

文章编号:1001-4179(2025) S1-0333-08

引用本文:张晓艺,杨柳,卢广毓,等.基于大模型分析的水利网络安全监管平台研究[J].人民长江,2025,56(增1):333-340.

基于大模型分析的水利网络安全监管平台研究

张晓艺¹, 杨柳¹, 卢广毓², 梁 锋³, 刘 闯⁴

(1. 水利部信息中心, 北京 100053; 2. 黄河水利委员会 信息中心, 河南 郑州 450003; 3. 奇安信网神信息技术(北京)股份有限公司, 北京 100044; 4. 广西大藤峡水利枢纽开发有限责任公司, 广西南宁 530299)

摘要:水利网络安全是推动数字孪生水利健康发展的重要保障之一。为解决当前水利网络安全监管的痛点难点,提升监管效能,提出并构建了水利网络安全监管平台。从业务场景出发,采用数据采集、基础能力、业务应用三层架构,开展平台安全分析、业务管理、综合态势分析能力建设。基于大模型强大的安全分析知识,该平台可全面、准确、自动生成研判结论,实现水利网络安全主动监测与全局态势分析,并对告警及事件进行全过程闭环管理。该平台创新技术应用,实现了安全监管与水利业务的深度融合。相关成果提升了水利行业网络安全联防联控水平,为数字孪生水利建设提供了技术支撑,可为行业级网络安全监管提供经验借鉴。

关键词:水利网络安全; 监管平台; 大数据建模; 大语言模型; 全局分析

中图法分类号: TP309.2; TV21

文献标志码: A

DOI:10.16232/j.cnki.1001-4179.2025.S1.065

0 引言

随着信息技术深入发展,网络安全形势日益严峻^[1-4]。党的十八大以来,以习近平同志为核心的党中央高度重视网络安全工作^[5]。国家相继出台《中华人民共和国网络安全法》《中华人民共和国数据安全法》《中华人民共和国个人信息保护法》《关键信息基础设施安全保护条例》等网络安全相关法律法规,明确要求维护网络安全空间主权和国家安全、社会公共利益,保护公民、法人和其他组织的合法权益^[6-7]。

水利是涉及国家关键信息基础设施的重要行业之一^[8],相关系统一旦遭受破坏,将会对国家安全、国计民生、公共利益造成严重危害。随着数字孪生水利建设全面深入推进^[9-12],水利业务对信息化依赖程度越来越高,网络安全面临新的风险和挑战,通过网络安全监测分析、态势研判、预警处置等开展行业监管的必要

性日渐凸显。目前,国内外已有一系列关于网络安全监测、感知、态势分析等技术的研究。Szwed 等利用模糊逻辑结合认知图技术,获取网内重要资产间依赖关系,从而进行恶意行为分析^[13]。Balta 等提出的基于数字孪生的网络攻击检测,可以用于检测受控瞬态业务行为异常^[14]。Eckhart 等提出建立信息系统副本,在虚拟并行环境下对网络态势进行全面分析^[15]。常利伟等研究了基于卷积神经网络的多源融合网络安全模型^[16]。白荣华提出并设计了多源一体化政务网络安全监测平台^[17]。徐波等提出数字孪生水利工程网络安全风险分析和保障体系^[18]。简玲等基于统一的数据采集传输规范、数据存储共享机制及多源分析算法共识机制,提出多源融合的大数据网络安全态势感知平台架构^[19]。王晨飞等提出基于构建行为画像的网络安全态势感知机制^[20]。段咏程等研究了基于RSAR 的随机森林网络安全要素提取^[21]。张克君等

收稿日期:2024-07-30;接受日期:2024-09-29

基金项目:国家重点研发计划项目(2021YFB3900600)

作者简介:张晓艺,女,工程师,硕士,主要从事水利信息化工作。E-mail:zhangxiaoyi@mwr.gov.cn

研究了基于 PSP-TSA 模型的网络安全要素识别方法^[22]。王婷婷等提出一种基于差分 WGAN 的网络安全态势预测机制^[23]。

从监测侧的卷积神经网络多源融合建模、随机森林、PSP-TSA 模型要素提取、基于数字孪生等攻击检测,到分析侧的模糊逻辑认知图技术、信息系统副本建立,再到感知侧的统一采集传输规范、存储共享机制、行为画像、差分 WGAN 等技术应用,总体而言,网络安全检测感知与态势分析能力已得到一定发展。但基于大模型技术的网络安全分析研究较少,且难以结合行业级网络安全监管实际解决痛点难点问题。因此,本文结合水利行业监管实际,开展基于大模型分析的水利网络安全监管平台研究,充分利用垂域大模型提升行业安全监测分析、综合研判、态势评估总体效能,保障新阶段水利高质量发展,推进数字孪生水利建设。

