主题：高管视角、课题摘录

交易技术前沿

2025年 第7期 总第67期

——网络安全专刊（第3期）

ITRDC 证券信息技术研究发展中心（上海）

P02 大模型在证券行业应用中的安全问题探索与实践
国投证券 许彦冰

P06 证券行业安全建设：构建智慧安全生态，护航业务稳健发展
国金证券 王洪涛

P09 数字化时代基金公司的网络安全
鹏华基金 聂连杰

内部资料 免费交流
准印证号（K）0671

交易技术前沿

2025 年 第 7 期

总第 67 期

总编

邱　勇　蔡建春

副总编

王　泊

本期“安全网络专刊”：

执行总编

道　晟　刘政言

责任编辑

张　涛

运营：

证券信息技术研究发展中心（上海）

网络安全创新实验室

主管、主办：

上海证券交易所

目录

01 高管视角

P02 大模型在证券行业应用中的安全问题探索与实践
许彦冰/国投证券股份有限公司

P06 证券行业安全建设：构建智慧安全生态，护航业务稳健发展
王洪涛/国金证券股份有限公司

P09 数字化时代基金公司的网络安全
聂连杰/鹏华基金管理有限公司

02 安全实践

P14 证券行业邮件安全运营实践
张浩哲、邢骁、李彤恩、徐文娟/西部证券股份有限公司

P23 证券行业移动应用软件安全与合规管理
宋士明、叶飞、姜玥/南京证券股份有限公司

P30 安全运营韧性构建：基于信创态感的智能化安全运营探索实践
丁安安、夏英杰、赵川/国联民生证券股份有限公司

P35 数据合规驱动下的全域风险监测体系构建与实践研究
—— 基于数据处理活动的安全运营创新路径
陆滢、徐正伟/华安基金管理有限公司

P38 基于大模型的全渠道信息流统一管控
李剑戈、陶昆、达其双、吴敏嘉/中信建投证券股份有限公司

P42 钓鱼邮件演练的实践与探索
张沁怡、崔毅然、吴鹏、陈其乐/上海证券有限责任公司

P45 智能体驱动的API安全风险管控研究与实践
徐承文、程际桥、吴琪、刘义卓、杨启/长江证券股份有限公司

P51 基于某次大型攻防演习应对零日漏洞攻击的威胁狩猎实践
林宝晶/奇安信网神信息技术(北京)股份有限公司
钱钱/中国航天系统科学与工程研究院

03 新技术应用

P56 基于零信任架构的安全访问和智能协同研究与实践
黄辉、崔荣涛、韩宇/湘财证券股份有限公司

P61 零信任体系在证券业数据安全领域的探索与实践
俞建群、邬晓磊/东方证券股份有限公司
何艺/北京持安科技有限公司

04 思考和观点

P68 韧性数字安全体系研究与建设
侯亮/国泰海通证券股份有限公司

P73 大模型应用场景安全思考与实践
钟蓉、吴佳伟、李鹏、曹杰、温志强、郑煜/兴业证券股份有限公司

05 热点解读

P79 **小程序安全解决方案**
张华/腾讯云计算（北京）有限公司

P85 **金融行业IPv6规模化部署顶层设计与落地策略**
葛锐、张绍峰、陈政、沈鑫尧/互联网域名系统北京市工程研究中心有限公司

P89 **金融行业勒索病毒防御评估研究**
施勇、张涵/上海霞安信息科技有限公司

P93 **2025RSAC大会解析：众声汇聚，共探全球网络安全新趋势**
江爱军、王伟涛、盛浩月/奇安信科技集团股份有限公司

06 2023年度实验室优秀课题摘录

P99 **基于可信内存指令序列检测技术的漏洞（包含未知漏洞）攻击防护能力**
刘嵩 胡广跃 王磊/光大证券股份有限公司
刘磊 刘志明/奇安信科技集团股份有限公司

P105 **挂图作战在证券行业应用实践研究**
华仁杰 沈嗣贤 徐俊超 鞠叶 任思豫/东吴证券股份有限公司
张海龙 金亮成 杨云云/金证金融科技（北京）有限公司

P111 **零信任架构下的业务系统敏感数据保护实践**
王洪涛 刘宏 马晓鹏 黄施宇/国金证券股份有限公司
何艺/北京持安科技有限公司

P118 **东方证券企业终端数据安全解决方案探索**
邬晓磊、甄明达/东方证券股份有限公司
焦健、汤华晟/数篷科技（深圳）有限公司

P124 **基于漏洞情报的拟态防御技术实践**
邢骁 蔡子豪/西部证券股份有限公司
薛辛/北京长亭科技有限公司

01 高管视角

P02 大模型在证券行业应用中的安全问题探索与实践
许彦冰

P06 证券行业安全建设：构建智慧安全生态，护航业务稳健发展
王洪涛

P09 数字化时代基金公司的网络安全
聂连杰

大模型在证券行业应用中的安全问题探索与实践

许彦冰 | 国投证券股份有限公司

摘　要： 随着金融科技数字化转型的深化，生成式人工智能，特别是大语言模型（LLM），正成为驱动证券行业数字化创新的核心引擎。在赋能投研、客服等业务场景的同时，其带来的数据安全、合规风险、模型滥用等新型安全挑战，已成为我们必须正视并加以管控的风险。本文系统性地剖析了大模型应用当前面临的主要风险，并结合自身的研究和实践，提出了一套覆盖大模型应用全生命周期的安全防御体系。该体系明确以“上线时的大模型安全测试”和“运行时的大模型安全网关”为核心抓手，旨在构建针对大模型输入输出内容的可控、模型攻击的可防能力，为大模型技术在证券行业的安全、合规引入与规模化应用，提供一套行之有效的参考方案。

关键字： 大语言模型、数据安全、合规风险、安全网关、内容安全

一、引言

当前，证券行业正处在数字化转型的关键攻坚期，以大语言模型（LLM）为代表的人工智能技术，已不再是“可选项”，而是决定未来核心竞争力的“必选项”。从智能投研、策略辅助到合规风控、智慧办公，大模型展现出的巨大潜力正驱动各业务条线积极探索与布局。然而，我们需要认识到，技术的加速落地正将一系列前所未有的安全风险暴露在我们面前。2023年三星公司敏感代码泄露、数家银行600余万条客户数据被用于大模型训练导致泄露等事件，已为全行业敲响警钟。更为重要的是，国家网信办等监管机构相继出台的《生成式人工智能服务管理暂行办法》等法规，划定了不可逾越的合规红线。因此，如何平衡业务创新与安全风险，在享受技术红利的同时，构建一套与之匹配的、强有力的安全防护与治理体系，确保技术应用始终处于安全、可控、合规的轨道之上，已成为摆在我们面前最为紧迫的问题之一。

二、大模型安全现状与挑战分析

（一）大模型应用现状调研

为了解大模型应用在行业内的基本情况，我们对业内包括头部券商及部分金融机构在内的机构进行了摸底调研。结果指出，行业整体虽处于积极探索期，但在安全管控上普遍存在不足。

大模型应用情况与场景：绝大多数机构（超过90%）已在非核心业务中试水大模型应用，但核心业务的渗透仍极为审慎。应用场景高度集中于四大领域：智能客服与营销（100%提及），作为降本增效的直接入口；投资研究与分析（70%提及），辅助处理海量信息；内部知识管理（70%提及），提升内部运营效率；软件开发与IT运维（70%提及），赋能科技部门。这反映出当前应用仍以“辅助”和“提效”为主，尚未深度触及交易与风控环节。

核心驱动力：推动大模型应用的根本动力源于战略布局及竞争焦虑。抢占技术高地、探索新业务模式的战略布局（50%提及）和来自管理层的战略要求（40%提及）是两大主因。这表明，大模型的应用已是“一把手工程”，安全部门必须跟上战略步伐，提供有效支撑。

风险共识：在风险认知上，全行业高度一致，数据安全风险被视为压倒性的首要挑战（100%提及），对核心数据资产的泄露风险抱有极大忧虑。其次是合规风险（70%提及），担心生成内容无法满足日益严格的监管要求。模型自身风险（60%提及）和应用安全风险（40%提及）同样是关注焦点。在应对手段上，“电子围栏”“内容审核”等传统思路被频繁提及，但体系化的防御能力普遍有所欠缺。

（二）大模型应用的主要安全风险剖析

1、合规与法律风险

首先，模型输出内容一旦触碰监管红线，例如生成包含“承诺收益”的营销话术，或传播未经证实的市场谣言，将直接触发监管红线。其次，内容合规是硬性要求，任何违背社会主义核心价值观的输出，都将引发严重的品牌声誉和法律风险。最后，数据跨境、个人信息保护等均有明确法律规定，需要按照要求妥善处理。

2、数据安全风险

在模型训练与微调阶段，若使用包含客户身份、交易、持仓等信息的业务数据，或敏感数据脱敏不彻底的，将在大模型应用中埋入“地雷”，随时可能在后续应用中被引爆。在

模型使用阶段，风险敞口更大：一方面，员工可能在日常工作中将未公开的投研报告、并购项目资料、自营盘策略代码等高价值敏感信息输入给互联网大模型；另一方面，模型自身也可能被恶意攻击者通过巧妙提问，诱使其吐出训练数据中的敏感信息。

3、模型自身安全风险

大模型自身存在一定的应用风险，一是“模型幻觉”问题，在投研、客服等场景下，看似合理的错误信息可能直接误导投资决策或引发客户纠纷。二是偏见与公平性问题，模型可能固化甚至放大训练数据中的偏见，导致歧视性的客户服务或产品推荐，引发合规与公平性问题。三是滥用与对抗性攻击问题，攻击者可利用提示注入（Prompt Injection）等高级攻击，能直接穿透应用层的传统防护，是必须严防死守的新技术风险。

4、应用及基础设施安全风险

大模型应用并非空中楼阁，其依赖的软件技术栈同样存在风险。从底层的PyTorch、TensorFlow等深度学习框架漏洞，到上层的API接口权限控制不当、速率限制缺失，再到Web应用本身存在的各类传统漏洞，构成了完整的攻击面。我们需要将大模型应用纳入现有应用安全管理体系，补充针对性的安全要求和测试用例进行测试，确保大模型应用系统本身的安全性。

三、大模型安全防护目标与策略

基于以上对大模型的安全风险分析，从输入输出环节的敏感数据泄露、违规内容生成，到模型层面的提示词注入攻击、训练数据投毒，再到应用系统的稳定性风险，这些问题不仅可能导致不符合监管要求，更会直接影响业务正常运转与用户信任。在此背景下，亟需明确清晰的安全防护方向与路径，为大模型安全应用提供指引，因此确立如下总体安全目标与策略：以合规性为根本底线，以业务赋能为核心目标，构建“底线不可破、价值能落地”的大模型安全防护体系，核心确立两大支柱策略。

内容安全：对模型全流程处理（输入、加工、生成、输出）的内外部信息实施闭环管控，结合多模态内容审核技术，确保信息合法合规、敏感数据（如业务机密、用户隐私）零泄露，杜绝违规内容生成与传播。

模型安全：保障模型本身（算法、参数、训练数据）及配套应用系统的安全性，能够有效抵御已知攻击（如提示词注入、数据投毒）与未知风险（如模型逃逸、隐性偏见滥用），确保模型服务持续稳定、功能可靠。

为落实上述策略，主要抓手是建立“上线前的安全测试”和“运行时的实时防护”相结合的闭环管理机制，具体落地为“大模型安全测试”和“大模型安全网关”两大核心能力建设。

四、大模型安全防护体系设计

为实现以上目标，充分应对挑战，我们参考国际主流AI安全框架、行业主流大模型安全服务商实践经验并结合自身情况，设计了一套覆盖大模型应用全生命周期的安全防护体系。

（一）国际主流AI安全框架参考

构建有效的大模型安全防护体系，需要借鉴业界成熟的安全框架，以便形成我们的设计思路，国际主流AI安全框架如下：

NIST AI RMF(美国国家标准与技术研究院AI风险管理框架)：该框架提供了顶层治理的抓手。其“治理-映射-测量-管理”的闭环流程，指导我们如何建立权责清晰的AI风险治理结构，如何系统性地识别、评估和管控风险，确保安全工作与业务目标对齐。

OWASP Top 10 for LLMs(开放式Web应用安全项目LLM十大风险)：这是应用安全的行动指南。它将大模型应用面临的主要技术风险进行了清晰归纳，为我们制定安全开发规范（SDL）、开展代码审计和渗透测试用例提供了具体、可落地的技术参考。

MITRE ATLAS(针对AI系统的对抗性威胁知识库)：该框架是红队与威胁情报工作的“弹药库”。它从攻击者视角出发，系统梳理了针对AI的战术、技术和流程，帮助我们模拟真实的攻击，检验防护体系的有效性，做到“知己知彼”。

（二）行业主流安全供应商能力调研

构建大模型整体安全防护能力，依靠证券机构自身的能力还不够全面，还需要借助行业优秀安全服务供应商的能力。在对行业优秀供应商进行技术能力摸底后，我们认为各家供应商能力各有所长，要构建完善的大模型安全体系需要整合多家供应商的优势。

部分供应商的优势在于金融场景的深度实践和大规模评测能力，其百万级的测试场景库和与国家标准的对齐能力，对我们确保合规性有重要参考价值。另一部分供应商的核心竞争力是其强大的内容安全能力和顶尖的红蓝对抗服务，其多年的敏感内容治理经验和多模态识别技术是保障内容合规的有力支撑，专业的攻击队能对我们的模型和应用进行最严格的实战检验。

同时，一些传统内容安全领域的资深供应商，在语料库积累和动态策略运营方面经验丰富，尤其在一些细分内容风险的识别上有独到之处，可作为我们内容审核能力的补充。还有的供应商特色在于高度定制化的评测服务和专业

的备案支持，能帮助我们更高效地满足监管要求。此外，也有供应商产品成熟度高，具备大规模实时系统的工程能力，其方案经过了海量业务的验证，可实现快速部署，适合快速上线、验证效果的初期项目。

（三）大模型应用全生命周期安全风险及防控措施

在充分了解大模型的的安全风险及参考行业最佳实践后，我们制定了以下大模型应用全生命周期安全风险及防控措施框架。

	大模型应用生命周期安全风险及对应措施													
	模型预训练		模型微调		模型部署		应用开发		应用测试		应用备案		应用运行	
	安全风险	安全措施	安全风险	安全措施	安全风险	安全措施	安全风险	安全措施	安全风险	安全措施	安全风险	安全措施	安全风险	安全措施
内容安全	记忆敏感信息	数据脱敏		生成信息标识					篡改暴恐+社会主义价值观+偏见	大模型安全测试	对照网信办《生成式人工智能服务管理办法》	大模型服务备案	篡改暴恐+社会主义价值观+偏见	大模型安全网关
数据层面	训练数据泄露	数据脱敏	微调数据泄露	数据脱敏									敏感数据泄露：回应/展示	应用数据审计系统
模型安全	数据和模型投毒	大模型安全测试	模型投毒		模型窃取/泄露				OWASP LLM TOP 10；NVIDIA AI Red teaming；	大模型安全测试			提示注入、系统提示泄露	大模型安全网关
应用层面	PyTorch 工具漏洞	渗透测试	RAG投门	渗透测试	AI部署工具漏洞	HIDS	代码漏洞	代码审计 SAST	通用应用安全漏洞	IAST			web应用漏洞	应用安全防火墙
					CUDA Toolkit 漏洞	大模型安全测试	供应链漏洞	SCA+Xray	API滥用/权限过度代理	手工渗透			提示词注入、无限制消耗+高频调用	大模型安全测试
网络层面					网络分区/访问控制	权限最小化			测试环境暴露面				互联网暴露面	
物理层面					计算资源消耗攻击									

图1 大模型应用全生命周期安全风险与防控措施

大模型应用全生命周期各阶段可以采取的安全措施主要为以下：

模型预训练与微调阶段：此阶段的核心是数据安全治理。建立严格的数据分类分级和审批流程，所有用于训练的数据经过彻底的清洗和脱敏，从源头上杜绝敏感信息的“注入”。

模型部署阶段：此阶段是连接系统研发与运行的关键环节，重点在于确保模型资产与运行环境的安全性。为了应对AI部署工具及底层CUDA工具链的潜在漏洞，需部署主机入侵检测系统（HIDS）对服务器进行持续安全监测。为严防模型被窃取或泄露，必须遵循权限最小化原则，通过严格的网络分区和访问控制策略，确保只有经过授权的应用能够访问模型接口。

应用开发与测试阶段：此阶段是漏洞发现与合规验证的关键所在。除了常规的代码审计（SAST）、供应链组件分析（SCA），还需要开展大模型专项安全测试。该测试需由安全检测专业技术人员主导，模拟各类对抗性攻击，并对模型输出进行多维度的安全评估，不通过则不允许上线。

应用备案与运行阶段：此阶段是持续监控与纵深防御的重要防线。首先，需要按照监管要求完成大模型服务备案。其次，所有大模型应用的流量统一收口于大模型安全网关，实现实时监控、检测和阻断。

（四）大模型安全架构设计

结合大模型通用技术架构及我司的实际情况，我们将大模型架构设计为数据平台、模型服务、开发平台、应用形态、应用场景等五层。基于纵深防御思想以及上述大模型应用全生命周期安全风险与防控措施，需要在架构的每个层级部署实施相应的安全防御措施，其中在数据平台、应用形态、应用场景层可复用传统的安全措施如数据脱敏、数字水印等，这里不做过多赘述。在模型服务和开发平台方面，基于大模型应用的特点，需要进行针对性的安全措施加强，具体说明如下：

在开发平台层面，如上文所述，需要在内容安全和模型安全方面进行专项的测试，测试通过后方可进行上线运行。

在模型服务层面，需要对大模型的输入和输出进行安全控制，设计使用“大模型安全网关”来承担此重要防护功能。大模型安全网关是大模型应用运行实时风险管控的核心关卡，其输入和输出侧防护功能主要为以下：

输入侧防护：作为第一道防线，对所有用户输入进行深度检测，精准识别并过滤提示注入、越狱等恶意指令。同时，强制执行隐私数据扫描和脱敏策略，严防内部敏感信息被发送至模型。

输出侧防护：作为最后一道关卡，对模型返回的所有内容进行全面审查，拦截违背社会主义核心价值观、行业违规（如承诺收益）等内容，并对可能泄露的敏感数据进行二次屏蔽。对于不合规的输出，执行拦截、改写或安全提示等策略。

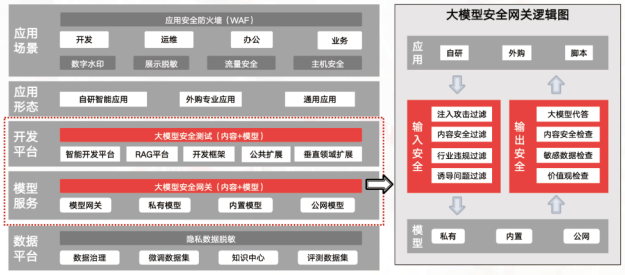


图2 大模型应用安全架构

五、成效与收益

通过实施上述大模型安全防护体系，在关键节点嵌入有效的防御措施，实现从被动响应到主动防御的转变，具体成效和收益体现在：

筑牢坚实的合规壁垒：体系化、全流程的内容审核与备案流程，将确保我们的大模型服务全面契合监管要求，切实规避合规风险，为业务创新提供稳定的政策环境。

构筑可信的数据防线：全流程的数据安全治理与网关的实时过滤，将极大降低核心数据资产的外泄风险，保护公司和客户的合法权益。

建立主动的威胁应对能力：安全测试与安全网关的协同，将使我们具备有效防御对抗性攻击的能力，保障模型服务的健壮性，防止其被恶意利用。

打造安全的创新基座：统一的安全基座将成为公司AI创新的“加速器”。无需在安全问题上重复投入，可以更聚焦于业务逻辑，从而在保障安全的前提下，更快、更好地推动数字化转型。

六、总结与展望

大模型已成为证券行业数字化转型的核心引擎，在优化投研效率、升级客户服务体验、提升办公效率等方面展现出显著赋能价值，但数据泄露、合规要求不达标、模型滥用等风险也成为其规模化应用的主要瓶颈。对此，本文结合实践构建了覆盖大模型全生命周期的安全防御体系，通过“上线时安全测试”前置风险筛查、“运行时安全网关”强化过程管控，协同实现输入输出内容可控与模型攻击可防，有效破解安全合规难题，为行业在风险可控前提下推动LLM从试点走向规模化落地提供了兼具理论与实操价值的路径。

未来，大模型在证券行业的安全应用需围绕技术、应用、行业三大维度进行突破：技术层面将零信任架构嵌入安全网关、结合联邦学习优化数据处理，进一步适配行业数据安全需求；应用层面依托成熟安全体系向智能风控、合规审计等核心业务延伸，同步动态更新防护规则保障业务合规；行业层面由监管机构牵头，联合头部券商与科技公司建立安全协作机制，制定应用标准与应急指南、共享威胁情报，同时持续关注深度伪造信息、模型偏见等新型风险，动态优化防御策略，最终实现大模型与证券业务的深度安全融合，为行业高质量数字化转型注入持久动力。

参考文献

1.国家互联网信息办公室. (2023). 生成式人工智能服务管理暂行办法.

2.国家互联网信息办公室. (2021). 互联网信息服务算法推荐管理规定.

3.国家互联网信息办公室. (2022). 互联网信息服务深度合成管理规定.

4.中华人民共和国. (2021). 中华人民共和国数据安全法.

5.OWASP Foundation. (2023). OWASP Top 10 for Large Language Model Applications.

6.National Institute of Standards and Technology (NIST). (2023). Artificial Intelligence Risk Management Framework (AI RMF 1.0).

7.MITRE Corporation. (2023). MITRE ATLAS (Adversarial Threat Landscape for Artificial-Intelligence Systems).

证券行业安全建设：构建智慧安全生态，护航业务稳健发展

王洪涛 | 国金证券股份有限公司

摘　要： 在数字化转型背景下，证券行业科技管理工作和系统建设秉持以“融合业务、平台赋能”的科技愿景，以“业务导向、高效敏捷、统一管理、稳进创新”的科技理念，结合证券业务“高实时性、高并发性、高复杂性、高价值性”的核心特性，落实做好企业网络安全体系建设。通过“网络 + 开发 + 安全 + 运营”协同防御打破数据与流程壁垒，利用数字孪生技术赋能全链路风险治理（优化 RTO/RPO、实现故障快速恢复与数据精细化防护），以长期安全规划锚定信创适配、抗量子密码等前瞻性方向，同时平衡 AI 技术赋能与风险防控（加固模型安全、优化基础设施、明确人机责任）；最终提出构建“预测 - 防护 - 检测 - 响应 - 优化”闭环的智慧安全生态，旨在保障业务连续性、降低安全事件损失、满足监管合规要求，支撑证券行业稳健发展。

关键字： 协同防御、安全韧性、数据全生命周期治理、隐私计算、大语言模型（LLM）、业务导向、AI模型安全、智慧安全生态

一、引言

证券行业作为金融体系的核心组成，直接关联巨额资金流转与市场信心，其“高实时性（毫秒级交易延迟即致损失）、高并发性（极端行情下海量访问）、高复杂性（多系统协同）、高价值性（牵动市场稳定）”的业务特性，决定了网络安全是行业运行的生命线。当前，随着数字化转型深化，AI、量化交易、云架构等技术广泛落地，网络攻击呈现“高级化、隐蔽化、自动化”趋势，传统单点防御已难以应对新型风险——亟需适配信创改造、抗量子密码等长期技术变革与强监管要求，还需抵御勒索软件、APT 攻击等外部威胁，解决 AI 模型安全、数据泄露、系统韧性不足等内生问题。在此背景下，构建系统性、前瞻性、动态化、有韧性的安全建设体系，成为证券行业抵御风险、保障业务连续、实现高质量发展的关键支撑，梳理证券行业安全建设的核心框架与实践路径。

二、构建“网络+开发+安全+运营”协同的体系化安全防御

证券业务的本质特性决定了其对网络安全近乎苛刻的要求：高实时性（毫秒级交易延迟可能导致巨大损失）、高并发性（极端行情下海量用户并发访问）、高复杂性（涉及交易、清算、风控、客户服务等多系统协同）以及高价值性（直接关联巨额资金与市场信心）。面对日益高级化、隐蔽化、自动化的网络攻击需要构建“网络 + 开发 + 安全 + 运营”协同的体系化安全防御。

（一）打破“数据孤岛”与“流程壁垒”

整体安全的核心是围绕“业务连续性”目标，打通网络防护、开发安全（DevSecOps）、安全运营（SOC）、运维管理（ITOps）的协同链路，实现“安全左移至开发环节、风险右防至运营末端”，避免单一环节防护脱节导致的整体防御失效。例如，开发阶段嵌入安全编码规范与自动化漏洞扫描，运维阶段联动威胁情报与实时流量监测，网络层实现“终端 - 链路 - 云端”一体化防护，形成“全流程、全要素”的安全闭环。

（二）锚定行业痛点的落地策略

协同机制建设，“威胁情报共享中心”模式，建立跨部门安全联防联动组，明确网络、开发、安全、运维团队的职责边界——如开发团队负责代码安全、运维团队保障基础设施韧性、安全团队统筹风险调度，通过“定期协同会议 + 自动化工单系统”实现高效联动。

技术底座整合，基于信创体系构建统一安全管理平台，集成攻击面管理、漏洞管理、威胁感知、挂图作战等功能，实现“一次采集、多端复用”，降低安全工具的重复投入与数据割裂问题。低侵入式安全设计，在核心交易系统中嵌入“安全能力接口”，如通过 API 调用统一安全服务，避免对业务流程的改造冲击，平衡“安全防护强度”与“业务运行效率”。

（三）从静态防御到智能预测

动态威胁检测与响应（DTR），大幅缩短威胁检测与响应时间（MTTR），最大限度遏制攻击影响，降低财产与声誉损失。通过AI驱动自动化响应可将证券业安全事件MTTR缩短60%以上。投资回报率（ROI）、运营效率提升、降低安全人力成本、满足监管对响应时效的要求。需整合现有工具，打

破数据孤岛，构建统一安全视图。安全知识图谱（SKG），整合碎片化安全数据，实现跨系统深度关联分析，快速评估事件影响范围、还原攻击链条、精准定位风险源头，为决策提供全局视角。提升风险管理透明度，支撑快速、精准的危机决策，满足监管对风险全面掌控的要求，是提升整体安全成熟度的关键基础设施。

零信任架构摒弃过时的“信任但验证”模型，在复杂混合IT环境中实施“永不信任，持续验证”。增强身份认证和访问控制，根据业务需求和风险评估，实施精细化数据访问控制，促进证券行业业务和数据安全创新，增强抵御内外部威胁能力，确保系统安全稳定，降低泄露和误操作风险。零信任架构满足日益严格的访问控制监管要求，是防范内部威胁和高级渗透的有效手段，平衡安全性与用户体验。

三、以安全韧性为核心，数字孪生赋能全链路风险治理

（一）从“风险抵御”到“快速恢复”

安全韧性的核心是提升系统在威胁冲击下的“可恢复性”— 通过优化RTO（恢复时间目标）、RPO（恢复点目标），结合数字孪生技术的“模拟 - 验证 - 回溯”能力，实现“威胁可观测、风险可验证、故障可回溯”，避免单一故障引发的业务中断（如核心交易系统宕机、重要系统数据泄露）。

通过数字孪生构建“安全场景数字镜像”—— 如模拟APT 攻击、勒索软件入侵的全流程，验证安全防护措施的有效性；或复刻核心交易系统架构，实现故障发生时的“秒级定位、分钟级恢复”，弥补传统安全验证“成本高、覆盖窄”的短板。

（二）技术与流程双轮驱动

构建“多活数据中心 + 异地灾备”架构，通过自动化灾备切换工具将RTO缩短至5分钟以内，RPO控制在秒级；引入BAS（安全有效性验证）技术，常态化模拟漏洞利用、恶意代码传播场景，验证防火墙、WAF、EDR等设备的防护有效性，定期输出“韧性优化报告”。

构建“网络资产数字孪生平台”，映射路由器、服务器、安全设备的实时状态，结合流量日志、告警数据实现“攻击路径可视化”，如通过孪生模型回溯钓鱼攻击的“邮件投递 - 终端入侵 - 数据外泄”全链路；在开发阶段搭建“业务系统孪生环境”，模拟高并发交易场景下的安全风险（如 API 越权、数据脱敏失效），提前发现业务逻辑漏洞。

（三）全生命周期的数据安全精细化防护

依托人工智能技术，实现数据“识别 - 分类 - 管控 - 监控”全流程精细化、动态化管理，解决传统方式“效率低、覆盖窄、难持续”的短板。智能数据识别，以 AI（大语言模型、视觉模型等）替代传统人工与规则识别，大幅提升敏感数据识别的准确性（可识别业务语义层隐含敏感信息）、覆盖范围（从结构化拓展至非结构化数据）与效率（全域梳理周期从“半年级”缩至“月 / 周级”），并生成敏感数据图谱。自动化打标与持续运营，AI 学习法规与行业标准，结合数据属性、监管要求等完成分类分级评分，高敏数据人工复核；同时动态监测数据敏感属性变化，自动触发重评与策略优化，定期输出“敏感数据分布与变动趋势报告”，避免传统集中梳理的低效问题。分级分类下的精细化管控，差异化加密，按数据安全等级（极高 / 高 / 中 / 低）实施分级加密，兼顾防护强度与业务效率。精细化访问控制，基于“数据等级 + 业务场景 + 用户角色”配置最小权限，AI 辅助识别异常访问并动态响应，所有行为可审计。静 / 动态脱敏，静态脱敏生成不可逆副本，动态脱敏实时按需处理（如“可用不可见”），AI 辅助调整策略。数据安全行为监控，实时分析AI识别大批量导出、跨境传输等高风险操作，结合多维度评分实时拦截（如锁账户、强阻断）。事后审计，分析访问异常（如越权、脱敏绕过），动态优化安全策略，AI辅助生成风险报告，形成风险管理闭环。

四、以安全规划为引领，锚定长期风险与技术变革

（一）统一“战略 - 技术 - 合规”的长期路径

安全规划的核心是结合证券行业特性（如高敏感数据、强监管要求、业务创新性），制定“3-5 年安全战略蓝图”，明确技术架构统一方向（如信创适配、抗量子密码布局）、组织能力建设目标（如复合型安全人才培养）、合规治理路径（如数据分类分级、隐私计算应用），避免“短期投入无序、长期能力断层”，也保证整体架构的稳定性和可扩展性，降低运维成本。

（二）构建标准化信息安全管理体系

在制定长期安全战略时，积极引入ISO27001信息安全管理体系认证，作为标准化、系统化的安全治理框架。ISO27001以风险管理为核心，通过“计划 - 实施 - 检查 - 改进”（PDCA）的循环模式，帮助证券机构建立全面的信息安全管理体系，涵盖网络安全、数据保护、业务连续性等多个维度。通过ISO27001认证，证券机构不仅可以提升内部安全管理水平，增强客户信任度。同时，ISO27001的持续改进机制与证券行业动态风险防控的需求高度契合，能够为长期安全规划提供坚实的实践基础，助力构建更加稳健、合规的安全生态体系。

（三）聚焦三大前瞻性方向

一是技术架构统一规划，信创安全方面，制定“服务器 - 操作系统 - 数据库 - 安全工具”的全栈信创适配路线图，完成核心交易系统的信创改造，同步嵌入商密算法

(SM2/SM3/SM4)；抗量子密码布局，跟进 NIST 后量子密码标准，在客群身份认证、交易签名等场景试点 CTRU、Kyber 等算法，避免“量子计算破解传统加密”的前向安全风险。二是合规与风险一体化规划，建立“合规 - 风险”映射体系，将《数据安全法》《证券期货业网络安全管理办法》要求拆解为“数据采集 - 传输 - 存储 - 销毁”各环节的安全控制点，如客户身份信息采用“口令 + 三方协同鉴别”，满足三级等保强身份认证要求；布局大模型安全，针对投研大模型、智能客服大模型，制定“训练数据脱敏 - 提示词注入防护 - 输出内容审计”流程，避免敏感交易策略、客户信息泄露。三是组织能力规划，培养“安全 + 业务 + AI”复合型人才，通过“内部培训 + 外部认证”提升团队对大模型安全、量子安全的认知；建立“安全创新实验室”，联合行业机构（如 ITRDC、证券期货行业网络安全创新实验室）开展技术预研，提前储备数字孪生、隐私计算等前沿安全技术。

（四）技术赋能与风险防控并重

AI 技术（如 LLM、MOE）已在证券行业智能客服、投研、量化交易等场景落地，但模型安全、合规等问题成新风险，需在技术赋能与风险防控间平衡。一是模型安全加固抵御恶意攻击，增强模型抗对抗攻击、数据投毒能力，防核心算法 / 模型窃取，严格控制模型访问与输入。保障交易、风控等核心业务中 AI 可靠运行，避免模型漏洞致损失，保护核心知识产权。二是AI 基础设施优化保障高效可靠运行，通过模型优化（量化、剪枝等）、边缘部署及 MLOps 体系，确保 AI 系统高效稳定，实时监控性能与数据漂移。满足欺诈检测等业务的高实时性需求，控制成本，保障系统稳定，实现 AI 投资可持续回报。三是人机协同与责任机制明晰边界，发挥各自优势，界定 AI（初级监控、标准化响应）与人类（复杂决策、责任认定）边界，培养复合型人才，提升风控 / 合规类 AI 模型可解释性，建立模型审计机制。明确“AI 为工具、责任在人”，满足监管对算法透明与审计的要求，规避算法偏见风险，解决人才瓶颈。

五、未来展望：构建智慧安全生态

（一）前沿技术融合创新

■量子安全密码学：评估并规划向抗量子密码算法迁移，探索量子密钥分发（QKD）在核心链路应用，应对未来量子计算威胁，保障长期安全。

■区块链与AI深度协同：构建基于区块链的可信数据共享平台（威胁情报、反欺诈、KYC核验），利用区块链实现AI模型训练与决策的透明审计，增强信任与合规性。

■智能安全决策中枢：推动SOC向整合网络、数据、AI风险的多维智能决策中心进化，实现风险自适应的动态防御体系。

（二）安全体系持续进化

践行“预测（Predict）— 防护（Protect）— 检测（Detect）— 响应（Respond）— 优化（Optimize）”（PPDRO）的闭环治理模式，与中国证券业协会《三年提升计划》的“主动、动态、纵深防御”理念深度契合。将AI深度融入PPDRO各环节，驱动安全运营智能化、自动化变革。

（三）安全文化、流程与人才

■安全左移：将安全要求嵌入系统设计、软件开发（DevSecOps）、数据架构（DataSec by Design）、AI模型开发（MLSec）的源头。

■持续风险管理：建立常态化风险评估机制，紧密跟踪监管政策（如生成式AI监管细则）变化，确保持续合规。

■打造敏捷安全团队：投资培养兼具技术深度（云、大数据、AI）、业务理解力（交易、风控、合规）和战略视野的复合型安全人才队伍，建立持续学习机制。

六、结语

证券行业安全建设并非单一技术的堆砌，而是“技术架构 + 流程机制 + 组织能力 + 文化理念”深度融合的系统工程。从“网络 + 开发 + 安全 + 运营”的协同防御，到数字孪生赋能的全链路风险治理，再到锚定信创、抗量子密码的长期规划，其核心逻辑始终是“以业务连续性为目标，以动态风险防控为核心，以技术创新为驱动”—— 既需通过 AI、数字孪生等技术提升防御与恢复能力，又需通过流程优化与责任界定规避技术应用风险，更需通过常态化运营与人才培养保障体系落地。未来，随着量子计算、区块链等前沿技术的融合，证券行业安全建设需进一步践行“PPDRO 闭环治理”理念，推动安全体系从“被动防御”向“智慧预测”进化，最终构建兼具韧性、合规性与创新性的智慧安全生态。这不仅是抵御当前风险的必然选择，更是证券行业在技术变革与市场波动中，实现长期稳健发展的根本保障。

数字化时代基金公司的网络安全

聂连杰 | 鹏华基金管理有限公司

摘　要： 随着数字化时代的来临，基金行业加速数字化转型，网络安全对于基金公司的稳健运营愈发关键。证券期货业网络安全相关法律法规的不断完善，《网络安全法》、《证券期货业网络和信息安全管理办法》等构建了严格的监管框架，以应对日益复杂的网络安全挑战。基金公司面临着数字化转型带来的诸多挑战，如数据安全风险增加、新技术应用带来的管理复杂度提升等，但同时也迎来利用先进技术强化安全防护的机遇。本文阐述鹏华基金管理有限公司在网络安全事前需完善系统设计、加强数据治理；事中通过应用网络安全技术进行实时监测；事后开展网络安全应急及舆情管理等措施。旨在构建全方位、多层次的基金公司网络安全管理机制，根据不断变化的网络环境，持续优化安全策略，为基金公司在数字化浪潮中筑牢安全屏障，确保投资者利益和金融市场稳定。

关键字： 网络安全、数字化转型、金融市场稳定、信息安全建设、安全管理流程、大数据、人工智能

一、引言

在数字化时代，基金行业正经历着深刻变革，数字化转型成为行业发展的关键驱动力。基金公司的业务运营愈发依赖信息技术，从交易系统、客户服务到风险管理等各个环节，信息技术的广泛应用显著提升了运营效率和服务质量，拓展了业务边界。然而，这种高度数字化的运营模式也使基金公司面临前所未有的网络安全挑战。

网络安全威胁的演变日益复杂和多样化，从传统的恶意软件、网络钓鱼攻击，到新兴的勒索病毒、人工智能驱动的攻击手段等，给基金公司的信息系统安全、数据安全和业务连续性带来了巨大风险。一旦发生网络安全事件，不仅会导致基金公司的财务损失、声誉受损，还可能引发投资者的信任危机，对整个金融市场的稳定造成冲击。

同时，证券期货业网络安全法律法规体系不断完善，监管要求日益严格。《证券期货业网络和信息安全管理办法》等一系列法规的出台，明确了基金公司在网络安全管理方面的责任和义务，要求基金公司建立健全网络安全管理机制，加强信息系统安全防护、数据保护和应急处置能力。在这样的背景下，公司肩负着至关重要的职责，需要从战略高度出发，制定全面有效的网络安全策略，确保基金公司在数字化时代的安全稳健运营。

二、数字化时代基金公司面临的网络安全形势

（一）数字化转型对基金公司的影响

1、业务模式的变革

数字化转型促使基金公司的业务模式发生了根本性变革。线上交易平台的广泛应用，使得投资者可以随时随地进行基金申购、赎回等操作，极大地提高了交易的便捷性和效率。智能投顾服务借助大数据分析和人工智能算法，为投资者提供个性化的投资建议，满足了不同投资者的多样化需求。基金销售渠道也逐渐从传统的银行、券商等线下渠道向互联网平台拓展，拓宽了客户群体和市场覆盖范围。基金公司与互联网金融平台合作，推出了创新的基金产品销售模式，吸引了大量年轻投资者。

2、信息技术的广泛应用

基金公司在数字化转型过程中，广泛应用了大数据、云计算、人工智能等先进信息技术。大数据技术用于收集、分析海量的市场数据、客户数据和交易数据，为投资决策提供支持，帮助基金经理更好地把握市场趋势和投资机会。云计算技术为基金公司提供了灵活、高效的计算资源和存储服务，降低了信息系统建设和运维成本，提高了系统的扩展性和可用性。人工智能技术在风险预警、客户服务等方面的应用，提升了基金公司的风险管理能力和客户体验。公司利用人工智能客服，能够快速响应客户咨询，解答常见问题，提高了客户服务效率。

（二）网络安全威胁的演变

1、新型网络攻击手段的出现

随着信息技术的发展，网络攻击者的手段也不断创新，出现了许多新型网络攻击手段。勒索病毒攻击近年来呈现出高发态势，攻击者通过加密公司的重要数据，索要赎金，给金融行业带来巨大的经济损失和业务中断风险。如某公司遭受勒索病毒攻击，公司核心业务数据被加密，导致交易系统瘫痪数日，不仅面临高额赎金支付压力，还对公司声誉造成了严重损害。人工智能驱动的攻击手段也逐渐兴起，攻击者利用人工智能技术分析公司网络系统的弱点，制定针对性的攻击策略，使得攻击更加精准和难以防范。

2、攻击目标的多样化

网络攻击目标不再局限于基金公司的交易系统和客户数据，而是呈现出多样化的趋势。除了传统的窃取客户资金、篡改交易数据等目标外，攻击者还将目光投向了公司的知识产权、商业机密和战略规划等重要信息。同时随着基金公司与第三方服务提供商的合作日益紧密，第三方服务提供商的网络安全漏洞也可能成为攻击者入侵基金公司的跳板，导致供应链安全风险增加。

（三）证券期货业网络安全法律法规及监管要求

1、相关法律法规概述

为了规范证券期货业网络和信息安全管理，保障金融市场稳定运行，我国出台了一系列相关法律法规。《证券期货业网络和信息安全管理办法》明确了证券期货业机构在网络和信息安全管理方面的基本要求，包括安全管理体系建设、信息系统安全防护、数据安全保护、应急处置等方面的规定。《网络安全法》强调了网络运营者的安全义务和责任，要求采取技术措施和其他必要措施，保障网络安全、稳定运行。《数据安全法》对数据安全保护进行了全面规范，明确了数据分类分级管理、数据安全风险评估、数据安全应急处置等制度。《个人信息保护法》则着重保护个人信息权益，规范个人信息处理活动，对基金公司在收集、使用、存储和传输投资者个人信息等方面提出了严格要求。

2、监管要求对基金公司的影响

严格的监管要求促使基金公司加大在网络安全方面的投入，完善网络安全管理体系。基金公司需要建立健全网络安全组织架构，明确各部门和人员的安全职责，制定完善的网络安全管理制度和流程。在信息系统建设和运维过程中，要遵循相关安全标准和规范，加强系统安全防护，定期进行安全评估和漏洞扫描。对于投资者个人信息保护，基金公司需采取严格的数据加密、访问控制等措施，确保个人信息的安全。同时，监管部门还加强了对基金公司的监督检查，对违反网络安全法律法规的行为进行严厉处罚，这进一步促使基金公司重视网络安全工作，不断提升网络安全管理水平。

