# 基于人工智能的通信行业网络安全新运营体系

## New Cybersecurity Operation System for Communication Industries Based on Artificial Intelligence

滕开清，曾哲凌，胡芳燕（中国联通福建分公司，福建 福州 350001）
Teng Kaiqing，Zeng Zheling，Hu Fangyan（China Unicom Fujian Branch，Fuzhou 350001，China）

**摘　要：**
在5G/云网融合背景下，通信运营商面临异构日志治理难、告警处置低效、AI攻击防御滞后三重挑战。提出基于人工智能的新型安全运营体系，该体系通过XDR架构与图神经网络（GNN）实现核心网、基站、终端日志的跨设备智能聚合；结合强化学习构建“AI初筛—专家精判”分级机制；沉淀SOAR自动化处置剧本。实测结果表明，AI模型将原始告警日均处理量从11 041条压降至18条，安全事件调查耗时从12 h压缩至15 min，运营效率提升48倍。

**关键词：**
网络安全；大模型；智能聚合；强化学习；自动化处置
**doi：**10.12045/j.issn.1007-3043.2025.11.009
**文章编号：**1007-3043（2025）11-0046-06
**中图分类号：**TN915
**文献标识码：**A
**开放科学（资源服务）标识码（OSID）：**

**Abstract：**
Under the background of 5G and cloud-network convergence，telecom operators face three major challenges：heterogeneous log governance difficulties，inefficient alarm handling，and delayed AI-based attack defense. It proposes a novel AI-driven security operations framework，which utilizes XDR architecture and Graph Neural Networks （GNN） to intelligently aggregate logs from core networks，base stations，and terminals，and combined with reinforcement learning，a “AI preliminary screening —expert refinement” hierarchical mechanism is constructed，SOAR playbooks for automated incident handling are codified. Test results demonstrate that AI models reduce daily raw alerts from 11 041 to 18，compress security incident investigation time from 12 hours to 15 minutes，and improve operational efficiency by 48 times.

**Keywords：**
Cybersecurity；Large language models；Intelligent aggregation；Reinforcement learning；Automated response

**引用格式：**滕开清，曾哲凌，胡芳燕. 基于人工智能的通信行业网络安全新运营体系[J]. 邮电设计技术，2025（11）：46-51.

## 1 概述

随着5G/云网融合的加速推进，运营商网络呈现全域连接、服务泛在化的特征。网络安全对运营商提出了三大挑战：一是数据治理困境，异构设备日均产生TB级日志，跨系统关联分析缺失导致有效威胁识别率难以保障；二是响应效率瓶颈，人工处置速度（200条/h）与攻防时效要求（分钟级响应）存在量级差距；三是新型威胁防御滞后，AI驱动的自适应攻击（如APT攻击、高级混淆攻击）使传统规则库漏报率持续提升。

在此背景下，国家出台相关法律法规对网络安全体系做出要求，其中《网络安全法》第三十一条、《数据安全法》第二十七条明确要求建立“主动防御、智能协同”的安全体系，因此运营商行业面临着合规压力，技术升级的需求迫切。为满足上述需求，行业内通过生成式人工智能的三大核心技术路径重构安全运营范式，为网络安全领域升级迭代提供了新路径。

多维数据融合技术：基于Transformer架构的日志解析模型，可自动识别多种异构日志格式[1]，通过注意力机制构建跨设备事件关联图谱，将威胁检出准确率

收稿日期：2025-09-15

从传统规则引擎的68%提升至92%[2]。

动态风险评估体系：利用强化学习算法实时分析攻击链上下文，自动生成风险等级评分[3]，并通过可解释性AI技术输出攻击意图分析报告，使安全人员决策效率提升80%[4]。

自动化响应闭环：结合自动化编排响应（SOAR）平台，生成式AI可根据预设剧本自动执行阻断攻击IP、隔离感染终端、触发漏洞修复工单等操作，将平均响应时间（MTTR）从人工处理的720 min压缩至15 min以内，实现从“被动处置”到“主动免疫”的能力跃迁[5]。