1 水利网络安全监管痛点分析

水利行业信息资产类型多、数量大,分布广,涵盖防洪、供水、生态等水利公共服务产品和服务供给等内容,涉及网络、业务、数据、工程等类型,网络覆盖水利部、流域管理机构、省、市、县等多级水行政主管部门,网络安全边界复杂,行业网络安全监管工作面临巨大的挑战,需要行业监管平台提供技术支撑,主要存在以下三方面问题。

1.1 主动监测能力不足

水利网络安全监管更多依赖沟通联络与信息通报机制,缺乏主动监测能力,无论上传下达如何高效通达,只能靠各单位主动报送异常情况,承担平时与战时指挥部角色的部网信办缺乏主动权,指挥调度被动且滞后。

1.2 数据集成共享难

虽然各地自行建有威胁感知能力,但能力各异,且信息口径不一,难以作汇总分析与横向比对。

1.3 宏观态势分析存在差距

缺乏对行业网络安全的宏观感知,难以对行业态势趋势进行整体把握与分析,在行业通盘考虑、宏观决策与调度优化方面存在一定差距,且缺乏可视化手段。

2 平台设计

2.1 总体架构设计

水利网络安全监管平台采集、汇总、处理行业单位网络安全相关数据、网络流量日志、安全日志、情报信息等。通过对其进行综合研判分析,形成告警及事件,并针对告警及事件进行通报预警、协同处置、指挥调

度、整改督办、闭环跟踪,完整攻击处置过程以场景化一站式数据链方式展示。平台通过可视化大屏呈现并分析行业整体网络安全态势。除来自互联网的攻击外,水利信息网^[24]内行业单位之间横向跨网异动、跨网异常访问或探测也是关注的重点。

平台总体架构如图 1 所示,主要分为数据采集层、基础能力层和业务应用层三层。

2.1.1 数据采集层

实现统一的数据采集,包括部直属单位、省级水行政主管部门网络流量、日志数据、资产信息、威胁情报,同时对接行业单位已建威胁感知平台及第三方采集器。

2.1.2 基础能力层

该层可分为数据预处理、数据组织、数据治理、数据服务。数据预处理对接入的结构化、半结构化、非结构化数据进行数据加载、清洗、消重、验证、抽取、融合、转换等处理,使数据格式相对统一,标识清晰。数据组织包含 5 类库,经过预处理的数据首先存入原始库,经过关键要素提取后形成资源库,再进一步提取、归并建立各实体之间的关系后形成主题库,再根据上层业务需求形成业务库,最后综合分析学习形成知识库。

该层提供数据接入监控、处理监控、接口管理、元数据管理、数据资源目录管理、数据质量管理等数据治理能力,及数据检索、查询、共享、订阅等数据服务能力。

2.1.3 业务应用层

该层又分为态势感知、协调指挥与可视化展示。通过资产管理、资产脆弱性安全管理、研判分析、告警协同管理、跨网安全监测等进行态势感知。通过事件管理、通报管理、指令协同、重要时期保障、资源治理管理等开展协同指挥。该层的最上层具备可视化能力,通过综合态势、资产态势、安全监测、行业联防联控 4 张大屏综合展示水利网络安全总体态势。

2.2 功能模块设计

2.2.1 安全分析能力

基于网络安全监管平台研发算法模型,实现信息搜集、漏洞利用、建立据点、权限提升、权限维持、横向移动、痕迹清除等完整攻击链覆盖,结合场景化分析提取关键线索(图 2),实现威胁监测与安全分析全覆盖。

2.2.2 业务管理能力

如图 3 所示,支撑日常、攻防演练及重要时期水利网络安全告警协同、事件跟踪、预警通报,实现行业内外协同、上下联动。平台监测到异常情况后,第一时间将告警信息或事件详情通报至受影响单位,与受害单