三、网络安全事前管理

（一）系统设计阶段的安全考量

1、安全架构设计原则

结合 SDLC（软件开发生命周期）的信息系统安全架构设计，需牢牢遵循“安全左移”“持续验证”“内生防御”三大核心原则，确保安全贯穿系统从需求到运维的全生命周期，开展工作需按阶段层层推进、深度落地：在需求分析阶段，不能仅停留在业务需求梳理，需同步联合业务、安全、技术团队开展安全需求挖掘，明确数据分级保护要求、权限管控规则，同时通过威胁建模识别数据泄露、注入攻击、权限越权等潜在威胁，最终形成包含安全目标、风险等级、防护指标的《安全需求规格说明书》，确保安全需求与业务需求同步立项、同等重视；进入设计阶段，需以三大原则为指导，将安全架构融入系统整体设计。

2、安全漏洞预防机制

建立安全漏洞预防机制是系统设计阶段的重要任务。在系统开发过程中，采用安全编码规范，要求开发人员避免使用存在安全风险的代码函数和编程习惯，减少因代码漏洞导致的安全隐患。例如，禁止使用容易引发缓冲区溢出的函数，对用户输入数据进行严格的校验和过滤，防止 SQL 注入、跨站脚本攻击等常见漏洞。同时，引入安全测试工具，在系统开发的不同阶段进行安全测试，如静态代码分析、动态漏洞扫描等，及时发现并修复安全漏洞。在系统上线前，进行全面的安全评估和渗透测试，模拟黑客攻击场景，检验系统的安全防护能力，确保系统在上线前不存在重大安全漏洞。

（二）数据治理与安全

1、数据分类分级管理

基金公司拥有大量的业务数据和客户数据，对这些数据进行分类分级管理是数据安全保护的基础。根据数据的重要性和敏感性，将数据分为不同的类别和级别，如公开数据、内部敏感数据、客户机密数据等。对于客户机密数据，如身份证号码、银行卡信息等，列为最高级别进行重点保护。针对不同类别的数据，制定相应的数据管理制度和安全策略，明确数据的存储、传输、使用和销毁等环节的安全要求。例如，客户机密数据在存储和传输过程中必须采用加密技术，确保数据的保密性；内部敏感数据的访问需经过严格的审批流程，限制访问人员范围。

2、数据加密与访问控制

数据加密是保障数据安全的重要手段。基金公司应对重要数据，特别是客户敏感数据和交易数据，在存储和传输过程中进行加密处理。在存储环节，采用数据库加密技术，对数据库中的敏感字段进行加密存储，防止数据在存储介质被窃取时泄露。在传输环节，使用 SSL/TLS 等加密协议，确保数据在网络传输过程中的安全性。同时，加强数据访问控制，建立完善的用户身份认证和授权管理体系。采用多因素身份认证方式，如密码、短信验证码、指纹识别等，提高用户身份认证的准确性和安全性。根据用户的角色和职责，为其分配相应的数据访问权限，实现对数据的细粒度访问控制，确保只有授权用户才能访问特定的数据。

（三）安全管理元素的融入

1、安全管理制度的建立与完善

建立完善的安全管理制度是基金公司网络安全管理的重要保障。制定涵盖网络安全各个方面的管理制度，包括信息系统安全管理、数据安全管理、人员安全管理、应急处置管理等。明确各部门和人员在网络安全工作中的职责和义务，确保网络安全工作的顺利开展。信息技术部门负责信息系统的日常安全运维，风险管理部门负责网络安全风险评估和监测，人力资源部门负责员工安全培训等。同时，根据法律法规的变化和公司业务的发展，及时对安全管理制度进行修订和完善，确保制度的有效性和适应性。

2、员工安全意识培训

员工是基金公司网络安全的第一道防线，加强员工安全意识培训至关重要。定期组织员工参加网络安全培训课程，向员工普及网络安全知识和技能，提高员工的安全意识和防范能力。培训内容包括网络安全法律法规、常见网络攻击手段及防范方法、数据保护意识、安全操作规范等。通过案例分析、模拟演练等形式，让员工深刻认识到网络安全的重要性，增强员工的安全责任感。例如，通过模拟网络钓鱼攻击场景，让员工亲身体验网络钓鱼的危害，提高员工识别和防范网络钓鱼邮件的能力。同时，建立员工安全行为考核机制，将员工的安全行为纳入绩效考核体系，激励员工自觉遵守公司的安全管理制度。

四、网络安全事中管理

（一）网络安全监测技术的应用

1、纵深防护体系建设

在数字化业务场景下，基金公司需构建多层级、全链路的纵深防御体系，整合态势感知平台、IPS、WAF、Anti-DDoS、HIDS 与零信任架构，形成立体化安全防护能力。Anti-DDoS 系统作为边界首道防线，部署于网络出口，通过流量清洗技术过滤 SYN Flood、HTTP Flood 等 DDoS 攻击流量，避免核心交易平台、客户服务系统因流量过载瘫痪，保障业务连续性。IPS 承担精准拦截职责，在边界及核心业务区节点部署，既能识别 SQL 注入、跨站脚本等攻击特征，又可自动阻断恶意流量。HIDS 聚焦主机层防护，部署于服务器、数据库服务器等核心设备，实时监测进程异常、文件篡改、权限变更等行为，填补网络层防护盲区。态势感知平台作为体系中枢，整合各系统日志、告警与流量数据，通过关联分析识别潜在风险，迅速定位攻击源头并发出预警。零信任架构贯穿防护全程，遵循“永不信任、始终验证”原则，对访问核心系统的主体持续校验身份、动态分配权限，即便外部威胁突破边界，也能阻止未授权访问，结合定期更新各系统规则库，全面抵御多变网络攻击。

2、安全信息和事件管理系统

安全信息和事件管理系统（SIEM）能够收集、整合来自不同安全设备和系统的日志信息和事件数据，进行集中分析和关联处理，帮助基金公司及时发现潜在的安全威胁。SIEM 系统通过建立安全事件模型和关联规则，对海量的安全数据进行筛选和分析，识别出真正具有威胁的安全事件。当 SIEM 系统检测到多个安全设备同时报告与同一 IP 地址相关的异常事件时，能够通过关联分析判断是否存在大规模的网络攻击行为，并及时发出警报。基金公司应建立完善的 SIEM 系统，实现对网络安全事件的实时监测、快速响应和有效处置。同时，利用 SIEM 系统的报表和分析功能，对网络安全态势进行定期评估和总结，为网络安全策略的优化提供依据。

（二）实时监测与预警机制

1、关键指标的监测

确定关键指标进行实时监测是建立有效预警机制的关键。应根据自身业务特点和网络安全风险状况，确定一系列关键监测指标，如网络流量、系统性能、用户登录行为、数据访问频率等。通过对这些关键指标的实时监测，及时发现异常变化，判断是否存在潜在的安全威胁。例如，当网络流量突然出现异常增长，远远超出正常业务需求时，可能意味着遭受了 DDoS 攻击；当某个用户在短时间内频繁尝试登录系统，且登录失败次数较多时，可能存在暴力破解密码的风险。对关键指标设置合理的阈值，当指标数据超出阈值时，系统自动发出预警信号。

2、预警信息的及时处理

当收到预警信息后，应建立快速响应机制，及时对预警信息进行处理。相关部门和人员在接到预警通知后，迅速对预警事件进行评估和分析，判断其真实性和严重程度。对于真实的安全事件，立即启动应急预案，采取相应的处置措施，遏制安全事件的发展，降低损失。例如，当检测到网络攻击行为时，及时阻断攻击源，对受影响的系统进行隔离和修复；当发现数据泄露事件时，立即停止数据传输，对泄露数据进行追溯和评估，采取措施防止数据进一步扩散。同时，对预警事件的处理过程进行记录和跟踪，总结经验教训，不断完善预警和应急处置机制。

五、网络安全事后管理

（一）网络安全应急响应

1、应急预案的制定与演练

制定完善的应急预案是基金公司应对网络安全事件的基础。应急预案应包括应急组织机构及职责分工、应急响应流程、处置措施、恢复计划等内容。明确在不同类型和级别的网络安全事件发生时，各部门和人员的具体职责和行动

步骤，确保应急响应工作的高效有序进行。例如，成立应急指挥中心，负责统一指挥和协调应急处置工作；明确信息技术部门负责技术层面的应急处置，如系统修复、数据恢复等；风险管理部门负责评估事件的风险和损失等。同时，定期对应急预案进行演练，通过模拟真实的网络安全事件场景，检验应急预案的可行性和有效性，提高员工的应急响应能力和协同配合能力。演练结束后，对应急预案进行评估和总结，针对演练中发现的问题及时进行修订和完善。

2、事件处理与恢复流程

在网络安全事件发生后，基金公司应按照应急预案迅速启动应急响应。首先，对事件进行快速评估，确定事件的类型、范围和影响程度。对于数据泄露事件，立即采取措施防止数据进一步泄露，如关闭相关系统端口、修改用户密码等；对于系统故障事件，尽快定位故障原因，采取修复措施。在事件处理过程中，及时向监管部门、投资者和其他相关方通报事件情况，保持信息的透明和畅通。事件处理完毕后，启动恢复流程，对受影响的系统和数据进行恢复和验证，确保系统能够正常运行，数据的完整性和准确性得到保障。同时，对事件原因进行深入调查和分析，总结经验教训，采取改进措施，防止类似事件再次发生。

（二）舆情管理与声誉修复

1、舆情监测与分析

网络安全事件发生后，舆情管理至关重要。应建立舆情监测机制，实时监测社交媒体、新闻媒体等渠道上关于公司的舆情信息，及时了解公众对事件的关注焦点和态度。通过舆情分析工具，对收集到的舆情信息进行分析和筛选，识别出正面、负面和中性的舆情。关注负面舆情的传播趋势和影响范围，分析其产生的原因和潜在影响。例如，当发现社交媒体上出现大量关于公司数据泄露事件的负面评论时，及时分析评论内容，了解公众的担忧和诉求，为后续的舆情应对提供依据。

2、声誉修复策略

针对网络安全事件引发的声誉危机，基金公司应制定有效的声誉修复策略。及时发布权威信息，向公众说明事件的真实情况、公司采取的应对措施以及后续的改进计划，以消除公众的疑虑和误解。通过公司官网、官方社交媒体账号等渠道，发布公开声明，表达公司对事件的重视和解决问题的决心。积极与媒体沟通，提供准确的信息，引导媒体进行客观公正的报道，避免不实信息的传播。同时，加强与投资者的沟通和互动，通过电话、邮件、在线客服等方式，解答投资者的疑问，安抚投资者情绪，恢复投资者信心。在事件处理过程中，注重采取实际行动改进公司的网络安全管理，提升公司的安全防护能力，并将改进成果及时向公众展示，逐步修复公司的声誉。

六、未来展望

（一）新技术在网络安全中的应用趋势

随着数字化时代的不断发展，新技术在基金公司网络安全领域的应用将呈现出更加广阔的前景。人工智能和机器学习技术将进一步深入应用于网络安全监测和防御。通过对海量的网络数据进行学习和分析，入侵检测系统能够更加精准地识别新型攻击模式，提前预警潜在威胁。块链技术在保障数据完整性和不可篡改方面的优势将进一步凸显，可用于构建安全的交易记录和数据存储系统。

（二）网络安全管理的持续优化方向

未来，基金公司的网络安全管理将朝着更加智能化、自动化和精细化的方向持续优化。智能化方面，通过引入人工智能和机器学习技术，实现安全策略的自动生成和优化。根据网络安全态势的实时变化，智能系统能够自动调整安全防护措施，提高安全管理的效率和准确性。当检测到某种新型攻击手段出现时，系统能够自动分析攻击特征，生成相应的防护策略，并及时应用到相关系统中。自动化方面，进一步加强安全运营流程的自动化，减少人工干预，降低人为错误带来的安全风险。从安全设备的配置管理、漏洞扫描修复到安全事件的应急处置，都可以通过自动化工具和脚本实现，提高安全管理的响应速度和处理效率。精细化方面，对网络安全风险进行更细致的评估和管理。深入分析不同业务环节、不同系统组件的安全风险，制定针对性的风险控制措施，实现对网络安全风险的精准管控。同时，加强与监管部门、行业协会以及其他金融机构的信息共享和协作，共同应对日益复杂的网络安全威胁，维护金融行业的整体网络安全稳定。

七、结论

在数字化时代，网络安全已成为基金公司稳健发展的基石。从网络安全事前的系统设计、数据治理及安全管理元素融入，到事中的网络安全监测技术应用与实时预警，再到事后的应急响应和舆情管理，基金公司需要构建全方位、全流程的网络安全管理体系。面对不断演变的网络安全威胁和日益严格的监管要求，公司必须充分发挥领导作用，积极推动网络安全技术创新和管理模式优化。通过持续投入资源、加强员工安全意识培训、完善安全管理制度和流程，不断提升基金公司的网络安全防护能力。同时，密切关注新技术在网络安全领域的应用趋势，提前布局，利用新技术为网络安全管理赋能。只有这样，基金公司才能在数字化浪潮中有效应对网络安全挑战，抓住数字化转型带来的机遇，保护投资者利益，维护公司声誉，确保金融市场的稳定运行，实现可持续发展目标。

02 安全实践

P14 证券行业邮件安全运营实践
张浩哲、邢骁、李彤恩、徐文娟

P23 证券行业移动应用软件安全与合规管理
宋士明、叶飞、姜玥

P30 安全运营韧性构建：基于信创态感的智能化安全运营探索实践
丁安安、夏英杰、赵川

P35 数据合规驱动下的全域风险监测体系构建与实践研究
—— 基于数据处理活动的安全运营创新路径
陆滢、徐正伟

P38 基于大模型的全渠道信息流统一管控
李剑戈、陶昆、达其双、吴敏嘉

P42 钓鱼邮件演练的实践与探索
张沁怡、崔毅然、吴鹏、陈其乐

P45 智能体驱动的API安全风险管控研究与实践
徐承文、程际桥、吴琪、刘义卓、杨启

P51 基于某次大型攻防演习应对零日漏洞攻击的威胁狩猎实践
林宝晶
钱钱

证券行业邮件安全运营实践

张浩哲、邢骁、李彤恩、徐文娟 | 西部证券股份有限公司

摘　要： 电子邮件是证券行业内外沟通的重要办公应用，承载了大量敏感信息和业务数据。同时，电子邮件系统也是恶意攻击者、攻防演练的重点关注应用，面临着诸如钓鱼邮件攻击、黑产欺诈等风险。本文章主要介绍了邮件安全基本知识，证券行业面临的邮件安全现状，创新性提出ESPF框架解决方案，并对具体工作如何以最佳实践落地进行探讨。

关键字： 邮件安全运营、社会工程攻击

一、邮件安全历史

在介绍邮件安全运营工作之前，我们先简单回顾邮件安全相关的历史。

1971年，Raymond Tomlinson[1]正在麻省理工学院攻读博士学位，他同时也在BBN公司工作。在那个时代，人们远程传递消息主要还是依靠写信和电话，而信件的消息快慢主要还是要靠快递员的奔跑速度；电话的问题则是，如果有人随时想分享一个消息给你，而你却没有办法时刻都待在电话旁边，那么消息就会被错过了。早在1969年，传奇的互联网前身ARPANet[2]已经被发明，四个网络节点在UCLA、SRI、UCSB、UTAH分别运行和组网成功，很多工程师都在为ARPANet开发各种RFC的协议。而Ray此时则萌生出一个想法，是否可以以一种简单的方式通过电子网络来传递消息（that's too complicated, we just want to send messages to people[3]）。虽然在当时计算机间互相通信以及同一个计算机之间不同用户之间传递消息都已经实现，但是他此时仍面临三个问题：1.他需要发送消息给某个远程计算机的某个用户，而不是这个计算机的所有用户；2.人们很难记住某个32位的比特数字（IP）到底代表了哪个计算机；3.当计算机数量和每个计算机的用户量增多时候，整个通讯录对象数量会快速爆炸。为了解决这个问题，Ray创造性的发明了“@”符号，如使用“用户名@主机名”，简洁且无歧义的告诉计算机向谁发送这个消息，“@”为冰冷的电子网络赋予了艺术性。

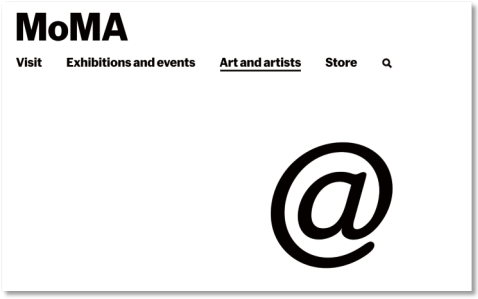


图1 纽约现代艺术博物馆（MoMA）将@列为设计收藏品

1978年，DEC销售代表Gary Thuerk在推销DEC-20操作系统，他通过ARPANet向列表中的数百名成员发送了商品促销邮件。当时的邮件程序还只能支持320个收件人地址，超出的部分溢出到了正文，而Gary也发现了这个问题，他把溢出的那部分地址又重新进行了发送，这也是历史上第一封垃圾邮件[4]。这封邮件引发了非常多的负面影响，如负责运行ARPANet的美国国防通信局给DEC提出了强烈的抗议。同时，这封邮件也造成了犹他大学的一个计算机宕机，因为在当时磁盘存储非常紧张的情况下，大量的垃圾邮件占满了磁盘空间。

20世纪90年代，美国在线(AOL)的聊天室很受大众欢迎，他也成功的引起了各路黑客的关注。早期针对AOL的攻击是Warez社区的黑客们，他们主要利用信用卡号校验算法的缺陷，随机生成假的信用卡号注册AOL账户。到1994年，一个网名为The Daytona的黑客发明了名为“AOHell”的工具，他能够免费注册AOL账户同时，还能够对其他用户发送垃圾邮件、聊天室踢人和密码窃取。当时他们通过假冒AOL员工发送伪造的链接，诱骗用户输入他们的信用卡账号和密码。他们把这种方式命名为“Fishing”，当时有一类电话飞客(Phreaking)采用古老的网络入侵技术盗用电话使用，他们在黑客中具有崇高的地位，包括沃兹尼亚克、嘎吱船长都是其中的一员，所以在AOLHell中为了致敬他们，将这种欺骗方式称为Phishing，这种攻击方式对邮件安全影响深远，Phishing的名称也一直流传至今。AOLHell是早期互联网脚本小子文化的代表性工具，他让不会编写代码的业余黑客能够实现破坏，这种文化影响一直持续到21世纪的前二十年。

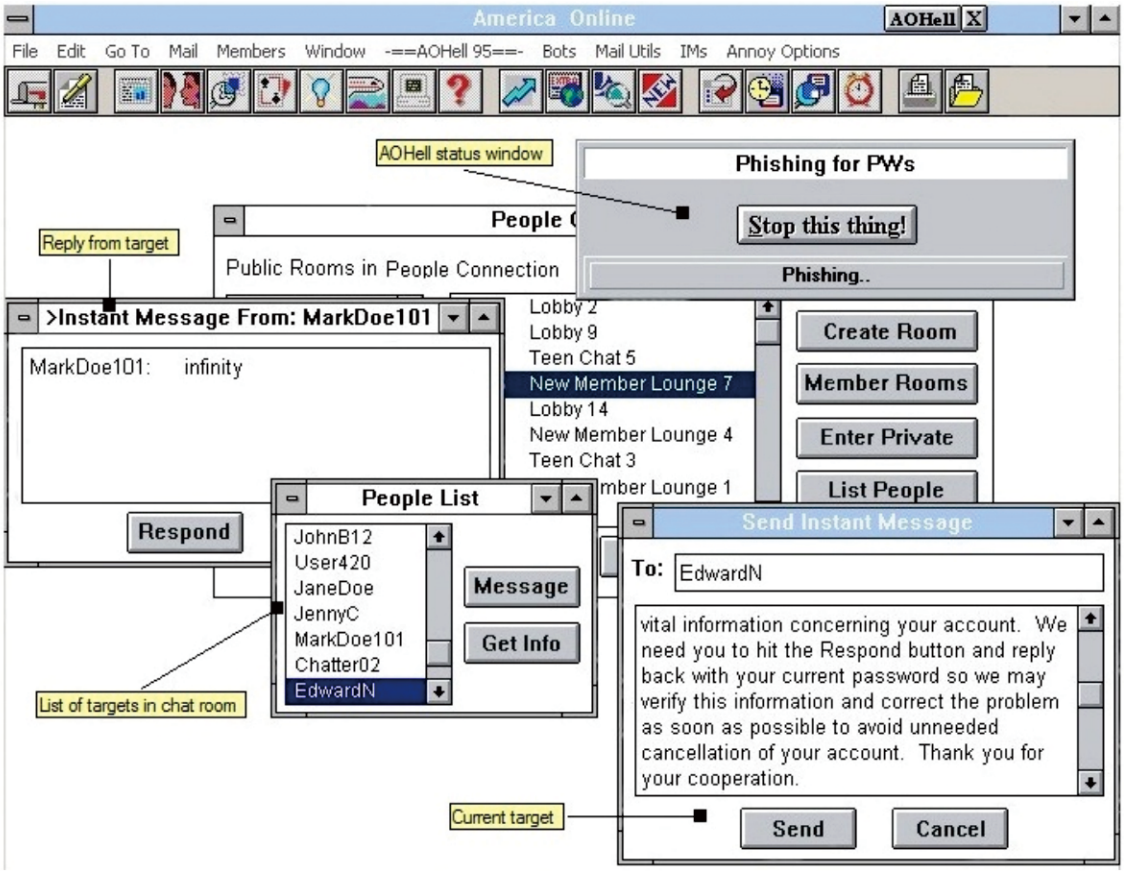


图2 早期的AOLHell界面

步入21世纪后，针对邮件的安全攻击伴随着互联网在逐步增多，我们将在如下章节中介绍现代的攻击方式方法，它们身上一般都带有现在的网络攻击影子。同时，随着大语言模型的发展，钓鱼邮件的文字内容逐渐多样化且更贴近正常的邮件，攻击者针对特定公司、个人使用大语言模型帮助编写更加难以识别的电子邮件，能够更好规避传统的可疑特征。经DarkTrace调查[5]，2023年的前两月这类“新型社会工程攻击”增加了135%，使用新型社会工程攻击技术的钓鱼邮件占比达到38%。与ChatGPT的广泛应用趋势相同，生成式人工智能已经为攻击者提供了更加快速和规模化的攻击。

二、邮件安全威胁

邮件安全是证券期货行业各单位经常遇到的一类信息安全风险，对于大部分的公司而言，可能并不会都能碰见复杂的APT攻击，但是一定会遇到邮件相关的信息安全问题。此类攻击造成的实质损失可大可小，但即使是非常微弱的损失，仍然会给信息技术部门带来比较负面的影响，更不用提针对高级管理人员的攻击，给信息技术部门带来的压力。

我们在这里列举了攻击者利用对邮件的攻击以期待实现的目标，当然攻击者一般攻击直接目标成功后，会有连续的进一步操作。我们这里仅讨论以邮件为第一攻击目标的情况：

（一）获取邮件控制权

邮件系统中，存在着大量的公司机密文件、数据，通常也包括部分的系统口令密钥，同时邮件也作为很多系统找回口令的途径之一。通过控制邮件后，能够重置或找回其他系统密码，从而获取其他系统登录权限。或利用被控邮件在公司内部横向传播病毒邮件、钓鱼邮件等，以作进一步的攻击行为。攻击者通常通过外网的直接暴力破解、内部攻击的密钥碰撞，以获取邮件的访问和控制权。

（二）通过钓鱼形式获取用户账户或主机控制权

钓鱼攻击是常见的邮件攻击形式，攻击通常通过恶意链接、恶意附件及二维码图片这三种类型，诱导收件人点击邮件中的恶意链接、打开附件或扫描二维码。进而植入木马或间谍程序，窃取敏感数据、银行账户、邮箱账号密码等重要信息，或者在设备上执行恶意代码实施进一步的网络攻击活动。

攻击者也常将邮件作为恶意程序传播的一个重要途径，通常以附件（如"xxx 应聘者个人简历.doc"、"员工薪酬统计表.xls"等）或正文内嵌 URL 链接（如"在线文档"）的形式传播，同时此类恶意软件一般也经过免杀等处理。

（三）鱼叉攻击

鱼叉攻击是一种高度针对性的网络钓鱼攻击手段，专门针对特定个人、组织或企业，通过伪造看似可信的邮件诱骗受害者泄露敏感信息或执行危险操作。鱼叉攻击在整个邮件安全攻击行为中占比很低，但成功率和危害性极高，尤其在大型攻防演练期间鱼叉攻击往往是突破企业防线的一个重要入口。其中两个典型的亚型包括：捕鲸攻击，针对于高价值人群的攻击行为，通常这类邮件并不会大范围投递至目标公司，以绕过安全人员的异常监测，攻击者往往通过社交媒体等各类渠道收集信息后选中目标受害人，向其发送定向的攻击邮件，目标对象则包括企业高管、财务负责人等等。BEC商业欺诈攻击，攻击者访问公司电子邮件账户并冒充所有者的身份启动欺诈，目的在于获取信任、拿取数据、商业欺诈以获取更多的金钱收益。

（四）高级威胁

当今瞬息万变的邮件威胁环境下，攻击者技术与手段持续演进，新型攻击层出不穷。二维码形式的邮件攻击在近些年频发，攻击者通过发送一封内嵌二维码图片但毫无附件、URL等传统安全风险的邮件，将攻击的方向从传统 PC 端转移至 BYOD领域，收件人通过手机扫码后被植入恶意软件或被欺诈骗取钱财。生成式AI则催生了伪造公司通知、模仿工作沟通邮件等新型手段，钓鱼邮件的精准率大幅升级，也更具隐蔽性与迷惑性，增加了企业安全识别的难度。此外邮件攻击作为APT攻击中的关键入口和传播媒介，攻击者的攻击手法隐蔽性极高，加以社会工程学的利用，邮件往往成绕过企业边界安全防护的“软肋”。

（五）内部威胁

企业邮件内部威胁同样不容忽视，这类威胁来自公司内部人员因故意或疏忽导致的邮件系统滥用、数据泄露或安全攻击行为。无论是故意的数据泄露还是由于疏忽导致的安全事件，内部人员的行为都可能给企业带来灾难性的后果。常见的威胁类型包括：外发敏感数据、传递商业机密、共享邮箱密码或使用弱密码导致的账号未授权利用、邮箱被控后的公司内横向安全风险等。员工可能因缺乏安全意识或对安全政策的不满而成为安全漏洞的制造者。

三、证券期货业邮件安全外规要求

我们梳理了外部法规、标准中对邮件安全相关的要求，主要内容包括如下：

1.《证券期货业信息安全运营管理指南》

5.3 安全管理最佳实践：可采用实战化的方式检验员工的安全培训效果，如钓鱼邮件、社会工程学等方式；

6.2.2 网络安全管理：在关键网络节点部署恶意代码和恶意邮件防范措施，对恶意文件和钓鱼邮件进行检测、拦截；

10.3 数据安全管理最佳实践：宜从终端、网络层面，建立全面的数据交换监控体系。如终端 DLP、网络 DLP 和邮件 DLP，同时联动其他设备如上网行为管理、堡垒机，构建分析规则，对可疑数据行为进行告警，一线运营人员跟踪确认。

2.《证券期货业信息系统审计指南 第5部分：证券公司》

表C.2 5.16：是否定期检查防病毒网关和邮件防病毒网关的恶意代码库的升级情况并进行记录。

表G.2 1.4.7、表H2 1.4.7、表I2 1.4.7、表J2 1.4.7：在读取移动存储设备上的数据以及网络上接收文件或邮件之前，先进行病毒检查。

3.《证券期货业网络安全等级保护测评要求》

8.1.3.4.2 测评单元（L3-ABS1-15）：应在关键网络节点处对垃圾邮件进行检测和防护，并维护垃圾邮件防护机制的升级和更新。

4.《证券公司信息隔离墙制度指引》

第七条（三）：对可能知悉敏感信息的工作人员使用公司的信息系统或配发的设备形成的电子邮件、即时通讯信息和其他通讯信息进行监测；

5.《证券公司网络安全攻防演练内部参考手册》

应制定点面结合的邮件安全防护体系，首先在邮件网关后部署附件及链接检测沙箱，用于过滤恶意附件和钓鱼链接，其次使用邮件攻击溯源系统场景功能，预测攻击方钓鱼方式，精准发现威胁，并结合情报平台打造情报联盟，继而做好应对措施，且聚焦账号安全，防止出现“内对内钓鱼”即一个邮箱失陷后横向钓鱼，结合网关提前拦截邮件从而抵挡绝大部分攻击方的针对性邮件攻击。

6.其他检查项

除了以上内容，每年的安全检查中，同样也会对邮件的自身安全防护进行检查，主要的重点包括：

a.查阅邮件系统，确认是否设置密码策略，包括密码复杂度、密码有效期等；是否设置双因子登录认证等安全防护措施；

b.检查离职人员的邮箱账户是否删除；

c.检查邮件系统的自动转发功能是否关闭；

检查是否对批量下载邮件、异常时段登陆的行为进行监控。

四、邮件安全运营实践

制定有效的邮件安全运营策略和应对措施，可以增强企业邮件安全的防御能力。我们将邮件安全的防御手段分为4个领域34个方法，并满足证券行业《JR/T 0112-2014证券期货业信息系统审计规范》《JR/T 0146.5-2016证券期货业信息系统审计指南 第5部分：证券公司》和常见的证券期

货业现场安全检查中常见的合规问题，并形成邮件安全防护框架Email Security Protect Framework（ESPF）。整体防护框架如下所示：

表1 邮件安全防护框架

领域			
安全配置（SM）	恶意内容识别（EI）	异常行为管理（AD）	邮件安全情报（TI）
对邮件服务器、邮件安全网关进行正确合规的配置，有助于降低大部分的外部攻击危险。	对投递的邮件进行检测，识别可能存在的恶意内容并进行告警和阻断。	建立邮件行为基线，对于异常的偏离行为及时进行告警和处置	通过收集外部、内部的邮件情报信息，建立企业威胁模型，有效识别内外部威胁。
实践方法			
安全配置（SM）	恶意内容识别（EI）	异常行为管理（AD）	邮件安全情报（TI）
1.关闭邮件服务器互联网端口 2.关闭邮件服务器匿名账户 3.使用特定客户端访问 4.设置密码策略 5.开启双因子认证 6.开启 SPF 协议检查 7.开启 DKIM 校验 8.开启 DMARC 校验 9.关闭邮件自动转发 10.设置邮件定期归档 11.分离发件和收件IP 12.配置拦截提醒功能 13.限制邮件批量发送上限	1.配置恶意文件沙箱 2.对链接和二维码开展检测 3.配置组织级恶意字典库 4.发件人标头校验 5.配置标头发件人过滤 6.开启图片 OCR 识别 7.高价值人群保护	1.设置登录提醒 2.显示账户上次登录信息 3.显示账户异常通知和锁定 4.异常时段登录和发送邮件 5.异常地点登录和发送邮件 6.频率控制 7.告警批量投递失败 8.常用语言 9.定期检查离职人员邮箱账号删除情况	1.识别潜在的攻击者 2.收集攻击技术情报 3.开启邮件信誉检查 4.开展钓鱼邮件演练 5.开展邮件安全攻防

（一）安全配置（SM）

SM1、关闭邮件服务器互联网端口

可配置邮件服务器上关闭对互联网开放的25和110端口，改用SSL加密的465和995端口提高邮件传输的安全性，降低垃圾邮件和恶意伪造邮件的转发投递。在邮件服务器上游部署的邮件安全网关用于承担SMTP的业务交换和端口暴露的风险，同时启用SMTP流量限制可在一定程度上抵抗邮件DOS/DDOS攻击，保护邮件服务的业务连续性。

图3 基于发件人的SMTP流量限制

在实际的落地中，我们将邮件服务器对互联网的25、110、465和995端口关闭，仅保留邮件安全网关的25端口对互联网开放，用于接收邮件后转发邮件服务器，部署方式如下。

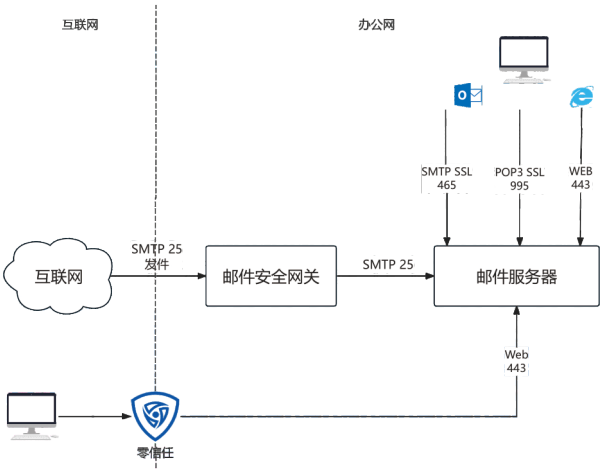


图4 不开放邮件服务器互联网端口的部署架构

此类部署的优点包括：

1.降低邮件服务器的互联网暴露面，消除了大部分互联网攻击。不会因为邮服的0Day/1Day漏洞导致从互联网攻破邮服；

2.邮件安全网关仅开放SMTP收信，但不提供认证功能，可以完全杜绝外部对邮件账号的密码破解；

3.收敛了邮箱的访问通道，将邮件的Web访问集成到零信任系统和企业微信后台（也通过零信任通道）。即使外部攻击者获取公司内部邮箱账号，从互联网尝试访问邮箱仍然需要攻破零信任。

但实施此项工作，通常会面临使用便捷性、使用习惯改变的困扰。证券行业一般遇到的典型问题包括：

1.研究所有大量使用邮件日历的需求，他们希望能将日程可以直接Book到客户的邮箱客户端；

2.部分同事并不喜欢从归档服务器中查询历史邮件，他们已经习惯于在本地邮件客户端中保存历史邮件并进行查询；

3.部分重要通知邮件，使用人员希望能够在电脑、手机上收到明确的提醒消息。

以上问题，可以通过对邮件系统本身的优化，如支持日历功能、优化归档文件查询、提供多功能通知通道等方法，基本的使用需求是完全可以满足。

SM2、关闭邮件服务器匿名账户

在匿名设置下不需要提供用户名和密码进行身份验证便可利用SMTP服务进行邮件发送，攻击者可以伪造任意的发件人完成恶意邮件的投递。关闭邮件服务器的匿名账户功能，且仅对上游的邮件安全网关中继开放，阻断公司邮件服务器被未授权利用。

SM3、使用特定客户端访问

员工使用邮箱的场景包括在公司内、家里或是任意地点，访问途径包括手机、电脑，更多的则是在互联网环境中随时查看邮件，但这样也随之带来了安全风险。主流的邮件服务商提供了自带的邮件客户端在提供便利的同时也增加了访问安全性，使用零信任、企业微信工作台集成邮件系统也是一种访问方式，避免将邮件登录页面直接暴露在互联网上的安全隐患。

但如SM1内容，在禁止互联网访问的情况下，如果需要使用客户端，推荐仅允许在办公网环境使用。目前主流的邮件系统的客户端与服务端的认证仍然是使用HTTPS，也就是仍然有HTTPS服务需要开放，我们不认为在目前情况下，此类服务的开放能够让安全运营人员放心。

SM4、设置密码策略

邮箱因需要通过互联网使用且是企业沟通的重要途径，暴露出去的风险相对而言更高，严格的密码策略可以有效地防范账号密码暴力破解的风险。同时在使用客户端登录时使用专用授权码的形式可进一步提高账号密码的安全性，严格的策略可设置为认证成功一次即失效，有效期时长3个月自动失效并重新认证。

SM5、开启双因子认证

开启邮箱双因子认证，是提高邮箱账户安全的非常有效的方法。完成账号密码登录后需通过手机验证码、微信扫码等形式完成第二次的认证，保证即使在账号密码泄漏的情况下可二次认证阻拦。严格的策略可同时设置当账号在新设备、非常用IP登录时的强制二次认证。

SM6、开启SPF协议检查

SPF（发件人策略框架）是一个电子邮件认证协议，能够防止未经授权的发件人代表特定的域名发送邮件。通过配置SPF校验策略可以拦截大量的伪造邮件，当一封邮件投递过来时首先校验该域名的SPF配置与实际的发件IP是否符合，如不在允许地址范围内则阻断本次投递。SPF防护的基本过程如下：

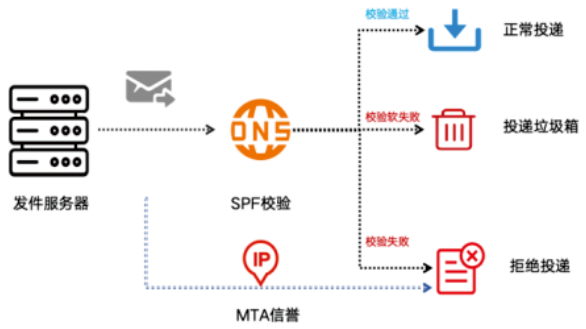


图5 SPF校验和MTA信誉的处置流程

当 b.com 邮件服务器收到一封邮件，其发送主机 IP 为 1.2.3.4，发信方在 Smtp.From 字段声称是 s@a.com 时，b.com 邮件服务器会去查询 a.com 的 SPF 记录来验证发件人是否伪造。若该域无 SPF 记录，那就意味着发件人可随意伪造；若存在SPF记录且设置允许此IP（1.2.3.4）的主机发邮件，服务器便认定邮件合法；要是该IP不在允许范围内，多数情况下这封邮件会显示代发标识或被投递至垃圾箱。

从实际的运营来看，攻击者通常会使用全新的域名同时不启用任何的SPF配置，或使用一个正常域名的不存在子域名作为发件人（例如hr.qq.com）且该域名不启用SPF配置，导致这类邮件会被投递至企业内部。为了应对发件人伪造绕过SPF校验的攻击，在发件人校验环节同时设置MTA信誉检查，结合威胁情报信息判断发件IP是否为恶意地址，如果命中规则则拒绝连接。除威胁情报外，通过公司的内部安全运营结合外部威胁预警，手工添加发件IP黑名单并持续动态更新，提前阻断。

SM7、开启DKIM校验

DKIM是另一种验证发件人身份的技术，通过使用DKIM校验可以有效地遏制冒充邮件地址。RFC 5585、6376和5863中指定的DKIM由雅虎的DomainKeys和思科的已识别互联网邮件这两个历史提议合并而成。通过DKIM技术为发件人提供了一种对传出邮件进行加密签名的简单方法，并将签名（及其他验证元数据）添加到邮件头“DKIM-Signature”。发件人将公钥发布在DNS中，使任何收件人都能轻松检索密钥和验证签名，如果邮件来源拥有发件人组织的私钥，则被默认授权代表他们发送邮件，而验证失败的则被拒绝。

SM8、开启DMARC校验

DMARC 是邮件身份验证技术中最新的一种，专门用于解决 SPF 和 DKIM 存在的不足。与其他两种技术不同，DMARC 会对邮件的Header From 进行身份验证，同时包括其他两种技术之前执行的相关检查。DMARC 优于 SPF 和 DKIM 的方面还包括：

1.确保所有可用身份（HELO、MAIL FROM 和/或 DKIM 签名域）与 From 信头一致（完全匹配或从属）

2.为发件人域所有者提供了一种方法，让他们能够为收件人指定必须如何处理失败邮件的策略

3.为发件人域所有者提供了一种反馈工具，让他们能够获得有关任何失败邮件的通知，以轻松识别网络钓鱼活动或 SPF/DKIM/DMARC 策略分发中的错误

DMARC 校验在进一步严格阻拦恶意邮件的同时最大限度地减少误报。

SM9、关闭邮件自动转发功能

将邮件自动转发功能关闭，避免恶意文件通过自动转发功能在企业内部大规模扩散，或者攻击者将自动转发地址更换为攻击者自己的电子邮箱地址以窃取企业数据。

SM10、设置邮件定期归档

邮件归档将历史及实时产生的邮件按照公司业务需要和合规的需求设置规则和策略，从邮件服务器或邮件客户端备份到指定的存储位置，以便于长期保留、检索和管理。归档的目的一方面是减轻邮件服务器的负担提高性能，同

时确保重要的邮件和信息能够在需要时方便地被访问和恢复。归档在满足国家法规或行业规定的同时，假使发生邮件数据损坏时能及时通过归档信息恢复业务，保障业务连续性。

SM11、分离发件和收件IP地址

邮件收件IP因DNS发布绑定不可避免地会对所有人开放，发件IP则更多的是公司专线出口地址或者邮件云厂商的专用地址集。将收发件IP分离在一定程度上防止因信息泄漏而被仿冒。同时在公司边界防火墙配置仅允许来自邮件服务器/邮件网关IP的出站连接，在邮服SMTP 会话期间使用 MAIL From 命令时检查发件 IP 地址是否匹配公司员工可信地址范围，防范中继垃圾邮件的攻击。

SM12、配置拦截提醒功能

对部分易出现误拦截的策略，配置拦截提醒功能，通知用户相关主题邮件被拦截，并提供自助解封方式，避免因为策略设置过严情况下拦截较多用户正常工作邮件，减轻运营压力。

在实践运营中，对于检测出确定的恶意文件一般不作提醒，一般应用于对标题、内容关键字的拦截规则。可以通过由用户转发被拦截通知的邮件到信息安全邮箱进行解封，或配置拦截邮件内容嵌入解封链接，对接邮件安全网关接口，在用户点击解封链接后自动将拦截邮件放行。

SM13、限制邮件批量发送上限

对于对内发送的邮件，单个用户发送的人员一般会有限制，除特定账号外，正常员工很少会给全员发送邮件。设置单个员工单日发送的邮件上限、单个员工单日发送的目标收件人上限，能够有效减少恶意攻击者控制公司少量账号发送恶意邮件时的扩散范围。

（二）恶意内容识别（EI）

EI1、配置恶意文件沙箱

部署在边界的邮件安全网关会内置多种对恶意样本的检测引擎来实现对外部恶意样本的初步检测，同时使用邮件安全沙箱技术作为第二道病毒检测补充，动态分析邮件附件，包括加密压缩包的拆解、附件多环境下的动态分析、内嵌URL链接的实时扫描等。反病毒引擎识别到恶意样本后，直接隔离拦截，其他无法判定风险的则统一投递至沙箱进行二次分析。邮件安全沙箱日均可处理分析千余次，识别阻断了大量的恶意URL和文件。