本研究立足运营商行业网络安全场景，致力于解决异构系统日志关联分析低效、安全运营人力成本高、新型攻击防御乏力3个方面的核心痛点。通过多源数据智能聚合的XDR技术架构实现告警聚合降噪，沉淀SOAR自动化处置剧本以压降安全防护应急响应时间，使用强化学习构建未知威胁检测能力。研究成果不仅为运营商行业构建主动防御体系提供技术路径，也通过探索专属告警调优，实现了安全运营效率提升。

## 2 人工智能赋能网络安全运营体系

### 2.1 技术架构

人工智能赋能安全运营体系架构如图1所示。针对网络安全攻防对抗与安全管理运营的特点，研究生成式大模型上层应用的构建方法，重点构建平台层、模型层、服务层和应用层的网络安全大模型智能研判决策原型系统示范应用。

a）在平台层，需要建立一个完整的大模型管理平台，包括模型训练、评估、部署、监控等功能，支持不同类型的大模型应用场景。

b）在模型层，基于不同的应用场景和数据特点，设计和训练高效的大模型。上述大模型不仅能够理解和生成自然语言文本，还能够进行安全漏洞检测、恶意代码分析、攻击行为识别等任务。

c）在服务层，示范应用需要提供全方位的大模型安全服务，包括漏洞扫描、威胁情报分析、攻击预警、事件处置等。

d）在应用层，需要将大模型应用于网络安全的不同领域，如网络安全事件管理、安全审计、安全培训等。

通过研究和应用网络安全大模型的系统架构，能够更好地应对网络安全的挑战，提高网络系统的安全性和可靠性。

安全模型采用MoE多专家架构，利用DeepSeek、ChatGLM、LLaMa、Qwen、BertMistral等多个开源基座，分别针对不同子任务进行预训练和微调（比如Web流量检测、自然语言对话、场景识别和任务规划、告警关联定性等）。其中DeepSeek-R1-32B蒸馏模型负责自然语言对话和部分场景识别规划，其他多个模型共同组成垂直领域模型。为了降低模型幻觉，达到模型输出可控和能力提升的目的，采用下面几种方法：一是持续给模型微调更多高质量数据；二是结合RAG等手