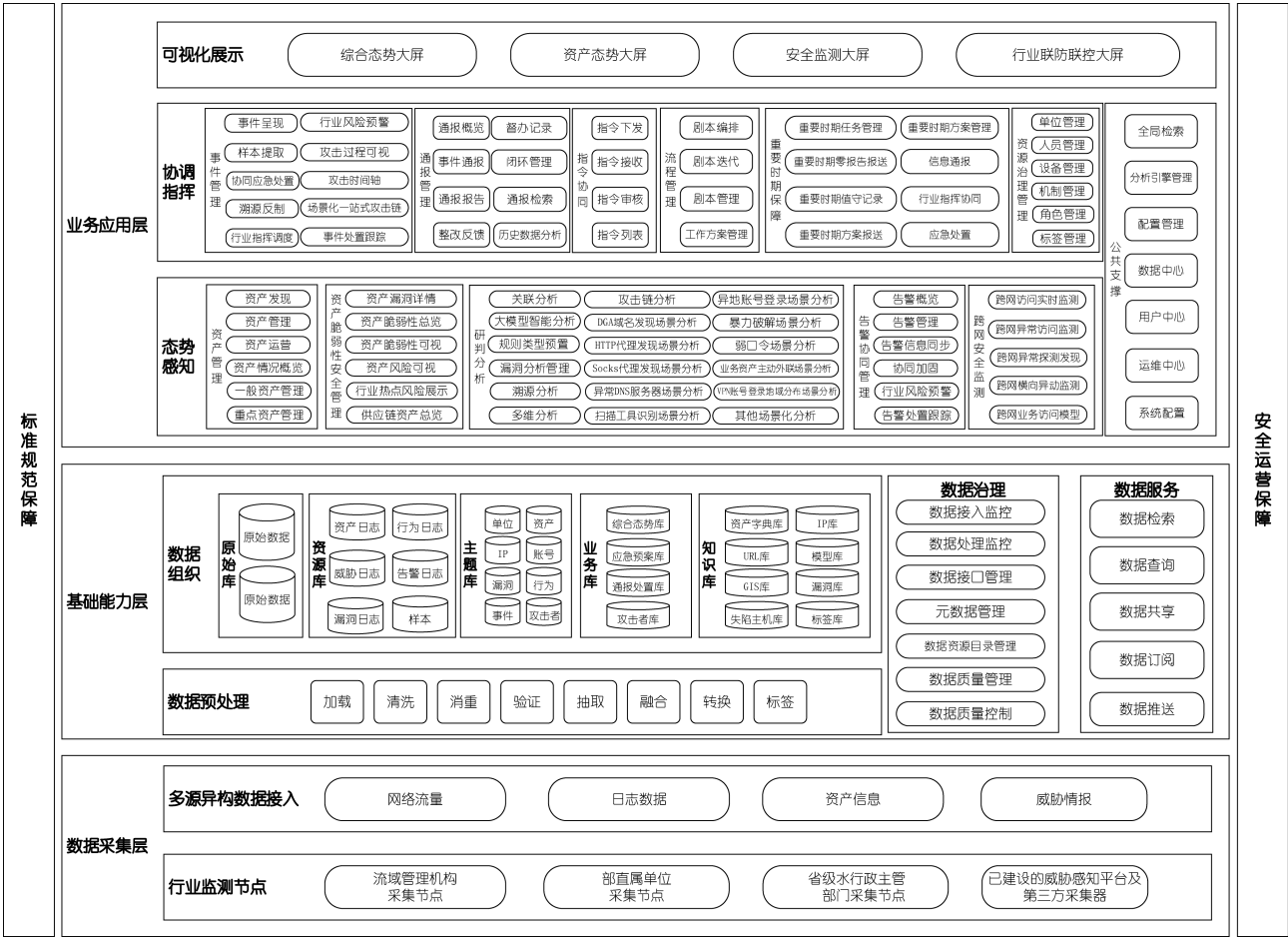


图 1 水利网络安全监管平台总体架构

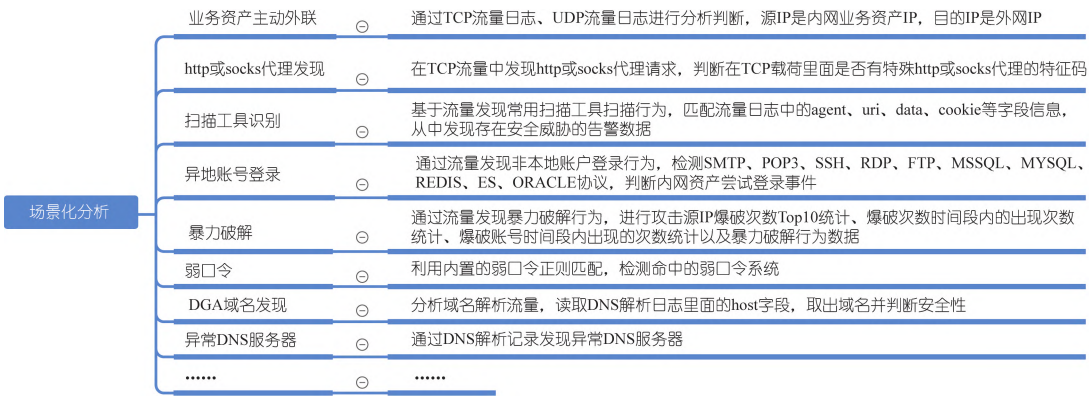


图 2 场景化分析示例

位协同研判,跟踪指导风险处置,支撑溯源反制,实现对应急响应流程的全方位支撑。对各类资源统筹调度、协同联动,并及时将发现的问题隐患在行业内通报预警,实现一处报警,处处设防,一处威胁,处处处置。同时,针对行业自行监测、上级主管部门预警通报、第三方供应商提供的威胁情报、高危网络或系统漏洞、供应链攻击风险、僵尸蠕毒分布情况等进行实时信息共享,提前预警,指导行业全面排查整改,事前防御,在可能的攻击发生之前即有效处置,整改加固。对于日常