URL	风险级别	文件 SHA-1
http://107.175.31.187:80/8808/isd/mygirlobegoodwithentierprocessididwithhersh...	✕	5ECA0441C41C2C99EBFB3608BD4FB0BC7034F79B
https://url8y.com:443/	!	C5F5C00C55140CDDDD198C32E5639AB41C0166DB
https://url8y.com:443/asy2xr	✕	D021B46C74E131124A7B4C3B6B004FF8E38D5395
http://url8y.com:80/	!	96424FA2CC3FC44572201F2D061D875F11E69F2A
http://url8y.com:80/asy2xr	✕	A37F7D419208A05AF68B914081F6E68C60804E76
http://url8y.com:80/_vti_inf.html	!	2139002FC21C6B61282C274A9D4774A62E8E740C
http://148.66.59.162:80/client.exe	!	277BA097894E5D5CC37548E17AEFFF784233601F
http://43.129.233.99:80/mtuser.dxm	!	7DC7F2B8CB0FD7CDF1CBC9D3E126CAFD7B69F9B1
http://43.129.233.99:80/winword.exe	!	38348397585B7AF818CF0B26643EAC1C7D5F8D97
http://43.129.233.99:80/wwlib.dll	!	E278B1B215C8284A7FB476A5EEDDACCDFFB7DA8F
http://43.129.233.99:80/wig.ppt	!	
https://euunclaimedpymt.com:443/main.js	✕	

图6 邮件安全沙箱分析识别

安全沙箱有效地拦截住了大量的免杀病毒，所谓道高一尺魔高一丈，攻击者的手段也在试图去突破沙箱防护，包括压缩包加密、恶意附件的动态远端加载、恶意文件/进程延时执行绕过沙箱分析、文件中转站/云盘链接跳转下载等。对于沙箱无法分析的邮件通过主题插入警示提示收件人加强对发信的判断，或在特殊时期直接进行隔离，由管理员研判分析后手动释放。即使邮件投递到收件人且被点击，在终端侧的终端防病毒软件可以进行二次的落盘扫描拦截。为了将恶意邮件在边界网关处完成拦截，绕过邮件安全网关病毒检测的但在落盘时被识别的样本特征会同步手动添加至邮件安全网关威胁列表，列表的信息来源还包括外部威胁情报、安全预警等，持续动态更新。定期对网关上截获的危险病毒或恶意代码进行及时分析处理，并形成书面的报表和总结汇报，是针对高级攻击的有效手段。

EI2、对链接和二维码开展检测

恶链URL攻击、二维码攻击是近期高发的钓鱼邮件攻击，需要在邮件安全网关上通过Web信誉和动态沙箱分析识别恶链并直接进行阻断。

对于高频出现的钓鱼网站特征，可以设置更加严格的组合策略，包括；

1.当邮件正文中包括链接+“邮箱账号”+“立即修改”等特征时拦截相关邮件。

对原链接进行替换，同时采用动态click on对点击链接的动作进行实时确认。

关键字
<a\s*[^>]*?href\s*=[^>]*(https?|ftp|file)
邮箱#邮件#账户#账号
点击#立即#马上#违规#异常#升级#点这里#点我#尽快#密码#点此#立刻#验证#及时#备案

图7 高频钓鱼网站拦截策略

3.对于近期频发的二维码钓鱼攻击，当邮件内容包括“补贴/薪酬/个税”等相关内容且包括图片编码时，执行严格的策略命中拦截，通过人工研判加发件人白名单的机制减少误拦率。

人力资源和社会保障政务服务平台
国家社会保险公共服务平台

2024 年财政个人劳动补贴领取通知

根据国家财政部、国家税务总局、国家市场监督管理总局、工商行政管理局联合下发《2024 年财政个人劳动补贴》现已开展。

1、此项补贴政策仅限于 1 年内工资扣税满 3 个月以上的中国公民，收到通知后，请及时进行申请登记，逾期视为弃权领取。

2、申请成功后登记的个人银行账户会收到一笔劳动补贴金，具体金额根据个人 1 年内工资扣税总额多少计算（纳税越多可领取的补贴金额越多）。

3、使用微信扫一扫下方二维码进入国家人力资源和社会保障官方小程序办理。

图8 二维码钓鱼邮件样本

4.对邮件正文内容为空，仅存在包含二维码的文件、图片的邮件进行阻断，此类邮件一般为钓鱼邮件。

EI3、配置组织级恶意字典库

企业收到的邮件内容具有多样性，比如官方通知、合作方业务往来、工作内容交流等。很多场景下员工会预留或对外提供办公邮箱，这也就提高了信息泄漏的风险，容易被攻击者精准投放。为了对内容伪造邮件进行防范，除了使用邮件安全网关的反垃圾引擎识别垃圾邮件外，更多的是添加各类内容过滤规则，消息匹配部分包括标题、正文、标头等，过滤内容包括常见的营销邮件（广告、课程推销等）、钓鱼邮件（密码重置、发票、人事告知等）、欺诈邮件（税务、补贴、财务等）等。内容关键字来源于威胁情报中提供的热点高频词汇、日常邮件运营中频发的邮件样本、时事热点关注内容等，同时根据监测到的样本内容举一反三，将高频关键词的近义词或相近主题同步加入过滤，形成了一套公司内部的过滤字典库，并进行动态补充调整。由于不同公司的业务不同，因此公司应对自己所在组织单独维护一套恶意字典库，该类字典库可以应用于邮件正文、邮件主题。

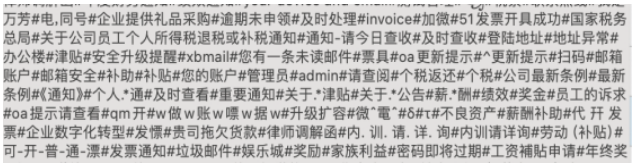


图9 公司维护的恶意字典库

EI4、发件人标头校验

从有些垃圾邮件样本来看，主题是毫无意义的字符，通过主题关键词几乎不能过滤类似的钓鱼邮件，但是通过设置发件人与电子邮件标头发件人的校验策略，当发件人与标头内的不一致时可以有效地阻断大量的类似垃圾邮件。

发件人	电子邮件标头 (发件人)	电子邮件主题
zmav@piyw.com	"eva" <zamvr@piyw.com>	陪
xwjgw@piyw.com	"rosie" <zdsbq@piyw.com>	取
ynkweqoli@kmfs.com	開＝未兑=票 +VX:liurui0rui <13066176976@163.com>	IBPP
objzdfmy@kmfs.com	埽-標3%-专漂-13% Vr\n亻言：liurui0rui <13066176976@163.com>	GFaJQSes
bgtloqjj@kmfs.com	普通3%票=专佣13%票Vr\n亻言：liurui0rui <13066176976@163.com>	IuqJBo

图10 发件人标头不一致样本

EI5、配置标头发件人过滤

主题通过添加特殊符号的拼接通常能够绕过关键词过滤规则，但由于攻击者常使用“人力资源部”“财务部”等发件人标头来提高邮件的可信度，因此在过滤规则中添加相关常见标头的过滤规则，即使主题、正文发生变化，仍能完成钓鱼邮件的隔离。

图11 发件人标头伪造样本

EI6、开启图片OCR识别

部分钓鱼邮件将欺诈和钓鱼内容直接放置于图片内以规避针对邮件内容的检查，通过邮件安全网关的图片OCR识别能力，识别图片中文字内容或二维码，并匹配组织级的字典库和恶意链接的扫描，能够及时发现该类恶意攻击。

EI7、高价值人群保护

利用业务电子邮件入侵（BEC）诈骗，攻击者访问公司电子邮件账户并冒充所有者的身份启动欺诈电汇。攻击者通常利用高管的身份诱骗目标汇钱到攻击者的账户。BEC诈骗也称为中间人诈骗，通常以定期发送电汇至国际客户的企业为目标，可能会涉及使用恶意软件和社交工程攻击手段。依照趋势科技的统计，在整个2023年BEC检测总数增长了16%。在攻击者视角中，公司的高管也更具有应用系统的高权限账户。

通过对单位内部的高价值用户（如CEO、高管、关键业务部门）的邮件收发行为配置单独的深度安全扫描与检测，从异常邮件收发行为的角度，如受保护发件人的姓名与可疑域结合使用，对高价值的邮箱提供进一步的保护。

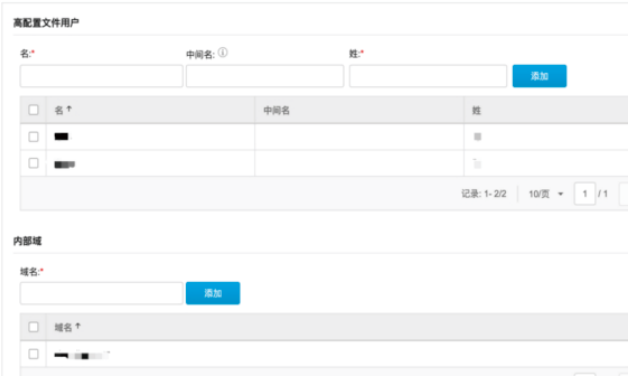


图12 BEC欺诈对高价值用户的防护策略

（三）异常行为管理（AD）

AD1、设置登录提醒

通过短信、IM、邮件等一种或多种方式通知用户登录消息，登录消息内容包括登录IP地址、登录时间、登录结果，及时发现异常登录。

AD2、显示账户上次登录信息

在Web邮箱首页或特定客户端中，显示账户上次的登录IP地址、登录时间。

AD3、显示账户异常通知和锁定

对账户的异常登录、失败登录等进行记录，对多次登录失败账号进行锁定，提示用户和邮箱管理员的异常登录告警。

Ad4、异常时段登录和发送邮件

对单个用户定期分析用户登录习惯，形成用户的邮件登录、发送行为基线，对非常用时间，如凌晨时的登录邮件行为进行告警，必要时可进行二次认证或禁止登录。

AD5、异常地点登录和发送邮件

对单个用户不同地理位置登录邮件进行提醒，对存在短时间内在不同城市、国家登录情况进行告警，必要时可进行二次认证或禁止登录。

AD6、频率控制

面对类似批量发送的钓鱼邮件攻击，需要从发件人、收件人、发件IP、时间、发件IP位置、主题等维度设置邮件的投递频率限制，包括：

1.基于单个发件人IP或发件人地址在一定时间内发送超出设置的最大邮件数量时，则阻断该IP或地址的连接。

2.基于邮件主题当同样主题的邮件在一定时间内超出设置的最大邮件数量时，则阻断该主题的邮件投递。

3.基于单个发件人在特殊时间段，如凌晨时候，对多人发送相同主题邮件，进行告警。

4.基于单个发件人发件IP在国外时批量发送超过阈值的同主题邮件进行告警。

此类规则可以根据企业当前情况扩展，同时也可以基于当前组织用户的邮箱收发规律，通过机器学习方法建立行为基线。

图13 相同主题频率控制

AD7、告警批量投递失败

针对通过爆破收件人地址的方式进行批量发送钓鱼邮件的行为。当收件人地址不存在时，邮件网关的投递是失败状态。通过设置投递失败的预警信息可以及时的阻止进一步的攻击行为，当大量投递失败的预警通知后，意味着此时公司邮箱正遭受着邮件攻击，可第一时间进行封堵处置。

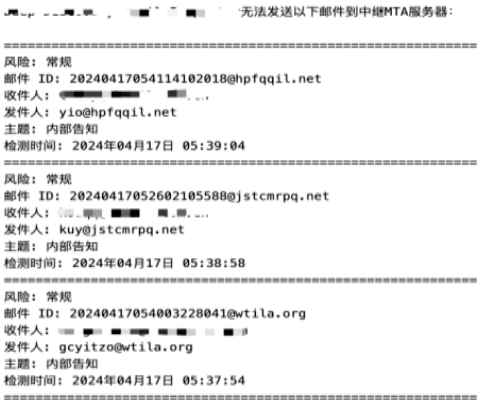


图14 批量投递失败告警信

AD8、设置常用语言

对于部分公司员工账号，进行邮件联络的语言比较固定，对非常用语言进行隔离或投递到垃圾箱可以有效减少部分恶意打扰。

AD9、定期检查离职人员邮箱账号删除情况

建立公司离职流程，及时清除已离职人员的邮箱账号等，避免无主账号的失控风险。

（四）邮件安全情报（TI）

TI1、识别潜在的攻击者

跟踪多种外部情报和威胁预警，筛选外部恶意邮件发件人地址/域名、IP作为可能的攻击者，并将发件人、发件地址、IP加入黑名单。对公司内部已拦截的邮件进行分析，提取恶意邮件的IP、发件人地址、主题，统计高频攻击信息封禁域名、同C段地址等，提前完成可疑拦截封堵。

TI2、收集攻击技术情报

跟踪近期常见的攻击手段，对时下爆发的钓鱼文本内容、钓鱼技术方式进行收集和分析，并将内容、技术等提炼为检查和拦截规则，制定公司内部的防御策略。

TI3、开启邮件信誉检查

开启发件人、发件邮箱地址的信誉检查，根据信誉情况对正在发送垃圾邮件、钓鱼邮件的邮箱、IP等进行阻断操作。

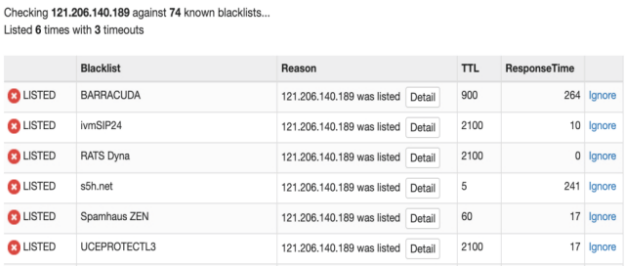


图15 MTA信誉检查

TI4、开展钓鱼邮件演练

钓鱼邮件演练是外规要求，也是切实可以提高公司员工防范钓鱼邮件能力的手段，公司可通过定期组织邮件安全教育培训及内部钓鱼邮件演练，传授钓鱼邮件识别技巧、剖析典型案例、讲解攻击趋势，尤其针对新型的二维码扫码、恐吓欺诈手段等，降低员工失陷概率。

对于证券行业，可通过外采专业钓鱼邮件演练服务开展相关工作，能够实现非常好的宣传效果和切实教育效果。当然，也可以借助类似Gophish 搭建的钓鱼平台发送批量内部钓鱼邮件进行演练。钓鱼演练中，可追踪员工点击仿冒的链接和提交敏感数据情况，以“模拟攻击”检验员工防范意识，可对点击者二次强化培训。经安全教育与模拟演练双重强化，员工邮件防范意识得以提升。即便有少量“漏网”邮件送达，收件人也能及时上报公司安全团队，由其处置并优化安全策略，封堵恶意样本，持续加固企业邮件安全防线。

TI5、开展邮件安全攻防

针对公司的邮箱系统，需要定期开展渗透测试和攻防演练，可以发现邮箱配置存在的缺陷、问题，同时也能够及时发现、验证近期邮箱系统存在的漏洞。一般可将此类工作纳入公司渗透测试、攻防演练工作中，针对邮箱系统，开展专项工作。

五、未来挑战

邮件攻击手段层出不穷，呈现出攻击者多人对抗防守方少数人的局面。以往邮件攻击多以获取用户信息、终端权限，突破边界内网为目的，然而随着企业安全防御体系逐步完善，内网突破难度剧增。当下，欺诈黑产因“效率优势”占比日增，将攻击阵地转向企业防护薄弱的移动设备端，通过诱导收件人微信、支付宝扫码等手段骗取钱财，虽未侵入公司网络内部，但危害极大且影响恶劣。这种打一枪换一炮的攻击模式，使基于特征的扫描拦截难以奏效。因此除强化员工安全意识教育外，加强移动设备侧邮件收信防范成为邮件运营防护的新关键。

此外，生成式人工智能等技术的革新，也给企业邮件安全防护带来全新挑战。包括DeepSeek等众多大语言模型的出现，有可能对邮件安全进行重新定义。对于攻击者来说，通过大语言模型不断对发送的邮件内容进行变形，使得钓鱼邮件越来越难于区分；而大语言模型也对邮件安全的防护也带来了新的能力，对语义的识别智能，能够帮助安全运营人员更简单和有效识别传统钓鱼邮件。

参考文献

1.互联网名人堂Raymond Tomlinson
https://www.internethalloffame.org/official-biography-raymond-tomlinson/

2.纪念ARPANET诞生50周年：互联网发展史
https://www.edu.cn/xxh/focus/rd_xin_wen/202001/t20200102_1703268.shtml

3.[The Verge] Ray Tomlinson, the inventor of email: ‘I see email being used, by and large, exactly the way I envisioned’
https://www.theverge.com/2012/5/2/2991486/ray-tomlinson-email-inventor-interview-i-see-email-being-used

4.Reaction to the DEC Spam of 1978
https://www.templetons.com/brad/spamreact.html

5.How Phishing Attacks Are Becoming Harder to Identify
https://www.darktrace.com/blog/email-attack-trends-how-phishing-attacks-are-becoming-more-sophisticated-and-harder-to-identify。

证券行业
移动应用软件安全与合规管理研究

宋士明、叶飞、姜玥 | 南京证券股份有限公司

摘　要： 随着移动互联网技术在证券行业的深度融合，移动应用软件（APP，含小程序）已成为证券公司服务客户、拓展业务的核心渠道，其安全与合规水平直接关系到投资者权益、公司声誉及行业稳定。然而，证券APP在带来便捷性的同时，也面临着日益严峻的安全与合规挑战。本文系统性地阐述了一套贯穿APP全生命周期的安全与合规管理体系，在保障业务稳定运行的同时，切实保护投资者合法权益，促进行业健康发展。

关键字： 移动应用安全、合规管理、个人信息保护、隐私威胁管控、APP备案、检测认证

一、引言

移动互联网的飞速发展深刻改变了金融服务的模式，在证券行业，移动应用软件（Mobile Application Software, APP，包括其轻量化形态如小程序），已成为证券公司服务投资者的主要阵地。APP为投资者提供行情查询、证券交易、理财购买、资讯获取等操作，承载了大量的核心业务功能和敏感数据。据上海证券交易所最新年度《统计年鉴》，近年移动端交易金额占比普遍在85%以上并持续增长，APP的用户体验、运行稳定性直接影响着证券公司的市场竞争力与客户黏性。

然而，证券APP在提供便捷服务的同时，也成为网络攻击者和金融黑灰产觊觎的目标。证券APP通常处理包括用户身份信息、银行账户、交易记录、持仓情况等在内的高度敏感个人金融信息，一旦发生数据泄露或系统被攻击，不仅会导致投资者财产损失，更会严重损害证券公司的声誉，甚至引发系统性风险。近年来，针对APP的攻击手段不断翻新，从传统的漏洞利用、拒绝服务攻击，到利用仿冒APP进行钓鱼欺诈、通过API接口进行非法交易，甚至采用人工智能技术进行深度伪造（如AI换脸）等。同时，第三方软件开发工具包（SDK）的广泛应用在加速APP开发的同时，也引入了新的安全和隐私风险，部分SDK可能存在恶意行为或安全漏洞，成为数据泄露的潜在源头。

与此同时，国家也相继出台并实施了《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》等一系列重要法律法规。金融监管机构如中国证监会、中国人民银行等，以及网信办、工信部、公安部等部门，也针对金融行业特别是证券行业的APP发布了更为严格和细化的监管要求与行业标准，例如《证券期货业移动互联网应用程序安全规范》（JR/T 0192-2020）、《证券期货业移动互联网应用程序安全检测规范》（JR/T 0240-2021）、《个人金融信息保护技术规范》（JR/T 0171-2020）等。监管机构通过常态化的检查、通报、处罚等手段，督促证券公司切实履行安全与合规主体责任，合规要求的“高标准、严监管、强处罚”态势给证券公司的APP管理带来了前所未有的挑战。

二、APP安全与合规管理体系

构建科学有效的APP安全与合规管理体系是证券机构应对安全挑战、满足监管要求、保障业务连续的基础。一个成功的管理体系能够将安全与合规要求从被动的检查项转变为主动的安全控制，从而有效控制风险并持续创造价值。

（一）证券行业APP安全与合规挑战

证券行业的APP在安全与合规方面面临着多重挑战。首先，监管合规要求日益严厉且动态变化。证券行业需同时遵循《数据安全法》、《个人信息保护法》及多项行业规定，法规交叉叠加且不断更新。监管机构的常态化抽查，如隐私政策合规性审查，对机构的合规管理能力提出了更高要求。其次，网络攻击与黑灰产风险持续升级。APP的交易接口、数据库等易成为攻击目标，DDoS攻击、API滥用、仿冒APP、AI换脸诈骗等手段层出不穷，金融黑灰产链条化运作更加大了攻击的规模和频率，严重威胁用户资金与机构声誉。最后，敏感个人信息的隐私保护难度大。证券类APP涉及大量高敏感数据，且在第三方SDK风险管控以及数据跨境传输合规等方面也存在实践挑战。这些挑战共同构成了证券APP安全与合规管理的复杂背景。

（二）证券行业APP主要监管部门

图1 证券行业APP主要监管部门

证券APP的安全与合规受到多个监管部门的共同监督，包括证监会、网信办、公安部、市场监管总局和工信部等，各监管部门的支撑机构及处罚措施如图1所示，形成了一个多元化的协同监管格局。近年来，监管机构对金融类App的违规行为保持高压态势，通报和处罚力度不断加大，无论是网信体系（如中央网信办、地方网信办）还是工信体系（如工信部、各地通管局），都在持续开展专项行动，针对违规收集使用个人信息、侵害用户权益等问题进行通报、约谈、责令整改甚至下架处理。近期发生的金融行业个人信息泄露事件，以及个别证券同业APP因违规收集个人信息被通报如图2所示，给证券公司敲响了警钟。这表明监管执法态度坚决，检查日趋严格，合规已成为不可逾越的红线。

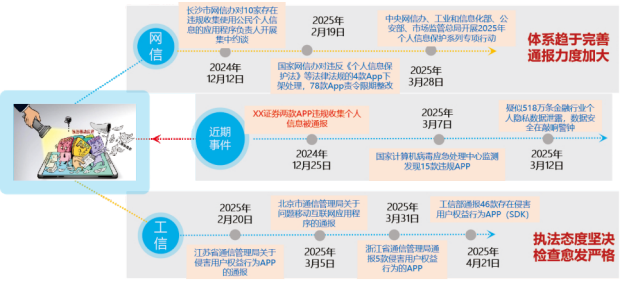


图2 监管部门对金融类APP违规的通报

（三）APP安全与合规建设路径

面对复杂的挑战和严格的监管，证券机构应采取系统化、分阶段的方法推进APP安全与合规建设工作。第一阶段是“建立安全合规工作小组”，这需要成立跨部门的安全合规工作小组，梳理监管要求。第二阶段是“现状评估及快速合规”，通过对现有APP进行管理和技术层面的调研与测试，识别差距，制定并执行整改方案，形成遗留问题清单，目标是达成“基础性保障”。第三阶段是“体系化整改与认证”，目标是实现“建设性完善”。基于前两阶段的成果，结合遗留问题清单，进行更深入的差距分析和系统性问题整改，并完成APP的安全检测、安全加固和监管要求的安全认证。第四阶段是“构建安全与合规长效机制”，目标是确保“长效持续落地”。这要求建立完善的APP安全与合规管理体系，将“基于隐私的设计”（Privacy by Design）原则融入持续合规开发与测试流程，对已获认证的APP进行持续监督与维护，建设APP的内生安全能力及相应的人才培养体系，真正推动安全与合规流程的落地，形成长效机制。

（四）APP安全与合规管理体系构成

一个完整的APP安全与合规管理体系，应以国家法律法规（如网络安全法、数据安全法、个人信息保护法）和相关标准规范（如国家标准、金融标准、证券行业标准）为外部依据。体系内部应包含明确的安全与合规管理制度、清晰的组织架构、规范的工作流程以及配套的安全与合规平台工具。该体系需要贯穿APP的全生命周期，从需求/设计、开发、测试、发布到运行的各个阶段，全面覆盖APP合规要求、APP安全防护、个人信息保护、备案检测认证等核心领域，形成一个系统化、全方位的管理框架，如图3所示。

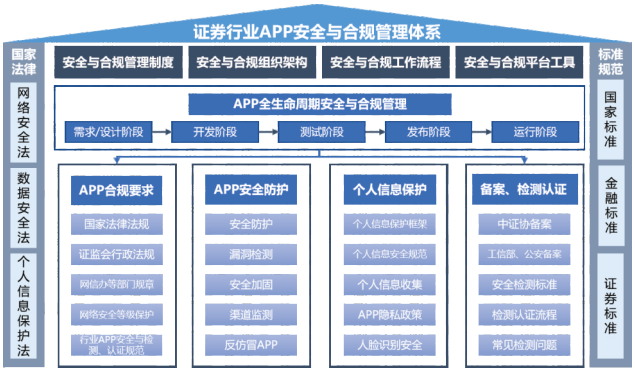


图3 证券行业APP安全与合规管理体系

安全与合规管理制度：管理制度是体系有效运行的保障。证券机构应建立健全一系列与APP安全与合规相关的管理制度，并可将其归纳为几大类。例如，在网络安全及应急响应制度方面，要有顶层的《网络安全管理办法》，以及配套

的《网络安全事件应急管理办法》、《应急处置预案》、《APP应急操作流程》和相应的演练记录等。在数据安全及个人信息保护制度方面，需制定《数据安全管理办法》、《个人信息保护管理规范》，以及针对生物特征数据、第三方数据共享、用户账户注销与数据删除等的专项规范或指引。在系统运维及开发安全制度方面，应有《系统测试及上线管理办法》、《信息系统运维管理办法》、《软件开发安全规范》、《第三方SDK引入与管理规范》、《API接口安全防护管理规范》等。这些制度共同构成了APP安全与合规管理的规则基础。

组织架构：有效的组织架构是确保管理体系落地的关键。通常，证券机构党委（或党组）是最高决策层，下设网络和信息安全领导小组以及信息技术治理委员会。管理层则设有网络和信息安全工作小组，并应成立专门的APP安全与合规工作小组，负责统筹协调。执行层面，需要明确各团队职责：安全团队负责平台工具建设、制度规范制定、安全审核支持、渗透测试等；互联网金融团队（或相关业务/技术部门）负责安全需求分析、功能设计、运维应急；研发团队（含外部开发商）需遵循规范开发测试、管理开源组件；测试团队（含外部开发商）负责安全测试需求对接与执行。此外，应设立风控合规专员角色，负责梳理确定合规要求、审核隐私政策等。业务、投保（投资者保护）、合规、风控、审计等部门也需深度参与，形成多方协同、职责清晰的组织保障。

工作流程：清晰的工作流程能够确保跨部门协作顺畅高效，形成有效的风险防线。在APP生命周期的早期，业务部门提出需求，互金团队进行安全需求设计，APP架构师审核APP安全需求设计，安全团队与风控合规专员提供安全与合规支持，进行审核并解答疑问。开发团队依据安全设计进行开发。测试团队执行安全功能测试，渗透测试人员（内部或外部红队）进行深入渗透测试。安全团队在此过程中提供全流程的安全赋能和技术支撑，并对发现的问题进行反馈。以上构成了第一道防线。同时，风控合规部门、审计部门等作为第二道、第三道防线，对流程进行监督和控制。投保部门负责接收用户关于仿冒APP的投诉，需要与安全团队、业务部门协同处理。整个流程由APP安全与合规工作小组统筹，需要依赖跨部门合作。

平台工具：技术工具是提升APP安全与合规管理效率和能力的重要支撑。构建一体化的APP安全平台，可以整合各类安全合规工具能力，如开发测试环节的源码扫描（SAST）、交互式扫描（IAST）、制品库管理、隐私检测、发布环节的应用加固、版本管理、变更管理等。该平台应实现对安全工具的统一管理调度和统一用户界面。更进一步，平台可以提炼不同厂商工具的能力，对上提供抽象且统一的安全合规能力；实现工具能力的整合与关联分析，针对特定问题进行联合验证或交叉验证，自动关联不同工具的分析结果。通过将这些工具覆盖到APP开发、测试、发布、运行的各个阶段，能够显著提升安全防护和合规检测的自动化水平和深度。

三、APP监管合规要求

（一）APP安全规范

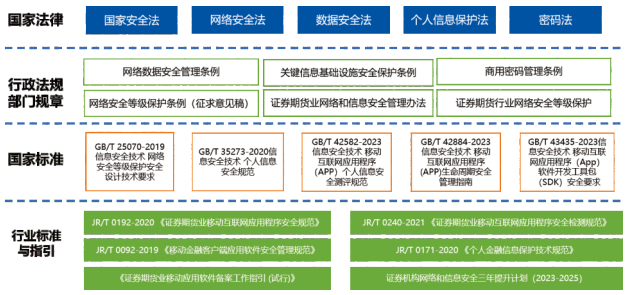


图4 APP安全规范

APP安全方面涉及的规范体系层级分明、内容广泛，如图4所示。顶层是国家法律，如《网络安全法》、《数据安全法》、《个人信息保护法》、《密码法》等，这些是所有网络活动的基础准则。其次是行政法规和部门规章，如《网络数据安全管理条例》、《关键信息基础设施安全保护条例》、《商用密码管理条例》、证监会的《证券期货业网络和信息安全管理办法》和证券期货行业网络安全等级保护相关要求等，对法律进行了细化。再往下国家标准（GB/T）提供了具体的技术和管理指引，如等级保护设计要求（25070）、个人信息安全规范（35273）、APP个人信息安全测评规范（42582）、APP生命周期安全管理指南（42884）、SDK安全要求（43435）等。最后是金融行业标准（JR/T）和协会指引，如证券期货业APP安全规范（0192）、安全检测规范（0240）、移动金融客户端安全管理规范（0092）、个人金融信息保护技术规范（0171）以及APP备案工作指引等。此外，《证券机构网络和信息安全三年提升计划》也是重要的行动指南。

（二）APP个人信息保护规范

在APP个人信息保护规范中，上述国家法律和行政法规是基础。除此之外，四部委联合发布的《App违法违规收集使用个人信息行为认定方法》（191号文）是重要的执法依据。工信部近年来发布《工业和信息化部关于开展APP侵害用户权益专项整治工作的通知》和《关于开展纵深推进APP侵害用户权益专项整治行动的通知》，持续开展APP侵害用户权益专项整治。网信办也出台了针对儿童个人信息保护、算法推荐等的专门规定。这些构成了个人信息保护监管的主要政策框架。

国家标准（GB/T）为个人信息保护提供了详细的技术和管理规范。其中，《个人信息安全规范》（GB/T 35273）是核心基础标准。针对APP场景，《收集个人信息基本要求》（GB/T 41391）定义了“最小必要”原则，《个人信息安全测

评规范》(GB/T 42582)提供了评估方法。其他相关国标还包括告知同意实施指南、SDK安全要求、预置应用安全要求、应用商店审核指南、平台处理规则等。

在行业标准方面，金融行业的《移动金融客户端应用软件安全管理规范》(JR/T 0092)和《个人金融信息保护技术规范》(JR/T 0171)也至关重要。此外，电信终端产业协会(TAF)和全国信息安全标准化技术委员会(TC260)发布的团体标准和实践指南，也提供了非常有价值的参考。

(三)APP行业监管要求

根据证监会行业监管要求，针对APP安全合规专项工作的检查通常会关注以下要点：是否已将APP安全合规内容纳入信息技术管理体系；是否按要求完成了在行业协会的备案登记；是否通过了证券期货业App安全认证；是否定期开展了APP客户端和服务器端的安全性检查；是否进行了个人信息保护专项检测。特别是在个人信息保护方面，检查会关注：APP是否明确说明了收集使用个人信息的目的、方式和范围；是否存在未经用户同意就收集使用个人信息的行为；是否违反必要原则，收集了与服务无关的个人信息；是否清晰公示了个人信息使用规则；以及是否建立并公布了有效的投诉、举报方式。这些都是监管高度关注的核心合规点。

同样，对于投资者个人信息保护的检查，行业监管也有明确的关注点：是否建立健全了个人信息管理机制和岗位职责要求；处理个人信息时，是否明确告知了目的、方式、范围和隐私政策，并关注是否存在超范围收集和使用的情况；是否存在违规截取、留存、提供客户信息的行为，或允许第三方违规存储使用的情况；在公司网络安全边界外处理个人信息时，是否采取了数据脱敏、加密等保护措施；通过短信、邮件等非自主渠道发送敏感信息时，是否进行了脱敏处理；对于生物特征识别的应用，是否进行了必要性和安全性的风险评估，是否强制用户使用某一生物特征或将其作为唯一认证方式。这些要点直接关系到投资者权益保护的落实情况。

四、APP安全关键技术与实践

本章将整合讨论证券行业APP在其全生命周期中应采用的关键安全防护技术，以及个人信息保护这一核心合规要求的具体技术实践。安全与合规并非孤立，而是相互交织、共同构成了APP稳健运行的基础。

(一)APP全生命周期安全防护

APP的安全防护应贯穿其整个生命周期，形成闭环管理，这种“安全左移”和“全程覆盖”的理念是现代应用安全的核心，如图5所示。

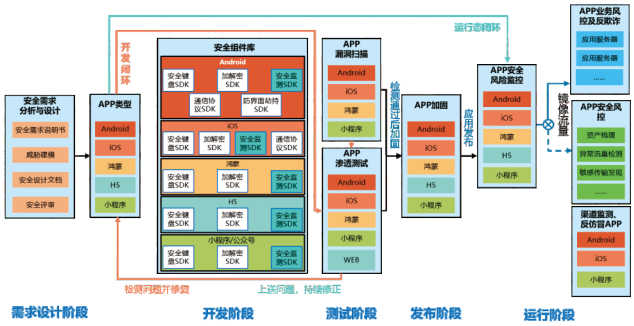


图5 APP全生命周期安全防护

1、开发阶段安全防护

在APP开发阶段，建立健全的安全防护机制至关重要。目标是确保App免遭逆向分析、篡改和调试等攻击，保护核心代码、核心数据，维护交易环境安全。关键技术包括：防逆向保护，通过加密、加壳、虚拟化等技术增加破解难度；防篡改保护，利用完整性校验、交叉校验等技术防止恶意代码注入和逻辑篡改；防调试保护，采用反调试技术对抗内存dump、hook等高级攻击；防泄露保护，通过透明加解密保护配置文件、数据文件等。同时，要加强通信协议保护，采用加密、完整性校验、白盒加密、身份验证等技术保障数据传输安全；要积极推进国密算法改造，在关键环节使用国产密码算法；并完成IPv6改造，确保在IPv6环境下安全策略同等有效，并符合中国人民银行IPv6安全合规要求。这些措施的有效落实将有助于从源头解决APP应用完整性、输入保护、运行环境、密码存储与传输、会话管理等常见安全问题。

2、测试阶段安全测试

安全测试是验证安全防护措施有效性、主动发现并修复安全漏洞的关键环节。全面的安全测试机制能够带来多重益处：显著提升开发效率并降低修复成本，因为在开发生命周期的早期发现并修复漏洞，其成本远低于在生产环境中修复；有效增强业务连续性与用户体验，通过提前消除可能导致服务中断或性能下降的安全隐患；强力保护用户敏感数据与交易安全，防范黑客攻击、数据泄露、欺诈交易等风险。安全测试通常包括前期沟通(确定范围、目标、时间、流程)、漏洞扫描(APP脱壳、漏洞检测、验证、分析)、渗透测试(信息收集、执行测试、清理痕迹等)、结果输出(报告编写、讲解、成果确认)等步骤。

3、发布阶段安全加固

在APP通过所有安全测试、准备正式发布之前，进行安全加固是提升其在真实运行环境中抵抗攻击能力的最后一道重要屏障。APP加固是以自动化的方式通过加壳、混淆、加密、签名校验、进程保护等技术手段，在防范APP/SDK被代码反编译、完整性篡改、动态调试、本地数据明文窃取等方面提供针对性的防护措施。由于不同移动平台的特性差

异，加固方案也需有所侧重：Android加固侧重阻止代码被反编译、阻止篡改二次打包、阻止动态调试注入、资源加密和环境检测；iOS加固侧重增加逆向难度、阻止篡改二次打包、阻止动态调试注入、字符串加密和越狱检测；鸿蒙加固类似Android；SDK加固则侧重保护其核心代码不被逆向分析和调用逻辑不被篡改；H5和小程序加固主要针对JS代码进行混淆、加密、防调试和防篡改。

4、运行阶段安全监控与反欺诈

APP上线后需持续监控风险，从被动防御转向主动感知。此阶段的核心目标是实时监控APP的运行环境、用户行为以及交易流程中的异常风险。技术实现上，通常需要一个综合性的监控平台，其数据来源多样：通过在APP客户端嵌入探针（SDK）、在H5页面或小程序中部署JS探针、在后端服务的网关层面部署插件或采集镜像流量获取API调用和网络通信数据。这些数据被汇集到后端的风险分析引擎，利用机器学习等技术，对设备环境（如代理、位置欺诈、注入攻击）、访问行为（如高频访问、自动化工具）、数据传输（如敏感数据泄露、重放攻击）等进行实时监测和分析。同时，建立资产与漏洞风险的关联机制，管理资产安全风险。监控需覆盖APP及其相关的微应用、H5、小程序等全渠道客户端以及API流量层面，并建立前后端数据比对等可信机制。目标是判断访问是否来自正常设备、使用正常手段、来自正常渠道，访问频率行为及报文是否正常，从而全面发现异常，为风险处置提供依据，保障APP安全稳定运行。

5、与业务风控联动

技术层面的运行时安全监控若能与业务层面的风险控制或反欺诈系统有效联动，将能极大提升整体风险防控的效能。将移动端技术维度的安全监测平台与应用风控系统或业务反欺诈系统相结合，帮助实现业务风控对全渠道安全事件的精准捕捉、分析及处置。这种联动采用“客户端”与“平台端”结合的模式。客户端侧，监测SDK作为风险感知探针，探测到风险后可直接传递给APP应用层，使其能迅速响应，如阻断高风险操作。平台侧，APP安全监测平台将丰富的监测点信息智能编排成结构化的安全事件，推送给业务风控平台。业务风控平台基于风险等级统一决策并执行响应策略（如二次验证、预警、阻断）。优化的安全事件策略也可反向下发至客户端SDK驱动响应。本质上，APP安全监测平台充当风险信息的传递者和分析者，业务风控平台基于这些信息统筹执行防护策略，共同构建坚固的业务风险防护屏障。

6、渠道监测与反仿冒APP治理

近年来，仿冒APP是证券行业面临的一类重要的安全风险，严重损害用户利益和公司声誉。建立常态化的APP渠道监测与反仿冒治理机制至关重要，可以确保用户通过正规渠道下载使用。建议依托专业的安全服务机构进行APP全渠道监测，目标是形成全面精准的安全风险画像。好的监测服务应具备：广泛的渠道覆盖，监控应用市场、搜索引擎、社交平台等，涵盖Android、iOS、小程序、公众号等全形态；精准的识别能力，基于多维度特征（名称、包名、签名、图标、界面相似度等）结合黑白名单和机器学习进行判断；7*24小时的持续数据更新与监控；以及高价值的数据输出，自动关联合法APP，人工复核疑似风险APP，提供精准全面的情报，支撑后续处置工作。

（二）APP个人信息保护实践

在数字化时代，个人信息保护不仅是法律法规的强制要求，也是证券机构赢得用户信任、实现可持续发展的关键。

1、个人信息保护管理框架

构建APP个人信息保护体系，需秉持战略性思维，在满足合规要求与支撑业务发展之间寻求平衡。顶层设计应确立清晰的个人信息保护方针和目标，并建立一个覆盖组织保障、制度流程、技术工具三维度的综合治理模式。该模式必须贯穿于个人信息从收集、存储、使用、加工、传输、提供、公开到删除的整个生命周期。基于此顶层设计，应构建个人信息保护管理框架，搭建完备的组织架构，出台详尽的规章制度，规范APP个人信息收集、存储、使用和共享流程，定期开展APP安全与隐私合规评估，全面排查信息安全隐患。以降低数据外泄风险、确保处理行为合法为目标，搭建完备的组织架构，同时出台详尽的规章制度，如个人信息保护政策、事件应急方案、信息泄露管理规范、隐私政策设计指引等。个人信息保护管理措施包括影响评估、合规评估、安全防护、审计，整个过程应遵循PDCA循环，持续改进，如图6所示。

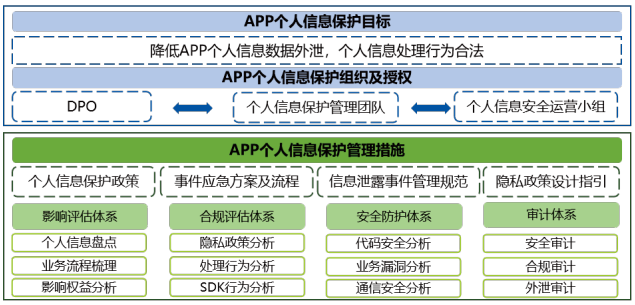


图6 APP个人信息保护实现路径

2、隐私合规检测

为确保持续符合个人信息保护要求，证券机构应建立针对APP个人信息保护的常态化隐私合规检测机制。考虑到APP版本迭代快、功能复杂，建议采用“自动化检测工具＋人工分析验证”相结合的模式。自动化工具可以高效完成基础性检查，例如：静态分析进行APP漏洞扫描和组件成分分析

（识别第三方组件）；动态分析在沙盒环境中模拟用户操作（自动界面遍历），记录APP的行为（如权限调用、数据传输、数据存储），并与预设的合规规则比对。在此基础上，由专业人员进行人工审核，重点分析截屏录屏、调用权限等APP事件以及隐私政策合规性等，从而实现检测效率与深度的平衡。

3、隐私威胁管控

在APP个人信息保护中，隐私威胁管控是一个重要环节，特别是针对第三方SDK的隐私威胁行为。核心目标是实现对隐私威胁行为的“可见”与“可控”。常见的隐私威胁行为主要包括个人信息的违规采集、个人信息的违规外发以及利用APP的热更新/动态加载技术来逃避常规隐私威胁检测。针对这些隐私威胁行为，管控策略主要围绕两类权限展开：一是终端个人信息采集权限的监控与管理，二是网络通信外发权限的监控与管理。基于此，可以实施两种策略：对APP隐私威胁行为进行监测审计，发现问题；或者更进一步，实施阻断管控，直接阻止违规行为的发生。

五、APP备案与检测认证

证券行业APP符合监管要求并通过必要的备案与检测认证，是其得以合法合规运营的基础保障，监管部门高度重视并持续推动备案与检测认证工作。2023年6月，证监会办公厅印发通知，进一步强调了证券期货经营机构需加强移动应用软件安全管理，落实备案和检测认证要求。随后，2024年8月，多地证监局也下发了类似通知，明确要求辖区证券期货经营机构报送APP备案和检测认证的计划，要求在2024年底前完成至少50%，2025年底前完成所有面向投资者APP的备案登记和检测认证工作。2024年12月，证券、期货、基金业协会发布了《备案工作指引（试行）》，进一步规范了APP的备案流程。这些文件表明，完成APP备案和检测认证已成为一项硬性的合规任务，同时也是机构向市场和投资者展示其安全与合规能力的有效途径。

（一）APP备案

根据国家和行业监管要求，证券APP在上线提供服务前或上线后一定时间内，必须完成一系列备案手续，这是其合法合规运营的必要前提。备案工作主要涉及向三个不同的主管部门进行登记，完成这三项备案是证券APP合规上线的基础性工作，缺一不可。

1、证券行业备案（向中国证券业协会备案）

此项备案依据《证券期货业移动应用软件备案工作指引(试行)》进行，旨在加强行业自律管理。流程主要通过证券期货业移动应用软件备案系统线上备案。完成首次备案后，备案信息并非一成不变，当APP的关键信息（如版本重大更新、功能调整、开发商变更等）发生变化时，应及时更新备案信息。

2、工信部备案（向工业和信息化部备案）

此项备案依据《互联网信息服务管理办法》进行，旨在落实APP主办者的主体责任。备案通常通过APP的网络接入服务提供者（如服务器托管商、云服务商）或分发平台（针对小程序、快应用等形态）的企业侧备案系统，向APP主办者（即证券公司）主体注册所在地的省级通信管理局提交申请。管局审核通过后，APP将获得唯一的ICP备案号（针对网站）或APP备案号。按规定，此备案号应在APP的启动加载页、设置页的“关于”或“版本信息”等显著位置进行标注。

3、公安部备案（向公安机关备案）

此项备案依据《计算机信息网络国际联网安全保护管理办法》进行。要求在APP获得工信部备案号（即服务开通）后的30日内，登录“全国互联网安全管理服务平台”（www.beian.gov.cn）进行备案。备案成功后，系统会生成公安备案号，同样建议在APP的适当位置（如“关于”页面）进行标注。公安备案旨在落实网络运营者的网络安全保护义务。

（二）APP安全检测认证

证券期货业APP安全检测认证是监管机构为了确保APP达到一定的安全标准、保障投资者信息和资金安全而设立的重要合规要求。该项工作的主要依据标准是《JR/T 0240—2021 证券期货业移动互联网应用程序安全检测规范》和《App违法违规收集使用个人信息行为认定方法》（四部委191号文）。安全检测工作由多家获授权机构执行，而安全认证工作则由中证信息技术服务有限责任公司（中证技术）负责实施和管理。总体流程大致如图7所示：送检机构（即证券公司）首先按照中证技术发布的申请资料说明准备材料并提交申请，受理后由指定的检测机构进行型式试验（即技术检测），检测机构出具检测报告后，由认证机构结合报告、文件检查和现场检查结果做出认证结论，通过后颁发证书。

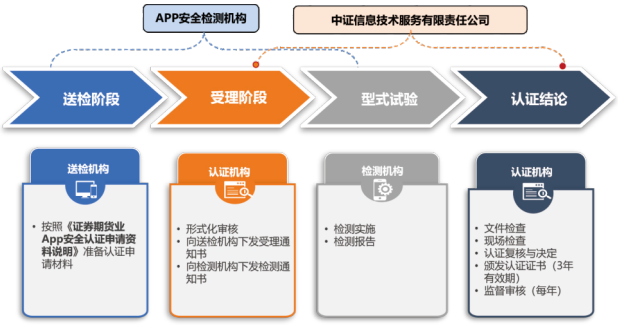


图7 APP安全检测认证流程

证书的获取与维持并非一劳永逸。认证证书的有效期通常为3年，在有效期内每年都需进行一次证后监督审核（监审），证书到期前需申请续证审查。这一机制强制机构必须建立和维持一个长效的安全与合规管理机制，而不是为了认证而进行一次性的突击整改。

从过往的检测认证实践来看，有几个领域是问题的高发区。统计数据显示，个人信息保护相关的问题占比最高（约40%），其次是网络通讯安全（约18%）和移动终端安全（约17%）。这些问题的根源往往在于开发设计阶段对法规标准理解和落实不到位、隐私政策的制定与调整涉及多方协调导致整改困难、对第三方SDK的安全与合规性管控不足，以及在产品设计中过于侧重用户体验而牺牲了一部分安全性考虑等。

为了能够更顺利地通过检测认证，建议机构从三个阶段着手：在开发阶段，就要强调规范管理，严格遵守安全设计和编码规范，加强内部的安全自测，高度重视个人信息保护要求，并建立严格的第三方SDK引入审查和管理机制。在送检阶段，要与检测认证机构保持良好沟通，确保按要求提供完整准确的测试材料和稳定、符合要求的测试环境。在检测与整改阶段，一旦收到问题报告，要认真分析、充分理解整改要求，制定切实可行的整改方案，并就技术实现细节与检测机构积极沟通，力求高效、准确地完成整改，避免因反复整改而延长认证周期。

六、总结

针对证券行业APP面临的安全与合规挑战，本文系统性地阐述了一套贯穿APP全生命周期的安全与合规管理体系，梳理了APP监管合规要求及APP安全关键技术，帮助推动安全与合规从被动的外部要求转变为内生的核心能力。此外，整理了APP监管备案及安全检测认证流程，分析了认证常见问题并提供改进建议。证券经营机构应当将安全与合规要求深度融入业务流程和技术架构，才能系统性地提升APP的安全与合规水平，赢得投资者信任，实现稳健和可持续发展。

参考文献

1.GB/T 43435-2023，信息安全技术 移动互联网应用程序（App）软件开发工具包（SDK）安全要求[S]. (National standard GB/T 43435-2023, Information security technology—Security requirements for mobile internet application (App) software development kit (SDK)[S].)