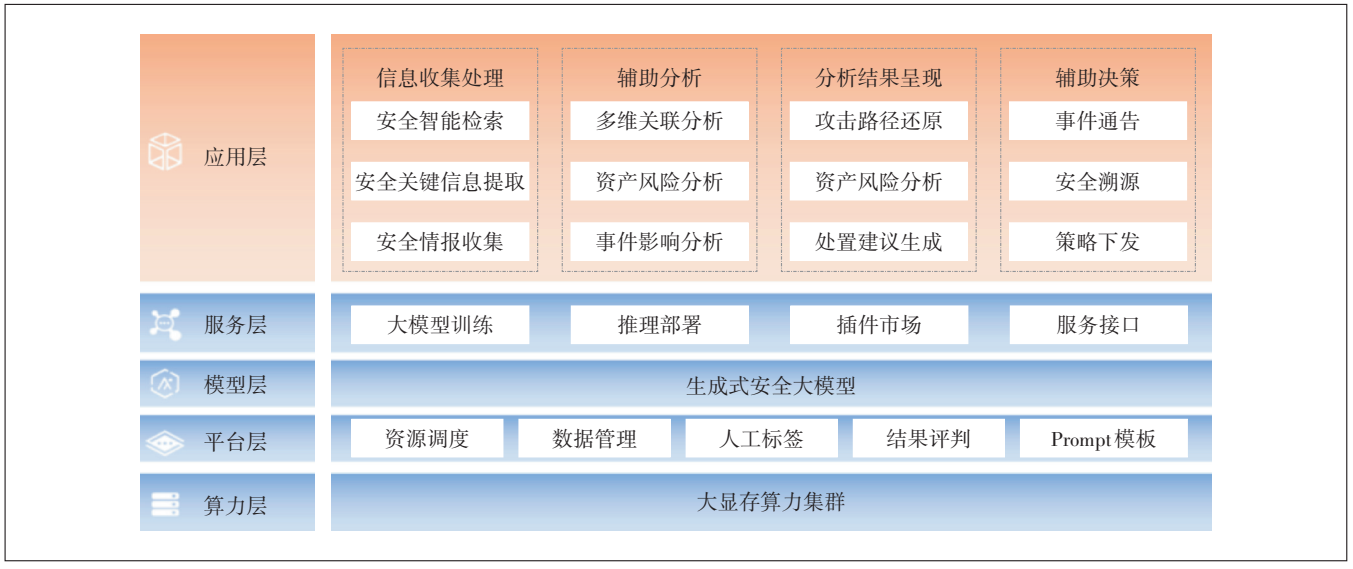


图1　人工智能赋能安全运营体系架构

段丰富模型知识，控制输出效果；三是增加参数量较小的专家模型，突破某些领域的认知瓶颈，这种方式可以持续降低模型幻觉和挖掘模型能力，可控且迭代快速，可以做到一周更新1~2个场景识别模型[6]。

## 2.2 核心技术能力

### 2.2.1 多源数据智能聚合

依托可扩展检测响应平台（XDR）构建全域数据采集层，实现网络侧与终端侧等异构安全设备日志的全量接入与标准化解析。通过图神经网络（GNN）构建跨设备事件关联图谱，基于时序特征匹配、资产脆弱性关联等算法，将孤立告警转化为场景化安全事件，以此解决“数据孤岛”问题[7]。

XDR告警智能聚合技术实现了按照安全语义对告警进行融合，结合当前网络拓扑、资产、漏洞、补丁、弱配、白业务访问关系等，引入了几百种的细粒度告警融合策略，其中部分策略如表1所示。

**表1** 告警融合策略示例

| 类型 | 融合策略 | |
|---|---|---|
| 网络侧告警归并技术 | 多对一 | 多个源IP爆破我方一个资产，这里的主要关注对象是我方受攻击资产，所以以资产为视角进行聚合 |
| | 一对多 | 内网某个资产沦陷后横向扫描多个内网资产，这里的主要关注对象是发起扫描的资产，因为它大概率已沦陷 |
| | 一对一 | 黑客持续攻击我方一个资产，过程中可能利用了多种手法，执行了不同的恶意命令 |
| 终端侧告警归并技术 | 传统病毒查杀 | 只有病毒查杀告警，有病毒路径、病毒名称、病毒文件Hash几个关键因子。类似于网络侧告警的聚合逻辑，分为Hash唯一和病毒类别唯一2种情况 |
| | 高级威胁类 | 高级威胁类告警是以“进程树”的形式呈现给用户的，具备一定时间、空间上的唯一性。所以进程树的溯源、生长过程本身就是降噪的一部分，将环境中离散的行为信息关联为一棵树，让离散行为具备更高层次的语义 |

在融合过程中，XDR告警智能聚合技术通过因果举证、弹性聚合，突出高危攻击、多源告警消减与融合能力。它不仅具备对告警进行payload相似度、单个攻击源对多个目标发起的相似攻击、多个攻击源对单个目标发起的相似攻击等维度的深度聚合归并能力[8]，而且针对同一个攻击行为，将不同安全设备产生的多条安全告警关联融合成同一条告警；还可将同一个安全事件的不同阶段分散在端和网的告警聚合到同一条事件内。融合后的安全告警举证页面可以展示不同数据来源的举证字段内容[9]。