工作通知,通过选定工作通知对象后下发指令,平台可查看通知详情,未读、已读和已回复情况等。在攻防演练及重要时期,平台根据战况实时动态优化协防布局,实时指挥,各级单位协同响应,联合作战。

2.2.3 综合态势分析能力

如图 4 所示,通过宏观、中观、微观视角,综合呈现与分析行业网络安全态势及变化趋势,横向评价对比分析,纵向洞悉网络安全的过去、现在和未来,清晰掌握安全因果关系,体现行业总体态势和风险隐患,对于

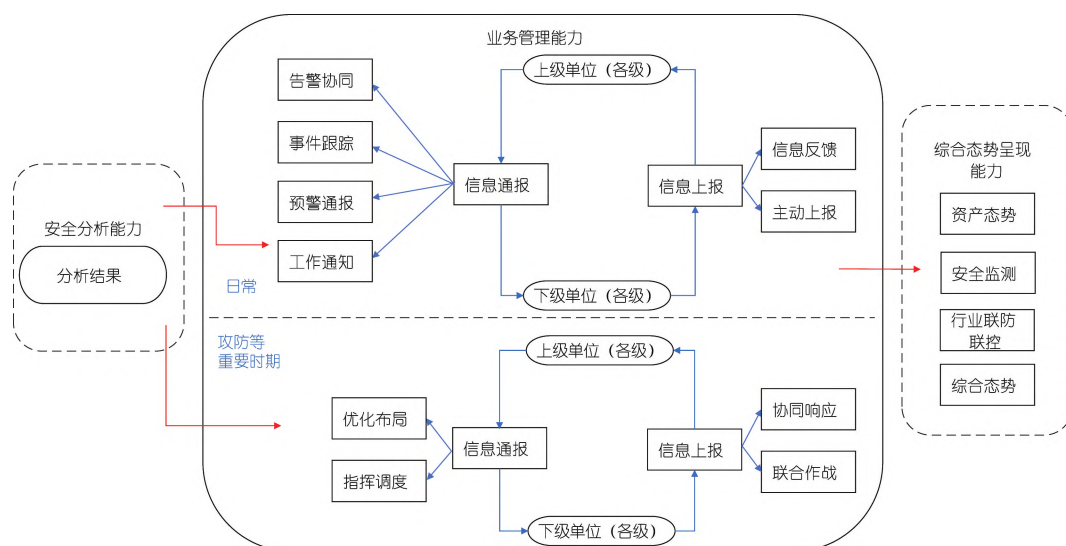


图3 业务管理能力示意

发现的问题进行重点干预,全盘指挥。

综合态势包括4张大屏:资产态势大屏统计行业单位资产数量、IP数量、设备数量等,支持对资产被攻击次数等进行排名;安全监测大屏从全局视角进行多维度综合评价,体现行业单位综合安全系数,对攻击类型、攻击手段、告警来源分布、攻击结果等进行分析统计;行业安全联防联控大屏对行业协防事件处置情况进行分析,统计事件处置单位、事件处置率、已处置事件等;综合态势大屏通过对资产、漏洞、告警、事件、通报等多维度评估,形成总体安全系数评分,支持以全局视角分析流域管理机构、部直属单位、省级水行政主管部门等总体态势。

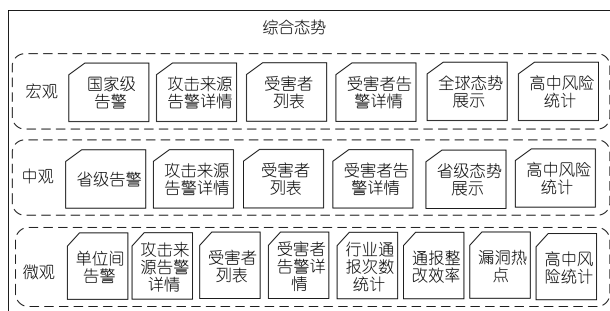


图4 综合态势示意

3 关键技术

3.1 大数据建模技术

平台基于大数据技术构建了统一的网络空间安全逻辑数据模型,能够将汇聚的海量数据进行整合,对网络空间领域多源异构数据进行数据融合与价值挖掘。

3.1.1 安全分析算法模型

平台利用大数据分析技术对告警规则进行训练,

同时以实战攻防对抗为目标进行网络攻击样本特征训练,分析出高精度网络攻击行为。通过剧本编排自定义建模模式,不断迭代精度更高的算法模型(图5),贯穿大数据建模分析的全生命周期,构建一套“用数据说话,用数据决策,用数据创新”的数据驱动机制,实现威胁全方位感知。