2.中华人民共和国网络安全法[EB/OL]. (2016-11-07)[引用日期]. http://www.npc.gov.cn/npc/c30834/201611/0d5ba00cdc3c4493a71620fca5897500.shtml. (Cybersecurity Law of the People's Republic of China[EB/OL]. (2016-11-07)[Access date]. http://www.npc.gov.cn/npc/c30834/201611/0d5ba00cdc3c4493a71620fca5897500.shtml.)

3.中华人民共和国数据安全法[EB/OL]. (2021-06-10)[引用日期]. http://www.npc.gov.cn/npc/c30834/202106/7c97ef437577419386cd3419356c1828.shtml. (Data Security Law of the People's Republic of China[EB/OL]. (2021-06-10)[Access date]. http://www.npc.gov.cn/npc/c30834/202106/7c97ef437577419386cd3419356c1828.shtml.)

4.中华人民共和国个人信息保护法[EB/OL]. (2021-08-20)[引用日期]. http://www.npc.gov.cn/npc/c30834/202108/a8c4e3672c74491a80b53a1720dabb89.shtml. (Personal Information Protection Law of the People's Republic of China[EB/OL]. (2021-08-20)[Access date]. http://www.npc.gov.cn/npc/c30834/202108/a8c4e3672c74491a80b53a1720dabb89.shtml.)

5.JR/T 0192-2020，证券期货业移动互联网应用程序安全规范[S]. (Financial specification JR/T 0192-2020, Security specification for mobile internet applications in securities and futures industry[S].)

6.JR/T 0240-2021，证券期货业移动互联网应用程序安全检测规范[S]. (Financial specification JR/T 0240-2021, Security testing specification for mobile internet applications in securities and futures industry[S].)

安全运营韧性构建：基于信创态感的智能化安全运营探索实践

丁安安、严星宇、夏英杰、赵川 | 国联民生证券股份有限公司

摘　要： 在信息技术飞速发展、网络安全威胁不断涌现的当下，构建具备技术韧性的安全运营体系，已成为保障网络空间稳定运行的关键路径。本文结合我单位的实际运营经验，系统梳理了当前安全运营面临的复杂挑战，提出以信创态势感知平台为核心，打造覆盖数据全生命周期的智能化运营体系。通过融合人工智能与大数据分析能力，实现安全事件的智能识别、精准定位与闭环处置，显著降低对外部技术依赖，提升供应链安全水平。实践表明，该模式在证券行业具备可复制性与推广价值，为构建全栈可信的安全运营能力提供了可行路径。

关键字： 信创态感，安全运营韧性，网络威胁应对，智能化运营探索

一、引言：安全运营韧性构建——时代背景与核心命题

（一）网络安全发展形势

当下，网络安全发展形势呈现出复杂且严峻的态势。随着数字化进程的加速，网络攻击手段日益多样化和高级化。黑客组织不再局限于简单的恶意软件传播，而是采用更为隐蔽的APT攻击，长期潜伏在目标网络中窃取敏感信息。同时，云计算、容器等新兴技术的广泛应用，使得网络攻击面不断扩大，安全防护难度剧增。此外，国际网络空间的竞争与冲突不断升级，网络战成为国家间博弈的新领域，关键信息基础设施面临前所未有的安全威胁。在此背景下，传统的安全防护体系已逐渐难以满足需求，我们需要研究构建更加智能、高效、具备技术韧性的安全运营模式来应对日益复杂的网络安全挑战。

（二）安全运营韧性构建的价值体现

具备技术韧性的安全运营体系，是指安全运营工作依托可信替代的平台组件、自动化本地闭环流程、多源异构数据替代机制以及持续更新的知识沉淀库等内生能力，即使在遭遇外部技术供应中断、高级持续攻击等极端场景下，仍能持续、稳定、有效地完成检测、响应、恢复等核心功能。

从国家层面看，构建具备技术韧性的安全运营体系，是保障关键信息基础设施稳定运行的前提。对企业而言，降低对单一技术来源的依赖、提升供应链安全能力，可有效应对外部服务中断、技术封锁等极端场景，同时根据自身业务特点灵活定制安全策略，提升运营效率与风险应对能力。

（三）研究重点：信创态感驱动下的智能化安全运营实践

态势感知平台作为网络安全领域的重要基础设施，需要实时感知和全面掌握网络安全攻击态势，其平台可信建设尤为重要。通过将信创技术与态感理念相结合构建智能化安全运营体系，该体系利用信创环境下的先进技术，如大数据、人工智能等，对海量的网络安全数据进行实时采集、分析和挖掘，实现对安全威胁的精准识别和预测。同时，基于态感平台分析结果智能化地调整安全策略，实现自主联动响应和处置，提升安全运营的效率和精准度。

二、传统安全运营困境与韧性构建的迫切性

（一）核心技术供应链安全风险造成的平台韧性不足

在网络安全领域，核心技术环节对于国外技术的过度依赖构成了体系中的关键潜在风险点。当前，在芯片、操作系统、数据库等关键信息技术领域，全球供应链呈现一定的集中化态势。这种集中化态势在全球地缘政治经济环境的不确定性的情况下，可能导致关键技术供应链出现波动或中断风险，比如，其潜在漏洞的发现、响应和修复周期可能受制于外部因素，一旦被恶意利用，对网络系统安全的威胁更为直接且难以有效溯源处置，同时也使得数据安全治理的透明度和可控性面临挑战，增加了关键信息在存储、处理和传输过程中被未授权访问或泄露的潜在风险。

因此，突破核心技术瓶颈，构建可信的技术供应链，是提升国家网络空间安全韧性和保障关键信息基础设施安全的战略基石。

（二）系统断点和协调缺口造成的运营韧性不足

在实战中，安全运营响应会因为系统层与运营层两个维度的联动缺失而显得韧性不足。

在系统层面，态势感知平台与周边关键系统之间的集

成度不足，形成了“数据孤岛”，制约了响应效率。例如，与CMDB（配置管理数据库）的联动缺乏自动化机制，导致安全告警无法快速、准确映射到资产责任人，延长了事件确认与指派时间；与IT流程平台的对接不完善，则使得事件处置难以实现线上化闭环管理，不仅增加了运营成本，也影响了审计与复盘的有效性。在运营层面，当安全团队与运维、业务部门之间的职责边界模糊、信息同步不及时，会导致应急响应过程中决策链条过长、动作协调不一，会导致跨团队、跨部门的协同机制存在明显短板。此外，传统基于静态规则的联动模式缺乏对新型攻击的适应性，难以在复杂攻击场景下实现智能决策与自动化动作执行，反映出整体安全运营体系在弹性与自适应能力上的不足。

（三）安全运营韧性构建战略意义与需求

在网络安全运营语境下，技术韧性已成为与国家网络安全战略、企业业务连续性同频共振的核心目标。其战略意义体现在：通过底层硬件、平台软件到运营流程的可信设计，能够在供应链受限、高强度攻击等场景下依旧保持检测、响应、恢复、演化的持续运行，防止因技术断供或数据泄露引发的系统性风险。从需求维度看，数字化业务的高并发、高依赖特征使威胁暴露面指数级放大，唯有构建具备技术韧性的安全运营体系，方可依据自身业务演进灵活调整策略、快速适配新型攻击手法，并凭借多源替代、知识沉淀和内生闭环显著降低对外部单点技术与产品的依赖程度，为网络安全能力的可持续迭代奠定坚实基础。

三、多层级韧性构建：从平台到运营

（一）平台层可信构建：信创态感平台

信创态势感知平台是我单位安全运营的核心载体，依托国产化硬件与软件深度整合搭建了新的态势感知引擎。该平台有效实现了企业安全运营全流程的自主管理能力与风险把控能力提升，构建起从风险识别、分析研判到应急处置的完整闭环，不仅充分契合证券行业对安全运营平台“高可靠、高可用、高安全”的核心要求，更为行业关键业务的安全稳定运行筑牢技术根基。

1、信创基础设施

平台的基础设施层采用全栈国产化组件，构建了稳定高效的运行环境。

硬件层面基于国产高性能处理器架构进行服务器部署，该架构具备优异的计算性能与内置安全特性，支持国密算法加速，有效保障了数据处理效率与加密安全需求。

操作系统选用符合金融行业安全标准的国产操作系统。该系统经过深度优化，具备完善的安全审计、进程隔离与恶意程序防护机制，为上层应用提供坚实的安全底座。

数据存储采用支持高并发、分布式事务处理的国产关系型数据库。该数据库能够高效承载安全运营平台日均数千万级日志数据的存储与分析任务，并通过数据加密、细粒度访问控制等机制确保敏感安全数据的安全存储。

中间件层面使用高可靠国产中间件。它为平台各模块间的通信与数据交互提供稳定支撑，其负载均衡与故障转移能力是保障平台持续稳定运行的关键选用东方通中间件，为平台各模块间的通信与数据交互提供稳定可靠的支撑，其具备的负载均衡与故障转移能力，确保了安全运营平台的连续稳定运行。

2、自主态势感知引擎

态感平台内置引擎采用国产化编程语言开发，能深度适配国产化组件，充分释放硬件性能。它具备强大的IAAS层适配能力，可针对不同操作系统和硬件做适配封装，并通过引入容器技术containerd，屏蔽底层差异，提供通用信创适配能力。

态势感知引擎日志处理能力出色，通过分布式计算架构，每秒可实时解析与关联分析8000条以上日志。内置证券行业业务场景特征库与分析模型，可以结合业务特性精准研判安全事件。

（二）能力层韧性设计：智能分析与高效响应

对于传统类型攻击以及个人攻击者所发起的单次攻击，常规规则检测方式往往能够应对自如，实现有效识别与防范。但对于隐蔽性、长链路攻击行为，则检测难度较大，需要将各安全设备的告警进行串联，通过关联分析形成攻击长镜头，这不仅是态势感知平台履行核心职能的必然要求，更是衡量整体安全运营体系防护深度、响应效率与决策智慧的关键指标，能够全方位彰显安全运营的综合水准。我单位目前初步构建了“检测-分析-处置-溯源”的安全运营能力体系，实现了从被动防御向主动狩猎的战略转变。

1、多源异构数据智能融合与时空上下文关联

我司以信创态势感知平台为安全运营底座，接入了全流量威胁分析探针、IPS、WAF、沙箱、威胁情报、DNS解析等多元数据。原始流量数据经网络流量分析设备处理后，产生的异构告警日志输入到大数据流式消息队列。由于攻击者入侵常采用一系列关联手段，所以需按时空上下文关联事件。平台结合APT知识图谱本体结构，导入预定义攻击范式化模型及设备与攻击模式等的映射文件，利用大数据键值关联，快速将多源异构海量告警范式化为统一格式安全事件。同时，基于攻击链模型，把同一相近时间段、针对同一目标的一系列攻击事件，按时间序列和攻击链阶段整合生成攻击链，自动化整合一定周期内针对同一目标资产的多个安全事件，还原更准确全面的攻击场景。

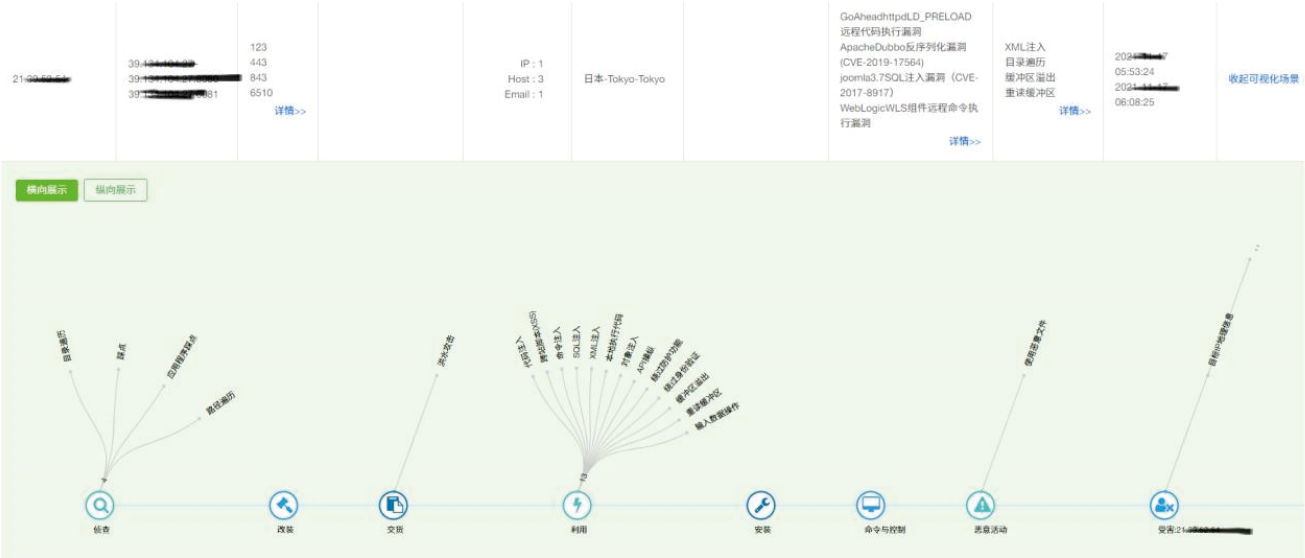


图1 安全事件整合攻击链示意图

2、攻击组织特征关联计算

借助大数据的快速关联能力，分析模型平台能够自动化地依据档案库知识，对安全事件展开情报拓线工作。同时，基于事件扩充后的威胁语义，开展团伙自动化归因关联计算，进而达成对事件以及攻击链的团伙归因。完成关联计算后，平台会把归因结果以标签形式，标注在对应的攻击链与安全事件上；还可以同步展示团伙标签以及所有相关威胁语义，为研判人员提供参考依据。

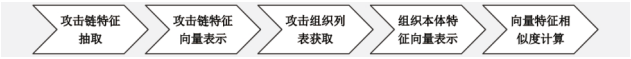


图2 攻击组织特征关联计算流程图

3、智能化分析与攻击结果判定

为高效应对海量多模态数据场景中，攻击组织关联的高危安全事件的快速识别难题，我们以攻击组织本体为基础，搭建起上下文感知计算框架。同时借助大数据流式计算技术，对多模态数据进行范式化解析，完成上下文信息的采集、关联以及特征相似度计算，从而从海量威胁告警里迅速锁定攻击组织相关的高危事件。

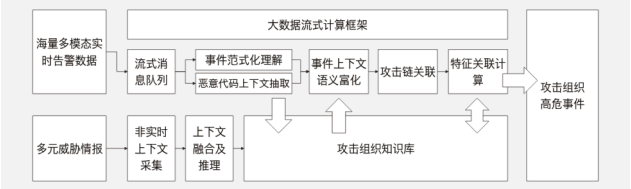


图3 基于上下文感知计算的攻击组织追踪方法总体框架

失陷资产判定是基于攻击识别环节所输出的威胁事件，综合考量其攻击阶段以及攻击结果，进而判断本地资产是否存在被攻陷的情况，判断标准包含单攻击阶段和多攻击阶段两种。由于部分事件存在置信度低、误报等问题，系统进行推断时不仅要按事件置信度、结果及白名单规则预处理数据，还需依据事件类型和攻击方向，对推理后的失陷场景做逻辑验证，以提升准确性。在实际运营场景中，失陷资产判定是生成运维事件的前提，预判引擎据此可调用第三方接口对资产实施一键封堵、钉钉告警、邮件通知等操作，以便于安全人员及时进行处置。

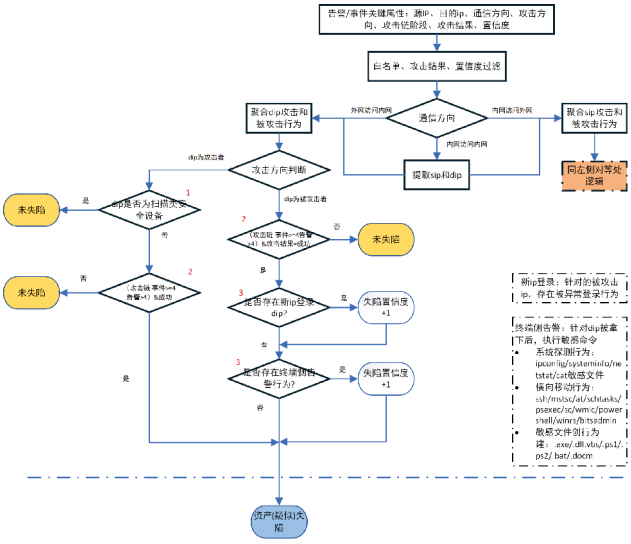


图4 失陷资产判定流程

（三）运营层韧性提升：流程与知识双驱动

1、流程化处理助力安全运营效能提升

在日常安全管理工作中，我单位结合实际场景逐步探索出一套相对完善的运营流程。态势感知平台凭借自主研判能力，对接入的多元数据进行事件分级（一般、较大、重大）与分类（如扫描探测、远程代码执行等）。针对一般事件，通过钉钉和邮件每日输出聚合告警，避免信息冗余；较大和重大事件则实时发送告警，确保及时响应。对于高置信度攻

击源IP地址，平台可自动调用SOAR剧本封禁。

安全运营人员收到告警后进行人工分析，若为误报，则在数据源安全防护设备和态势平台调整检测规则、加白降噪；若为真实事件，通过跨团队协同机制进行人工封禁。针对系统本身存在的缺陷或漏洞，安全运营人员在OA中推送漏洞修复流程，修复后进行复测验证。

事件处置完成后，安全人员及时总结处置经验，优化策略并将经验总结纳入团队知识库。通过构建完整的事件闭环管理体系，从数据接入到分析处置，再到经验总结与体系优化，有效地提升了安全运营工作的效率。

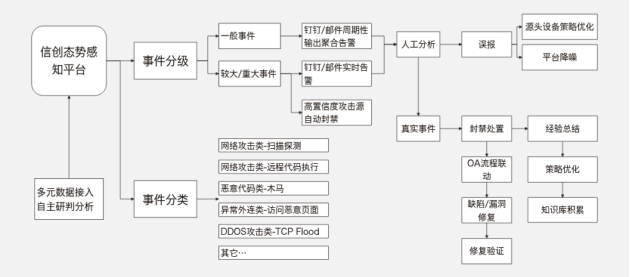


图5 国联民生证券安全运营流程

2、知识库的积累与能力复用

为切实提升安全运营韧性与整体效能，我们积极运用工作知识库开展系统性建设与实践。一方面，依托知识库对安全运营知识进行全面积累与动态更新，涵盖安全策略、应急预案、风险案例等核心内容，形成标准化、结构化的知识资产，为安全运营提供坚实支撑。另一方面，借助知识库搭建内部协作平台，通过技能共享、案例研讨、群策群力等机制，推动安全团队人员整体技能提升与经验传承。同时，通过定期的内部培训与实战演练，将知识库资源转化为团队实战能力，确保安全团队具备快速响应与协同处置能力，逐步实现每一位专职安全人员都可以参与到安全运营工作中来的目标。

四、实践探索：信创态感智能化运营的落地实施

（一）信创态势感知平台建设

在组件选型阶段，我单位结合证券行业安全运营的业务需求与技术特点，建立了多维度的选型标准。优先选择通过金融行业信创适配认证的产品，确保组件的合规性与稳定性；同时注重组件的性能指标，如服务器的并发处理能力、数据库的事务响应速度等，以满足安全运营的高负载需求。

整体平台建设分为三个阶段逐步推进：

第一阶段完成基础设施的部署与调试，搭建基于国产CPU和操作系统的基础环境，部署国产数据库和中间件，实现信创环境的基本运行。

第二阶段进行态势感知引擎的建设与集成，基于信创环境完成引擎的部署与性能优化，将日志采集、事件分析等核心功能模块与信创组件进行深度整合，实现安全运营的核心流程在信创环境下的顺畅运行。

第三阶段开展平台的功能验证与性能测试，通过模拟高并发日志输入、复杂攻击场景等方式，验证平台的稳定性与可靠性，针对测试中发现的问题进行优化迭代，最终实现平台的全面上线。目前，该平台已上线运行近1年，系统运行平稳符合预期。

（二）智能化能力的构建与优化

在智能化安全运营能力构建中，我们逐步探索落地了智能IP管控、漏洞驱动防护、钓鱼邮件防护、威胁情报AI研判四项基本能力：

（1）智能IP管控：基于威胁情报与实时流量分析，自动封禁高危IP，同时设置分级解禁机制：高频高危攻击源地址24小时后自动解禁；普通攻击源地址6小时后自动解禁。

（2）漏洞驱动的WAF防护：态势平台对接第三方漏洞库，自动解析漏洞特征并生成WAF自定义规则，经灰度验证后部署。在正式防护规则更新包推出前即实现对特定漏洞的快速精准防护。

（3）钓鱼邮件防护：平台接入了邮件沙箱投递和检测日志，一旦发现大量可疑邮件被正常投递，经过分析研判可自动启动邮件防护联动机制，自动更新邮件网关拦截规则（如拦截主题包含“管理员密码”邮件）。同时将告警信息推送给邮件系统管理员，对漏过的钓鱼邮件进行统一处理，避免员工误点击。

（4）威胁情报+AI研判：系统接入多个情报源，AI分析引擎可自动解析恶意IP特征（如历史攻击记录、关联域名），生成结构化研判报告并推送给运营人员，缩短人工研判时间并提升准确率。

此外，我们还探索大语言模型的嵌入，将业务系统训练的本地大语言模型扩展至安全场景，微调使其支持安全日志解析、历史事件知识库关联查询，提升事件调查效率。

（三）自主联动响应实践

基于信创底座的智能响应闭环，实现安全响应全流程自动化。通过可视化Playbook编排，将专家经验固化为标准处置流程。

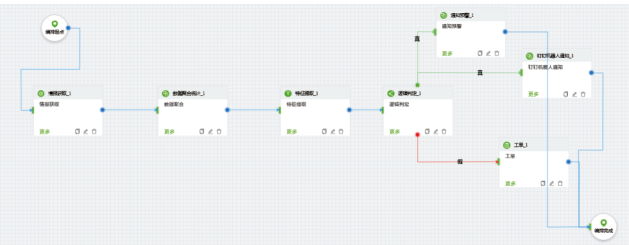


图6 可视化Playbook编排

流量探针与防火墙联动时，采用即插即用插件架构，支持多种第三方设备接口，实现IP封禁、流量牵引等动作的秒级响应。

（四）行业态势感知对接与联动

在信创生态下参与构建跨机构的行业协同防御网络，攻克传统安全体系“数据割裂、响应孤立”痛点。国联民生证券按照行业网络与信息安全态势感知平台建设要求，根据统一规范完成相关网络安全数据的按时报送。具体系统部署方面，在隔离区部署专用前置服务器，对接内部信创态势感知平台，通过标准化接口实时采集并预处理日志、告警及攻击链分析结果。数据经聚合后，按监管要求格式封装，通过加密通道自动上传至行业态势感知平台，保障传输效率与合规性。

（五）网络安全运营视角下的资产管理

在数字化安全运营的攻防实战中，动态精准的资产管理是构建主动防御体系的核心支点。传统IT资产管理主要聚焦硬件生命周期与静态配置记录，安全运营则需要构建以态势感知平台为中枢的"全域资产动态治理体系"。其核心优势在于：态势感知平台天然具备多源数据融合、风险关联分析与实时响应能力，能够整合网络流量、设备日志、漏洞情报等数据流，形成覆盖"暴露面识别-资产关联-风险量化"的完整闭环。

全域资产动态治理体系

图7 全域资产动态治理体系

从安全运营视角出发，资产管理体系需重构为三大核心维度：

（1）互联网暴露面资产：聚焦互联网可访问资产的全生命周期管理，包括域名、公网IP、开放端口及对应的服务类型，通过调用防火墙/WAF/负载均衡等设备接口实现每日暴露面变更的自动同步。

（2）内网核心资产：以CMDB为基础，通过自动化工具持续采集服务器、数据库、中间件等资产的IP、MAC、操作系统、应用组件及责任人信息，重点解决传统CMDB数据滞后问题，实现“资产变更-CMDB更新-态势感知同步”的快速联动。

（3）业务应用资产：深度解析应用架构中的组件依赖关系、数据流向及权限模型，通过内部应用安全管理平台管理开源组件漏洞，结合流量分析技术绘制应用服务调用链，精准定位脆弱环节。

五、结束语

本研究在信创环境下构建并验证了具备技术韧性的安全运营体系：通过国产化硬件、自研态感引擎与CMDB、流程平台的深度对接，实现了暴露面-资产-责任人-工单的全链路自动化闭环；结合多源替代组件与知识沉淀机制，确保在情报缺失或供应链波动场景下检测、响应、恢复、演化四大功能持续运行。实践表明，该体系不仅显著降低了外部依赖，也提升了关键信息基础设施的可持续防御能力。未来，将进一步融合大模型与知识图谱，持续优化韧性指标，为证券行业乃至关键信息基础设施的安全运营提供可复制、可扩展的范式。

数据合规驱动下的全域风险监测体系构建与实践研究——基于数据处理活动的安全运营创新路径

陆滢、徐正伟 | 华安基金管理有限公司

摘　要： 在数字经济时代，数据合规已成为驱动数据安全治理体系变革的核心力量。随着《数据安全法》《个人信息保护法》等法规深入实施，数据安全建设正从传统技术防护向“合规要求+业务需求”双轮驱动转型，治理范畴从IT基础设施延伸至数据全生命周期。为应对这一变革，我们研究构建了覆盖“三纵三横”的全域风险监测体系。这项以合规为基础、以风险为导向的治理模式，通过将合规要求嵌入数据处理全流程，不仅降低了合规风险，更为组织释放数据要素价值提供了核心保障，标志着“合规即竞争力”的新范式正式形成。这种多维度的全域风险监测框架，也实现了从被动合规到主动风险控制的战略转型。

关键字： 数据合规、数据全域风险监测、数据全生命周期治理。

一、引言

（一）合规驱动下的数据安全新变革

在数字经济蓬勃发展的当下，数据已成为核心生产要素。据国家互联网应急中心统计，2024年涉及数据违规的网络安全事件同比增长37%，凸显出构建系统化合规监测体系的紧迫性。随着《数据安全法》《个人信息保护法》等法律法规的深入实施，数据安全治理正经历从技术防护向合规治理的范式转型。数据安全建设，从单一安全事件驱动升级为"合规要求+业务需求"双轮驱动，尤其在金融、政务等领域，数据跨境流动、供应链合规等新型监管要求倒逼组织重构安全体系。网络安全边界从传统信息系统延伸至数据全生命周期，涵盖数据采集合法性、加工处理合规性、流通共享安全性等多维度，覆盖 "人 - 系统 - 数据" 三维度的全域风险监测框架，为组织实现从被动合规到主动风控的转型提供理论与实践指引。

（二）合规驱动下的数据安全治理新范式

1、监管环境的演进与治理逻辑重构

新一代数据安全法规体系呈现 "全链条规制 + 精准化监管" 特征，在金融领域《跨境数据流动安全评估办法》要求组织建立覆盖数据出境全流程的风险防控机制；政务领域则通过《政务数据分类分级指南》推动数据资产的精细化管理。这种监管转型促使组织安全建设发生三大质变：

驱动模式双轨化：从传统单一事件响应升级为 "合规要求 + 业务需求" 的双轮驱动模型。将合规要求内化为业务流程的一部分，同时将业务需求嵌入到合规框架的设计中，寻求二者的最佳结合点。合规工作不仅是为了“避害”（避免损失），更是为了“趋利”（创造价值、赋能业务）。这是现代组织实现可持续、高质量发展和有效风险管理的必然选择。它要求深刻转变思维模式、优化治理结构、重塑流程、拥抱技术，并最终将合规从成本中心转变为价值创造伙伴。合规的边界转化为业务的跑道，让监管的约束力与创新的驱动力成为组织前行的双翼。

治理范畴全域化：安全边界从 IT 基础设施延伸至数据生命周期全环节。当前数据价值超越硬件/软件，成为攻击核心目标与合规监管焦点。安全需覆盖数据从“生”到“死”的全旅程，实现 “数据流到哪里，防护就跟到哪里”。安全责任从IT部门扩散至数据所有者、业务部门、第三方合作伙伴等全链条角色。全域化治理不是成本，而是数字化稳定生存的磐石。

技术要求体系化：监管环境的演进对数据安全治理提出了更高的要求，促使治理逻辑发生重构。信创技术体系与数据安全的深度融合，是适应这一趋势的必然选择。通过构建完善的信创技术体系、创新应用数据安全技术、建立融合的技术标准与规范以及技术创新生态建设，能够有效提升数据安全保障能力，实现信息技术的自主可控，促进数字经济的健康发展。

2、数据安全治理的价值重构

在监管环境与技术演进的双重驱动下，数据安全治理的价值坐标系正在重绘。合规不再是组织发展的 "紧箍咒"，而是激活数据要素价值的 "金钥匙"。当金融系统通过合规体系实现效率与安全的双重提升时，它验证的不仅是一种治理模式，更是数字经济时代的基本商业逻辑 —— 在合规的地基上，才能搭建起数据价值释放的摩天大楼。未来，"合

规即竞争力" 将从创新理念演变为商业常态，推动数字经济进入安全与发展并重的新境界。

二、立体化审计体系的三维构建策略

（一）外部监管合规对接机制

立体化审计体系的三维构建策略，结合外部监管合规对接机制和区块链存证技术，为组织的合规运营和审计工作提供了全面、高效的解决方案。通过纵向深度穿透、横向协同联动和技术赋能创新三个维度的构建，能够实现对组织运营的全方位审计；完善的外部监管合规对接机制，加强了组织与监管部门之间的沟通与协作；基于区块链存证技术的审计证据链构建，保障了审计证据的真实性和可靠性。在未来的发展中，随着技术的不断进步和监管环境的变化，立体化审计体系还需要不断优化和完善，以更好地适应组织发展和监管要求，为数字经济的健康发展提供有力保障。

（二）动态智能巡检系统建设

动态智能巡检系统与立体化审计体系深度融合，实现数据共享与功能协同。智能巡检机器人采集的数据为审计工作提供实时、准确的一手信息，辅助审计人员发现潜在风险；立体化审计体系的风险评估结果可指导智能巡检系统优化巡检策略，重点关注高风险区域与设备。通过两者协同，组织能够提升风险防控能力，降低运营成本，增强合规管理水平，在激烈的市场竞争中赢得优势。立体化审计体系的三维构建策略结合动态智能巡检系统建设，通过研发智能巡检机器人与构建融合 AI 算法的风险评估模型，为组织审计工作带来智能化升级。在数字经济与严格监管的时代背景下，该体系能够有效应对组织面临的复杂挑战，提升审计效能与风险防控能力。未来，随着技术不断进步，还需持续优化完善该体系，以更好地适应组织发展与监管要求，为组织高质量发展提供坚实保障。

（三）人员行为精细化审计

建立IAM（身份访问管理）超级平台，实现账号全生命周期管理,部署UEBA用户实体行为分析系统，识别异常登录轨迹（如凌晨访问敏感数据），记录细粒度操作日志，涵盖数据查询条件、导出文件MD5、API调用参数等关键要素。通过异常行为识别模型，可检测出凌晨敏感数据访问、跨地域高频登录等10多类风险场景。账号全生命周期管理覆盖率100%，离职账号权限回收时效从 48 小时缩短至 2 小时。

三、数据行为分析的技术突破与创新应用

（一）全域监测对象体系构建

为突破传统安全监测的局限性，我们着力构建一个覆盖“三纵三横”的立体化监测网络。系统在人员管理方面对研发、运维、数据分析等不同岗位进行精细划分，建立起清晰的“角色-权限-数据”映射矩阵，实现以身份为核心的访问控制。数据监测范围全面覆盖结构化数据库、非结构化文档以及实时API数据流等多种形态，确保各类数据资产均得到有效监管。在场景使用方面，监测体系贯穿于数据采集、存储、加工、共享等全生命周期环节，最终形成了一张横跨多种数据形态、纵贯全部处理流程的全场景监测网络。

（二）三维风险检测引擎架构

为构建全面有效的数据安全防护体系，我们致力于开发一个融合合规、业务与技术视角的多维度检测模型。在合规维度，系统通过自然语言处理技术与规则引擎，自动执行隐私协议一致性验证，精准比对实际数据使用范围与隐私声明的匹配度，确保操作合规。在业务维度，我们通过业务模型建模与异常检测算法，对用户操作进行场景偏离度分析，能够有效识别诸如财务数据访问中的非业务性异常操作。在技术维度，则依托流量分析与安全基线监控，主动发现数据流转过程中的漏洞与API接口的未授权调用风险。此外，模型的一项关键创新在于应用了自然语言处理技术，实现了对法律合同条款的自动解析，并将其与数据使用实践进行智能比对，从而将合规审查能力提升至新的水平。

（三）大数据平台治理创新实践

在技术防护层面，我们构建了多层次的数据安全控制体系。通过部署数据血缘追踪系统，利用图数据库技术对数据加工链路进行可视化呈现，成功将数据溯源时间从小时级压缩至分钟级，极大提升了链路追溯效率。同时，在模型训练环境中引入数据沙箱监控机制，严格实施数据使用白名单管控，有效防范训练数据的泄露风险。此外，我们还建立了API调用画像功能，通过分析调用频率、时间窗口和参数特征等多维指标构建异常检测模型，显著降低了API滥用带来的安全风险。

四、可视化安全运营体系的构建与实践

（一）全景式运营大屏架构设计

可视化安全运营体系的构建与全景式运营大屏架构设计是提升企业安全运营水平的重要手段。通过数据整合与治理、安全运营流程优化、可视化展示与交互设计等方面构建可视化安全运营体系，结合科学合理的硬件架构、软件架构和安全保障架构设计全景式运营大屏，能够实现安全运营数据的直观展示、高效分析和精准决策。在实践中，可视化安全运营体系和全景式运营大屏已在众多企业取得显著成效，未来随着技术的不断发展，它们将在企业安全运营管理中发挥更加重要的作用，为企业的安全稳定发展提供有

力保障。

（二）关键技术支撑体系

可视化安全运营体系的构建与实践，离不开关键技术支撑体系的有力保障。机器学习算法、深度学习技术、多因素身份认证技术、基于角色的访问控制（RBAC）技术等，从数据处理与分析技术实现对海量安全数据的高效处理，到可视化呈现技术将数据转化为直观易懂的可视化信息，再到交互实现技术提升用户操作体验，以及安全保障技术确保系统的安全性和稳定性，这些关键技术相互配合、协同工作，共同推动可视化安全运营体系的发展和完善。随着技术的不断进步，关键技术支撑体系也将持续创新和优化，为企业的网络安全运营提供更强大、更可靠的技术支持。

（三）典型应用场景实践

某支付公司，整合银行核心系统交易数据、第三方支付平台数据、用户行为数据（如登录地点、时间、操作频率等）以及外部威胁情报数据。利用大数据采集工具，实现多源数据的实时采集、存储。采用机器学习算法对交易数据进行分析，建立异常交易检测模型，例如通过随机森林算法识别交易金额、频率、时间等维度的异常模式。在全景式运营大屏上，展示交易数据的实时趋势，如不同时间段的交易笔数、交易金额变化曲线；构建动态的交易关系网络图，直观呈现资金流向和账户之间的关联关系。当检测到异常交易时，大屏以醒目的颜色和动画效果进行预警，同时展示该交易的详细信息，包括交易时间、金额、涉及账户等。安全运营人员可通过 WebSocket 技术实现与大屏的实时交互，点击异常交易预警信息，获取更多关联数据进行深入分析。利用语音识别技术，通过语音指令查询特定账户的交易历史、相似异常交易案例等信息。一旦确认是欺诈交易，系统自动触发响应流程，冻结相关账户，同时生成详细的事件报告，供后续调查和处理。部署可视化安全运营体系后，异常交易识别准确率提升至 98%，平均响应时间从原来的 30 分钟缩短至 2 分钟，每年成功拦截欺诈交易数十万起，挽回经济损失数亿元。同时，管理层可通过大屏实时掌握全行交易安全态势，为风险管控策略制定提供有力支持。