同时，XDR告警智能聚合技术引入“大模型+小模型”协同架构，大模型负责自然语言告警的上下文关联分析，小模型专注二进制恶意样本精准分类，辅以智能体工作流实现告警聚合。大小模型级联分层架构如图2所示。

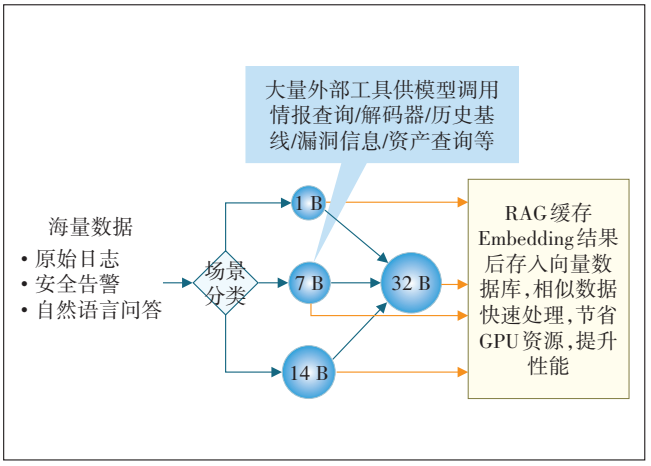


**图2** 大小模型级联分层架构

大小模型级联的核心是解决性能与效果平衡问题。大模型采用一问一答的方式，速度慢，如DeepSeek回答简单问题要数秒到十几秒的时长，这种吞吐性能无法满足实时检测研判的需求。因此，考虑构建一个级联分层架构，如图2所示，采用大小模型混合部署的方式，前端使用参数量较小的模型处理简单的任务，速度更快；复杂任务以及综合结论由参数量大的模型来完成，这样基本实现了针对实时流量、邮件和告警的研判。比如针对流量检测的大模型，经实际测试，通过8卡4090可支持10G流量的检测[10]。

最终实现将全网海量日志聚合为少而精准的安全告警，安全日志数量消减比例达到95%以上。同时，可进一步将安全告警聚合为真实准确的安全事件，实现从上百万条的安全日志聚合到数十条的安全事件。

### 2.2.2 SOAR自动化处置

SOAR是针对网络中产生的安全事件触发剧本执行的技术，共有2种执行方式，一种是“手动执行”方式，即填写相关信息后触发剧本执行；另外一种是自动触发方式，当有事件触发剧本执行后，系统会自动进行剧本执行。执行过程中会进行剧本规则匹配，例如事件类型、规则ID等，当匹配到对应剧本后，该剧本会进入任务队列，任务管理模块定期轮询任务队列中的待执行任务，并将待执行任务同步到应用管理模块，应用管理模块根据剧本中具体的动作以及应用资

源，通过对应的调用方式（如API、SSH、命令行等），调用具体的实例执行动作。最终，由任务管理模块将动作执行的情况同步到事件列表中，完成事件状态的更新[11]。

本次研究利用AI智能体的自主规划和工具调用能力，结合实战化的网络安全知识图谱，总共构建了4个SOAR剧本，实现了攻击事件的自动化研判与处置。通过人机对话接口，可实现告警查询、策略下发等操作，将安全运营流程从“多系统切换操作”简化为“自然语言指令交互”，减少人工操作步骤。此外，AI模型基于攻击结果自动匹配预设响应剧本，实现了从“人工处置“到“机器自主响应“的范式转变，成功构建了“检测—研判—响应“三级联动机制[12]。实测数据显示，在10万条日志事件中，AI模型可自动筛选出300条真实威胁事件，其中286条由平台自主响应，仅14条需专家介入，关键场景响应时间压缩至5 min以内。专家介入的事件还会快速沉淀为规则，持续优化自动化响应能力，全面提升实战攻防的效率和韧性。