在通用算法模型基础上,充分结合水利实际业务特点,基于批流画布技术,开发与业务深度融合的算法模型^[25](图6),提高网络安全监测的主动性与针对性。以不同视角,实现内外齐管。从内部脆弱性视角,实现对系统漏洞、应用漏洞、配置不当等各类内部脆弱性情况全方位实时感知;从外部攻击视角,实现对外部扫描、暴力破解等各类外部攻击的全方位实时感知,实时掌握攻击源和高频被攻击资产。

3.1.2 Spark 计算模型

在大数据分析处理过程中,平台基于Spark通用计算模型建立分布式迭代计算机制。Spark提供SQL接口、MLlib集群学习算法、GraphX库等用于海量结构化数据处理及深度数据分析等。基于Spark机器学习的大数据分析可以对接多种数据来源,包括经数据治理后的Hive表、数据表、K-V表等,利用机器学习算法(图7)对多源数据进行处理、转换、计算、聚合等,实现网络安全态势分析与呈现,并支持实时动态数据同步。

3.1.3 ES 引擎

ElasticSearch是一个可高度扩展的检索与分析引擎,能够快速、近实时地存储、查询和分析海量数据。除了原生接口,定制开发了基于Spark的SQL访问接口。ES将数据存储在索引文件中,并将索引分成若干



图5 算法模型示例

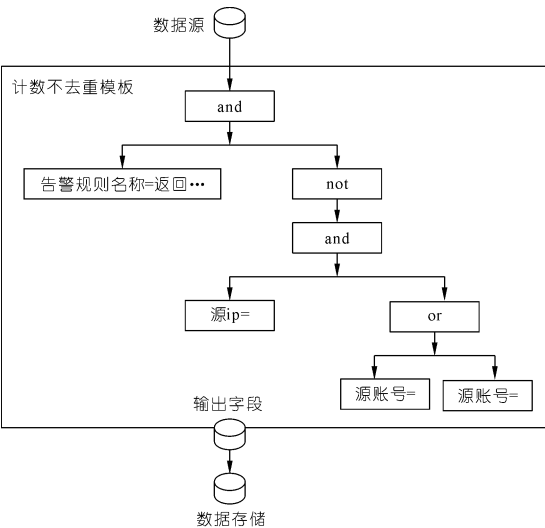


图6 专业业务深度融合的算法示例

个分片存储在不同的节点上。索引允许 ES 快速定位某个特定记录,又可有效地分布存储和管理数据,对海量数据检索起到至关重要的作用,基本可实现数 10 亿日志数据查询检索时间不超过 3 s。

3.2 大语言模型技术

在大数据建模与分析基础上,平台基于人工智能大语言模型技术形成安全大模型。大语言模型(Large Language Model,LLM)是基于海量文本数据训练的深度学习模型。结合水利网络安全场景,安全大模型在威胁检测、指挥决策方面发挥作用。大模型通过千万级流量样本测试,采用 NLP 等算法,具备语义分析引擎,通过对全网流量日志和安全日志的分析,能够检测 Web 应用、内网穿透、异常访问等传统安全场景攻击

以及 0Day、混淆绕过等未知威胁,能够基于行业日常安全告警分析及业务特点进行模型学习优化。大模型具备自然语言交互能力,具有自动解读安全事件、分析安全日志、多维度自动化统计查询、全网资产漏洞排查、解读 HTTP 告警数据包、解读威胁情报、解读恶意文件、总结网络安全总体态势、解读安全专业词汇等功能。

3.2.1 自注意力机制

Transformer 模型的核心能力在于自注意力机制。自注意力机制着重关注序列内部元素之间的关联关系,通过计算各元素之间的关联度,让模型在输入序列上下文中更精准地找到关键部分,即找到“注意力”。自注意力机制包含查询(Query)、键(Key)、值(Value) 3 个部分,对于序列中的每个元素,模型都会生成对应的 Q 、 K 、 V 向量,通过计算 Q 与 K 之间的点积,得到关联度得分,通过 softmax 函数进行归一化,用此得分加权 V ,得到输出。

为了使模型能够同时关注序列的不同位置,Transformer – Decoder 采用了多头注意力机制。在多头注意力中,模型会并行执行多个自注意力层,每层关注序列的不同部分,然后将各层输出结合起来完成最终输出。

3.2.2 神经网络构架

神经网络是构建大模型的基础,基于大数据技术,通过构建和训练深层神经网络模型,从海量数据中学习和提取高级抽象特征,直至具备安全感知、分析与决策能力。将网络空间概念拟合至现实空间,将现实空