五、实施路径与落地保障体系

（一）分阶段实施策略

为确保数据合规监测体系有序落地，我们制定了分阶段实施策略。在规划阶段，首要任务是采用成熟度模型全面评估组织现状，并建立清晰的合规差距分析矩阵，为后续建设明确方向。进入建设阶段后，将遵循“IAM权限体系升级→全链路日志治理→用户行为分析”的递进路线图，确保各系统模块间的协同联动与能力叠加。运营阶段则致力于组建一支跨领域的复合型团队，融合业务、技术与法务视角，从而显著提升风险识别的精准度。最终的优化阶段聚焦于持续迭代风险监测模型，并通过建立常态化的红蓝对抗演练机制，推动整个监测体系在实战中不断完善与进化。

（二）保障机制建设

为确保数据合规与安全运营体系的有效落地，需建立多维度的保障机制。企业在组织层面应设立专职的数据安全委员会，构建起由董事会、管理层与执行层共同组成的三级责任体系，确保权责清晰。在制度层面建立健全一套契合外部法律法规的内部制度文件，形成完整的合规体系，让所有数据处理活动都能有章可循，并得到严格执行。在技术层面搭建自主可控的技术底座，着力推进核心技术组件的国产化进程，建议其替换率不低于70%，以夯实安全根基。在人才层面，应推行“双认证”机制，要求安全运营人员不仅掌握通用安全知识（如持有CISSP认证），还需精通数据合规（如具备数据合规官资质），从而打造复合型人才队伍。

六、结论与展望

数据合规驱动下的全域风险监测体系通过将合规要求嵌入数据处理全流程，构建完善的风险监测技术体系和组织管理保障机制，实现了安全与发展的平衡。实践表明，该体系能够有效降低组织的数据合规风险，提升数据流通效率，为组织释放数据要素价值提供了有力保障。随着数字经济的不断发展和技术的持续创新，数据合规治理面临着新的挑战和机遇。未来聚焦 AIGC 数据合规治理、量子计算环境下的数据加密技术等前沿领域的研究，将推动数据安全治理向智能化、自适应方向演进，更好地适应数字时代的发展需求，为数字经济的健康发展保驾护航。在数字经济加速发展的背景下，这种以合规为基础、以风险为导向的治理模式，将成为组织释放数据要素价值的核心保障。

基于大模型的全渠道信息流统一管控

李剑戈、陶昆、达其双、吴敏嘉 | 中信建投证券股份有限公司

摘　要： 这是对线上营销服务一次全新的尝试和探索，展现了对当前互联网环境下客户体验和信息合规管理的深刻理解。在信息泛滥时代，如何有效管理和利用信息流变得愈发重要。中信建投证券结合自身情况，针对触达渠道多、信息过载等现状，通过金融科技驱动，在信创安全可控架构下，借助大模型、大数据等技术，构建以“监管控”为核心的全渠道信息流统一管控平台，实现信息流全流程数字化管理。在安全合规前提下，保证服务质量与客户满意度。

关键字： 全渠道信息流、统一管控、大模型、文本相似度

一、引言

近些年，随着互联网技术的快速发展，特别是移动互联网、大数据、人工智能等技术的广泛应用，证券公司的运营模式和服务方式发生了深刻的变革。线上营销和服务逐渐成为证券公司获取客户、服务客户、推广业务的重要手段，成为证券公司提升竞争力、拓宽市场、增强客户粘性的关键法宝。搭建标准化、平台化、精细化的线上客户运营体系，推动传统营销模式向数字化智能营销模式转变，助力提高客户开发和服务效率，实现客户服务体验提升，已经成为行业共识。

与此同时，证券公司的数字化智能营销模式，依托线上和线下全渠道触达客户，对客户进行线上营销和服务时，也必须承担起保护投资者合法权益、防止过度营销、保障客户隐私安全等责任。证券公司有责任也有义务对触达客户的全渠道信息流进行安全、合理、高效的管控，从而保证客户体验，有效防范风险，维护投资者合法权益。

二、证券公司线上运营现状

当前阶段，证券公司的线上运营具备渠道多样化、营销和服务场景丰富、客户数量庞大等特点。头部券商客户数量已达千万级，依托线上+线下的全渠道，针对不同客户群体、不同需求，提供各式各样的服务，全渠道每天触达的客户数量超过百万级，每天发送的信息数量达到亿级。

同时，线上运营也面临一些问题和挑战：运营团队和系统众多，互相之间存在割裂的情况；面临既要精准触达客户，又要避免过度打扰的双重挑战；缺少统一管控，渠道资源利用率有待提升。由于团队众多、运营策略众多，不同团队的运营策略，对同一客户的触达频率和内容可能存在交叉情况，从而引发重复触达、过度触达的风险。这些问题具有很高的隐匿性，极难发现，单靠流程和制度约束难以完全避免。如图1所示。

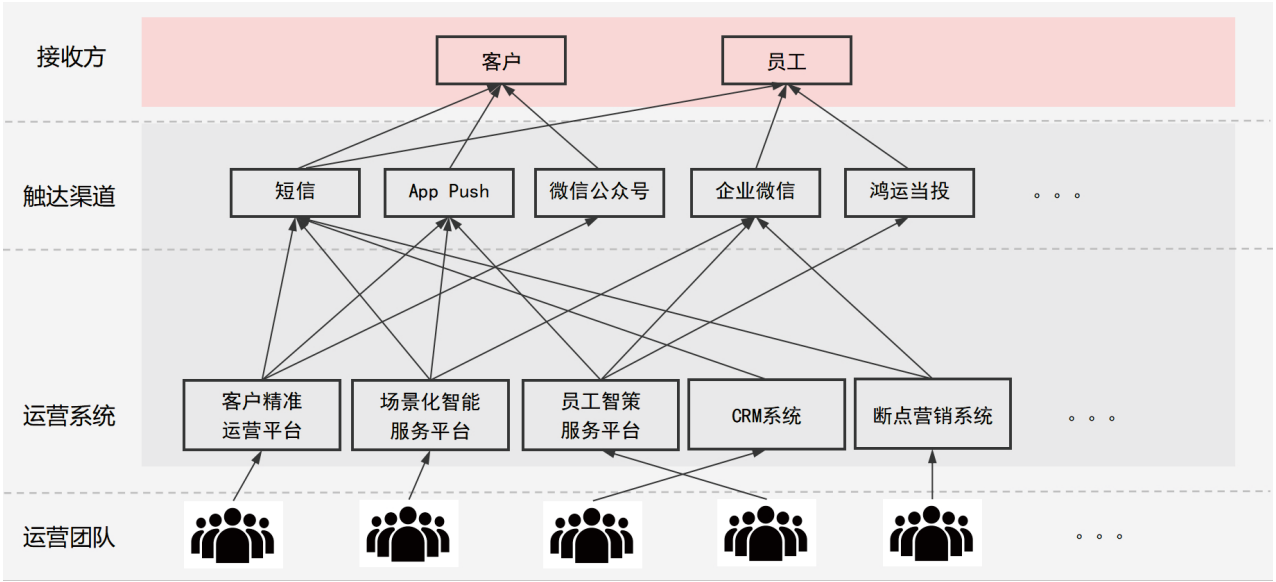


图1 线上运营现状

三、全渠道信息流统一管控的实践与探索

全渠道信息流统一管控的目标，是要实现对客户和员工两类对象的信息流统一管控，逐步将各系统各渠道触达客户和员工的信息流汇总，形成管理视角的消息监控、管理，从而实现全渠道触达信息流的统一管控。在触达客户和员工时，我们必须要有所“克制”，要通过技术手段进行合理的频率控制。既要实现精准触达，又要避免过度打扰、重复触达，提高渠道信息流触达的精准度和质量。

中信建投证券积极尝试和探索通过金融科技驱动，在信创安全可控的架构下，借助大模型、实时数据处理及大数据等技术，构建一个安全、高效、智能的全渠道信息流统一管控平台，实现全渠道信息流的统一化、数字化、智能化管理，帮助公司在竞争激烈的市场中，更好地管理和利用渠道信息流，提高线上营销和服务的质量和效果，提升客户满意度。

（一）全渠道信息流统一管控体系

通过构建以“监、管、控”为核心的公司层级信息流管控平台，实现渠道信息流监测、管控、发送、跟踪等全流程管理。将当前分散在不同系统的渠道信息流统一接入，实现信息流数据的集中和统一。通过一系列的管控策略，进而实现统一管理和监控，做到合理管控，提升管理效率。如图2所示。

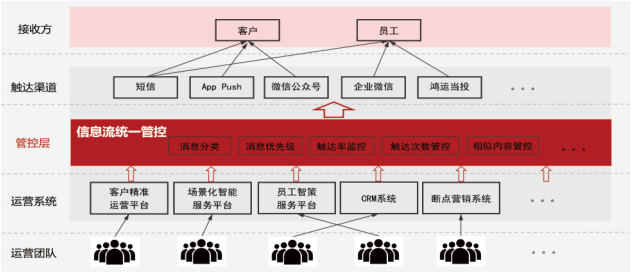


图2 全渠道信息流统一管控体系

（二）系统主要功能

全渠道信息流统一管控的本质是渠道信息流匹配，主要包含以下几个要素：

1.触达对象。包括客户、员工两类。

2.触达渠道。包括短信、微信公众号、智能外呼、自建App消息、通达信客户端、企业微信、移动CRM等等。

3.管控要素。包括优先级划分、管控策略、质量控制、数量控制等等。

根据渠道信息流生命周期，系统主要分为负载层、消息接收层、消息管控层和数据分析展示层等。消息接收层对消息内容做分类和校验，判断优先级和消息类型后流转到消息管控层；消息管控层包含一揽子管控策略，根据策略规则进行实时计算和判断，实现有效拦截或放行通过。数据分析展示层可视为渠道信息驾驶舱，能以全局的视角查看实时数据、历史数据、策略拦截数据等。如图3所示。

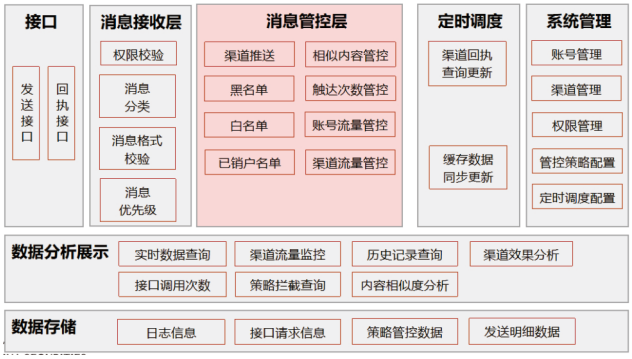


图3 系统核心功能图

系统的主要核心功能如下：

1、大模型内容相似度管控

通过引入大模型搭建智能底座，借助大模型在自然语言处理能力等方面的优势，实现事中管控和事后预警，对消息内容进行相似度计算，减少相似内容的发送，提升精准性。

（1）事后预警

对已经发送的消息，离线分析各系统和渠道的内容相似度，形成分析结果和预警报告，及时反馈给相关运营团队，进而进行调整和优化。

（2）事中管控

对将要发送和正在发送中的消息，准实时计算和当前已发消息的内容相似度，根据管控策略规则，有效进行拦截。

在自然语言处理和信息检索领域，经常需要通过文本相同相似比较，来判断两个或多个不同文本间的相似度，以达到相应的应用目的。但现有的文本相同相似，比较方法效率低且准确度不高。所以我们需要借助大模型的大数据量的模型计算能力，且利用对自然语言的推理能力。基于私有化部署的Deepseek，实现了文本的相似度相对准确的计算。如图4所示。

（1）输入大模型少样本提示工程。

（2）结合各渠道发送的文本内容和大模型提示工程（需要不断优化提示工程才能得到相对准确的结论），提交到大模型。

（3）根据大模型计算结果，如超出预先设定的阈值，则进行管控。

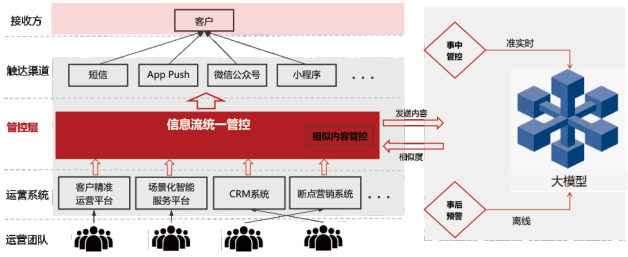


图4 基于大模型的相似度分析

2、消息优先级智能管控

根据不同的业务种类和场景，通过指定消息高、中、低三种优先级和通知类、服务类、营销类等消息类型，对触达客户和员工的消息进行分类分级。如图5所示。

(1) 正常情况下，根据接收到的消息请求及指定的优先级、消息类型参数，自动分配到默认对应的消息通道。

(2) 消息出现积压时，进行智能判断，结合消息请求数、各个通道忙碌情况，自动合理分配通道和流量，确保资源合理、高效利用。

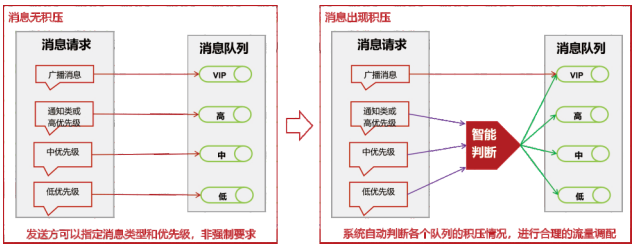


图5 消息优先级管控实现机理

3、流量管控策略

从渠道、系统、人员三个粒度，进行发送数量的管控。

(1) 渠道管控。每个渠道，可以限定每天或者某段时间内的最大发送数量。

(2) 系统管控。每一个接入的系统，可以限定每天或者某段时间内的最大发送数量。

(3) 人员管控。每一个触达对象，可以限定每天或者某段时间内的最大发送数量，避免过度打扰。

4、名单管控策略

包括白名单、黑名单拦截、已销户拦截。

(1) 白名单。默认情况下，管控策略适用所有对象，如果有对象需要特殊对待，将对象加入到白名单中，将不再受管控策略约束。

(2) 黑名单。对于主动退订消息或其他原因，不希望再接收渠道消息时，将对象加入到黑名单中，将不再收到任何渠道消息。

(3) 已销户拦截。按照《中华人民共和国个人信息保护法》的要求，对已销户客户信息停止除存储和采取必要的安全保护措施之外的处理，并减少对已销户客户的打扰，个人信息处理者应当避免向已销户的客户发送不必要的营销信息、广告等，以免打扰其生活和工作。系统会自动判断客户是否销户，自动进行拦截，不再触达客户。

5、渠道信息流数字化

(1) 统一监控和查询。支持实时查询当天消息，离线查询历史消息，对所有消息能够进行全方位的监控，帮助各业务团队统筹管理、分析和监控各渠道信息流数据，辅助决策。

(2) 策略拦截查询。支持方便快速地查询和追溯被策略拦截的消息和详情，从而对后续的营销和服务策略进行优化改进。

(3) 信息流数据分析和挖掘。将全渠道信息流数据全部统一、集中存储到大数据平台，然后结合渠道的回执状态、消息埋点、渠道转化、用户反馈等数据，结合机器学习及数据模型，进行深度分析和挖掘，为后续的策略制定和优化提供有力的数据支持。根据不同特点和喜好的用户，有针对性的精准选择匹配合适的触达渠道，在合适的时间点进行触达和服务，提升渠道资源的利用率和效果。

6、管控策略可视化配置

所有管控策略都支持可视化配置，可以方便、灵活地进行开关控制、参数配置。

(三) 系统架构

全渠道信息流统一管控平台的系统架构，如图6所示。

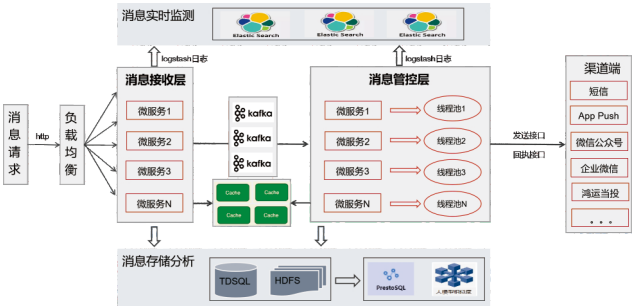


图6 系统架构图

系统基于信创+容器化部署，集成大模型构建智能底座，为管控策略提供智能化支撑，核心技术特点如下：

1、落地大模型在线上运营合规管控领域的运用

大模型通过语言处理技术，结合提示工程和参数调优，已经具备了一定的对券商业务的理解能力。提示工程有效提升了模型响应效率，并提供灵活有效的方法，使模型能够在不调整核心参数的情况下处理多样化的业务需求。而指令微调能够更深入地优化模型以满足特定需求和上下文，在实际应用中能够达到较高的准确性和可靠性。新消息到达时，大模型通过设定的特定提示词完成包括意图识别与槽位抽取在内的消息内容自然语言理解，迅速计算新消息与已推消息的相似度，根据阈值判断拦截或放行，减少无效和相似触达，提高触达精准度和质量。

2、信创架构下的实时计算，确保高性能

信创架构下确保自主、安全、可控，从硬件服务器，到操作系统，再到数据库和缓存中间件，全部符合信创标准。我们必须在保证消息发送性能的前提下，实现各类管控策略，不能因为增加的管控策略影响了现有的消息发送速度和性能。其中高性能的实时计算依赖缓存中间件，中信建投证券

与信创厂商深度合作共研，在性能验证阶段进行反复测试和优化迭代，不断提升缓存中间件（CacheDB）性能。最终实现核心性能数据对标开源缓存中间件（Redis），部分性能甚至超越Redis。如表1所示。

表1 信创缓存中间件性能测试数据

场景	性能要求	原理	数据深度	单步长	优化前	优化后	创新
推送重复判定	0.01s	根据hsahcode生成key，类型为 set	2000w	10k	6s	0.008s	Redis 协议提供的 set 集合计算只支持 key 运算。信创缓存优化后增加命令，支持流式数据和现有 key 比较，无需删除和生成临时 Key，从而突破 Redis 固有限制。
拦截过滤	0.08s	sdiff 查找	500w	1k	1s	0.003s	
限流	0.05s	操作原子化，适用 lua 脚本	3000w	1k	1.18s	0.048s	

3、分布式微服务部署，保证系统高可用

系统采用分布式微服务的部署架构，保证系统高可用，支持快速横向扩展。消息处理能力超过1万条/秒（QPS ＞ 1W），单日峰值消息发送量突破1.5亿人次，并且支持快速弹性扩展。在保证管控策略有效执行的同时，不影响正常的消息发送速率。

四、全渠道信息流统一管控的价值

（一）降本增效，有效提升合规管理

通过整合资源，将分散的渠道信息流统一接入，实现集中和统一。通过实现一系列的管控策略，进而实现渠道信息流统一的管理和监控，做到合理管控，提升管理效率。通过合理的、有效的拦截，减少重复发送、相似发送，从而降低成本，减少资源浪费，解决了当前运营过程中面临的问题，有效控制合规、投诉风险，提升渠道资源的管理效率和利用率。

（二）实现大模型技术与业务场景相结合

大模型发展到现阶段，其技术能力已经取得了长足的进步，并在部分领域得到了广泛的应用。然而，大模型在证券金融领域，如何与实际业务相结合一直困扰着大家。我们充分发挥大模型的特长，在全渠道信息流管控这个具体领域，进行尝试和探索，找到新兴技术与业务应用场景的结合点，促进了技术与业务的深度融合，充分发挥金融科技的力量。

（三）探索数据价值，促进业务创新

通过将渠道信息流数字化，可以对海量的渠道信息流数据进行深入分析和挖掘，可以更加精准地了解客户需求、偏好，优化线上营销和服务流程，提高服务质量并降低风险。同时，深挖渠道信息流数据价值，可以发现用户趋势的细微变化甚至预测趋势，挖掘新的业务机会，为业务拓展提供有力支持。

这是中信建投证券对线上营销和服务一次全新的尝试和探索，展现了对当前互联网环境下客户体验和信息管理挑战的深刻理解。在信息泛滥的时代，有效管理和利用信息流，不仅关乎企业的运营效率，更直接影响到客户的满意度和忠诚度。中信建投证券结合自身情况和对全渠道信息流的理解，充分借助大数据、大模型等科技力量，与实际应用场景相结合，探索全渠道信息流的统一管控，提高触达质量，提升客户满意度！

参考文献

1.单玉波,张鹏翔. 数字金融背景下加强信息科技风险管理研究[J]. 农业发展与金融,2024(6):87-89. DOI:10.3969/j.issn.1006-690X.2024.06.041.

2.丁建强,翁航,闫珵珵.证券公司经纪业务向财富管理转型的路径探讨[J].金融纵横, 2020(5):10.

3.ZIEGLER D M, STIENNON N, WU J, et al. Fine-tuning language models from human preferences[EB]. arXiv preprint, 2019, arXiv: 1909.08593.

4.黄峻,林飞,杨静,等. 生成式AI的大模型提示工程:方法、现状与展望[J]. 智能科学与技术学报,2024,6(2):115-133. DOI:10.11959/j.issn.2096-6652.202424.

5.张玉. 中小金融机构数智化转型中的挑战及对策研究[J]. 西部金融,2024(4):76-80.

6.文本数据相似度确定方法和系统[P].中国发明专利, CN118211588.2024-06-18.

7.Tom B.Brown.Language Models are Few-Shot Learners.San Francisco:OpenAI,2020.

钓鱼邮件演练的实践与探索

张沁怡、崔毅然、吴鹏、陈其乐 | 上海证券有限责任公司

摘　要： 在数字化时代背景下，网络安全正面临前所未有的挑战，其中钓鱼邮件已成为公司信息安全的重大威胁。为切实提升员工防范钓鱼邮件的能力，破解安全培训形式化的难题，笔者推动“宣贯-演练-培训/奖励-考核”闭环体系在钓鱼邮件演练中的落地实践，将安全培训从“走过场”转向“务实化，切实为业务发展筑牢安全防线。本文从目标设定、方案设计、执行过程以及效果评估四个维度，着重分享公司在钓鱼邮件演练方面的实践经验，以供同业参考。

关键字： 钓鱼邮件演练、员工安全意识、闭环体系

一、概述

在数字化迅猛发展的今天，网络安全问题日益严峻且复杂。随着网络攻击技术的持续进步，钓鱼邮件已经成为威胁企业信息安全的一大隐患。这些邮件通常以高度隐蔽的伪装，利用员工的疏忽和好奇心，诱使其点击恶意链接、下载恶意附件或泄露敏感信息。一旦员工不慎落入陷阱，可能给公司带来财务损失、信誉损害以及数据泄露等严重后果。为切实提升员工防范钓鱼邮件的意识，强化其识别和应对的能力，破解安全培训形式化的难题，笔者推动“宣贯-演练-培训、奖励-考核”闭环体系在钓鱼邮件演练中的落地实践，将安全培训从“走过场”转向“务实化，切实为业务发展筑牢安全防线。本文从目标设定、方案设计、执行过程和效果评估四个维度，着重分享我司近年来在钓鱼邮件演练方面的实践经验。

二、安全意识宣贯

日常安全意识宣贯是闭环体系的基础环节，旨在通过常态化、多维度的信息传递，让钓鱼邮件防护知识融入员工工作习惯，从源头降低被攻击风险。

笔者固定每周通过公司企业微信的“信息安全小课堂”栏目，以生动的图文形式普及安全知识 —— 内容不仅涵盖钓鱼邮件防范，还延伸至账号密码安全、公共 WiFi风险等各类职场高频安全场景，让抽象的防护知识更易理解、更易传播。在年度网络安全宣传周期间，还会进一步丰富线上形式，通过主题短视频、上线轻量化安全小游戏等，以更具趣味性的方式强化员工记忆。

三、钓鱼邮件演练

（一）明确演练目标

钓鱼邮件演练的核心目标是提升员工对钓鱼邮件攻击的识别能力和响应水平，发现员工信息安全意识的薄弱环节。具体目标包括：

1、检验员工是否具备识别常见钓鱼邮件特征的能力。例如识别伪装的发件人、伪造的官方网站链接以及内容上的诱惑性。

钓鱼邮件的发件人地址通常与公司高层、部门、合作伙伴或重要客户的邮箱地址极为相似，邮件内容可能包含诸如“紧急提醒”、“福利领取”、“退税通知”等紧急或吸引人的措辞，旨在诱使员工点击邮件中的链接或下载附件。通过钓鱼邮件演练，可以测试员工是否能够识别这些异常特征，并通过其他途径验证邮件的真实性，从而避免被虚假信息所欺骗。

2、评估员工在遭遇可疑邮件时的行为反应，例如是否点击链接、下载文件或泄露敏感信息。

在实际演练过程中，员工可能由于安全意识不足或好奇心驱使，轻易点击邮件中的链接或下载附件，甚至在伪造的网页上输入账号、密码等敏感信息。通过模拟各种真实的钓鱼邮件场景，观察员工在不同情境下的行为表现，可以准确评估他们的安全意识和应对能力。

3、分析不同部门及岗位员工的安全意识差异。

针对不同部门和岗位的员工，我们设计并发送了具有不同主题的钓鱼邮件，以此来分析各部门及岗位间在安全意识上的差异，为后续的培训工作提供数据支持。鉴于各部门和岗位的工作性质、接触信息的种类以及面临的安全风险各不相同，例如，以公司高层名义向普通员工发送的邮件，从常识来看似乎不太可信；然而，如果以IT部门的名义发送关于OA系统密码修改的通知，钓鱼邮件则可能成功。通过收集和分析演练中的数据，我们可以观察到不同部门和岗位的员工在识别和应对钓鱼邮件方面的表现差异，从而能够制定出更有针对性的培训计划，有效提高员工的信息安全意识。

（二）方案设计

一次成功的钓鱼邮件演练，前期必须制定详尽的演练方案。这包括但不限于精心挑选钓鱼邮件的主题和内容设计、精确选择发送对象、确定邮件的发送周期、制定敏感数据的处理方法，以及后续的培训计划。此外，演练方案在执行前需要通过公司内部的严格审核，以确保既能达到教育员工的目的，又不会引起不必要的恐慌或部门间的冲突。

1、钓鱼邮件设计

应尽可能地模仿真实的钓鱼邮件，例如模拟系统异常登录通知、薪资调整通知/补贴发放、电子发票下载等场景。邮件正文应采用公司官方邮件的正式格式、语气和字体样式，发件人的邮箱地址应使用与公司域名相似的后缀，并利用显示名欺骗技术，使员工误认为邮件来自公司内部的相关部门或特定员工。对于虚假的福利和中奖信息邮件，可采用夸张的措辞、引人注目的奖品以及紧迫的领取时间限制等手段，以吸引员工的注意力。

2、邮件发送对象及周期

考虑到实际办公场地的布局，同一部门的员工通常会被安排在同一楼层的相邻区域工作。如果在同一办公区域的员工在短时间内连续收到相同主题的钓鱼邮件，他们可能会相互讨论，这将导致钓鱼邮件的点击率和后续操作的成功率降低，从而影响演练的效果。因此，在规划邮件发送周期时，我们会结合员工所属部门的楼层信息，调整钓鱼邮件的主题和发送批次，以提高钓鱼邮件的识别难度。

3、敏感数据处理

尽管钓鱼邮件演练并非真实的黑客攻击，但员工点击链接并上传其真实账号、密码和其他个人信息的风险依然存在。为了保护个人信息，方案中应明确指出，仅收集与演练效果评估直接相关的基础数据，例如邮件查看次数、邮件链接点击次数、网页提交动作以及相关数据的占比分析。所有涉及员工实际填写的账号、密码和个人信息等敏感数据均不会被保存。演练结束后，所有演练过程中产生的相关数据将被彻底清除，确保数据不留痕迹，防止任何形式的二次泄露或滥用。

4、通知

在演练开始前，通过OA系统通知员工公司将进行钓鱼邮件演练，但不透露具体时间，以此确保员工在知晓演练事项的前提下仍保持应有的警惕性。同时，通过公司的信息安全意识宣贯平台，强化对钓鱼邮件防范的风险提示。

（三）演练执行

执行阶段构成了整个钓鱼邮件演练活动的核心，必须确保流程的规范性、数据的准确性以及反馈的及时性。

1、邮件发送与监控

利用专业的钓鱼邮件模拟平台向目标群体发送定制化的邮件，并实时追踪邮件的送达率、员工的点击率以及输入敏感信息的比例等关键指标。在挑选钓鱼邮件模拟平台时，需考虑其功能性、稳定性和安全性。平台应提供通用的钓鱼邮件模板，同时支持自定义模板，能够依据预设规则精确地向目标群体发送邮件，并实时记录每位接收者的操作行为，包括但不限于邮件的送达时间、员工打开邮件的时间、是否点击了链接、点击链接的次数、是否下载了附件、是否在模拟页面上输入了敏感信息等。这些数据将为后续的评估和分析提供关键依据。

2、实施钓鱼邮件演练通知机制

对于未能识别钓鱼邮件的员工，应准备一个警示页面，用于进行钓鱼邮件演练的告知和风险提示。当员工执行了高危操作，例如点击邮件中的链接并提交个人信息后，系统应自动跳转至提醒页面，告知员工他们当前访问的是公司设置的模拟钓鱼邮件演练页面。同时，页面应详细说明点击链接并提交信息的行为存在风险，并列举钓鱼邮件的危害、如何识别钓鱼邮件以及正确的应对措施等内容。

3、技术支持与答疑

成立专门的钓鱼邮件演练支持团队，在演练期间为员工提供即时的咨询解答和应急处理，以保障演练活动的顺畅进行。在演练过程中，员工可能会遭遇多种问题，例如验证邮件的真伪、反馈邮件无法打开、点击链接后页面无法正常显示、对警示信息内容感到困惑等。支持团队需迅速回应员工的咨询需求，提供精确的解决方案，并安抚那些反应较为激烈的员工。同时，团队应密切关注邮件发送和数据统计情况，以便及时发现并处理潜在的突发状况，如邮件发送失败、系统故障、数据异常等，确保演练能够依照既定计划顺利推进。

（四）评估演练效果

演练结束后，必须对收集到的演练数据进行深入分析，以持续提高员工的安全意识，并进一步完善公司的安全策略。

1、数据分析

统计钓鱼邮件的打开率、链接点击率以及提交敏感信息的比例，并绘制趋势图以直观展示演练结果。通过深入分析这些数据，可以全面了解员工在演练过程中的整体表现情况，以及随着演练的多次进行或常态化开展，员工的钓鱼邮件识别和应对能力是否有所提升。例如，如果某个特定主题的钓鱼邮件链接点击率及提交敏感信息的比例显著偏高，通过分析邮件主题的特点，可以揭示公司员工相对感兴趣且容易受骗的邮件类型。此外，从数据中发现员工已经擅

于识别某类钓鱼邮件，这将有助于信息安全部门调整后续钓鱼邮件的策略，以增强演练的实际效果。

2、部门与岗位安全意识分析

通过部门和岗位的差异化统计，识别出安全意识相对薄弱的员工群体，并据此制定个性化的培训方案。对不同部门和岗位的员工在安全演练中的行为数据进行细致的对比分析，特别关注那些邮件打开率和链接点击率较高、以及提交敏感信息比例较大的部门或岗位，如图1所示。针对这些安全意识不足的员工，深入探究其背后的原因，例如工作性质可能要求他们频繁处理电子邮件，从而忽略了对发件人信息的仔细审查。基于这些分析结果，设计专门的培训计划，例如在培训中融入钓鱼邮件案例分析，并在后续的安全演练中向这些员工发送类似主题的钓鱼邮件，以增强他们的安全防范意识。

图1 岗位统计示例图

3、持续改进措施

依据数据分析结果及最新的网络威胁趋势，调整演练内容，并坚持定期执行。演练的目标不仅在于识别员工在安全意识方面的不足，更关键的是通过演练发现公司安全策略的潜在问题或改进空间，以便及时进行解决和优化。定期组织钓鱼邮件演练，建立常态化的安全培训体系，确保员工持续保持警觉。同时，密切关注网络威胁的最新发展，及时更新演练内容，保证演练活动与时俱进，有效防范新型钓鱼邮件攻击。

四、奖励和培训考核

对于在演练过程中，把钓鱼邮件事件上报给相关部门或询问是否为钓鱼邮件的员工，将记录其姓名和所属部门。待演练结束后，为这些具备高度安全警惕性的员工发放小礼品，以示鼓励，提高员工积极性。

同时，开展我们钓鱼邮件演练的目的是为了提升员工的安全意识和保护公司的信息安全，因此对中招员工的名单仍将保密。演练结束后，将对这部分员工单独进行针对本次演练的钓鱼邮件专项培训和考核（此举仅用于检验培训内容的学习效果，并非与绩效考核相关）。其中，培训内容不仅涵盖通用的钓鱼邮件防范知识，还会聚焦本次演练中暴露出的关键问题与关注点进行有针对性的讲解和培训。

五、总结

钓鱼邮件演练是公司信息安全保障链条中至关重要的一环。它以实战模拟的方式，让员工在接近真实的场景中深刻认识到钓鱼邮件的危害，掌握有效的防范技巧。通过持续开展演练，公司能够形成一套成熟、完善的安全应对机制，使员工在面对实际网络威胁时能够迅速做出反应，降低安全风险。同时，演练也为公司不断优化信息安全策略提供了实践依据，有助于公司及时调整和改进安全防护措施，以适应不断变化的网络安全形势。只有将钓鱼邮件演练常态化、制度化，才能真正实现公司信息安全的可持续保障，为公司的稳定发展保驾护航。

智能体驱动的API安全风险管控研究与实践

徐承文、程际桥、吴琪、刘义卓、杨启 | 长江证券股份有限公司

摘　要： 针对证券行业API安全面临的高敏感性、高实时性和强监管性要求，提出了一种智能体驱动的API安全风险管控体系。通过融合大模型与多智能体协同技术，实现了API资产的智能识别与全生命周期管理、敏感数据的精准分类与动态分级防护，以及网络攻击的高效精准检测，自动化落实监管合规要求，提升数据安全治理能力，有效保障业务连续性与数据完整性，为证券行业构建自适应、智能化的API安全防护能力提供了实践路径参考。

关键字： API安全、智能体、大模型、数据安全

一、引言

在当代的数字生态系统中，应用程序编程接口（API）已经从一个简单的开发工具，演变为现代软件架构的中枢神经系统，促进了无数平台和服务之间无缝的数据交换与功能实现。这种普遍性使得API成为恶意行为者的主要攻击目标，形成了一个既广泛又复杂的攻击面。证券行业作为国民经济的核心组成部分，在数字化转型与智能化升级的过程中，API安全事件频发。例如，2021年某证券交易平台因API权限控制漏洞，导致数据泄露，影响了数百万投资者的隐私安全；2023年某金融科技公司因API密钥泄露，导致黑客通过恶意调用获取用户交易数据，造成了数千万元的经济损失。这些事件表明，API安全风险不仅影响企业的数据安全和业务连续性，还可能引发系统性金融风险，对国家经济安全构成潜在威胁。

传统的安全措施，通常基于静态签名和预定义规则，在面对复杂的、多阶段的API攻击时日益显得力不从心。行业急需一种有效的API安全风险管控手段，帮助行业机构保护敏感数据，防止客户信息、商业机密等重要数据被窃取和滥用；维护系统的完整性，避免因API漏洞导致系统被入侵、篡改或破坏；确保服务的可用性，保障业务的连续性和稳定性，减少因安全问题引发的服务中断或性能下降；保证业务活动遵守相关法律法规和行业标准，避免因数据泄露或违规而面临的法律风险和巨额罚款，维护企业的声誉和品牌形象。

二、证券行业API资产管理的基础性地位与挑战

在证券行业的数字化架构中，API资产是连接数据、服务与应用的核心枢纽，是业务创新和智能化转型的基石。证券行业又因其高敏感性、高实时性和强监管性，对API安全提出了更高要求。在大模型时代，证券行业的API安全风险呈现出新特征：一是金融数据的高敏感性和高价值性使其成为攻击者的重点目标；二是大模型驱动的智能金融系统增加了API调用的动态性和不可预测性；三是监管合规性要求（特别是《数据安全法》《个人信息保护法》）对API安全提出了更高标准。

API资产管理的有效性，直接决定了上层安全策略能否精准落地、安全风险能否被全面覆盖。然而，在证券行业复杂的IT环境中，API资产管理面临着巨大的挑战，其复杂性和动态性给安全管理带来了前所未有的困难。

传统的API发现依赖于分析已声明的规范（如Swagger/OpenAPI）或被动流量监控。这些方法常常会遗漏未记录的端点，存在以下诸多弊病：

（1）覆盖不全：由于业务快速迭代、多团队并行开发、第三方系统集成等原因，大量API未经统一登记便上线使用（“影子API”），或在业务下线后未被及时停用（“僵尸API”）。这些游离于安全管控之外的资产，成为了最易被攻击者利用的薄弱环节。

（2）信息滞后：API的功能、参数、认证方式等信息变更频繁，导致安全团队基于过时信息制定的策略形同虚设。

（3）分类粗糙：缺乏统一和精细的分类标准，无法有效区分API的业务重要性和数据敏感性，导致安全资源无法向高风险资产倾斜，防护效率低下。

针对传统的API管控平台的缺点，本文提出了一种基于大模型的API资产智能识别技术，并以此为基础构建了一个由AI智能体驱动的全方位API资产管理和审计体系，从根本上解决API资产“看不清、管不住”的难题。

三、智能体驱动的API资产识别与全生命周期管理

（一）基于大模型的API资产智能识别技术

为实现对API资产全面、精准、实时的识别，本文采用一种融合了平台扫描与大模型深度分析的创新技术路径。首先，利用大型语言模型（LLM）的模式识别能力，分析请求和响应样本，自动推断API协议规范，例如推断REST参数规则，甚至是gRPC接口定义，提升识别未知API的效率和准确性。其次，利用AI模型结合多个数据点（例如来自Swagger描述的代码注释和实际的流量行为模式）来进行复杂的聚类分析。利用丰富的上下文信息和AI大模型理解能力，可以极大地增强API识别的准确性和全面性，为所有后续的API安全风险管控提供坚实的基础。

在技术实现中，首先在网络关键节点以旁路或串联的方式获取全量的API通信流量。通过对这些流量进行解码和分析，从海量的网络报文中初步筛选和识别出具备API通信特征的数据流，完成对业务系统中所有在用API资产的初步发现。这一步骤可以解决传统方法依赖人工登记而导致的覆盖不全问题，确保资产发现的广度。

在初步发现的基础上，利用大模型的自然语言处理（NLP）和模式识别能力，对采集到的API请求与响应报文样本进行深度分析。大模型的能力也超越了简单的URL模式匹配，能够像人类专家一样“理解”API的协议规范、功能语义以及数据结构等内在逻辑。比如通过分析请求方法（GET/POST）、头部字段（Content-Type）、报文结构（JSON/XML）等特征，大模型能够自动推断出API遵循的协议规范（如RESTful、GraphQL、gRPC等），并能通过分析API的URL路径、参数名称（如getUserInfo、queryTradeHistory）以及响应数据中的字段，结合上下文信息，精准理解该API的业务功能是用于“用户信息查询”还是“下单交易”。大模型通过解析请求和响应体中的数据结构，识别出关键的参数字段、数据类型和嵌套关系，为后续的敏感数据识别和攻击面分析奠定基础。

识别的第三部分是分类与标注。利用大模型的分析结果，使用AI智能体对API资产进行多维度的自动化分类和标注。基于深度语义理解的结果，AI智能体能够将发现的API自动归类到“用户管理”、“行情服务”、“资产查询”等具体的业务域中。在此基础上，AI智能体能进一步结合从API通信流量中统计到的调用频率数据，来评估每个API的业务活跃度和其在系统中的重要性。通过对API传输数据的深度分析，特别是对其中包含的客户身份信息、金融数据等高价值内容的精准识别，AI智能体还能为这些API自动打上“高敏感”、“中敏感”、“低敏感”等不同等级的敏感性标签，从而为后续的安全防护提供精准依据。

通过这一系列智能化的识别、分析与分类技术，能够构建一幅动态、精准、多维度的API资产全景地图，极大地提升API识别的准确性和全面性。

（二）AI智能体驱动的API全生命周期管理框架

清晰的资产清单是前提，动态、闭环的管理才是目标。本文将设计并实现一个覆盖API“发现、登记、监控、审计、弃用”的全生命周期管理框架。该框架的核心优势在于AI智能体的自主学习能力。它不仅是流程的执行者，更是流程的优化者。通过持续分析API的演化模式和管理过程中的数据，AI智能体能够不断优化其识别模型和管理策略，动态地更新API资产清单，确保整个管理过程的完整性和实时性，真正做到“资产一本账，动态全掌握”。

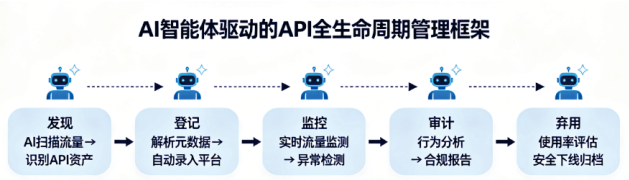


图1 AI智能体驱动的API全生命周期管理框架

任何新上线的API一旦产生流量，就会被API安全管控系统自动发现和识别。AI智能体随即会将其详细信息（包括URL、功能、敏感性等）自动登记到统一的API资产库中。系统将持续监控每个API的活动状态和配置变更：通过监控API的调用量、响应时间、错误率等性能指标，及时发现异常或已停止服务的“僵尸API”；当API的参数、认证方式或返回数据结构发生变化时，AI智能体能够自动检测到这些“漂移”，并与资产库中的基线进行比对，生成变更告警，发送给安全人员审计确认；对于长期不活跃或已确认下线的API，系统会将其标记为“待弃用”状态，并启动废弃流程，从资产库中归档或移除，完成管理的闭环。

四、构建证券行业定制化的API安全风险管控平台

证券行业的API安全管控，不能照搬通用的互联网安全方案。行业业务场景具有高敏感性、高实时性和强监管性的显著特征：高敏感性要求平台具备对客户身份信息、金融数据等核心资产的深度保护能力；高实时性要求攻击检测和响应必须在不影响交易执行的前提下完成；强监管性则要求平台的设计和功能必须满足国家法律和行业法规的合规要求。

因此，适应证券行业的API安全风险管控平台，将不仅仅是大数据、大模型、AI智能体等功能的堆砌，而是一个集敏感数据保护、高级攻击检测和业务威胁防范于一体的综合性、智能化作战平台。