以互联网攻击业务系统自动联动封锁效果为例，该剧本实现了互联网业务系统对攻击的自动封锁，它通过处理网络中的源IP地址（SRCIP），判断其是否为攻击源，并采取相应的处理措施。当发现某个IP地址可能存在异常行为时，系统先通过GPT进行研判，判断其是否为误报。如果不是误报，则进一步判断该IP是否在白名单中。白名单IP通常是指经过授权或可信的IP地址，对白名单IP，系统直接发送群聊消息提示相关人员即可。对于非白名单IP，系统会判断其是否为攻击源IP。如果是，则进一步判断是否为IPTV地址。最后，根据该IP是否为基站IP，决定是临时封禁还是永久封禁，并发送群聊消息通知相关人员。该剧本可应用于实时监控和处理网络中的异常IP流量，保障网络安全。互联网攻击业务系统自动联动封锁剧本如图3所示。

2.2.3 强化学习构建未知威胁检测

针对威胁检测，传统检测方式主要依赖特征库、固定阈值设置等方法。其中，特征库规则是指传统安全工具（如WAF）依赖预定义规则（如正则表达式、签名库）来检测攻击。但攻击者可通过路径变异（如将“/actuator/health”改为“/actuator/.env”）、参数混淆（如Base64编码）或请求头伪造等方式绕过规则。固定阈值设置以单IP请求频率为例，这种方式难以区分正常流量洪峰与攻击行为，易产生误报或漏报。并且传统