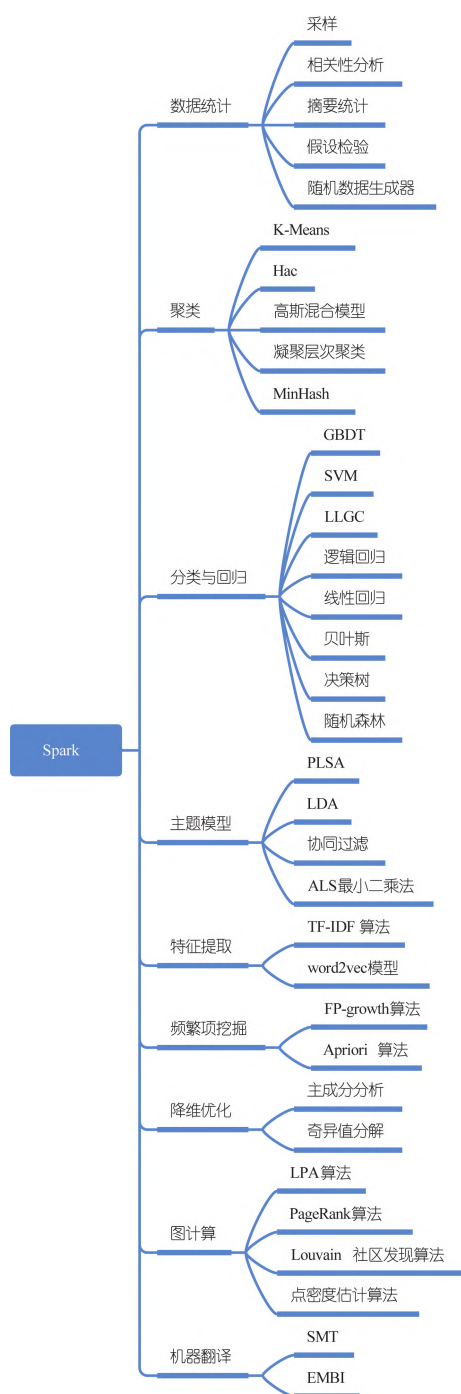


图7 Spark 机器学习算法

间从技术视角拆解为网络空间概念,并将网络空间不同数据归并至相应主题。深层次神经网络可抽象为输入层、隐藏层、输出层3层。平台构建起由海量数据作为输入,网络空间主题作为隐藏层进行数据抽取与分析,直至拟合出具备安全分析能力模型作为输出的大规模神经网络。

3.2.3 预训练与微调

安全大模型核心在于自动开展研判分析,它可以从监管平台日志信息出发作出研判,如提取需要专业

运营人员研读的五元组、请求头、请求体、响应头、响应体等信息,GPT 模型利用强大的安全监测、分析与运营知识,全面、准确、自动化生成研判结论。该研判结论通过自然语言输出,简单易懂,降低了传统安全监管的人工成本与时间成本。

安全大模型是采用“预训练 + 微调”方式训练得到的,其基座大模型基于 transformer - decoder 深度学习结构,使用海量通用数据和安全领域大数据预训练得到一个预备可用的大模型,再基于这个基座大模型进行微调,形成具有安全感知与态势分析能力的垂域大模型。大模型预训练数据来自安全大数据,包括但不限于安全日志、安全情报、安全告警、安全文档、安全知识库等,以及自然社会科学通识数据,使用自监督学习方法,支撑大模型学习语言的一般特征和结构及丰富的语义表达,实现安全大模型的对话交互。如图8所示,在预训练基础上对模型进行微调,与预训练阶段的无监督学习方法不同,微调阶段采用有监督的专家标注方法。标注的过程采用“一标二审三抽检”方式。实现效果上,大模型通过对监管平台产生的告警进行GPT 研判,为每条告警打标签,包括是否误报、是否为真实攻击行为、攻击行为是否成立等,从而大大消减告警量。

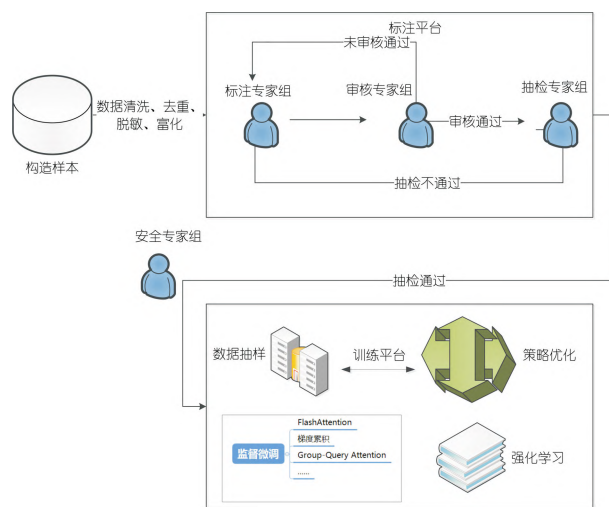


图8 预训练与微调示意

3.3 全局分析技术

3.3.1 态势评估技术

网络安全态势评估技术需要以网络安全监测为驱动,以安全威胁线索为牵引,对网络空间安全相关信息进行汇聚融合,将多个安全事件联系在一起进行综合评估,实现对整体网络安全状况的判定与安全态势分析。对安全事件尤其是对网络空间安全相关信息进行汇聚融合后所形成的“人”“物”“地”“事”关系的多维