（一）核心功能一：敏感数据的精准定位与分类分级

1、基于大模型语义理解的敏感数据识别

传统的API安全管控平台一般基于正则表达式来进行敏感数据的识别，其误报率和漏报率较高，且难以适应金融术语的多样性。例如，一串数字可能是一个信用卡号，或者

是一个手机号，也可能是一串无意义的测试ID。一个典型的敏感信息误报如图2所示，API平台把13位数字识别成了手机号，并当作高敏感信息告警。但结合上下文分析可以得知，这只是一串随机的测试代码，只是刚好与手机号的格式相同。

图2 敏感信息泄露误报

为减少这种类型的误报，上下文语义的分析是关键，这正是大模型的长处。通过分析数据出现的上下文，例如参数名、周围的数据结构、响应体的数据内容以及API的业务功能等，大模型能够准确地区分敏感信息（如明文的真实卡号）和无害的测试数据。这种上下文感知能力可显著提高敏感数据识别的准确性，减少由误报引起的安全告警噪音。

大模型语义理解能力在经过证券行业语料的微调后，能够对API请求与响应中的数据流进行深度内容分析，精准识别和标注各类敏感数据字段，如身份证号、银行卡号、股东代码、持仓数量、交易金额等，即便是这些字段被复杂地嵌套在JSON或XML结构中。

2、动态数据敏感性分级体系

可基于《证券期货业数据安全风险防控 数据分类分级指引》（GB/T 422275-2023）对证券行业内涉及的各类数据进行系统性的识别和评估，明确各类数据的敏感程度。通过精细化的数据分类分级，识别哪些数据最为核心和敏感，进而可以有针对性地部署安全控制、访问权限管理和数据加密措施，有效保障证券行业数据的安全与合规性。

```
{
  "tag": "个人投资者基本信息",
  "level": "3",
  "description": "一般指参与证券期货市场交易活动的自然人基本信息，包含姓名、证件类型代码、证件号码、出生日期、联系地址、移动电话、国籍、个人标识等信息。",
  "sensitiveItems": [
    "军官证",
    "出生地",
    "出生证",
    "国籍",
    "地址",
    "姓名",
    "婚姻状态",
    "居住证",
    "年龄",
    "往来台湾通行证",
    "往来港澳通行证",
    "性别",
    "护照号码",
    "民族",
    "港澳居民居住证",
    "港澳居民来往内地通行证",
    "生日",
    "第二代身份证",
    "证件类型",
    "退休证",
    "个人识别码"
  ]
}
```

图3 个人投资者基本信息

本文采用结构化元数据定义模式对数据分类进行标准化描述。如“个人投资者基本信息”数据类别可用如图3的模型来描述。该数据单元被定义为三级敏感性标签，其内涵为证券期货市场自然人参与者的身份标识信息集。此类数据包含可直接或间接识别特定自然人身份的多维要素，具体涵盖：基础标识信息（姓名、证件类型代码、证件号码）；生物特征信息（出生日期）；社会属性信息（联系地址、移动电话、国籍）；辅助识别信息（个人标识等）。敏感性项目枚举清单则采用最小化原则进行设计，包含出生地、出生证、国籍、地址、姓名等24类敏感要素。该分类体系通过结构化描述语言实现了数据敏感性等级的标准化映射，既符合《个人信息安全规范》中对个人敏感信息的界定标准，又通过机器可读的JSON格式实现了分级规则的可计算化表达，为后续动态分级防护策略的制定提供了元数据支撑。

本文利用上述多维数据的描述方案构建了一个基于上下文感知的动态数据敏感性分级体系，不仅静态地考量数据资产本身的固有敏感属性（如前述基于内容的分级），更将其置于具体的业务操作环境中进行动态权重调整。

具体而言，数据敏感性的动态评估模型主要耦合三个关键维度：

（1）数据本体敏感性：即数据字段的固有敏感级别，是数据分级体系的静态基础。

（2）API业务场景风险：不同业务场景的风险系数差异显著。例如，同为“客户姓名”字段，在仅供内部风控人员使用的后台管理API中，其潜在风险与暴露面可控，可被评估为较低风险等级；若同一字段出现于面向公众匿名访问的开放查询接口中，其面临数据爬取、滥用和泄露的风险急剧升高，必须以此动态上调其情境风险值。

（3）调用上下文元数据：分析API调用的上下文信息，包括但不限于请求者的角色权限、访问时间、地理位置、网络环境（如内网/公网）、操作行为（如批量查询或单次访问）等，形成动态的风险置信度评估。

数据敏感性动态分级决策过程的核心逻辑如图4。

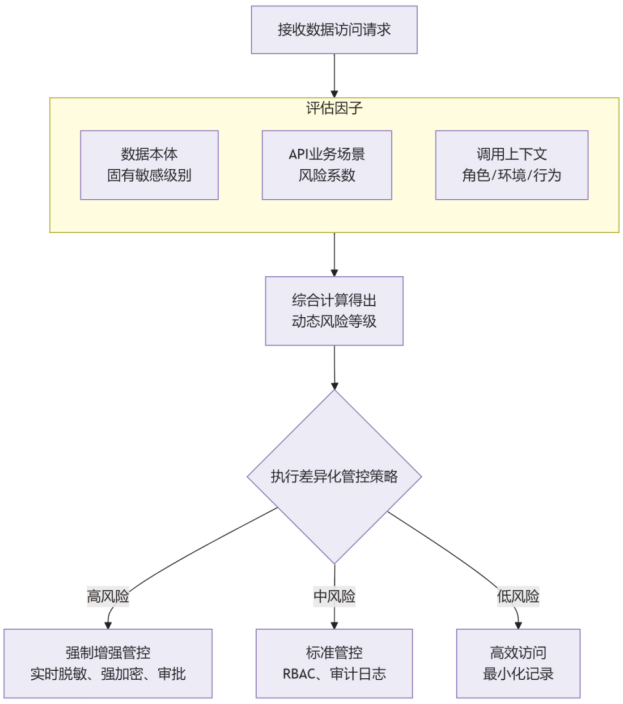


图4 API敏感数据动态分级决策过程

以上精细化、多因子加权的分级体系，将成为API安全风险管控平台实施差异化数据治理策略的核心决策依据。它将驱动系统智能地实施细粒度的安全控制：对于高风险场景下的敏感数据强制执行强加密与实时脱敏；对中风险场景实施基于角色的访问控制并记录完整审计日志；而对低风险场景则允许高效访问。从而最终实现从“以数据为中心”的静态防护到“以场景为驱动”的动态自适应防护的范式转变，在保障数据安全的前提下，优化数据的流通与利用效率。

（二）核心功能二：网络攻击的高效与精准检测

传统的API安全管控平台主要依赖于预定义的安全规则（如SQL注入、XSS的攻击模式特征库）、频率阈值（如每分钟API调用次数限制）和基础的身份认证与授权（如API密钥、OAuth）来提供防护。然而，对于证券行业对实时性、准确性和数据完整性的高要求而言，这种传统模式的局限性被进一步放大。首先，在瞬息万变、毫秒必争的金融市场中，API承载着海量的用户交互、数据同步、支付处理等核心业务，其任何一个环节的延迟或中断都可能带来巨大的业务损失或系统风险。其次，面对日益高级的、低慢性的、以及针对业务逻辑的API攻击（如撞库、批量爬取、欺诈交易、权限滥用），这种静态、基于规则的防御机制显得捉襟见肘。传统模式不仅误报率高、难以检测未知威胁，更关键的是无法提供证券行业所必需的快速响应速度和动态适应能力，无法有效保障金融业务的连续性和安全性。另外，维护规则库的巨大成本也与行业追求高效运维的目标相悖。

为应对这一挑战，本文提出在现有API安全管控平台之上，构建一个AI智能体增强层。通过引入多智能体协同分析框架，赋予API安全管控平台动态学习、智能推理和主动响应的能力，实现从“被动规则匹配”到“主动智能狩猎”的范式转变，显著提升API攻击检测的有效性与精准性。

AI智能体增强层的核心是一个由多个专用AI智能体组成的协同分析系统，其工作流程与平台集成关系如图5所示：

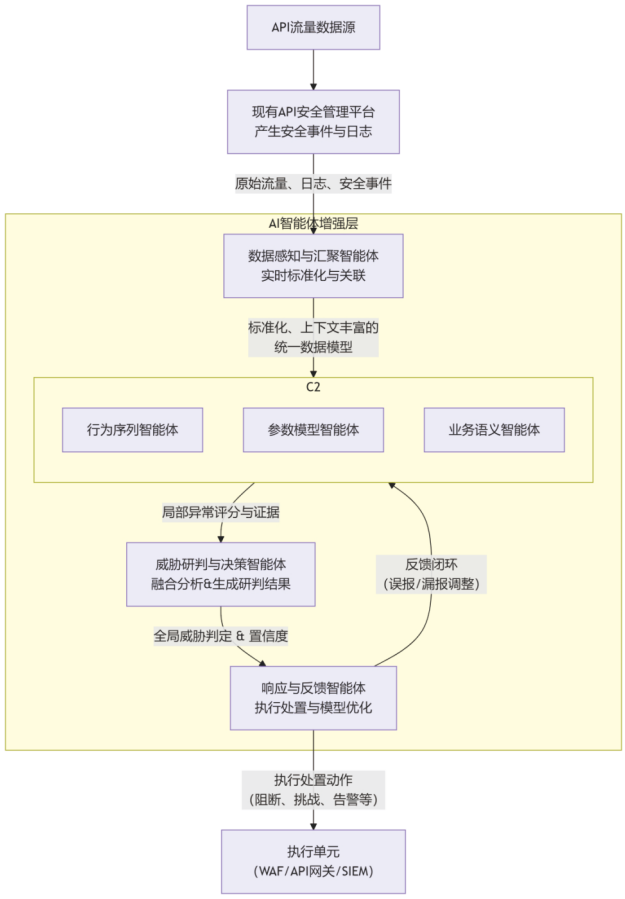


图5 AI智能体核心架构与流程

1、数据感知与上下文增强智能体

角色：数据感知与汇聚。

流程：作为“感官神经”，持续从现有API安全平台、API网关、应用日志、身份管理系统等数据源摄取原始数据。包括但不限于：原始API请求/响应、安全事件日志（如WAF告警）、性能指标（如响应延迟）、用户上下文（用户角色、登录历史、设备指纹）。其核心任务是对多源、异构的数据进行标准化、清洗和关联，为一个API请求构建一个包含“谁（Who）、在何时（When）、从何处（Where）、以何种方式（How）、做了什么（What）、结果如何（Outcome）”的完整上下文画像。

输出：一个富含上下文信息的、标准化的统一数据模型。

2、多模态协同分析智能体组

角色：多个专用分析智能体并行工作，从不同维度审视API请求。

流程：

A 行为序列智能体

焦点：分析用户行为序列的异常。不只看单次请求，而是将一个用户会话（Session）内的API调用序列视为一个整体。

技术：采用长短期记忆网络（LSTM）或Transformer模型，学习正常用户的行为序列模式（例如，“登录->查看商品详情->添加购物车->支付”是一个常见序列）。检测是否存在异常序列，如跳过关键步骤、高频访问非热门API、异常时间操作等。

B 参数模型智能体

焦点：分析API请求参数和负载的异常。

技术：结合有监督（如对已知攻击模式的分类模型）和无监督学习（如自动编码器-Autoencoder），识别出偏离正常分布的、潜在的恶意负载（即使是未知攻击变种）。

C 业务语义智能体

焦点：分析业务逻辑层面的异常。

技术：基于知识图谱或业务规则模型，理解API背后的业务含义（如“一个账户每秒最多只能尝试转账5次”、“一个新用户注册后通常先完成身份验证而非直接提现”）。识别出如批量爬取、积分作弊、欺诈提现等违反业务规则的恶意行为。

输出：每个智能体针对同一个API请求，输出局部异常评分和证据摘要。

3、威胁研判与决策融合智能体

角色：威胁研判与决策

流程：该智能体作为“大脑”，接收所有分析智能体提交的局部结果。采用基于加权平均算法，综合考量各智能体的评分、置信度以及它们的历史准确率，得出一个全局威胁判定和最终置信度。例如，若一个请求的参数模型评分略高，但行为序列和业务语义评分均正常，则可能被判定为低风险；反之，若三个智能体均报高分，则几乎可确认为高危攻击。

输出：生成最终的、高可信度的安全事件，及详细的研判依据。

4、智能响应与闭环学习智能体

角色：响应与反馈

流程：根据决策智能体的判定结果，自动执行或推荐相应的处置动作，构建一个持续的反馈闭环。动作可根据威胁等级进行差异化配置，例如：彻底阻断、人机验证、请求限速、仅记录告警、发送至SOC平台供安全分析师研判。安全分析师的最终裁决（如确认是攻击、或标记为误报）会被实时反馈给系统。所有AI模型利用这些反馈信号进行在线学习或增量学习，不断调整和优化自身参数，实现检测能力的自我进化。

输出：执行处置动作，并完成模型的迭代优化。

通过引入上述AI智能体增强层，API安全防护体系实现了从“单一、静态、基于规则”到“协同、动态、基于智能”的升级。该架构不仅显著提升了针对隐蔽和高级API攻击的检出率（有效性），并通过多智能体协同研判与反馈闭环极大降低了误报（精准性），有效应对了该行业对极致实时性与安全性的双重挑战。

（三）核心功能三：证券行业强监管性的设计实践

证券行业的强监管性不仅体现在交易业务本身，更深度融入数据安全、隐私保护、身份治理与操作风险等方方面面，呈现出全面性、穿透性和审慎性的显著特征。传统的通用型API安全管控平台难以充分适配这些细粒度的行业监管要求。

证券行业的强监管性在非交易类API管控上主要体现为以下三大核心诉求：

（1）数据安全与隐私合规：API行为必须严格遵循《网络安全法》《数据安全法》《个人信息保护法》以及金融行业数据安全标准。要求对数据，特别是高敏感度的客户信息，实施基于分类分级的精细化访问控制，并确保其在传输、存储、销毁全生命周期中的安全，严防数据泄露、滥用和违规出境。

（2）身份认证与访问控制：要求建立并严格执行最小权限原则和职责分离原则，确保API的每一次调用背后都有一个明确的、经过严格认证和授权的身份，并能有效防范权限冒用、账号共享和越权访问。

（3）可审计性与可溯源性：要求能对任何数据访问和操作行为进行完整、不可篡改、可快速追溯的审计。

1、面向行业强监管的AI智能体架构设计

为满足行业对API管控的核心诉求，本文设计三大核心AI智能体，将监管的抽象条文转化为持续运行的数字化管控能力，如图6。

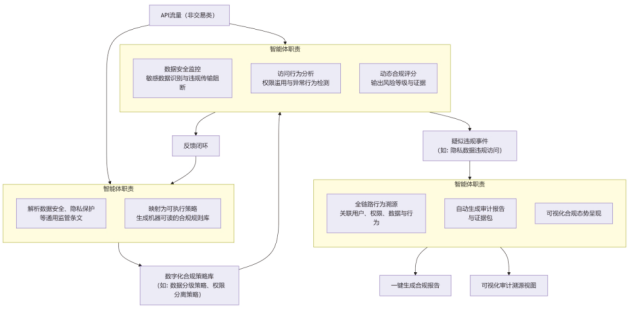


图6 面向证券行业强监管的AI智能体协同架构

（1）监管规则数字化智能体

角色：监管规则数字化。

流程：利用自然语言处理（NLP）技术，解析《数据安全法》《个人信息保护法》及金融行业标准中的关键条款，并结合企业内部的数据分类分级政策，自动生成可执行的管控规则。

输出：结构化的数字化合规策略库。例如：

```
数据策略：IF (DATA_FIELD == "身份证号码") THEN (LOG_FIELD must be MASKED) AND (DESTINATION_IP must be DOMESTIC)
权限策略：IF (USER_ROLE == "实习生") THEN DENY (ACCESS to API_PATH containing "/customer/delete")
分离策略：PREVENT (USER who has "敏感数据导出"权限) from (REQUESTING "用户权限申请"API)
```

（2）实时合规审计智能体

角色：实时合规审计。

流程：对API流量进行深度实时分析，7x24小时审视所有API调用，专注于数据安全和权限合规。其核心能力包括：

A 敏感数据智能识别与监控：基于NLP和模式识别，动态发现流经API的客户姓名、证件号、手机号、资产证明等敏感信息，并依据策略监控其是否被违规访问、明文传输或试图向境外发送。

B 用户行为异常检测：建立每个用户、每个服务账号的正常访问基线（时间、频率、数据量、访问模式）。实时检测偏离基线的异常行为，如：内部员工非工作时间批量下载客户清单、运维账号访问与其职责无关的业务数据API等。

C 权限滥用分析：检测违反职责分离原则的API调用序列，或权限提升后的异常操作。

输出：实时产生数据安全风险事件与身份治理风险事件，及上下文证据。

（3）智能溯源与报告智能体

角色：智能溯源与报告

流程：将技术证据转化为管理语言，满足审计与合规汇报需求。针对风险事件，自动关联其背后的用户身份、授权记录、访问的数据资产、完整操作序列，并生成清晰的可视化图谱，定位问题根源。

输出：

A 自动化合规报告：根据内部审计或监管报送的固定周期与模板，自动聚合一段时间内的API安全态势、风险处置情况、合规性评分等，生成数据安全合规月报或自查报告。

B 合规态势可视化大屏：为管理层提供全局视角的API安全与合规态势大屏，直观展示核心敏感数据的分布、访问热度、风险趋势和整体合规度。

2、智能体赋能下的监管价值体现

引入AI智能体后，文本化的法规和制度可转化为可自动运行、持续监控并不断优化的数字规则，实现了监管要求的技术化管控，从而确保了合规执行的刚性和一致性。这一转变将提升数据安全治理的主动性和精细化程度，企业得以从被动响应转向主动发现，从粗放管理迈向精细管控，真正实现了对敏感数据流动的“看得见、管得住、说得清”，从而筑牢数据安全的底线。同时，可提升企业应对内外部审计和检查的效率与可信度，快速、准确、全面地提供证据材料，充分证明在数据安全和权限管理方面的尽职尽责，从而有效构建并维护企业良好的合规声誉。

五、总结与展望

通过智能体驱动的API安全风险管控研究与实践，构建了一个集“智能资产发现、深度风险检测、自适应防护、动态响应”于一体的、闭环的API安全风险管控体系。该体系不仅能够帮助企业有效保护核心敏感数据资产，防止客户信息和商业机密被窃取滥用，还能确保核心业务系统的完整性与可用性，保障业务的连续稳定运行，同时助力企业高效地满足日益严格的法律法规和行业监管要求，规避潜在的法律风险与巨额罚款，维护企业的品牌声誉。

证券行业面临的API攻击类型多样且变化迅速。展望未来，通过构建一个能够自我学习、自我调整、自我进化的自动化、智能化、可扩展的防御体系，不仅能防御已知攻击，还能预测未来的威胁，在复杂的生态系统中发现隐藏的风险，并最终达到可自我修复的水平，为保障我国证券行业的数据安全和业务稳定，贡献出坚实而深远的力量。

基于某次大型攻防演习应对零日漏洞攻击的威胁狩猎实践

林宝晶 | 奇安信网神信息技术(北京)股份有限公司

钱钱 | 中国航天系统科学与工程研究院

摘　要： 零日漏洞属于未知威胁的一种，鉴于现有安全设备防护均是检测发现已知威胁攻击，导致现有的防御体系缺少未知威胁防护能力，为了应对未知威胁或者高级持久攻击，需要一种新的方式来弥补当前防护体系的缺陷。本文通过介绍攻击过程特点、聚焦后渗透阶段的异常行为的威胁狩猎方法，从而尝试有效地应对未知威胁的攻击。并通过具体某一次实战攻防演习中应用威胁狩猎的方法，捕捉到演习中的零日漏洞攻击，为后续安全防守团队构建未知威胁防护能力提供参考。

关键字： 零日漏洞、威胁狩猎、未知威胁、高级持久性威胁

一、零日漏洞难题

经过数十年的发展与演变，现有网络安全设备均是通过已知攻击特征检测发现攻击行为，简言之，当前的防御能力基本构建于已经披露的漏洞和已知的攻击行为，而针对未知威胁尚无检测与防护能力。零日漏洞（以下简称0Day）是指软件或系统中未被公开的、未被厂商知晓的安全漏洞，是未知威胁的一种，由于其引发的攻击伤害性巨大，被称为网络安全行业中的“核弹”。从防守角度看当前市场上缺少有效的0Day防护安全系统。所以，0Day攻击已经成为网络安全防守方面临的主要难题之一。

二、零日漏洞应对之道

正如ATT&CK模型或者杀伤链模型中所阐述，任何攻击都是面向过程的，由一系列的“攻击单点”所组成。从以往攻击队的技战术分析，0Day攻击行为一般都是发生在建立据点阶段或初始访问阶段。而攻击者进入到受控网络，更多采用一些常规攻击手段开展横向移动与隐藏。这就为防守方发现零日漏洞攻击提供了解决思路。随着近两年攻防演习的开展，我们摸索出一套零日漏洞防御的方法。目前该方法建立在如下的2个模型之上——后渗透阶段检测模型和威胁狩猎模型。

（一）后渗透阶段检测

如前文所述，高级攻击行为很难在第一时间被网络安全设备发现，所以安全团队应聚焦渗透后的攻击行为检测发现（如下图所示），可以帮助防守方尽早发现来自外部的高级攻击行为。

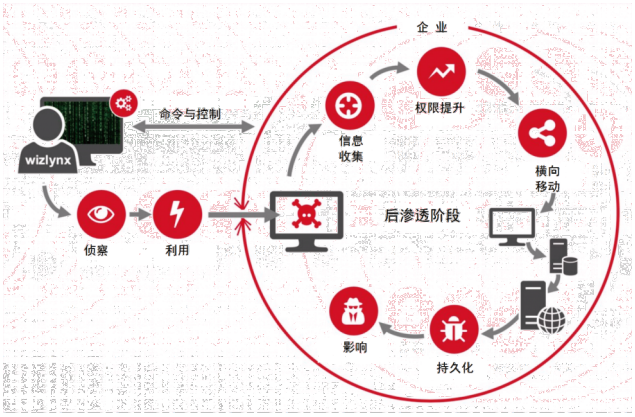


图1 后渗透阶段检测理论模型

从图1中可以看出，在侦察、利用阶段，主要是攻击者从外部进入到内部的一个过程，这个过程一般会比较短暂。而攻击者在企业内部的活动时间会更长。依据ATT&CK理论，攻击者为了达到攻击目的，在内部会开展大量的信息收集、权限提升、横向移动、持久化、数据渗出、影响（勒索加密）等一系列操作。这样就势必在内部留下各种蛛丝马迹，如产生webshell、创建新的进程、服务或账号等信息。防守方应充分地利用网络中的安全设备所产生的数据，发现攻击者在网络内的活动，从而反向分析、溯源到真实的攻击源。

（二）Sqrrl威胁狩猎

虽然我们知道要关注后渗透行为的检测与发现，但是在发现异常事件或现象后，安全防守团队需要一个结构化的理论来支持其分析行为，这就是现在比较热门的话题——威胁狩猎。为了应对日益变化的网络安全威胁，威胁狩猎成为传统安全防护技术与安全运营的有效补充。所谓威胁狩猎，就是人为驱动，主动并迭代搜索网络、端点侧的

数据，旨在发现那些已绕过系统内自动化检测工具的恶意、可疑或风险行为。威胁狩猎是由安全从业人员在实践中总结出来的一种安全工作方式，网安人员逐渐意识到传统的防御手段发现攻击行为的时效性较为滞后，为了尽早发现已入侵到内部网络的攻击行为，安全人员发现，可以利用已掌握的数据和异常线索，进而假设内部存在某种攻击行为，利用数据分析技术验证并找到真实的攻击行为。

威胁狩猎已存在多年了，但它是一个新的技术话题。并且，我们要认识到威胁狩猎不是一种技术，而是一种方法。在威胁狩猎理论方面，许多厂商都发布了自己的威胁狩猎模型，本文介绍一种结构化威胁狩猎的方法——Sqrrl威胁狩猎环。

虽然Sqrrl公司现在已经不复存在，但是Sqrrl公司是最早在威胁狩猎领域开展研究的公司之一。威胁狩猎存在一定的不确定性，所以为了避免威胁狩猎无效（无功而返），制定一个常规的威胁狩猎流程就显得十分必要。Sqrrl公司构建了威胁狩猎环模型（如图2所示），Sqrrl威胁狩猎环包括4个阶段，该模型不仅可以指导威胁狩猎分析人员的战术执行，从而尽可能迅速地、有效、有序地开展威胁狩猎活动。威胁狩猎环更为有利的一点是通过不断的有效迭代执行威胁狩猎环，从而不断促进威胁狩猎流程自动化地不断发现新的威胁。

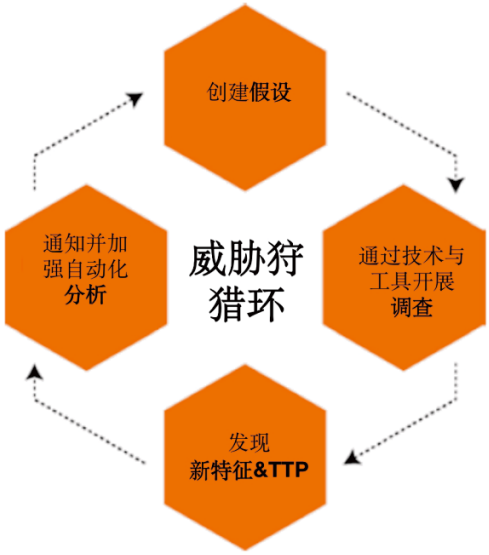


图2 Sqrrl威胁狩猎环

接下来，我们详细解释一下Sqrrl威胁狩猎环的四个阶段：

第一阶段：假设。一次威胁狩猎始于IT环境中可能正在进行的某种活动而提出一个假设或猜想。该假设即由威胁狩猎人员综合考虑多种因素后提出的可验证的一种想法，这些因素包括可信的情报、威胁情报、企业的数据基线以及威胁狩猎人员自身的从业经验等。该假设被用于威胁狩猎人员后续调查的场景，并通过收集各种数据检验该假设是否成立。

第二阶段：调查。一旦假设被创建，威胁狩猎人员就应该通过各种工具和技术手段来跟踪假设，验证假设是否成立。因此调查阶段的核心目标就是证明假设成立或否定假设。一般来说，在调查过程中，首先要确定为了验证假设所需的数据集合，然后通过各种技术手段和分析工具来分析这些数据，最后得出假设是否成立的结论。比较常见的假设是企业内部已经存在了某种攻击行为，然后威胁狩猎人员希望利用网络中的现有数据来分析出这种攻击行为是否存在。而这种攻击行为很可能代表一种新的攻击手段或者攻击技战术，所以威胁狩猎人员在这个过程可能发现新的攻击行为的特征，并确认新攻击行为的攻击路径，通过总结这些特征和攻击路径，可以进一步揭示攻击者的战术、技术和流程(TTP)。即使在这个阶段，威胁狩猎人员没有发现异常或攻击行为，威胁狩猎人员也可以通过调查过程排除内部存在特定战术或失陷的行为。本质上，这一步是“证明或驳斥假设”的过程。

第三阶段：发现新特征&TTP。威胁狩猎人员使用特定工具或者技术对调查过程中所发现的特定攻击行为或异常数据进行技术分析，从而提炼出攻击者的攻击特征或者攻击技战术。威胁狩猎人员在本阶段的发现是威胁狩猎成功的关键标志之一。也就是说，威胁狩猎人员在本阶段总结出所发现的攻击行为的新型恶意代码特征或者攻击者所利用新的技战术手段。对于Sqrrl威胁狩猎环来说，本阶段的成果，在威胁狩猎环中发挥承上启下的作用。所以，对于Sqrrl威胁狩猎环来说，威胁狩猎不仅仅是要发现未知威胁/新威胁，更重要的是完善已有防御体系中的检测及防护能力，这就需要威胁狩猎人员提取出未知威胁的攻击特征或者TTP，然后才能转化成现有设备的自动化检测/防护的策略。

第四阶段：通告并加强自动化分析。一次成功的威胁狩猎一定为后续的安全通告和自动化分析能力奠定基础。一旦威胁狩猎人员在威胁狩猎过程中找出一套检测新威胁的数据查询分析过程，安全运营团队就可以通过脚本/程序实现自动化检测。这样做有两个好处，其一可以丰富完善防御体系的自动化检测能力，其二避免浪费威胁狩猎团队的时间做重复性的胁狩猎活动，让威胁狩猎团队聚焦新的狩猎活动。威胁狩猎另一个成果是针对发现的新威胁形成安全通告，安全通告内容不仅包括新威胁的基础信息，同时还包括新威胁检测机制与防护机制，例如SIEM规则或IPS检测特征。而这需要威胁狩猎人员对其所要保护的网络环境有一定的了解，才能针对发现的新威胁编制适当的安全通告。所以当威胁狩猎人员在威胁狩猎中有新发现时，记录并利用好这些新发现对提升整体的防御能力具有十分重要的意义。

三、Sqrrl威胁狩猎环应用

（一）应用背景

随着国家网络实战攻防演习不断深入，0Day已经成为攻击过程中一个重要的突破手段。在演习的总体方案中制定了0Day应对措施——聚焦后渗透阶段的威胁狩猎，已经成为一个主要的策略。

（二）基于异常产生假设

一般来说，0Day漏洞利用之后总是有一些异常现象，威胁狩猎人员可以对业务的敏感性来产生威胁狩猎的假设。攻防演习期间所监控数据中出现了大量的加密HTTP流量，可以基于这个异常来产生一个威胁狩猎的初步的假设——“难道我们的内部主机已经被攻击队入侵了？”。

为了初步验证假设，威胁狩猎团队可以先围绕资产属性信息做出初步判断，如果是高危的资产，如X软统计报表系统。则可以依赖攻防领域的经验，设定更进一步的假设：攻击队已经利用了X软的0Day漏洞入侵到我们内部网络。产生该假设的前提其一是在攻防演习前安全团队已经对X软做了版本升级；其二X软报表中间件应用广泛，攻击队每年都会储备X软报表的0Day漏洞。

（三）调查分析

基于上述产生的假设，威胁狩猎团队应开展调查必要的活动。首先在数据方面，本次威胁狩猎应需要如下的基础数据：

■异常流量的内部IP地址的资产信息

■异常流量内部IP前12小时的所有流量数据

■异常流量内部IP前12小时的主机活动日志

首先，威胁狩猎团队对感觉异常的流量占展开分析——确认是否属于标准的协议。从本次收集的网络数据包可以分析出，其不是一个标准的HTTPS协议，但是采用了HTTPS的标准通信的443端口传输数据，大概率是一个利用443端口做C2隧道的通信。

另外，威胁狩猎团队利用自身经验，马上对前12小时的流量做了分析，通过工具过滤分析聚焦该主机与外部疑似攻击IP的通信流量。在起始通信流量的payload发现了一个有特征的二进制处理过的内容，通过拷贝上述内容，存成一个二进制的文件，然后利用HEX工具查看（如图3所示）：

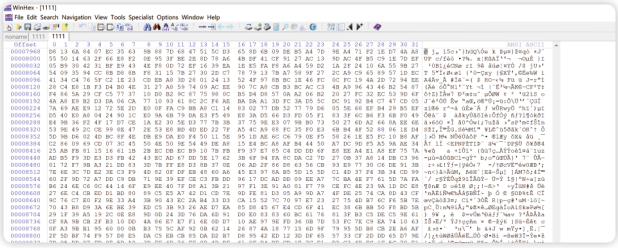


图3 流量中包含的gzip文件流

通过工具把二进制文件存储成.bin文件，然后利用文件识别工具，发现是一个gZip类型文件，通过Gzip解压后，发现对应的文件是一个Java的可执行文件（class类型文件），解压后的部分内容如图4所示。狩猎团队的成员基本上确定这是一个Java反序列攻击行为。

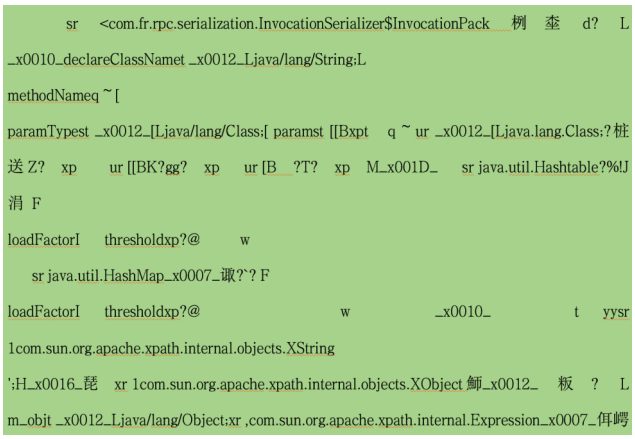


图4 gzip流解压后的class文件内容

为了进一步验证自己的假设，威胁狩猎团队分析目标主机的进程执行相关的数据，在时间戳相近的位置，发现主机上三条疑似信息查询（如图9所示）的命令执行：

■Ipaddr

■Ping 8.8.8.8

■Whoami

上述这些线索基本上验证了威胁狩猎团队的之前的假设，这台主机已经被攻击队入侵。为了进一步确认这个入侵是一个0Day攻击，威胁狩猎团队在受攻击主机上提取数据，进一步确认0Day漏洞有关的信息。从主机中间件的日志结合攻击流量时间戳发现了如下的请求：

图5 主机中间件的日志信息

从图5可以看到一个类com.fr.decision.extension.report.api.remote.RemoteDesignResource.onMessage类，其中包括了messageListener.handleMessage()接口，调用了deserializeInvocation(var1, var2)方法，并且进一步确认了这里的var1变量是上面传的GZipSerializerWrapper.wrap(InvocationSerializer.getDefault())。通过检查GZipSerializerWrapper.deserialize方法，截止到这一步，基本确认了该系统存在一个未知的反序列化漏洞。

（四）提取特征或TTP

在提取特征或者TTP之前，小H首先通知了应急小组的人员马上对受害目标主机采取隔离措施，避免攻击者利用这个目标主机进一步扩大攻击范围。同时让安全运营团队

尽快分析攻击源是否与其他主机有通讯，确认内部是否存在更多的失陷主机。

小H通过阅读前面的class文件的源代码，小H从反编译的伪代码发现源代码中的几个关键问题：

（1）cube参数是一个加密的字符串，其加密方式采用AES，密钥为34522cea8z276c89。

（2）同时，在cube加密之后，还采用了Base64的编码处理。

从上述反编译的代码可以发现，攻击者对web请求的返回数据进行了加密处理。采用加密手段Base64编码与Rot18编码之后返回。这样，小H获得了加密流量两端的解密算法以及与之对应的密钥之后，编写程序快速、全面对攻击者IP与受害主机之间的网络流量网络数据包解密并验证，进一步证实所怀疑的攻击行为以及攻击者在内部的所有操作。攻击者通过加密流量做了如下的操作：

■上传内存马

■获取主机的信息（IP addr、whoami）

■尝试出网（ping 114.114.114.114）

■主机提权

■获取中间件信息（hibernet以及X软配置信息）

■获取X软访问令牌

■尝试数据渗出

小H全面分析后，对于本次的攻击行为获取到如下的特征：

（1）0Day漏洞入口：本次的0Day漏洞为X软组件的根目录，存在反序列漏洞，具体的入口点为http://xxx/webroot/decision。

（2）内存马特征：通过前面的分析，、依据前面反编译的源代码发现了Java内存马的文件为Fr_memshell.class。同时，并且小H依据获得的Fr_memshell.class文件（通过二进制存储获得）生成了对应的SHA256和MD5值：

■MD5：2c30ce56568d76fbe31ad5a81343027f

■SHA256：39b91f4d83d93d4e4dbfbb2d64a1f1d82affc7639930901ad370c88e9b9caf8d

（3）恶意IP：通过威胁狩猎分析，攻击的IP地址以及C2的控制服务器均为221.216.117.204。

（五）通告并加强自动化分析

依据提取到的攻击TTP信息，小H将上述的信息通知到安全运营团队与应急团队，并将漏洞的信息通知到X软供应商，希望其尽快修复加固漏洞。同时，在内部也围绕自动化防御开展如下的工作：

■在WAF防护规则方面，增加/webroot/decision/remote/design/channel的访问过滤的拦截策略，保证在漏洞修复隔离该漏洞入口的访问；

■将内存马Hash值增加到HIDS系统的黑名单，防止后续受入侵主机的此类内存马的运行，同时手工运行HIDS的主动全网内存马扫描，检查现有环境内是否存在感染此次内存马的主机；

■把本次发现的恶意IP添加到互联网防火墙黑名单。

四、小结

依据SANS 2020年度威胁狩猎报告，攻击者可以在受到感染的环境中平均驻留90天以上。而本次在实战攻防演习采用威胁狩猎方法对抗0Day漏洞攻击，从0Day漏洞利用到0Day漏洞攻击行为发现，时间间隔为8小时左右，可以看到通过威胁狩猎可以大幅压缩攻击者采用未必威胁行为在内部驻留的时间，也为我们探索未必威胁对抗积累经验。随着威胁狩猎方法、内部数据以及数据智能分析手段的不断完善，威胁狩猎将成为网络安全防御体系中重要的一环。

参考文献

[1]Sqrrl. Hunt Evil: Your Practical Guide to Threat hunting

Https://www.threathunting.net/files/hunt-evil-practical-guide-threat-hunting.pdf 2015

[2]Sqrrl Team. The Threat Hunting Reference Model Part 2: The Hunting Loop

https://www.threathunting.net/files/The%20Threat%20Hunting%20Reference%20Model%20Part%202_%20The%20Hunting%20Loop%20_%20Sqrrl.pdf 2015

03 新技术应用

P56 基于零信任架构的安全访问和智能协同研究与实践
黄辉、崔荣涛、韩宇

P61 零信任体系在证券业数据安全领域的探索与实践
俞建群、邬晓磊
何艺

基于零信任架构的安全访问和智能协同研究与实践

黄辉、崔荣涛、韩宇 | 湘财证券股份有限公司

摘　要： 本文探讨了证券行业在数字化转型过程中面临的网络安全挑战及零信任架构的应用实践，通过可信访问控制台（TAC）、应用代理系统（TAP）和客户端（TrustAgent）三大核心组件，实现动态细粒度访问控制、业务隐藏与端口收敛、自动化权限管理等功能。通过与IAM、CMDB、ITSM和OA等系统的深度集成，推动了灵活办公与数字化转型。

关键字： 零信任架构、自动化权限管理、动态访问控制、智能协同办公

一、前言

随着证券行业数字化转型推进，核心交易、风控等多类系统规模持续扩大，远程办公与分支互联让网络边界愈发模糊，传统安全架构面临严峻考验。基于 VPN 的远程接入存在“先连接后认证”缺陷，易遭 0Day 攻击和口令爆破，静态访问控制难以抵御横向移动与内部威胁。

为此，湘财证券以“身份为基石，数据为中心”构建零信任安全接入体系，通过端到端身份化管理实现统一识别与动态细粒度控制。结合用户身份、设备状态等多维度属性，依据业务分类分级策略持续验证安全状态、动态调整权限，并与IAM、CMDB、ITSM和OA等系统深度集成，打破网络位置依赖，践行最小权限原则下的持续验证与授权，有效防范风险，支撑多场景安全便捷访问。

二、证券行业零信任建设需求分析

针对零信任的安全建设，业内已有较多的实践案例。本方案将重点关注建设过程中的实际需求和场景，提出不同场景的解决方案，供同行参考。在业务远程访问方面，通常的需求如下：

1、传统VPN远程接入业务访问易用性较差

用户接入后常不清楚可访问的业务系统，需依赖收藏夹或文本记录地址，系统增多后体验更差。更关键的是，若终端失陷，攻击者易获取业务访问地址，带来较高安全隐患。

2、统一身份认证系统对接

企业广泛采用统一身份认证系统（IAM）实现精细化权限管理，而零信任架构要求用户无论位置均需严格身份验证授权。因此对接 IAM 单点登录（SSO）时，需重点保障 IAM 安全性，需将其置于零信任保护边界内，防止暴露于公网，避免成为攻击者直接目标，从而降低潜在安全威胁。

3、按需申请和授权，访问期限灵活调整

用户远程应用访问需按业务需求灵活设置权限期限，避免“永久授权”以降低滥用和暴露风险。但人工处理权限开通、控制与回收效率低且易疏漏，难适应大规模场景。因此需引入自动化权限管理，集成身份治理、审批流程与零信任结合，实现申请、审批、发放、调整及自动回收的闭环。

4、用户规模大，用户角色和应用场景复杂

证券公司用户类型多样，含正式员工、外包及合作伙伴等，不同角色访问需求差异大，接入场景涵盖互联网远程与办公网访问等。需结合业务敏感度、用户身份、设备状态、网络环境及行为特征等多维度信息，开展端到端的动态风险评估。

在此背景下，研究零信任架构与统一身份认证（IAM）、自建应用门户、持续认证等技术的深度融合，对解决证券行业安全挑战意义重大。构建以身份为核心的动态化、自动化零信任体系，可解决权限滥用等问题，适应复杂场景需求。

三、解决方案设计与建设

本方案零信任系统含三大核心组件：可信访问控制台（TAC）作为控制平面核心，实现应用控制、动态授权等能力；应用代理系统（TAP）为数据平面关键，代理业务应用并加密授权流量以收缩暴露面；零信任客户端（TrustAgent）是用户侧入口，提供认证、终端感知等功能。

以零信任为底座，融合身份、权限、终端及数据安全，采取“由外及内、分步实施”路径：先从互联网远程访问切入收敛暴露面，通过分批试点替代传统 VPN，再逐步扩展至内

网场景，实现全场景防护，兼顾安全性、可扩展性与体验，支撑证券业安全协同需求。

主要建设方案分为以下几个方面：

1、建设零信任基础设施架构

方案基于零信任架构，采用双数据中心冗余部署，各节点均以集群高可用模式部署，实现冗余接入与统一管理。TAC 控制台单数据中心采用集群双活部署，单站点故障不影响服务；双中心间为主备冗余，数据实时同步，支持故障时平滑切换。可信应用代理系统集群多活部署，具备智能路由能力，可动态适配接入环境。单点故障时流量自动调度，保障业务连续与服务高可用。