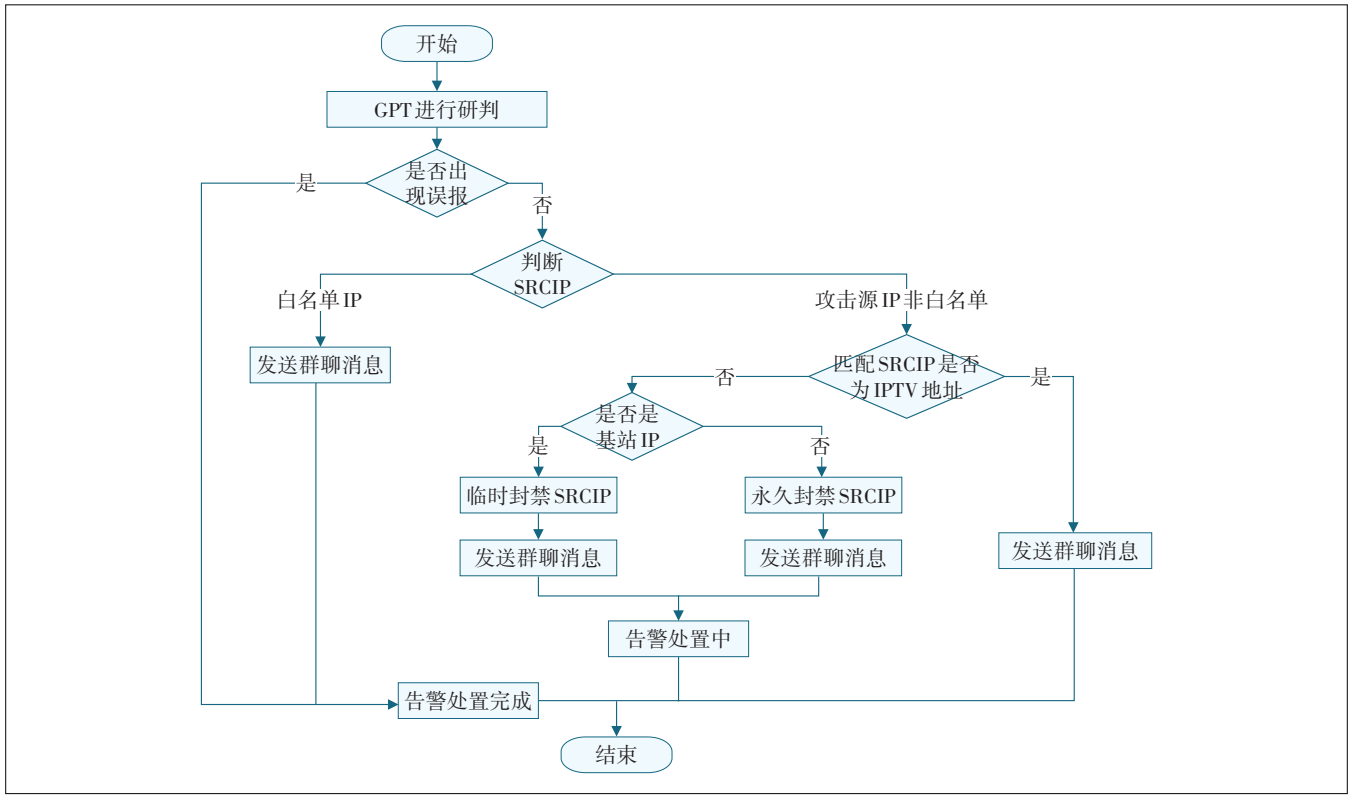


图3 互联网攻击业务系统自动联动封锁剧本

的检测方法仅能对已知威胁进行检测，存在未知威胁检测盲区，尤其对零日攻击或新型变种（如利用Spring Boot新特性漏洞发起的攻击）缺乏检测能力。

在已有安全AI的基础上，通过学习和推理来识别新出现的攻击手段，再结合网络中的安全告警对大模型进行微调，使AI模型能够具备HTTP流量理解能力、代码理解能力、攻防对抗理解能力和安全常识理解能力，使其类似一个攻防专家，能够对各种各样的威胁进行识别、告警，实现对未知威胁的动态检测，提升AI模型对定向攻击的溯源能力[13]。

通过持续强化学习，实现猎捕高绕过未知威胁（如Web 0day漏洞利用检测，混淆、绕过型攻击等）。通过将黑白Web流量在预训练及微调阶段用于大模型训练[14]，并基于安全专家研判经验对大模型检测和研判能力进行牵引，大模型具备了类似安全专家的HTTP流量、日志、代码理解和分析能力。它能通过自注意力机制关注到Web流量代码中的异常，判断代码意图，理解代码是否具有恶意特性，进一步基于上下文进行准确关联和综合研判。大模型针对Web流量具有良好检出效果。经多个高绕过的Web攻击数据集验证，相比传统正则规则和语法语义引擎，AI大模型的Web流量检测方法检出率明显提升，误报率明显下降[15]。

AI结合实际业务流判断告警是否为真实攻击的流程如下。

a）AI模型收到安全设备告警，显示系统遭受高危攻击。

b）AI模型结合原始数据包进行分析，重点分析请求数据包、响应数据包中的关键信息。

c）基于历史业务访问情况，给出研判结论：这是一次内部业务间的正常通信，并非真实攻击。

d）AI判定这是一次误报，后续相关人员二次研判也验证了该研判是真实准确的。

## 3 实测分析

### 3.1 实验概述

为验证“人工智能”技术对运营商行业网络安全运营能力的提升作用，从2025年1月至2025年4月，开展了为期4个月的实际测试。此次测试针对内网全部原始告警进行分析，测试数据主要包含终端告警数据、内网访问外网数据、互联网攻击数据、内网业务访问数据等，数据合计达89亿条。

### 3.2 告警聚合降噪效果

在测试期间（2025年1月—2025年4月），AI安全大模型共分析原始告警89亿条，自动研判分析定位出有效告警1.69万条（去掉加白及误报情况），告警降噪率达99.9%，日均待人工研判告警30条。