安全事件知识图谱(图9),是网络安全态势评估分析的关键。

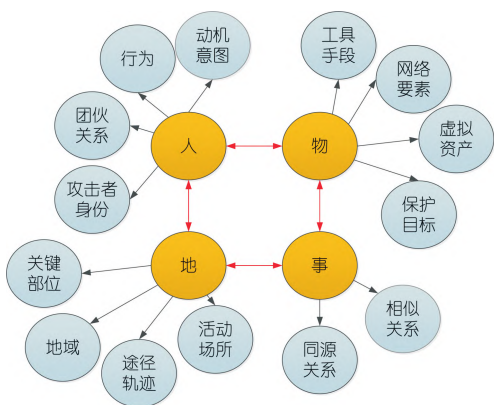


图9 多维安全知识图谱

3.3.2 攻击链分析技术

安全大模型输出研判分析过程和结果的同时,反推研判过程,将攻击过程按正向时间线顺序复现出清晰、完整的攻击链,便于统筹全局开展分析、处置、溯源,并举一反三进行通报预警。在攻击阶段呈现方面,如图 10 所示,平台通过 ATT&CK 模型反映整个攻击生命周期各阶段的攻击行为,有效描述攻击所使用的策略与手法。

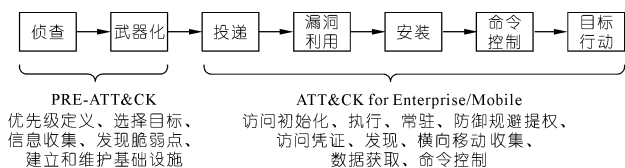


图 10 ATT&CK 模型

同时,平台通过钻石模型搭建了一个描述攻击过程的框架。该框架由至少4个互连的特征点组成:攻击者、受害者、基础设施和能力。特征点间的链接线表示操作者执行操作流程的一系列有序事件序列,线可以在各种活动之间进行关联和匹配,从任何一个特征点出发,能够观察到通过活动线与之链接的其他特征点的活动。

4 创新应用分析

4.1 与业务深度融合的安全分析

该平台基于安全大模型,对访问关系进行告警类型定性,在成千上万条威胁线索中快速自动分析识别出真正有风险的攻击或漏洞,厘清行业单位间正常业务访问关系,减少告警误报,通过学习,不断迭代知识,精准认定行业关键风险,并不断提升模型精度。更关键的是,针对具体业务访问流量、访问场景、访问操作

构建的与业务深度融合的算法模型,能够从业务实际出发,真正为水利业务提供支撑。目前,平台日均告警由过去成千上万条消减至几十条,同时结合大模型强大的研判分析能力,通过对话或者多轮追问等实现完整研判过程,大大提升监测分析的自主性、准确性与时效性。

4.2 跨网监测

在网络攻击中,攻击者通常会通过横向异动、登录突破、权限提升等过程,控制一台主机并将其作为跳板,水平移动至其他主机,通过获取的口令直接登录到目标主机,执行远程命令,从而达到最终目的。水利行业单位计算机与网络设备互联形成水利信息网,任何一点突破,攻击者都有可能进行横向探测与渗透以获取重要信息和相关权限。平台通过跨网监测能力,基于监测模型结合威胁情报对网内各单位交互流量日志进行风险分析,监测交互流量中的攻击行为与网络风险,识别风险涉及的漏洞、资产与访问关系,通过关联本地风险日志从全局角度分析安全威胁对本地和行业其他单位的安全影响,发现行业单位间跨网横向异动、跨网异常访问或探测等。

4.3 攻击链与态势分析

平台支持以场景化一站式数据链描述整个攻击过程甚至守方响应对抗过程以及过程中的每一个阶段,还原出主要环节、各阶段时间节点、动作、行为、攻防对象等相关信息,这有利于深刻理解攻防全貌及各阶段细节,对于场景复盘和防守方案优化迭代起到关键作用。此外,当将视角扩展至全行业,不仅可以整体把握行业态势,分析变化趋势,同时可以在行业单位之间进行横向比对,不断提高行业网络安全宏观感知和联防联控水平。

5 结语

本文通过构建水利网络安全监管平台,基于大数据建模、大语言模型、全局分析等关键技术,针对水利网络安全监管安全分析、业务管理、综合态势分析等业务实际需求,采用“预训练+微调”方式训练,形成了具有安全感知与态势分析能力的安全大模型。结合水利网络安全特点,提出了平台在算法模型与业务深度融合、跨网监测、攻击链分析等方面的应用,从而实现水利网络安全主动监测与深入分析,解决行业监管痛点。该平台特点如下:① 主动监测能力显著提升。基于网络安全监管平台,提升行业主动监测与态势分析能力,在传统信息通报与沟通联络基础上,建设主动感知能力,得以在行业指挥侧掌握主动权,从技术层面扎