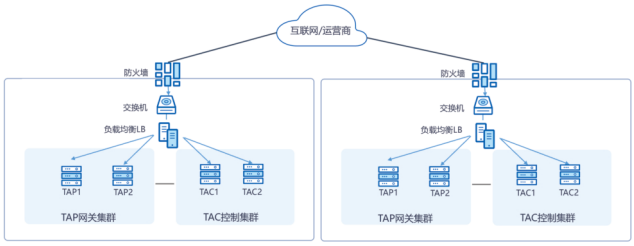


图1 零信任基础架构

2、收敛互联网应用系统访问入口

通过零信任可信应用代理系统（TAP），可全面保护互联网侧非对客应用资源，实现业务隐藏与端口收敛。所有访问请求经零信任客户端（TrustAgent）导流，通过 SPA 单包授权建立安全隧道，传输数据端到端加密。应用系统不再直连公网，仅允许认证合规用户接入 TAP，有效防未授权访问。非对客系统接入零信任网关，统一管控入口，显著缩小攻击面，降低暴露风险。

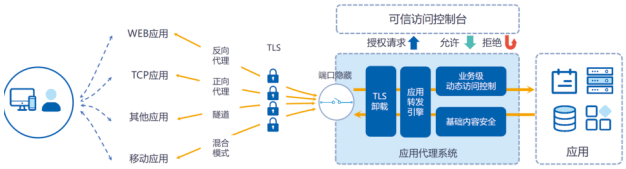


图2 应用系统访问示意

3、CMDB应用发布自动推送

为提升零信任环境下应用发布的管理效率与数据一致性，本方案将所有应用发布台账统一维护于CMDB（配置管理数据库）中。系统运维人员在ITSM平台发起应用发布流程，经审批后在CMDB中完成应用信息配置。该信息通过标准化API接口实时同步至OA系统和零信任平台，确保三方在应用清单、访问权限和系统状态等方面保持台账一致。通过自动化推送机制，避免了人工重复录入带来的误差与延迟，提升了应用上线和权限开通的协同效率。同时，该机制为后续的权限自动化管控、访问策略动态生成和安全审计提供了准确、可信的数据基础，实现了应用生命周期管理与安全访问控制的深度融合。

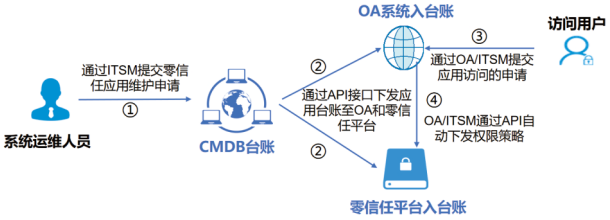


图3 应用发布维护

4、应用权限自助申请

为提升权限管理的效率与用户体验，本方案支持用户通过OA/ITSM系统自助发起应用访问权限申请。系统根据申请人角色、所属部门及申请的权限类型，自动匹配差异化的审批流程和策略模板，实现精细化的权限管控。申请经多级审批通过后，OA/ITSM系统通过标准化API接口将授权信息实时同步至零信任平台，自动完成访问策略的生成与下发，实现“申请—审批—生效”全流程自动化。该机制避免了人工配置带来的延迟与操作风险，确保权限分配及时、准确、可审计。同时，支持临时访问、项目制权限等场景，可设置有效期并自动回收，全面落实最小权限原则和动态授权理念，提升安全合规水平。

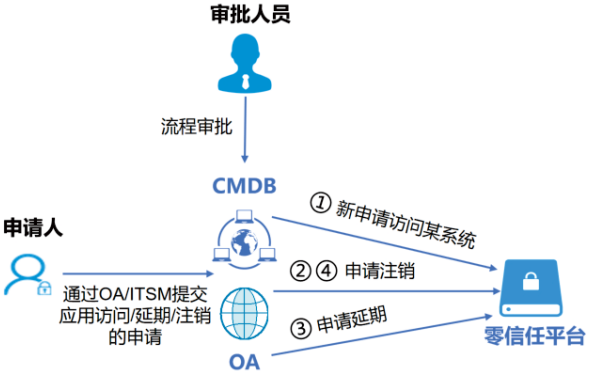


图4 全流程权限自动化配置

5、动态访问控制

在静态授权基础上，结合业务分级分类要求构建多维度动态访问控制策略。综合用户角色、终端状态、网络环境、访问时间、行为特征及应用敏感度等属性，建立与业务风险等级匹配的差异化规则。高敏感系统设更严条件，如仅限合规终端、特定时段访问并触发多因素认证。系统实时采集属性信息，经策略引擎动态匹配，自适应调整准入权限与认证强度，异常时即时降权或阻断。通过业务分级与持续风险评估，实现“按级授权、动态调整、最小权限”管控，平衡灵活性与安全性。

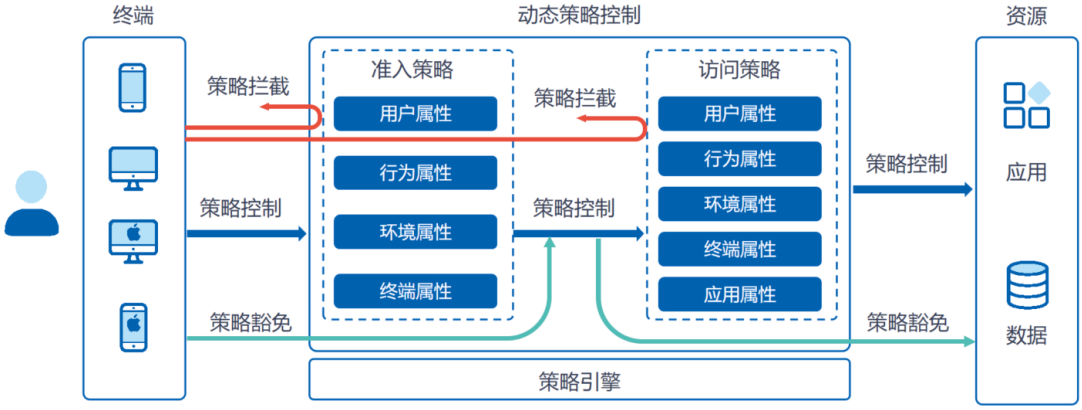


图5 动态策略控制

6、搭建零信任安全访问门户，建立统一身份认证机制

为提升用户体验与系统安全性，本方案构建统一的零信任安全访问门户，并与企业统一身份认证系统（IAM）深度集成，实现跨终端、多场景的便捷安全访问。

■PC终端访问场景：

在PC终端场景下，用户通过零信任认证页面登录后，自动跳转至自建统一门户Web页面，并通过单点登录（SSO）机制完成与IAM的对接，实现一次认证、无缝访问。登录过程中，系统将检测终端是否安装零信任客户端（TrustAgent），若已安装则无感拉起并建立安全隧道，确保用户可直接访问隧道应用资源。对于未安装客户端的用户，亦可以通过门户直接访问七层应用资源。门户页面集中展示用户已授权的应用系统清单及权限有效期，用户点击目标资源即可实现页面跳转与自动登录，提升访问效率与直观性。

■移动端访问场景：

在移动端场景中，通过在企业办公APP中集成零信任SDK，实现与零信任架构的深度融合。办公APP与零信任组件均对接IAM系统，用户在APP内完成一次登录，即可同步完成身份认证与安全接入。所有业务访问流量经由零信任通道加密传输，结合终端合规性检查与动态策略控制，实现对移动访问的安全管控。同时，通过零信任能力收敛办公APP的互联网暴露面，有效防范未授权访问与数据泄露风险。

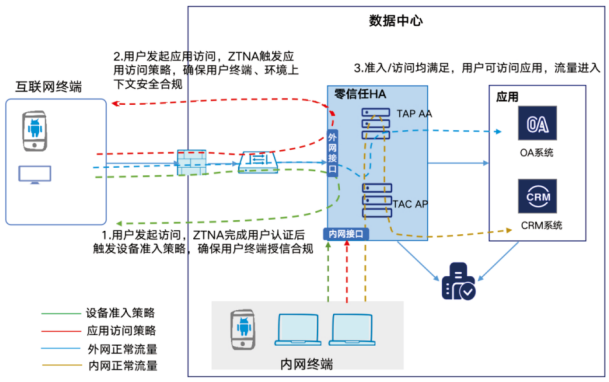


图6 终端访问场景

7、零信任边界对IAM的防护机制

通过零信任架构为用户提供统一的安全接入能力，实现内外网一致的身份认证、准入控制与访问策略管理。在此模式下，统一身份认证系统（IAM）不再直接暴露于互联网，无需开放任何公网端口，所有认证请求均通过零信任网关进行代理和转发。零信任平台与IAM之间采用OAuth 2.0协议的密码模式（Password Grant）实现安全对接，在确保单点登录体验的同时，将IAM系统置于零信任边界内部，形成逻辑隔离的保护层。该机制有效收敛IAM的暴露面，防止其成为外部攻击入口，避免因直接暴露导致的口令爆破、非法探测等安全风险，显著提升身份核心系统的安全防护能力。

四、零信任项目落地中的场景问题与解决方案

在部署零信任架构时，需充分考虑安全性、系统兼容性和人员操作习惯，以下为项目实施过程中的部分场景及相应解决方案：

（一）IAM部署在内网，需支持互联网应用的SSO认证

1、场景需求

由于安全要求，IAM系统未向互联网开放服务端口，但零信任平台及其他外部应用需要依赖IAM完成互联网用户的统一认证（SSO）。这种情况下，如何确保IAM系统的安全性同时支持外部应用的认证需求成为关键问题。

2、解决方案

将IAM作为受控应用接入零信任平台，用户访问IAM前需先通过零信任认证。具体措施包括使用OAuth 2.0协议的Password模式替代常用的授权码模式，以减少外部暴露的风险。采用双域名策略，内部系统继续使用原有的内网域名进行IAM服务访问，而零信任平台则使用独立的外部域名进行接入。这样不仅实现了内外网访问路径的隔离，还增强了IAM系统的安全性。此外，所有对外的认证请求均通过零信任网关代理，确保只有经过严格身份验证的用户才能与

IAM系统进行交互，从而避免了IAM直接暴露于公网带来的潜在威胁。该方案既保障了IAM系统的安全性，又实现了跨网络环境下的统一认证体验，有效支撑了互联网业务的SSO需求。

（二）改造企业现有门户系统，兼容零信任接入

1、场景需求

企业已建设统一门户系统，用户已形成稳定的使用习惯。若强制替换为零信任厂商默认门户，不仅影响用户体验，还将增加UI定制开发、多端适配及后续运维管理成本，不利于系统平稳过渡和长期运营。

2、解决方案

将企业现有门户作为受控应用纳入零信任体系，通过零信任平台进行安全发布。用户完成零信任客户端认证后，平台根据IAM下发的AppId和认证Code，自动触发跳转至自建门户页面，实现无感接入。门户后端与零信任控制中心通过标准API对接，完成身份令牌验证、访问策略同步和会话状态管理，确保整个访问流程的安全性与连续性。该方式保留了原有门户的界面风格与操作逻辑，避免用户重新学习成本，同时实现与零信任架构的深度融合，兼顾安全性、兼容性与可维护性。

（三）按需授权，权限到期前需提示用户

1、场景需求

为强化访问控制，零信任平台对每位用户、每个应用均实施细粒度的访问策略管理，避免“永久授权”带来的安全风险。同时，为提升系统可用性与用户体验，需在权限即将到期前主动提醒用户，防止因权限失效导致业务中断，影响工作效率。

2、解决方案

零信任平台提供查询用户应用授权有效期的开放接口，由企业自建统一门户系统调用并获取用户的权限状态信息。门户前端对授权剩余时间进行可视化展示，并在临近到期时通过弹窗、消息通知等方式主动提醒用户，支持续权申请操作，实现“到期前提醒—申请—审批—续期”的闭环管理。权限策略在审批完成后通过API自动下发至零信任系统，实现全生命周期的自动化管理。该机制支持按天、周、月或项目周期灵活配置访问期限，到期后自动回收权限，确保始终遵循最小权限与时效性原则，在保障安全的同时显著降低运维负担，提升管控精细化水平。

（四）应用自动维护发布

1、场景需求

用户在申请应用访问权限前，需由业务运维人员先行完成应用发布。为降低管理复杂度，零信任中的应用信息维护流程需与企业现有应用管理习惯和审批流程相融合，避免新增操作负担，确保与已有IT治理体系无缝衔接。

2、解决方案

零信任平台通过流程化方式实现应用发布自动化。业务运维人员在ITSM系统中填写应用名称、访问地址、端口、协议等信息，提交后进入审批流程。审批通过后，系统通过API自动将应用配置同步至零信任平台，完成在TAP网关的注册与策略生成，无需人工登录设备进行配置。当应用发生变更或下线时，同样通过流程发起，后端自动更新或注销，前端用户无感知。该机制替代了传统VPN时代手动整理白名单、批量导入的繁琐操作，显著提升运维效率，保障配置准确性，实现应用生命周期管理与安全访问控制的协同联动。

（五）基于业务分级的动态权限调整

1、场景需求

企业内部的不同业务系统根据其敏感度和重要性被分为多个级别，例如客户信息管理系统属于高敏感级。为了确保这些系统的安全性，需要根据业务的分级分类结果，结合用户角色、设备状态、网络环境等多维度属性，动态调整访问权限，以应对不断变化的安全威胁。

2、解决方案

在零信任架构中，针对不同级别的业务系统设置差异化的访问控制策略。对于高敏感级的应用，除了严格的身份认证外，还需满足特定的终端合规性和网络环境要求，并可能触发额外的多因素认证步骤。系统通过实时监控用户的访问行为、设备健康状况以及当前网络环境等因素，与预设的风险评估规则进行匹配。一旦发现异常活动或潜在风险（如异地登录、异常时间段访问等），系统将自动调整访问权限，可能包括临时降权、增加验证步骤甚至直接阻断访问。此外，所有变更均记录在案，便于后续审计和分析。该机制不仅提升了关键业务系统的防护能力，还保证了在正常业务操作下的高效性和灵活性，实现了安全与用户体验的最佳平衡。

（六）企业办公APP切换实践

1、场景需求

为实现移动端安全接入，需将零信任SDK集成至企业APP。由于APP更新依赖用户主动操作，无法强制即时升级，新旧版本将并行运行一段时间。因此，需设计兼容方案，确保新老用户均可正常访问业务，避免因架构调整影响用户体验。

2、解决方案

将企业APP作为零信任移动端接入的首个切换应用。新

版本APP集成零信任SDK后，通过专用内网域名经零信任通道访问后端服务，实现流量加密与身份动态验证。原有老版本APP仍通过原有公网接口继续访问，保持服务连续性。在此过渡期，后端其他系统可同时支持内外网双通道接入。随着其他互联网应用逐步收敛至零信任体系，企业APP新版本已具备完整适配能力，不会受后续网络调整影响。通过设置合理的过渡周期和用户引导策略，实现新老版本共存与平滑迁移，最大限度降低对用户的影响，保障业务稳定运行。

（七）零信任失效时的逃生机制

1、场景需求

虽然已采用了双中心架构来保障零信任的高可用性，但如果因软件故障导致整体不可用时，零信任作为核心访问控制组件，可能影响关键信息系统的正常访问。为保障业务连续性，需建立安全可控的应急逃生机制，确保在极端情况下运维与业务人员仍能访问核心系统。

2、解决方案

设计双路径逃生方案。对于纯七层应用（无需零信任客户端即可访问的应用），预部署一台Nginx作为应急访问节点，仅开启443端口并启用HTTPS加密。通过配置不同server_name区分系统域名，反向代理至纯七层的内网服务。平时Nginx服务关闭，仅在零信任平台全量不可用时，经审批启动服务并切换DNS解析，通知授权人员临时接入。

对于依赖零信任客户端的隧道访问应用，则采用云桌面作为备用通道。授权用户通过预设的隔离云桌面环境接入内网，实现对关键资源的安全访问，确保应急场景下的操作连续性与安全性。

五、方案建设效果

（一）收缩互联网暴露面，安全能力显著提升

基于零信任架构的端口隐藏与业务隔离机制，所有非对客互联网应用系统均通过可信应用代理系统（TAP）进行统一接入，彻底实现业务资源的逻辑隐藏。系统不再直接暴露公网IP与端口，仅允许经过身份认证和终端合规检查的用户通过零信任客户端建立加密隧道，与代理网关通信。这一机制大幅缩减了攻击面，有效防范端口扫描、暴力破解、0Day漏洞利用等外部威胁。同时，依托多维度动态访问控制策略，系统持续对用户身份、设备状态、网络环境、访问行为及应用敏感度进行综合评估，实施最小权限原则。一旦检测到异常登录、终端失陷或越权尝试，可即时触发权限降级、二次认证或会话阻断，实现风险的快速响应与闭环处置，显著提升整体安全防护能力。

（二）业务系统联动，推动灵活办公与数字化转型

本方案深度集成CMDB、OA、ITSM、IAM等企业核心系统，构建安全与业务融合的自动化运营体系。应用发布通过CMDB流程自动同步至零信任平台，权限申请经ITSM审批后策略自动下发，实现“申请—审批—授权—回收”全生命周期管理。用户可通过统一门户完成单点登录，访问已授权的各类资源，支持PC与移动端无缝接入。该模式不仅避免了传统VPN模式下人工配置、信息分散、权限滞后的弊端，还显著降低运维复杂度，提升响应效率。员工可在安全可控的前提下，随时随地高效访问所需系统，真正实现安全与效率的平衡，为远程办公、外包协作、分支机构互联等场景提供有力支撑，助力企业数字化转型纵深推进。

（三）细化访问权限，提高安全颗粒度

零信任架构打破传统“内网即可信”的边界思维，以身份为核心实施细粒度访问控制。针对不同业务系统按敏感等级划分防护策略，对高风险系统设置更严格的访问条件，如限定终端类型、访问时段、地理位置等。权限配置支持按需申请、临时授权与自动回收，避免“永久权限”带来的长期风险。通过终端感知技术，持续采集设备合规性、系统补丁、安全软件等状态信息，结合用户行为分析，实现动态信任评估。访问策略可根据实时风险评分自适应调整，如对高风险请求增加多因素认证或限制操作范围。该机制实现了从“静态授权”向“持续验证”的转变，将最小权限原则贯穿始终，有效防范内部滥用与横向移动，全面提升访问控制的智能化、精细化水平，为证券行业构建可落地、可扩展的新型安全防护体系。

零信任体系在证券业数据安全领域的探索与实践

周喆斌、邬晓磊 | 东方证券股份有限公司

何艺 | 北京持安科技有限公司

摘　要： 证券行业作为金融市场的核心基础设施，承载着海量敏感数据与关键业务系统，其数据安全直接关系到市场稳定与投资者权益。随着数字化转型加速，传统“边界防御”模式已难以应对内外部威胁交织的复杂环境，业务系统暴露面扩大、权限管理粗放、敏感数据泄露等安全风险日益凸显。零信任理念以“永不信任、始终验证、动态授权”为核心，成为重构证券业数据安全防护体系的关键路径。本文结合证券行业数据安全特性，阐述零信任体系的适配性与架构设计，重点分析动态身份认证、细粒度访问控制、敏感数据全生命周期防护等核心技术的实践路径，并通过案例验证其成效，为行业数据安全建设提供参考。

关键字： 零信任、应用网关动态访问控制、数据安全、敏感数据治理

一、引言

证券行业的数字化进程推动了业务模式创新，同时也带来了数据安全风险的集中爆发。一方面，客户身份信息、交易记录、资产数据等敏感信息在多系统间流转，泄露风险显著提升，其数据安全直接关系到金融市场稳定与投资者利益。另一方面，远程办公、第三方合作等场景的普及，使得网络边界逐渐模糊，传统静态规则化的防御逻辑彻底失效。监管层面，《证券期货业网络安全管理办法》、《证券公司网络和信息安全三年提升计划（2023-2025）》、《银行保险机构数据安全管理办法》等法规明确要求强化数据全生命周期安全管控，实现“权限最小化、操作可追溯”， 要求强化数据安全防护能力，实现“可知、可管、可控”。

然而传统基于网络边界的防御模式存在明显短板，办公内网直接暴露业务系统，易遭威胁攻击、权限管理粗放导致越权访问风险、0day/1day漏洞应急响应滞后，难以防范未知威胁。零信任体系以身份为核心，通过动态验证、最小权限、持续审计等机制，重构安全防护架构，成为证券业应对数据安全挑战的必然选择。

零信任理念最早由Forrester Research于2010年提出，其核心思想是打破对网络位置的信任依赖，将身份作为唯一信任基点，通过持续验证与动态授权构建安全边界。在证券行业，零信任体系的价值体现在三个方面，一是解决权限滥用问题，通过细粒度管控确保“数据可见即可控”。二是应对数据滥用挑战，实现数据生命周期的安全管理。三是满足合规要求，通过全链路审计支撑监管溯源。本文基于东方证券探索实践经验，系统阐述零信任体系的落地框架与关键技术，为证券行业提供可复用的建设思路。

二、证券业数据安全挑战分析

证券行业数字化转型加速，数据成为核心资源，应用场景持续扩展，其数据安全风险涵盖多元结构与全业务流程并行威胁，网络层面风险交织、攻防博弈且横向扩散，总体呈现“多维度、复杂化”特征，具体表现为：

1、多场景访问加剧边界防御失效

远程办公、移动展业、第三方运维等复杂场景的普及，使得访问来源从固定内网扩展至任意网络，攻击者可通过钓鱼邮件、恶意软件等方式突破终端防线，进而渗透核心系统。传统防火墙、VPN等边界设备基于静态安全策略，难以识别合法身份下的恶意行为，造成安全盲区导致数据泄露。

2、权限管理粗放导致越权风险突出

长久以来所采用的“一次授权、长期有效”静态权限管理模式，在访问控制机制上存在显著缺陷。一方面权限冗余现象突出，权限分配与职能的动态匹配度不足，权限生命周期缺乏管理，导致权限集合随时间推移呈现无序膨胀态势；另一方面，权限滥用风险加剧，访问控制粒度未实现最小权限原则，权限校验未与实时业务场景深度绑定，使得越权访问与非合规操作具备隐蔽实施的可能性。

3、敏感数据分布分散且流转频繁

证券机构的敏感数据涵盖客户基本信息、交易数据、核心业务数据等，这些数据存储于各系统、交易引擎、数据中台等多个节点，并在内部员工、合作机构、监管部门间高频流转，传统的防护模式难以实现精准管控。

4、合规审计能力不足难以满足监管要求

《网络安全等级保护基本要求（GB/T 22239-2019）》、《证券期货业数据安全管理办法》等法规要求对敏感数据操作进行全程记录与追溯，但部分证券机构的日志系统存在碎片化问题，访问日志、数据操作日志、身份认证日志未能关联分析，导致无法快速定位泄露源头。

表1 证券业数据安全挑战

传统安全表现	安全风险	影响形式
边界防御失效	访问突破传统边界防护	传统防火墙、VPN 等难以识别“合法身份下的恶意行为”，形成安全盲区
权限管理风险	静态授权模式的越权隐患	传统防护模式难以实现精准管控，易导致数据泄露风险
敏感数据管理困境	分布分散且流转频繁	敏感数据跨系统流转时缺乏动态保护，泄露风险高
合规审计能力不足	日志碎片化导致追溯能力弱	不满足《网络安全等级保护基本要求》等法规要求，合规风险高

上述挑战的核心在于信任静态化与边界模糊化，需要针对性解决以身份为唯一信任基点替代网络位置依赖，通过实时风险评估动态调整权限，结合全流程日志关联满足审计需求。利用零信任体系通过动态信任评估、最小权限管控、全链路审计三大特性，实现证券行业数据安全体系化安全防护手段。

三、证券业零信任体系架构设计

基于证券行业业务特性与安全需求，零信任体系采用“三横三纵”架构，实现“访问收敛、动态决策、数据防护”的一体化管控。

（一）横向架构：覆盖访问全生命周期

1、访问层：统一入口与流量收敛

通过零信任应用网关实现业务系统访问入口的集中收敛，所有访问请求均需经过应用网关代理，隐藏核心系统真实IP与端口。应用网关支持HTTPS双向加密与动态端口映射，有效抵御端口扫描与恶意探测。

2、决策层：动态信任评估与策略执行

决策层是零信任体系的核心，由信任引擎、策略引擎、分析引擎三部分组成。信任引擎实时采集用户身份、环境信息等数据，通过加权算法计算信任分值。策略引擎基于信任分值与业务场景预设策略，动态执行允许访问、阻断访问、二次认证、权限降级等动作。分析引擎通过机器学习识别异常访问模式，自动更新策略引擎的管控规则。

3、数据层：敏感数据全生命周期防护

数据层联动业务系统与数据中台，实现敏感数据的识别分类防护审计闭环。敏感数据识别基于正则表达式自动识别敏感字段。动态防护对敏感数据采用加密存储+访问脱敏，未授权用户仅能查看掩码后的数据并采用水印覆盖，下载文件自动嵌入用户身份水印。全生命周期审计记录敏感数据的创建、修改、传输、删除等操作，关联访问日志形成溯源图谱。

（二）纵向支撑：保障体系落地效能

1、身份治理体系

构建统一身份中台，整合业务系统的身份数据，实现一人一账号。同时，建立身份全生命周期管理流程，从账号创建、权限调整到账号注销全程可配置，避免僵尸账号与权限冗余。

2、合规审计体系

基于统一日志中台，实现访问日志、操作日志、安全日志的集中存储与关联分析，支持按用户、系统、时间等维度快速检索，自动生成基于人员用户的审计报告。

3、安全运营体系

通过对访问身份与被访问数据的有效融合，实时监控敏感数据访问动态，动态控制账号越权行为，实现安全风险的早发现、早处置，形成安全运营体系适配业务的动态变化。

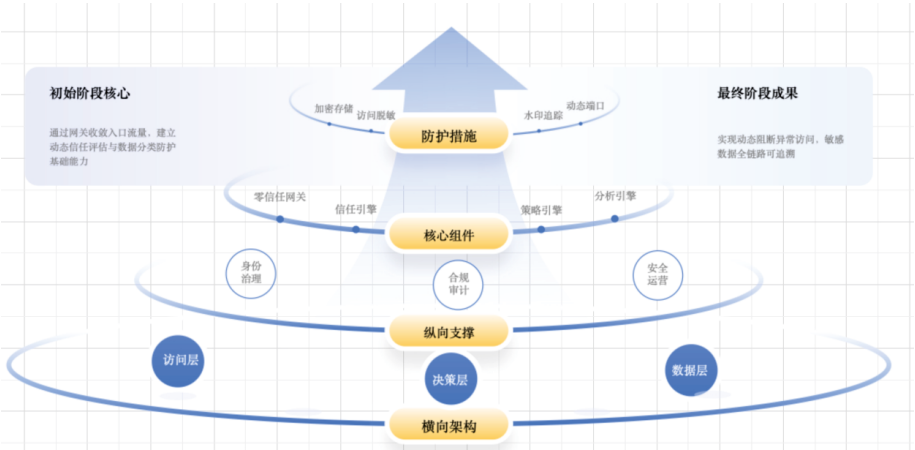


图1 零信任体系架构

四、关键技术实践路径

设计零信任安全防护体系时，需从身份授权角度搭建动态信任评估机制。通过构建“三横三纵”全链条信任体系，对每一层都进行持续验证和动态评估，实现“永不信任，始终验证”核心原则。遵循最小权限原则，各层面的安全功能都按照“最小必要”原则设计，既能有效缩小被攻击的范围，又能解决传统边界防护失效的问题。

结合东方证券的业务访问场景和网络架构特点，采用微服务架构实现数据中心高可靠部署，保障组件冗余和弹性扩展能力。深度整合东方证券内部的身份信息源，搭建以业务系统为核心的集中防御机制和数据保护体系。在用户端与业务端之间部署零信任应用网关，具备动态监测用户身份、按最小权限管控业务访问、识别敏感数据、给业务数据添加动态水印等功能，从而避免端口暴露、违规扫描、恶意攻击等安全风险，以及数据泄露的隐患。

下图是东方证券实际实践时的部署框架，该架构实现“三横三纵”全链条信任体系：业务转发层对应访问层的流量收敛，访问控制层对应决策层的动态评估，身份中心与策略中心对应纵向支撑的身份治理与合规审计体系。

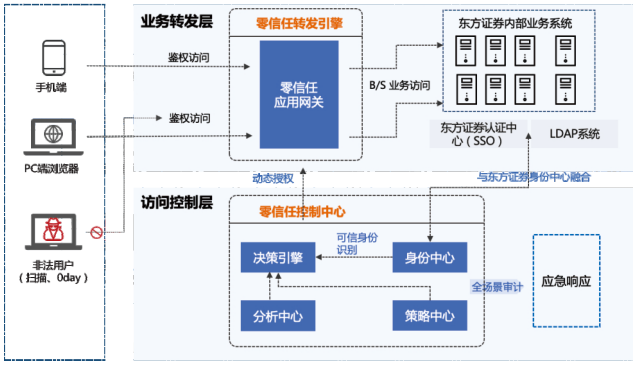


图2 东方证券零信任部署框架

零信任平台采用原生架构设计，基于微服务模式实现组件解耦，通过负载均衡实现高可用架构，确保系统弹性拓展与服务连续性。系统架构涉及的相关组件功能说明如下：

（一）横向架构：覆盖访问全生命周期

访问层：统一入口与流量收敛：零信任应用网关作为统一入口，实现流量集中收敛与访问流程重构。传统模式下，用户采用“先访问，再验证”流程，恶意攻击可直接触达业务系统并利用漏洞发起攻击。通过网关融合后，流程转变为“先验证，再访问”。用户访问业务系统前必须经网关完成可信验证，验证通过后由网关转发请求至业务系统，实现业务隐身，阻断恶意攻击直接访问路径，避免端口暴露与直接攻击风险，成为流量收敛与安全访问的核心屏障。

决策层：动态信任评估与策略执行：零信任应用网关是策略执行核心，承担动态信任评估与访问控制职责。网关通过拦截访问请求，基于访问者的身份权限、所属用户组、认证级别、终端环境等多维度信息进行持续验证，仅对符合条件的请求予以放行，践行“永不信任，始终验证”原则。作为零信任框架的策略执行点，网关将动态评估结果转化为具体管控动作，确保访问行为与信任度匹配，有效防范“合法身份下的恶意行为”，解决传统静态授权的粗放性问题。

数据层：敏感数据全生命周期防护：零信任应用网关联动业务系统与数据流转环节，实现敏感数据全流程防护。应用网关不仅检查来源请求合法性，还对业务系统返回数据进行深度检测，精准识别敏感数据获取行为，如未授权下载、越权查询。同时，基于动态策略提供脱敏对未授权用户展示掩码数据、为下载文件嵌入用户身份标识水印、访问与下载管控等能力，形成识别、防护、审计闭环，保障敏感数据在传输、使用环节的安全，从源头遏制数据泄露风险。

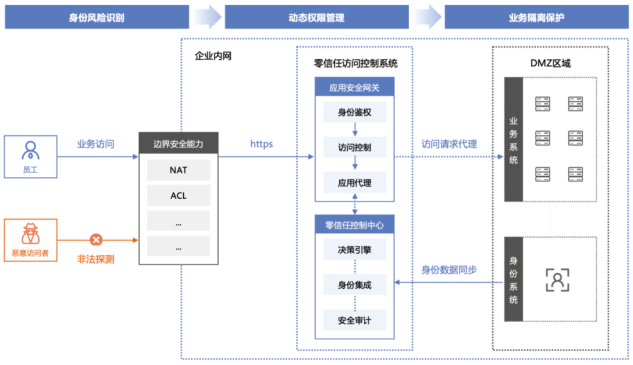


图3 横向架构

（二）纵向支撑：保障体系落地效能

身份治理体系：以精准身份识别为核心，构建全流程身份管控机制。零信任控制决策中心通过身份标签管理，将用户访问日志与身份信息深度绑定，实现访问者身份的快速定位与精准识别。同时，基于用户身份与业务需求，对必要使用系统的用户进行授信访问授权，并动态下发零信任安全策略，避免权限冗余与身份混乱，确保身份与权限的精准匹配。

合规审计体系：依托全量日志采集与关联分析能力，实现访问行为的全程可追溯。零信任框架记录所有访问行为及系统运行日志，不仅包含基础信息，更通过身份标签关联用户身份，按时间、日期、IP 位置、用户身份等多维度进行日志关联分析。结合零信任应用网关的动态校验机制，对用户每一次业务访问进行持续验证，根据风险行为自动执行放行、阻断或告警等动作，满足合规审计的溯源与风险处置要求。

安全运营体系：通过动态决策与威胁防护能力，保障系统安全与业务连续性。零信任安全控制中心整合风险事件、用户身份、访问权限、访问行为、访问敏感数据等因素，利用动态决策引擎进行综合动态评估，并联动零信任应用网关控制业务数据访问，实现安全与业务的动态适配。

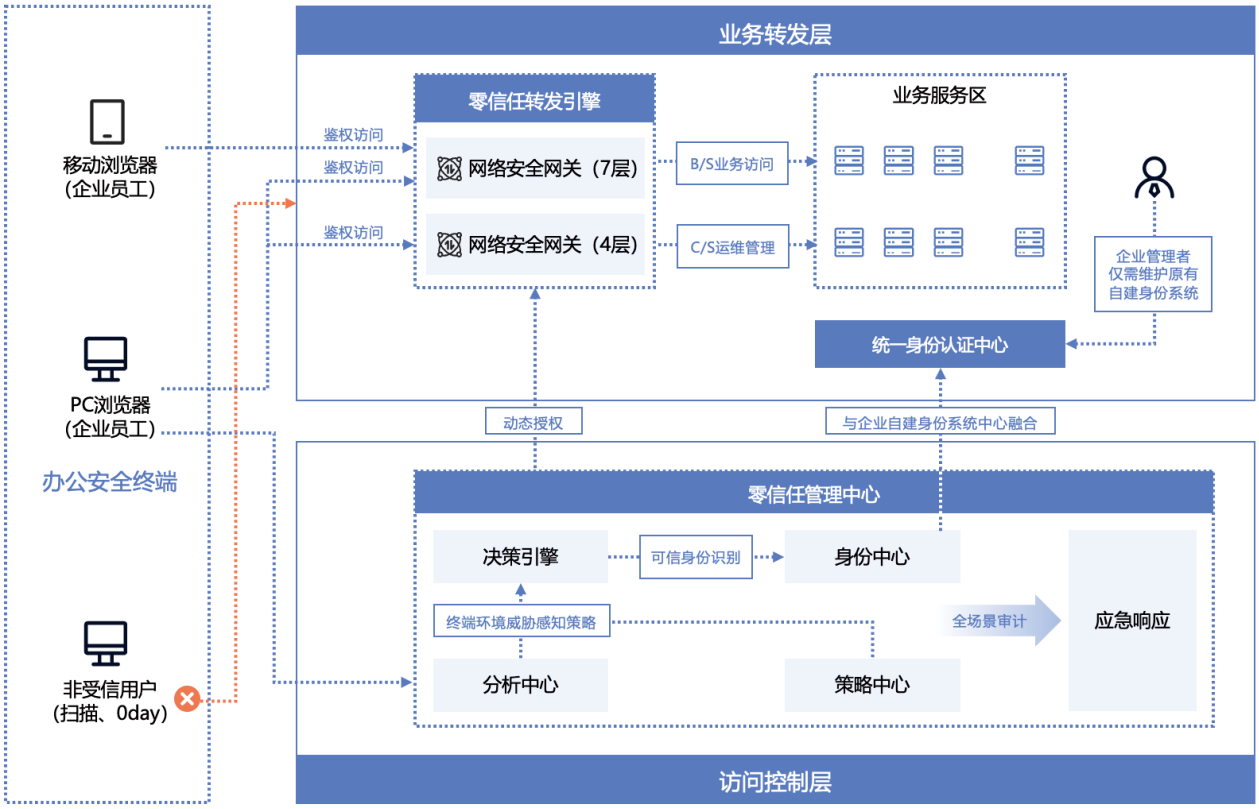


图4 纵向支撑

五、实践过程分析与解决方案

东方证券在零信任体系落地过程中，由于业务系统复杂、合规要求严格、用户场景多样等特点，实施过程面临多项技术挑战。结合东方证券实践经验，针对灰度推广、特殊应用管控、大文件处理、多认证体系适配四大核心难点，阐述解决思路与实践成效。

（一）灰度推广过程中业务可用性的思考

1、痛点分析：在零信任体系灰度推广阶段，全局修改DNS配置将流量引流至零信任应用网关，可能影响未参与灰度测试的用户及业务，违反业务连续性优先原则。此外，DNS服务器作为关键基础设施，其配置修改需周期长、灵活性低，难以满足灰度测试的快速迭代需求。

2、解决思路：采用局部定向引流策略，在不改动全局DNS的前提下，仅对灰度测试用户与业务系统实施流量拦截，实现影响最小化、测试精准化。

3、实践过程：hosts文件定向配置：针对参与灰度测试的PC终端，通过客户端工具推送批处理脚本修改本地hosts文件，将灰度测试人员的测试业务域名解析至零信任应用网关IP，非灰度测试人员以及非测试域名仍沿用原DNS解析路径，实现分级流量管控。零信任应用网关仅对hosts文件修改后的终端流量进行代理，其他终端流量直接放行至业务系统，避免干扰正常业务。

4、实践成果：成功覆盖大范围测试用户，完成了总部全员的推广；同时完成了30余个办公、业务系统的在应用网关的接入发布，均未对正常业务访问产生影响。验证了零信任应用网关动态认证、敏感数据识别功能的稳定性，为全量推广奠定基础。实现终端级精准引流，证明零信任体系对复杂网络环境的适配能力，验证了非侵入式部署模式的可行性，为证券行业边测试、边上线的安全建设提供参考。

（二）互联网访问来源对安全管控思考

1、痛点分析：东方证券在实施过程中存在部分Bypass应用需直接暴露在内部网络中。此类应用若被互联网访问，可能导致未授权渗透，直接封禁其网络访问又会影响内部工作效率，形成安全与效率的矛盾。

2、解决思路：采用双向防御机制，一方面阻断外部非法访问，另一方面限制内部访问范围，确保Bypass应用仅对可信主体开放。

3、实践过程：通过零信任应用网关动态访问策略拦截，在零信任控制中心配置来源IP+应用端口双因子规则，对Bypass应用的访问请求进行实时判定，仅允许内部IP段访问，拦截互联网IP公网地址。同时通过静态配置加固，服务器的防火墙配置文件中，添加互联网IP deny规则，仅允许放行经审批的内网IP，形成动态策略+静态防火墙的双重防护。

4、实践成果：测试发现能够成功阻断来自互联网的试

探性访问，实现动态规则与静态配置的协同防御，证明零信任体系对特殊应用的兼容管控能力。验证了最小权限原则在复杂场景的落地可行性，平衡安全防护与业务连续性。满足《证券期货业网络安全管理办法》中非核心系统需限制互联网暴露的要求，通过监管合规检查。

（三）大文件水印处理的性能优化

1、痛点分析：证券行业存在大量高敏感PDF文件，需通过动态水印实现溯源。在实际使用过程中发现大文件添加水印时会导致内存占用飙升且无法及时释放，可能会引发应用网关服务崩溃，影响业务连续性。

2、解决思路：采用流式处理+内存回收机制，避免一次性加载文件至内存，通过分块处理降低应用网关资源占用。

3、实践过程：采用分块水印算法将PDF文件按内存块拆分，逐块添加水印后即时写入磁盘，完成后合并文件，避免全量加载。同时动态释放应用网关内存资源，通过动态指针管理系统内存，每处理完一块数据立即释放占用资源，避免内存泄漏。

4、实践成果：突破大文件处理性能瓶颈，证明零信任体系对高负载场景的支撑能力。并验证了安全功能与系统性能的平衡设计，为证券行业大文件安全管控提供技术参考。

（四）多认证体系的动态适配

1、痛点分析：东方证券系统架构复杂，不同业务系统采用的单点登录方式存在差异。内部采用企业微信认证、Oauth2.0认证并存。若零信任应用网关仅支持单一认证方式，会导致用户体验割裂。

2、解决思路：构建域名认证方式映射机制，让网关根据访问的应用域名自动选择匹配的SSO方案。

3、实践过程：认证策略矩阵：在零信任策略引擎中配置应用域名认证方式对照矩阵序列，由零信任应用网关解析目标域名并查询矩阵序列，自动调用对应认证接口。

4、实践成果：验证零信任体系对异构认证系统的兼容能力，解决证券行业多系统认证碎片化问题。实现业务场景与安全策略的动态绑定，为复杂组织架构下的身份治理提供技术范式。

东方证券在实践过程中对上述难点的逐一攻克不仅验证了零信任体系性落地的可行性，更形成了一套适配证券行业的问题解决方法论，以业务连续性为前提，通过非侵入式改造、分层防御、性能优化等手段，在满足合规要求的同时，最小化对业务的影响。这为其他证券机构的零信任建设提供了可复用的实践经验。

六、零信任实践成效分析

零信任体系在东方证券业数据安全领域的实践，通过技术创新解决了行业长期存在的安全痛点，支撑了业务合规创新，为数据安全从合规驱动向价值驱动转型提供了可行路径。

（一）从被动防御到主动可控，重塑数据安全防护逻辑

传统边界防御模式下，证券业数据安全长期面临边界模糊后防护失效、权限静态化导致滥用、敏感数据流转不可控等被动局面。零信任体系通过动态验证、最小权限、全程审计的机制，将防护逻辑从依赖网络位置转向以身份为核心的动态管控，实现了三个关键转变。

通过零信任应用网关的流量收敛与“先验证后访问”机制，使核心业务系统与敏感数据从直接暴露转为隐身防护，从源头切断未授权访问路径，解决了远程办公、第三方合作等场景下的边界失效问题。通过动态信任评估与权限按需分配，打破一次授权长期有效的静态模式，让权限与用户身份、业务场景、实时风险动态匹配，从机制上遏制越权访问、权限冗余等内部风险。通过全生命周期防护闭环，实现敏感数据从识别、存储、传输到使用的全程可控，解决了传统模式下数据分布散、流转难管控的痛点。

（二）威胁响应效率显著提升，敏感数据异常行为实时阻断

零信任体系通过动态监控与联动响应机制，实现了对敏感数据异常访问行为的实时发现、精准分析，解决了传统模式下威胁识别滞后、响应流程繁琐的痛点。这一成效直接依托于分析引擎、决策引擎等核心产品组件的协同作用。从体系功能来看，分析引擎持续建模正常访问行为，当出现偏离模型的异常行为时，可实时触发风险预警，决策引擎基于预设策略与实时信任分值，自动判定风险等级并下达处置指令，零信任控制中心则作为中枢，联动零信任应用网关完成指令执行，形成监测分析决策处置的闭环响应链。有效降低了敏感数据因异常访问导致的泄露风险。

依托零信任应用网关组件协同实现的高效响应能力，强化了数据安全防护的动态性与主动性，让证券机构在面对内外部突发威胁时，具备了早发现、早处置的技术支撑，进一步筑牢了敏感数据的安全防线。

（三）满足监管合规刚性要求，降低合规风险

证券业作为强监管领域，《证券期货业网络安全管理办法》、《数据安全管理办法》等法规明确要求“实现权限最小化、操作可追溯、数据全生命周期管控”。零信任体系的实践成效直接呼应了监管核心诉求：

在权限管理层面，通过动态最小权限机制权限最小化要求，避免因权限冗余导致的合规隐患。在审计追溯层面，记录全量日志关联与全链路用户行为数据，满足操作可追溯的监管要求，使敏感数据操作的溯源从碎片化日志难以关联转为精准定位、快速响应。在数据防护层面，敏感数据的加密存储、访问脱敏、水印追踪等技术，满足数据全生命

周期安全管控的合规指标，降低因数据泄露导致的监管处罚风险。

七、总结与展望

证券行业数据具有敏感数据体量大、业务场景复杂、安全与业务深度绑定等特点，零信任体系的实践解决了身份信任的关键因素，更通过技术落地成效的闭环为全行业提供了可参考的建设路径。明确以身份为核心的防护主线，替代传统以网络为边界的思路，适配数字化时代边界模糊的现状，验证动态管控对复杂场景的适配性，证明通过技术手段能够实现风险实时评估、权限动态调整，为证券机构提供技术选型与落地参考。强化数据生命周期防护的必要性，推动行业从单点防护转向体系化管控，提升整体数据安全水位。

零信任体系通过“永不信任、始终验证、动态授权”的原则机制，有效解决了证券业数据安全的核心痛点，为行业数字化转型提供了安全保障。实践表明，该体系不仅能提升风险管控能力，更能支撑合规审计与业务创新，是证券业数据安全建设的必然趋势。

未来，随着人工智能、量子计算等技术的发展，证券业零信任体系将向三个方向演进。一是基于AI的预测式防御，通过机器学习识别潜在风险并提前干预，二是量子加密技术的融合应用，提升身份认证与数据传输的抗破解能力，三是与信创体系的深度适配，实现芯片、操作系统、安全组件的国产化替代，保障金融安全自主可控。

参考文献

[1]全国信息安全标准化技术委员会.信息安全技术网络安全等级保护基本要求（GB/T22239-2019）[S].北京：中国标准出版社，2019.