经过2轮告警优化，日均需人工研判的告警数量从11 041条降至18条，削减率达99%。人工研判告警数量削减情况如图4所示。

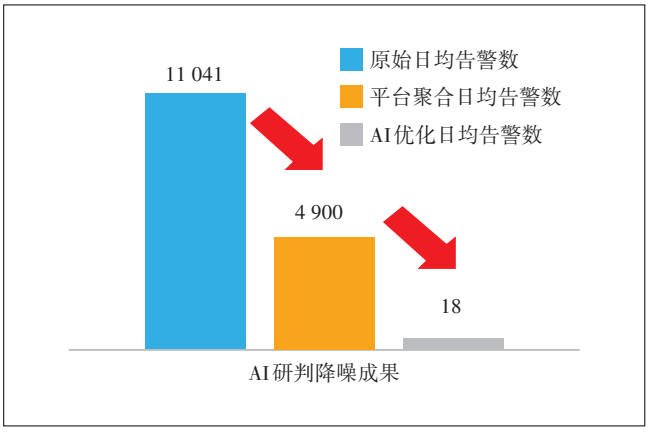


图4 人工研判告警数量削减情况

### 3.3 自动化研判效果

本次围绕4个关键场景进行SOAR剧本编排与实际测试。测试场景包含互联网攻击业务系统自动联动封锁、WebSHELL自动联动封锁处置、反弹类黑客工具自动联动封锁处置、服务器病毒自动联动封锁处置。测试结果表明，通过自动化分析处置，可将安全事件调查的全部工作时间由3~6 h压缩至5~10 min，原需多人处理的研判工作，现在仅需1人即可完成，提升了安全运营效率。在测试过程中，不断调整优化，以提升模型的可靠性，逐步实现事件闭环的自动化和高危场景的自动化提醒。针对AI研判准确度高且对业务无直接影响的安全事件（如互联网攻击事件），采用自动化闭环，运用模型进行监测—研判—处置；而对于对业务有影响的场景，则运用模型通过多重方式提醒专家，要求其限时研判处置，实现自动化提醒。

### 3.4 效果总结

针对安全事件，安全AI深度剖析攻击的完整过程、影响范围、攻击手段等详情，并呈现基于威胁建模、行为基线比对的分析思路，精准标记关键告警信息。通过时间线技术，安全AI将分散在不同异构设备中的告警日志进行串联整合，运用智能关联算法挖掘

各告警之间潜在的逻辑关系，实现网络端与终端检测数据的高效汇聚。在此基础上，通过智能过滤与聚类分析，安全AI能够自动消减重复、低价值的告警，显著提升告警质量，大幅降低人工研判的时间和精力成本，助力安全运营人员快速定位并处置安全威胁。

相较于已有防火墙、终端安全、态势感知设备所生成的简单告警内容，安全AI大模型的应用能够从多维度详细展示攻击过程。它可对源目的IP、威胁事件等进行综合分析，并结合历史事件数据，对攻击IP的其他攻击行为及所涉及的受害IP进行关联分析[16]。

此外，传统态势感知仅采用严格的匹配规则展示告警情况：如果攻击事件的五元组信息及威胁名称完全相同，则视为同一事件，系统会更新攻击次数和最新发生时间；若五元组信息或威胁名称不同，则作为新事件处理，生成新的告警条目。而本次采用的XDR+安全AI方案运用了聚合策略，系统会将不同的威胁名称聚合到同一安全事件中，例如将目录遍历攻击、Java代码注入攻击等不同攻击行为归类为“内网横向攻击”事件，这种基于攻击者行为的聚合方式，有效减少了告警条目数量[17]。

## 4 结论与展望

安全AI大模型的引入，系统性地革新了运营商网络安全工作模式，在风险预防、监测研判、调查处置、情报查询及溯源总结等核心环节实现了如下三大突破。

a）智能化升级驱动效能跃升。依托异构安全设备的日志关联分析，通过AI模型实现全量告警自动化精准研判，告警消减率达99%，研判及阻断准确率超过95%，安全事件调查耗时从人工处理的12 h以上大幅压缩至15 min以内，安全运营效率提升48倍。

b）技术融合构建智能联防体系。以网络安全垂直领域的生成式人工智能为核心，融合大数据分析与威胁情报能力，针对勒索软件攻击等典型场景构成自动化智能联防联控体系，显著提升对高级持续性威胁的检测、预警和处置能力。

c）运营模式转型破解行业痛点。通过AI自动化值守运营，推动网络安全工作从人力密集型向智能化、自动化方向转型，有效缓解专业人才短缺与经验传承难题，构建人机共智的新型安全运营体系。