实提升联防联控水平。② 实现威胁感知数据集成。在行业各节点部署统一监测设备,使得各节点数据接口一致,为进一步集成分析打牢基础。③ 宏观态势分析能力提到增强。利用平台综合态势分析能力,整体感知行业安全态势,结合时空分析,支撑行业宏观决策指挥调度。

参考文献:

- [1] 向夏雨,顾钊铨,曾丽仪.网络威胁情报共享与融合技术综述[J].网络空间安全科学学报,2024,2(2):2-17.
- [2] 董舟,谢碧云,李歆.政务外网信息安全管理策略初探[J].人民长江,2015,46(3):86-90.
- [3] 曹旭栋,黄在起,陈禹劼,等.安全漏洞库构建及应用研究综述[J].计算机学报,2024,47(5):1082-1119.
- [4] 曹俊启,黎伟,杨涛.南水北调中线工程信息化建设及安全防护对策[J].人民长江,2015,46(6):93-95.
- [5] 钱峰,张潮.以数据为核心开展网络安全攻防演练防守[J].水利信息化,2020(6):17-20,42.
- [6] 杨旭,李元杰,成萌.水利网络安全工作平台设计与应用[J].水利信息化,2022(6):59-64.
- [7] 付静,周维续,詹全忠,等.基于商用密码的水利重要数据点面结合安全保护方法[J].水利信息化,2024(1):1-5.
- [8] 付静.水利关键信息基础设施安全保护探索与实践[J].信息网络安全,2023,23(8):121-127.
- [9] 谢明霞.数字孪生水利内涵及应用场景研究[J].人民长江,2024,55(2):245-251,264.
- [10] 蔡阳.以数字孪生流域建设为核心构建具有“四预”功能智慧水利体系[J].中国水利,2022(20):2-6,60.
- [11] 罗斌,周超,张振东.数字孪生水利专业模型平台构建关键技术及应用[J].人民长江,2024,55(6):227-233.
- [12] 贺挺,李凤生,成建国,等.水利部数字孪生流域模型管理云平台设计及应用研究[J].水利水电技术(中英文),2024,55(2):1-15.
- [13] SZWED P, SKRZYNSKI P. A New Lightweight method for security risk assessment based on fuzzy cognitive maps[J]. International Journal of Applied Mathematics and Computer Science, 2014, 24(1): 213-225.
- [14] BALTA E, PEASE M, MOYNE J, et al. Digital twin - based cyber - attack detection framework for cyber - physical manufacturing systems [J]. IEEE Transactions on Automation Science and Engineering, 2024, 21(2): 1695-1712.
- [15] ECKHART M, EKELHART A, WEIPPL E. Enhancing cyber situational awareness for cyber - physical systems through digital twins [C] // 2019 24th IEEE International Conference on Emerging Technologies and Factory Automation. Zaragoza, 2019: 1222-1225.
- [16] 常利伟,刘秀娟,钱宇华,等.基于卷积神经网络多源融合的网络安全态势感知模型[J].计算机科学,2023,50(5):382-389.
- [17] 白荣华.多源一体化政务网络安全监测平台设计与应用[J].计算机应用与软件,2024,41(7):20-24,73.
- [18] 徐波,王昕.数字孪生水利工程网络安全风险分析和保障体系[J].人民长江,2023,54(11):242-250.
- [19] 简玲,叶天鹏,林祥,等.多源融合的大数据网络安全态势感知平台研究与探索[J].信息网络安全,2020(增2):139-143.
- [20] 王晨飞,徐李阳,李慧芹,等.基于构建行为画像的网络安全态势感知机制[J].计算机应用,2024,44(增2):118-122.
- [21] 段詠程,王雨晴,李欣,等.基于RSAR的随机森林网络安全态势要素提取[J].信息网络安全,2019(7):75-81.
- [22] 张克君,郑炜,于新颖,等.基于PSO-TSA模型的网络安全态势要素识别研究[J].湖南大学学报(自然科学版),2022,49(4):119-127.
- [23] 王婷婷,朱江.基于差分WGAN的网络安全态势预测[J].计算机科学,2019,46(增2):433-437.
- [24] 陈岚,詹全忠.水利信息安全管理平台研究与应用[J].水文,2011,31(6):67-69,85.
- [25] 张晓艺,戴逸聪.水利数据分类分级及安全保护技术[J].人民长江,2023,54(增2):232-237.
- [26] 白雪,王鸿元.大语言模型在网络安全领域的应用探索[J].电信工程技术与标准化,2023,36(12):23-30.

(编辑:郑毅)