实验平台与算法部署

实验平台搭建于仿真网络环境中，系统分为数据采集节点、模型分析节点与响应控制终端三部分。数据采集服务器选用 DELL PowerEdge R540，配置双路 Intel Xeon Silver 4210 处理器、64 GB 内存与 4 TB 存储，运行 CentOS 7.9 系统，并部署 Flume 与 Kafka 实现日志与流量数据的并发接入与缓存；模型分析服务器采用 NVIDIA DGX Station，搭载 Intel Xeon E5-2698 v4 处理器、256 GB 内存和 4×Tesla V100 GPU，部署 PyTorch 2.0、ONNX Runtime 与 TensorRT 环境，负责机器学习模型的加载与推理执行；联动响应终端包括 H3C SecPath

F1000 防火墙与安恒网神 SOC 5.0 平台，用于策略自动下发与告警联动。模型部署方面，系统采用 XGBoost 实现威胁分类与风险评分，异常检测模块使用 Isolation Forest，特征维度为 128，训练样本规模为 30 万条，推理引擎通过异步调度机制分别在 CPU 与 GPU 上并行执行，并由 Redis 队列与 Nginx 代理协调数据流，实现高并发、低延迟的数据处理与策略反馈 [10]。

4.2 性能评估与结果分析

为验证系统在实际网络场景下的处理能力与检测精度，构建覆盖扫描攻击、DDoS、暴力破解与正常流量的混合测试数据集，并从处理效率、分类性能、响应延迟等指标对系统进行量化测试。表 1 为实验数据汇总。

表 1 系统关键性能测试结果

指标类别	测试指标	标准值	测试值
数据处理能力	最大日志处理吞吐量 /(条 ·s^{-1})	≥30 000	36 800
	最大流量采集速率 /(Gbit·s^{-1})	≥5.0	6.2
模型推理性能	单条样本平均推理耗时 /ms	≤15	11.6
	GPU 批量推理平均吞吐率 /(条 ·s^{-1})	≥5 000	5 800
检测性能	攻击识别准确率 /%	≥95.0	96.4
	高危事件识别召回率 /%	≥90.0	92.7
	虚警率 /%	≤5.0	3.1
响应效率	风险评估结果可视化延迟 /ms	≤500	390
	联动策略响应平均延迟 /ms	≤1 000	710

表 1 中数据显示，系统在并发场景下的最大日志处理能力达到 36 800 条 /s，流量接入速率超过 6 Gbit/s，均优于行业中型态势感知系统的接入处理标准。推理模块中，单样本平均响应时间为 11.6 ms，批量 GPU 推理吞吐率为 5 800 条 /s，满足高频事件处理要求。检测性能方面，攻击识别准确率达到 96.4%，高危事件召回率为 92.7%，虚警率控制在 3.1%，验证了模型在真实场景下的鲁棒性与精度表现。响应模块平均策略下发延迟为 710 ms，态势展示延迟为 390 ms，均显著优于常规系统 1 s 级延迟控制阈值，说明该系统具备高实时性与联动执行能力，适用于安全运营环境下的实战部署需求。

5 结论

本文设计并实现了一种基于机器学习的网络信息安全态势感知系统，构建了以数据采集与预处理、威胁分析与评估、态势展示与联动响应为核心的三层架构体系。系统融合多源异构数据处理与智能建模技术，采用 XGBoost 与 Isolation Forest 等算法实现对攻击行为的高效识别与风险量化，构建了闭环式的智能防御机制。通过在实际网络环境下的部署与测试，验证了系统在处理能力、检测准确性与响应时效性等方面均达到甚至优于行业标准，具备良好的工程适配性与实用价值。研究结果表明，将机器学习方法嵌入态势感知系统架构中，能够有效提升网络安全事件的识别能力与联动响应效率，为构建智能化、自动化的网络安全保障体系提供了技术支撑与实验依据。

参考文献：

[1] 朱光剑 , 李阳冬 , 钱平 . 基于智能化网络模型的信息安全态势感知算法 [J]. 信息技术 , 2025(5):135-139.

[2] 覃珏斐 , 张寅虎 . 基于大数据技术的网络安全态势感知平台设计 [J]. 电子技术 , 2025,54(3):80-82.

[3] 王艺多 . 智慧城市建设中网络信息安全态势感知优化及其应用研究 [J]. 无线互联科技 , 2024,21(17):66-68.

[4] 王鑫 , 李坤 , 邱德熠 , 等 . 企业信息网络安全管理系统设计与实现 [J]. 微型电脑应用 , 2024, 40(8):150-154.

[5] 王丹 , 王大秋 , 顾俊杰 , 等 . 数字化驱动的核设施工控安全态势感知应用研究 [J]. 核标准计量与质量 ,2024(1):70-74.

[6] 林心怡 . 贝叶斯网络的供电企业通信系统信息安全态势感知方法 [J]. 长江信息通信 ,2023,36(12):135-137.

[7] 莫永华 , 陈昱希 , 何淼 . 企业网络安全态势感知系统设计与实现 [J]. 网络空间安全 , 2023,14(5):55-59.

[8] 王伟昊 . 基于网络安全的医院信息安全设计改造与分析 [J]. 信息与电脑 (理论版), 2022,34(23):225-227.

[9] 安锦程 , 王弢 , 崔君荣 , 等 . 基于工控安全标准的城市轨道交通平台建设 [J]. 信息技术与标准化 , 2022(5):96-99.

[10] 赵渊 , 严磊 , 聂凌云 , 等 . 网络安全监测标准应用实践 [J]. 信息技术与标准化 , 2022(5):125-128.

【作者简介】

曾倩倩（1984—），女， 山西太原人，硕士，讲师，研究方向：大数据分析、网络安全。

（收稿日期：2025-06-04 修回日期：2025-11-14）