该应用体系具备广泛的行业适配性，运营商行业均可基于现有框架拓展功能模块，实现对全业务场景的安全防护；地（市）、区县公司可灵活选择威胁检测、告警研判等核心功能，通过轻量化部署快速提升安全能力，为运营商行业网络安全建设提供可复制、可扩展的解决方案。

**参考文献：**

[1] 张伟，陈立，周涛. 基于深度强化学习的自适应网络防御决策模型[J]. 计算机学报，2024，47(1)：1-12.
[2] 方滨兴，时金桥，王忠儒，等. 人工智能赋能网络攻击的安全威胁及应对策略[J]. 中国工程科学，2021，23(3)：60-66.
[3] 周傲，吴杰，郑琳. 对抗样本在运营商AI检测系统中的攻防研究[J]. 计算机研究与发展，2024，61(2)：321-335.
[4] 刘艺，孙建国，李经纬，等. 面向云原生系统的智能告警根因定位技术综述[J]. 计算机学报，2024，47(4)：899-918.
[5] 周志华，李沐，张康. 运营商AI安全模型的持续学习框架[J]. 中国科学(信息科学)，2024，54(4)：567-580.
[6] 张伟，李强，王磊. 面向网络安全的大语言模型幻觉抑制方法研究[J]. 计算机研究与发展，2023，60(8)：1652-1665.
[7] 王刚，彭倩，段宏军，等. 基于人工智能技术的计算机网络安全防御系统的设计与实现[J]. 黑龙江科学，2024，15(18)：70-73.
[8] 林雪，马天宇，董方. 面向APT攻击的图神经网络溯源追踪模型[J]. 计算机研究与发展，2024，61(7)：1542-1556.
[9] 陈厅，潘剑锋，韦韬，等. 基于深度学习的软件漏洞挖掘研究进展[J]. 软件学报，2023，34(3)：1360-1386.
[10] 周涛，赵鑫，孙伟. 基于模型级联的实时网络入侵检测系统设计[J]. 软件学报，2023，34(6)：2105-2120.
[11] 刘鹏，王聪，马建峰. 基于深度强化学习的动态网络防御决策方法[J]. 软件学报，2023，34(11)：4987-5004.
[12] 吴晓倩. 基于人工智能技术的计算机网络安全防御系统设计[J]. 信息记录材料，2023，24(10)：67-69.
[13] 冷斌. 基于人工智能技术的计算机网络安全防御系统设计与实现[J]. 信息记录材料，2024，25(11)：91-92，95.
[14] 卢安文. 生成式人工智能：风险、监管与治理模式探究[J]. 重庆邮电大学学报(社会科学版)，2025，37(3)：113-121.
[15] 孙书魁，范菁，曲金帅，等. 生成式对抗网络研究综述[J]. 计算机工程与应用，2022，58(18)：90-103.
[16] 王健，方滨兴，刘欣然. 基于生成式对抗网络的网络威胁数据增强方法[J]. 计算机学报，2024，47(3)：580-593.
[17] 邬江兴，邹宏，张校辰，等. 网络空间拟态防御原理[J]. 中国科学：信息科学，2024，54(1)：1-28.

作者简介：
滕开清，工程师，学士，主要从事网络安全管理、安全运营、网络安全攻防技术等相关工作；曾哲凌，工程师，硕士，主要从事AI安全、网络安全攻防技术和网络安全应急响应等相关工作；胡芳燕，学士，主要从事网络安全管理、安全运营等相关工作。