[2]中国证券监督管理委员会.证券期货业网络安全管理办法[Z].2023.

[3]中国证券业协会.证券公司网络和信息安全三年提升计划（2023-2025）[Z].2023.

[4]ForresterResearch.ZeroTrustArchitecture:The NewSecurityModelfortheDigitalEnterprise[R].2020.

[5]沈昌祥.零信任架构与主动免疫可信计算[J].计算机学报，2022，45（3）：435-450.

[6]持安科技.零信任动态访问控制技术白皮书[R].2024.

[7]国家金融监督管理总局.银行保险机构数据安全管理办法[S].2024.

[8]中国信息安全测评中心.零信任安全能力成熟度模型[R].2023.

[9]王明远，李刚.证券行业零信任体系构建与实践[J].金融电子化，2023（6）：58-60.

04 思考和观点

P68 韧性数字安全体系研究与建设
侯亮

P73 大模型应用场景安全思考与实践
钟蓉、吴佳伟、李鹏、曹杰、温志强、郑煜

韧性数字安全体系研究与建设

侯亮 | 国泰海通证券股份有限公司

摘　要： 韧性数字安全强调“存活能力”而非“绝对安全”；追求“系统抗打击性”而非“攻击零发生率”；构建的是一种具备多层缓冲、动态调节、自主恢复与跨域协作能力的复杂系统安全结构。本文将系统性地论述韧性数字安全体系的理论基础、框架设计与建设、实践与应用、总结与展望等内容。通过理论分析与实践探索相结合，力图为当代信息社会建立一种能够应对不确定性、复杂性与极端事件的新型安全治理范式。

关键字： 韧性数字安全、体系研究与建设、新型安全治理范式、信息安全

一、韧性数字安全体系研究背景

在当代肿瘤医学中，治疗癌症的目标已从“彻底根除”转向“长期控制”，即将其转化为一种可以持续管理的慢性病。这一转变不仅是技术的进步，更是理念的深刻演化。从理念来看，现代医学抗癌至少经历了四次革命：第一次是以手术为核心的局部清除；第二次是通过放疗和化疗实现对全身病灶的打击；第三次是基因检测和靶向药物带来的精准治疗；第四次则是免疫疗法的兴起。尤为重要的是，免疫疗法所倡导的逻辑不同于传统的“直接消灭病灶”，它通过激活人体自身的免疫系统，使其具备识别与压制癌细胞的能力，从而形成一种系统性、内生性的防御机制。

这一医疗逻辑的深层转变，恰可比拟当前网络安全治理中正在发生的范式转型。随着数字技术在社会各个层面的广泛渗透，网络攻击的复杂性、隐蔽性和破坏性不断增强，传统的信息安全观念——如边界防御、边界模糊、静态隔离、威胁封堵——越来越难以应对不断演化的威胁。大规模勒索病毒、APT攻击、数字供应链污染攻击、物联网滥用等新型风险不断突破防御系统的边界，也使得边界越其模糊，而系统一旦崩溃，往往面临数据丧失、关键功能瘫痪、组织运营中断等灾难性后果。

在此背景下，网络安全的战略目标正从“防止被攻破”向“允许在被打击后仍能生存”发生根本转变，韧性数字安全体系应运而生。该体系不再一味追求系统的“不可穿透性”，而是强调系统在面对威胁和攻击时，能否快速识别、精准定位、有效隔离，并在最短时间内恢复关键业务运行，保持整体系统的稳定性与业务连续性。

韧性数字安全强调“存活能力”而非“绝对安全”；追求“系统抗打击性”而非“攻击零发生率”；构建的是一种具备多层缓冲、动态调节、自主恢复与跨域协作能力的复杂系统安全结构。

本文将系统性地论述韧性数字安全体系的理论基础、框架设计与建设、实践与应用、总结与展望等内容。通过理论分析与实践探索相结合，力图为当代信息社会建立一种能够应对不确定性、复杂性与极端事件的新型安全治理范式。

二、韧性数字安全体系理论基础

（一）韧性概念的源起与演化

“韧性”最早源于物理学领域，指材料在受力变形后恢复原状的能力。在20世纪后期，该概念被引入生态学、社会学和工程系统中，逐步发展出一套系统性韧性的分析框架。在网络和信息安全领域，“韧性”不再仅仅意味着恢复原状，而是指系统在遭遇攻击、故障、异常或突发事件时，能够维持其核心功能、快速适应扰动并逐步恢复能力的综合性能。

与传统的“信息安全”相比，网络韧性更强调的是应对失败的能力，而非单纯避免失败。信息安全的核心在于保密性、完整性和可用性，而韧性则将焦点转向了系统性的持续运行能力、功能退化后的承载力、以及资源调度与恢复策略的动态性。

因此，可以说，“韧性”并非替代“安全”，而是在现实威胁环境中对安全范式的一种必要补充和再定义。这也构成了韧性数字安全体系的理论根基：接受风险存在，重构系统设计。

（二）从“防御性安全”到“生存性安全”

传统网络安全强调边界设防、防火墙阻断、规则匹配与黑白名单。其逻辑基础是“攻击可以被预先阻断”，前提是“我们知道攻击来自哪里”。但随着威胁源高度多元化、攻击手法动态演化，防御策略逐渐陷入被动：一旦攻击手段绕过设定规则，系统几乎毫无抵抗能力。

韧性安全体系强调的是另一种逻辑：攻击不可避免，瘫痪无法彻底防止，但“生存”能力可以构建、可以优化、可以模拟测试。它反映了如下几种思维转向：

（1）从完美防护至允许受损但不崩溃；

(2)从静态规则至动态适应机制；

(3)从单点边界防御至系统性协同韧性；

(4)从攻击阻断至恢复保障与重建能力。

以此为核心，韧性体系不追求“绝对安全”，而是追求“相对生存能力最大化”，即在“已被攻击”的设定下，系统还能维持服务、限制蔓延、重建功能。

(三)韧性体系的四大理论支柱

在文献和实践中，成熟的韧性体系往往由以下四大理论支柱构成：

(1)冗余性——系统必须具备功能冗余与路径冗余。

(2)适应性——在不确定条件下，系统需具备根据环境变化自动调整资源和行为的能力。

(3)恢复性——即使在功能崩溃或被攻击的情况下，系统应具备以最快速度恢复核心服务的能力。它要求恢复路径明确、指令通畅、数据完整。

(4)可感知性——系统必须对内部状态与外部环境有持续、准确的感知能力。这包括对攻击、异常、流量变化等的实时检测与研判。

这四大理论支柱，相互配合、共同支撑一个系统“在混乱中仍能运行”的底层逻辑。本文提出的“分层防护”、“层层发现”、“主动防御”、“跨行跨业联动”、“极限生存”五大建设架构，正是对这四大原理的具体化、结构化与操作化落地。

(四)韧性安全体系构建的必要性

在现实网络环境中，任何足够复杂的系统都不可能保持永久的“完美运行”。无论防御策略多么精密、策略规则多么严密，总有一种攻击路径、操作失误或供应链缺陷能够突破设防。这种“不可避免的失败”，并非偶然，而是由系统的本质决定的。

从工程视角来看，现代数字基础设施正面临三大困境：

(1)系统边界不再清晰：随着云计算、物联网、移动终端的大规模部署，网络系统的边界趋于模糊，防御线越来越难以定义。

(2)攻击手段日益复杂：攻击者不再依赖单一手段，而是采取“低慢隐”方式，绕过规则检测，穿透多个安全层面。

(3)组织协作链条冗长：安全事件往往涉及多个部门、外部供应商、上下游合作单位，导致响应链条变长，决策滞后。

在这种条件下，继续寄希望于“零入侵”“零出错”是不现实的。安全系统不能只追求封闭式防御，而应像生命系统一样具备在失败后“控制损害、限制扩散、迅速恢复”的能力。这正是韧性数字安全体系建设的逻辑起点。

韧性数字安全建设不是抽象概念，而是对以下三个维度的具体“工程组织”：

(1)结构上的容错设计；

(2)机制上的恢复路径预设；

(3)快速响应链条。

更进一步说，韧性数字安全体系不是对现有安全架构的修补，而是对系统运行逻辑的整体再设计。它改变的不只是“用了什么工具”，而是“如何理解系统如何生存”。

当我们不再把安全理解为“屏障的坚固程度”与“业务的保驾护航”，而是“在攻击中系统能否站稳”的问题时，韧性便不再是理想主义附加项，而是数字安全的底线逻辑与工程相揉的和谐。

三、韧性数字安全体系框架设计与建设

在传统安全范式下，防护体系往往以边界设防、点状拦截为主，假设只要“前端不破”，整体系统便可安然无恙。但在现实中，复杂系统始终存在不可预测的漏洞与人为误差，攻击也逐渐呈现“多点突破、链条演进、隐蔽持久”的态势。在这一背景下，韧性数字安全体系必须构建一套多维联动、动态协同的结构机制，其核心即“五重架构设计”。

(一)分层防护：结构解耦，避免单点失效

韧性安全的第一原则是分层防护。该原则要求根据资产的关键程度、业务的耦合关系及潜在攻击面，对整个系统进行逻辑与物理上的分层与分区。在结构上形成“内外有别、等级分明、职责清晰”的防御纵深，在策略上制定各层独立响应与联动机制。其关键特征有：

(1)构建从互联网网络、互联网应用、互联网用户交互、互联网服务器、内网网络、内网应用、内网用户交互、主机、应用、数据等到身份的多重防线；

(2)各层独立部署安全策略，避免“一处突破，全局瘫痪”；

(3)引入微隔离技术，将关键服务模块细分成独立单元。

此类架构设计通过“结构解耦”与“功能最小化”，在提高攻击成本的同时，为系统提供了容错冗余基础。

(二)层层发现：连续感知，动态诊断

防护不能停留在封堵层面，更应强调实时发现与持续感知。“层层发现”要求在各个系统层级部署可持续运行的监测与检测机制，实现从外围异常行为感知，到核心数据访问分析的全景安全可视化。其关键特征有：

(1)建立覆盖全域的日志、告警、行为分析机制；

(2)采用人工智能与威胁情报驱动的智能分析模型；

(3)支持对未知攻击手法的“行为态势识别”与溯源能力。

系统必须像免疫系统一样，具备多层感知神经网络，不仅能“看到”攻击，更能“理解”其结构与走向。

(三)主动防御：预判攻击，打断链路

相较于被动响应，韧性安全更强调前置性防御，即在攻

击成功前实施打断与干预。这一策略不仅仅是“加强规则”，而是以对攻击链的认知为基础，通过构造“对抗机制”实现攻击链条的解构。其关键特征有：

(1)通过威胁建模与攻击图谱，识别潜在攻击路径；

(2)主动部署蜜罐、诱导系统与陷阱机制，误导攻击行为；

(3)实现对攻击早期阶段(如侦察、权限提升)的精准打击；

(4)构建主动防护运营。

主动防御的实质，是将攻击者置于不确定与受控状态，使其在入侵过程中不断暴露、消耗、误判，增加其攻击成本。

(四)跨行跨业联动：信息共享，共建集体免疫

韧性体系的第四个核心在于跨界协同能力。网络攻击往往呈现跨组织、跨行业的传导特性，孤立防守注定难以形成有效应对。因此必须建立一套跨主体、跨领域的情报共享与联动响应机制。其关键特征有：

(1)构建基于信任机制的信息共享平台，推动威胁情报标准化；

(2)建立跨域应急响应演练机制，提高集体应急效率；

(3)支持“单点受损、群体防御”的协同联动体系；

(4)威胁情报数据流转去敏化。

只有在整体生态层面形成协同作战体系，单一系统的韧性能力才能转化为社会级的系统性韧性。

(五)跨行跨业联动：信息共享，共建集体免疫

当攻击不可避免、崩溃难以阻止时，系统必须具备在“最坏情境”下依然能够维持最小运行能力的结构设计。极限生存并非妥协，而是一种“极端环境下的主动求生”。其关键特征有：

(1)明确“最小运行单元”，设计“最小功能集”；

(2)建立应急恢复机制，如离线切换、冷备自动接管、可信重建；

(3)引入“信任根”的快速再构能力，恢复系统可信基础。

这种生存逻辑类似医学上的“维生系统”：即便大面积组织功能丧失，仍通过核心机制维持生命体的最基本运行状态。

四、韧性数字安全体系实践与应用

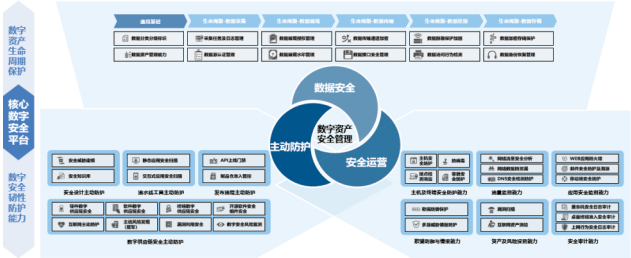


图1 韧性数字安全体系建设应用实例

(一)从结构原则迈向系统路径

韧性数字安全体系的构建，不是传统安全范式的延伸或修补，而是对“安全”这一概念本身的重塑与重构。前文所提出的五重架构设计——分层防护、层层发现、主动防御、跨行跨业联动、极限生存——在本质上指向的是系统自身组织形式、运行逻辑、环境适应力与危机承载力的全面跃迁。这五重架构设计构成一个互为支撑的有机整体，表现出高度的结构嵌合性、功能互补性与逻辑递进性。

(1)“分层防护”确立了韧性体系的结构性防御骨架，通过空间异构、职责分层与模块分区，打破了传统线性安全模型中的“单点依赖”与“平面脆弱性”，为系统建立了首道缓冲机制。这一原则奠定了体系韧性的“空间分布基础”。

(2)“层层发现”是“分层防护”基础上的动态感知机制延展，通过多级监测、异常行为建模与分布式响应，将安全能力嵌入系统的“神经末梢”，使系统具备持续感知、实时警觉与自我认知的能力。此处体现的是“从被动知觉到主动感知”的逻辑飞跃。

(3)“主动防御”在“感知能力”的基础上进一步引入了预判—干预—反制的动态安全逻辑，摆脱传统“事后响应”的应激模型，转而构建出具备前向识别与超前决策能力的攻防主动性。这一原则思想体现了韧性安全中“时间维度上的前瞻性演化”。

(4)“跨行跨业联动”则突破了个体系统的封闭边界，倡导构建面向多方协作、信息互通、联防共治的生态协同机制。这一架构设计基于现实中的复杂威胁演化趋势，回应了“没有一个系统是孤岛”的时代挑战，强调在数字共同体中塑造“集体韧性”。

(5)“极限生存”则是韧性体系的底层哲学支柱。它直面最极端的系统灾变场景，预设“系统不可避免会失败”，从而构建出灾难条件下维持关键功能、优先恢复路径与再信任重构机制的体系能力。这是从“保障不中断”向“保障不崩溃”的范式跃迁，是韧性理念区别于传统安全观的最核心断裂点。

五重架构设计，在逻辑上并非线性堆砌，而是构成一个动态闭环系统：

(1)分层防护创造结构隔离基础；

(2)层层发现植入动态认知感知；

(3)主动防御提升系统反应智力；

(4)联动协作扩大安全协同维度；

(5)极限生存则锚定系统生存底线。

它们共同织构出一个具备自组织能力、自我调节机制与非线性应对能力的韧性系统生态。这套体系不再诉诸“零风险幻想”，而是转向对“不确定性”的正视与制度化应对，其目标不是绝对防御，而是确保在任何攻击中“活下去”、“恢复来”、“适应变”。

在永恒的不确定性中建构秩序，是韧性安全体系的本质使命。“世界的本质是运动、变化和发展的”，而非静态和

永恒不变。传统安全体系的设想本质上是一种“静态的确定性控制”逻辑，它追求一种绝对封闭、永不出错、完全可控的秩序幻象。然而现实世界的技术系统深陷复杂性、相互依赖性与不断演化的对抗之中，任何单点的失败都可能迅速演化为系统性的瓦解。

韧性安全体系正是在这一历史条件下应运而生，它并不以消灭风险为目标，而是接纳不确定性、承认失败的可能性，并在这种不确定中寻求有组织的生存、有节奏的恢复、有方向的演化。它既是一种安全工程逻辑，更是一种辩证法的实践形式——在对抗之中发现秩序，在失败之后重构信任，在混乱内部塑造系统性稳定。

因此，韧性安全不只是“更复杂的防御”，它是数字世界中主动生存哲学的技术呈现，是从“安全神话”走向“动态现实”的范式革命。

（二）从理念到落地：韧性数字安全的工程化建设路径

如果说前述“五重架构设计”是韧性安全体系的逻辑骨架，那么如何将这一骨架转化为具备结构稳定性与动态适应性的现实系统，便是“工程化”建设必须回应的命题。尽管“韧性”作为安全治理的新范式已在理念上逐渐获得共识，但将这一理念转化为可执行、可验证、可持续演进的工程实践，仍是当下数字安全领域面临的重点难题。韧性安全体系并不等价于传统意义上的“加强防护”或“构建备份”，它要求在系统架构、组织运维、数据策略、风险处置机制等多个维度实现范式转换。这一转换不是简单的工具迭代，而是一种对系统性认知的重构，是对复杂性与不确定性主动承认并系统回应的工程建设路径。

1、工程化建设的三层结构：架构、机制与能力

韧性安全工程建设的核心在于从静态防御的结构逻辑走向动态调节的能力建构，其建设落地可划分为三层：

（1）结构架构层：设计上引入冗余、解耦、自治等工程原则，以构建具备“退化运行”与“渐进恢复”能力的基础架构。

（2）机制转化层：将理念层的“主动防御”“动态适应”“威胁共存”具体转译为触发式响应、行为感知、策略演化等机制流程。

（3）能力塑造层：通过训练、评估、演练及指标体系建设，使组织具备识别—吸收—恢复—学习的全过程韧性闭环能力。

韧性数字安全工程化不仅涉及“搭什么系统”，更重要的是“如何让系统不断适应”“如何让人组织持续学习”。

2、核心技术支撑的韧性转化路径

从工程化建设角度，韧性数字安全的建设落地依赖于多种关键技术的交互融合：

（1）动态资产可视化与拓扑映射：为“精准吸收”风险提供数据基础。

（2）基于行为的入侵识别与因果链重建：支持“过程理解”与“自我解释”。

（3）数字孪生环境下的灾害演练与模拟恢复：增强“演化式学习”。

（4）模块化重构与微服务弹性设计：支持“局部失败—系统存活”。

（5）零信任与最小权限模型的动态授权框架：实现“纵深控制”的内嵌化。

这些技术并非孤立堆砌，而应纳入系统性工程设计逻辑中，形成“技术—机制—能力”的闭环结构。

3、韧性安全建设的实践与应用

（1）制度层面：构建韧性安全的战略共识与治理机制

顶层设计：将韧性安全纳入组织战略全局，明确其核心地位。以制度创新推动治理逻辑转型，从“唯合规论”向“合规、动态适应、持续改进”的三轮驱动。

动态政策机制：建立弹性法规与安全标准，允许根据威胁环境和技术发展灵活调整，打破“一刀切”的刚性条框。

容错激励机制：设计容错机制和失败容忍度，减少惩罚性文化带来的创新障碍，激发组织内创新动力。

（2）组织文化与认知：培养韧性思维，推动理念内化

安全意识宣贯：开展系统的韧性安全培训，增强全员对韧性理念的理解和认同，塑造积极面对风险与失败的心理态度。

决策示范：管理者以身作则，推动从“零风险幻想”到“动态适应现实”的思维转变。

跨部门协同文化：建立跨部门沟通机制，促进信息共享和资源整合，打破“信息孤岛”，构建协同共治生态。

（3）技术与工程实践：打造韧性安全的技术体系和运维机制

渐进式技术迭代：采用模块化、可插拔的设计理念，支持系统逐步演化与升级，减少大规模改造的风险与成本。

持续监测与反馈：构建全链路、多维度的安全监测体系，结合人工智能和自动化技术，实现对威胁的动态感知与响应。

演练与实战：通过有效性验证、红蓝对抗和灾备测试等实践活动，检验韧性措施的有效性，推动安全能力的闭环提升。

（4）资源配置与激励机制：保障韧性建设的持续动力

长期主义：转变“短平快”观念，从局部收益转向全局韧性安全的长期价值主义，合理安排有限的资源。

绩效考核创新：设计与韧性目标相匹配的绩效指标，强调系统恢复力、业务连续性和风险管理能力的提升。

激励多元化：引入技术创新奖励、跨部门协作表彰等多样化激励手段，提升组织整体的韧性建设积极性。

（5）制度与实践的动态协同：建立韧性安全的持续进化机制

反馈闭环机制：形成制度设计—实践应用—效果评

估—制度优化的循环体系，确保韧性安全理念与实践的同步演进。

知识管理与经验积累：搭建韧性安全知识库和案例库，实现组织经验的沉淀与共享，避免重复错误，促进创新传承。

开放协同生态：推动与行业、学术界、监管和供应商伙伴的多方协作，构建共生共赢的安全生态圈。

以上策略从理论到实践，从制度到文化，从技术到管理，全方位破解“制度惰性”与“实践创新”之间的矛盾，实现韧性数字安全工程转化的质变。它们共同构成一个动态辩证的系统工程，推动组织在复杂多变的数字环境中稳健前行。

五、韧性数字安全体系总结与展望

（一）韧性数字安全体系出发点：在不确定性数字世界中构建可能的秩序

所有关于韧性的讨论，归根结底源于对世界本体的不确定性的承认。在这个意义上，韧性不是技术术语，而是存在论问题。它并不试图消除风险，而是承认风险常在，从而转向构建一个可以承受冲击、吸收扰动、并在扰动中保持连续性的系统性存在。

这种理解与现代科学技术在面对“复杂性”“模糊性”“非线性”时的范式转变高度一致：从确定论向演化论，从线性控制向动态调适，从封闭系统向开放系统。正如海德格尔的思想：“人并非主宰自然的主宰，而是驻留于存在的风暴之中。”韧性正是这种“驻留”姿态在数字安全世界中的具体实践。

（二）思想转译为系统建构：韧性体系的双向逻辑

思考的力量不在于提供工具，而在于提供思路。在将“不确定性中的秩序”转译为工程系统时，韧性体系体现出两个维度的逻辑：

（1）向下扎根：在底层架构中承认“脆弱性”是必然的，从而构建结构的缓冲、模块的替代性与联动的弹性；

（2）向上生长：在治理机制中接受“未知”与“突变”，发展出组织的自适应能力、人的反思能力、系统的学习能力。

这种上下互动的逻辑，构成韧性数字安全体系的“生成性机制”。

（三）理论与实践的张力：从理想到工程转化的实践矛盾

在将思想转化为工程现实过程中，韧性体系不可避免地遭遇诸多张力与悖论：

（1）预设与适应的矛盾：制度设计需要预设结构，但真正的韧性又要求能够脱离结构进行自由调整。

（2）稳定与变动的矛盾：安全往往追求控制与边界，但韧性必须承认模糊与渗透。

（3）控制与自治的矛盾：传统治理逻辑依赖中心化管理，而韧性更需要边缘智能与本地决策。

这不仅是实践困境，更是必须正面回答的命题。系统的生长需要克服自身的惯性，而这依赖于制度性自省机制的构建的能力。

（四）最终目标：系统的自觉、自省与自治

韧性体系的最终目标，不是通过外部工具加固系统，而是形成系统自身的认知能力与反思能力。也就是说：

（1）自觉：系统能够认识自身状态与演化趋势；

（2）自省：系统能够判断自身失效或惯性来源；

（3）自治：系统能够在最小外部干预下完成修复与重组。

这是一种“具有意识的系统”韧性安全体系建设雏形。它要求我们将安全从外部防御逻辑，转化为内部演化逻辑——不是构筑更高的墙，而是构筑更强的生态。

最后，我想用一句话以此总结——韧性，是人类对数字世界的敬畏，是系统对自身有限性的再认识，是在不确定性中持续创造秩序的内在力量。

大模型应用场景安全思考与实践

钟蓉、吴佳伟、李鹏、曹杰、温志强、郑煜 | 兴业证券股份有限公司

摘　要： 随着Deepseek、Qwen等开源大模型的发展，AI逐渐能被用来赋能证券期货业的业务场景，但与此同时大量从模型层到应用层开源工具的引入也给企业的安全防护体系带来新的挑战。兴业证券技术团队目前已借助大模型赋能多个内部业务场景，为了防范和消解新技术带来的风险，安全团队开始着手探索大模型的安全治理。

关键字： 大模型安全治理、安全开发、风险建模

一、大模型在金融单位常见落地场景及攻击面分析

依据《2025年中国证券业大模型应用跟踪报告》，大模型应用被广泛地使用在证券各个业务和技术部门。常见的落地形态有知识问答场景类、代码助手类及通用AI智能体类等。

（一）知识问答类

知识问答类应用主要采用RAG类技术。由于通用大模型不具备企业内或相关专业领域的知识，因此需要RAG类技术。该类应用十分常见，被广泛应用于财务、立项、开发、测试、运维、合规等场景用于规范解答等。一个RAG应用常见的架构如下所示，首先构造应用时需要先将企业内部的知识库文档切分（Chunk）后形成一系列高维向量(Embedding)，并存储在向量数据库里，然后用户登录后将用户的问题向量化，大模型优先去查询向量库里的回答给用户以提升模型回复的准确性。

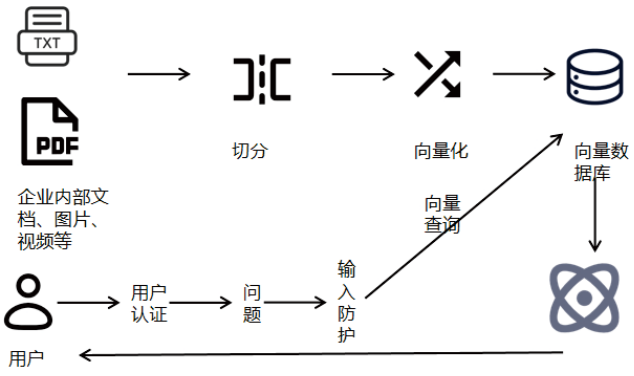


图1 知识问答类应用架构

知识问答类场景引入的新攻击面主要包含基础模型、向量数据库、开源RAG框架等。

表1 知识问答类应用架构

引入的新增攻击面	简介	案例	攻击方式
RAG框架	将信息检索（IR）与大型语言模型（LLM）的文本生成能力相结合的人工智能框架。当 LLM 需要回答一个问题或生成文本时，不是仅依赖其内部训练时学到的知识，而是先从一个外部知识库中检索出相关的信息片段，然后将这些检索到的信息与原始问题/指令一起提供给 LLM。	Haystack RAGFlow FastGPT RAG-AnythingLang Chain	a. 开源软件漏洞，多为各类未授权漏洞 b. 权限管理问题
向量数据库	向量数据库是一种专门用于存储、管理、查询、检索向量的数据库，可以把非结构化数据 embedding 为多维向量值，能够有效帮助 LLM 应用开发人员实现对非结构化数据的有效管理，提升 LLM 对非结构化数据内容的认知。向量数据库常被用于 LLM 应用训练、推理和知识库构建等场景。	OceanBase Milvus Qdrant Weaviate Chroma	向量数据库原生API 数据泄露风险
大模型	知识库场景下一般使用本地部署的模型		隐私泄露、鲁棒性缺陷（对抗攻击）、伦理偏见和合规风险

（二）代码助手类

代码助手类应用是目前较为成熟的大模型应用，一般支持编码辅助、脚本执行、配置MCP服务等功能。目前证券公司多以外购软件为主，形态包含传统IDEA插件或单独的IDE，一般为了数据安全考虑配合本地部署的模型使用。一个代码助手类应用常见的架构如下所示，用户通过安装客户端或客户端插件，输入开发需求，客户端调用大模型完成编码或代码补全，客户端也可配置远端MCP服务，实现CI/CD。

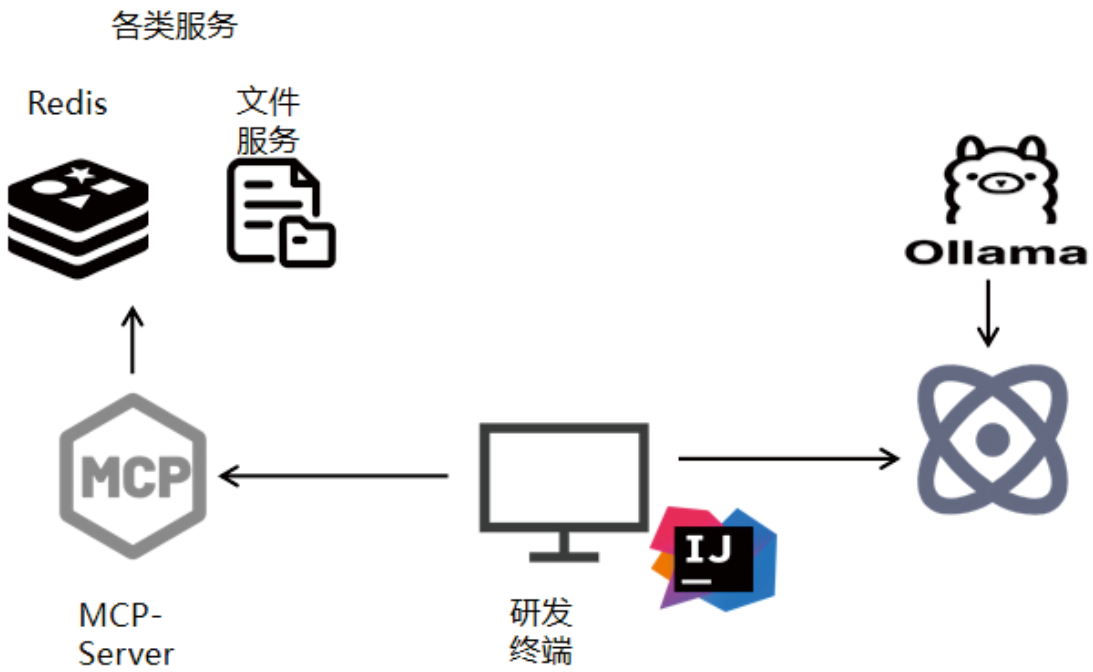


图2 代码助手类应用架构

代码助手的攻击面包含编码助手PC端客户端、服务端及MCP服务等，带来的安全风险包含但不限于数据泄露、客户端RCE、MCP服务投毒、秘钥失窃等。目前我司实践为自研IDEA/Vscode客户端插件--兴智研Copilot。该插件当前已提供了常用工程脚手架、数据库访问层代码生成、POJO相互转换、接口便捷定位等能力。通过集成本地部署的开源AI编码大模型，可根据当前代码文件及跨文件的上下文，自动生成行级/函数级代码、单元测试、代码注释等，此外还具备代码解释、智能研发问答、异常报错排查等能力。

表2 代码助手类应用主要新增攻击面

引入的新增攻击面	简介	案例	可能的攻击方式
PC 客户端	本地编码的开发 IDE	Cursor 腾讯 CodeBuddy 字节 Trae 阿里云的通义灵码	客户端 RCE
模型	代码助手场景下一般使用本地部署的模型		幻觉（生成不可靠内容）、隐私泄露、鲁棒性缺陷（对抗攻击）、版权争议、伦理偏见和合规风险
配置外部 MCP 服务	大模型调用工具服务通用 MCP 协议，可将传统的文件访问服务、数据库访问、接口访问都按照 MCP 通用协议改造以便 LLM 调用，目前市场上已将较多工具 MCP 化，也有金融服务商将自己的服务 MCP 化	大模型调用 excel 有 @negokaz/excel-mcp-server 调用 redis 有 @redis/mcp-redis 调用 kali 有 kali-mcp	MCP 投毒 RAG PULL 毯子攻击等

（三）业务AI智能体类

单体AI智能体类应用包含四个核心组件，推理和规划、工具、记忆以及检索和知识。AI智能体被广泛应用在证券业投顾、客服、营销、合规等场景。一个常见的单体Agent架构如下所示，用户完成基础认证后输入问题，前台调用智能体的功能进行理解查询、规划步骤、执行工具（如搜索）以及综合信息生成最终答案。

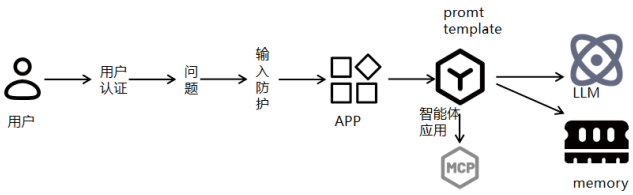


图3 单体智能体架构

构建一个单体AI智能体需要引入多项技术，涉及到多类开源框架。如果是复杂业务场景，可能还涉及到多Agent协同及调度。主要涉及到的攻击面如下表。

表3 业务类智能体主要新增攻击面

引入的新增攻击面	简介	案例	可能的攻击方式
应用	一般形态有插入原始业务网页、移动端 app 或 PC 端渠道的对话类机器人	原本的业务 app/web 等	web 攻击和移动端攻击等
大模型	依据业务的敏感度，可选择外部接口类、本地部署或自训小模型	阿里/腾讯等云厂商都提供大模型 API 服务	1.幻觉（生成不可靠内容） 2.隐私泄露 3.鲁棒性缺陷（对抗攻击） 4.版权争议 5.伦理偏见和合规风险等

引入的新增攻击面	简介	案例	可能的攻击方式
多模态	支持对于图像、视频、语音等格式输入的解析	Wisper StyleTTS2	1. 利用音视频文件投毒
工具/MCP服务	将传统业务接口 MCP 化改造，目前 java、python 等各语言都提供了接口转 mcp 服务的 sdk	Java: AI sdk Javascript: Python:	1.权限类攻击 2.开源组件漏洞 未授权服务 3.间接提示词注入
记忆	大模型的记忆可以分为短期记忆和长期记忆。基于对话历史为短期记忆，对用户输入次数有限制，引入记忆模块可以增强智能体的记忆功能	Mem0 MemoryOS Graphiti	1.上下文数据操控 2.数据泄露、丢失、损坏
开发框架	开发 AI 智能体的平台或框架，部分是语言驱动，如 Eino；部分是可视化拖拽，如 Dify。一般平台还支持 Agent 编排和大模型应用全生命周期管理等功能。	LangGraph Langchain Eino AutoGen Coze Dify n8n	1.web 漏洞 2.提示模板泄露 3.不安全的工具使用和代理功能
运维框架	对 LLM 应用的全链路追踪、调试、质量评估和持续优化的工具框架	Langfuse	1.web 漏洞

二、大模型应用风险治理

大模型应用带来了很多新增攻击面，对传统安全方向如开发安全、安全运营和数据安全都带来巨大的挑战。传统的应用开发流程为评估与设计、编码、测试及上线后运营四个阶段，大模型应用的建设分为设计、模型预训练（部分应用涉及）或者基模选择、应用开发、应用测试及应用上架等过程。

（一）威胁建模与安全设计

在安全需求评估与设计环节需要重构安全设计需求问卷，并针对AI应用的风险进行风险建模，该阶段有部分厂商或开源组织推出了AI应用的风险建模框架，如谷歌发布的针对整个人工智能开发全生命周期的SAIF框架（Secure AI Framework）、CSA发布的针对智能体的安全框架MAESTRO（Multi-Agent Environment, Security, Threat, Risk, and Outcome）、生成式AI攻击矩阵、NIST的AI风险管理框架和Databricks AI安全框架等。国内百度、腾讯也发布了对应的针对大模型的安全框架。以MAESTRO框架为例，MAESTRO框架在STRIDE、PASTA等传统风险建模框架的基础上进行了AI扩展，首字母分别代表着Multi-Agent Environment（多Agent环境），Security（安全），Threat（威胁），Risk（风险），and Outcome（成果）。MAESTRO 是围绕 Ken Huang 描述的七层参考架构构建，每一层都有各自的风险，同时也存在跨层的风险。使用该方法做风险评估，首先需要依据七层架构将系统分解为各层组件，然后依据特定层次识别风险并分析跨层之间的交互，识别跨层威胁，最后依据威胁制定缓解措施。依据框架和各层级具体攻击面分析完成风险建模后，可以参考云厂商各层级安全实践出具安全设计落地方案。MAESTRO框架要求为以下每一层的特定威胁制定缓解措施，对跨层的通信进行监控，构建多层安全防护架构。

表4 MAESTRO框架的分层风险

层次	层次名称	描述	威胁
7	智能体生态	智能体与应用程序和用户交互	受损智能体、智能体模仿、智能体身份攻击、集成风险、水平/垂直解决方案漏洞等
6	安全和合规性	这一垂直层贯穿所有其他层级，确保将安全性和合规性控制集成到所有 AI 智能体操作中	安全智能体中毒、智能体不遵守法规、智能体缺乏可解释性、智能体的偏见等
5	评估和可观察性	评估和监控人工智能智能体，包括跟踪性能和检测异常的工具和流程	可观测性泄露数据、受损的可观测数据、操控评估指标等
4	部署和基础设施	AI 智能体运行的基础设施（例如，云、本地）	受损的容器镜像、编排攻击、基础设施即代码操控、资源劫持、横向移动等
3	智能体框架	包含用于构建 AI 智能体的框架，例如用于对话式 AI 的工具包或集成数据的框架	受损的框架组件、后门、输入验证攻击、供应链攻击、框架未授权等
2	数据操作	为 AI 智能体处理、准备和存储数据的地方，包括数据库、向量存储、RAG（检索增强生成）管道	数据中毒、数据泄露、模型反转及提取、数据篡改、受损的 RAG 管道等
1	基础模型	构建智能体的核心 AI 模型。可以是大型语言模型 (LLM) 或其他形式的 AI 模型。	对抗性示例、模型窃取等

除了风险评估类框架，现行阶段在需求评估阶段通常也将各类应用的合规需求纳入考量。大模型应用目前需要满足的外规包含但不限于《互联网信息服务深度合成管理规定》、《生成式人工智能服务管理暂行办法》，涉及到国标如《GB/T 45654-2025 网络安全技术 生成式人工智能服务安全基本要求》等。具有舆论属性或者社会动员能力的，需要依据相关法律法规进行备案或登记，在设计应用时需要考虑合规要求的功能点在应用上架前要落实。

表5 合规要求

	要求	来源
个人信息保护	1. 不得收集非必要个人信息，不得非法留存能够识别使用者身份的输入信息和使用记录，不得非法向他人提供使用者的输入信息和使用记录。 2. 深度合成服务提供者和技术支持者应当加强训练数据管理，采取必要措施保障训练数据安全；训练数据包含个人信息的，应当遵守个人信息保护的有关规定。	《互联网信息服务深度合成管理规定》、《生成式人工智能服务管理暂行办法》
标识	1. 可能导致公众混淆或者误认的，应当在生成或者编辑的信息内容的合理位置、区域进行显著标识，向公众提示深度合成情况。 2. 应当按照《互联网信息服务深度合成管理规定》对图片、视频等生成内容进行标识。	
投诉	1. 应当建立健全投诉、举报机制，设置便捷的投诉、举报入口，公布处理流程和反馈时限，及时受理、处理公众投诉举报并反馈处理结果。 2. 深度合成服务提供者应当设置便捷的用户申诉和公众投诉、举报入口，公布处理流程和反馈时限，及时受理、处理和反馈处理结果。	

（二）安全建设

在完成整体方案设计后，开始开展智能体的开发。建设过程主要包含模型选择和应用开发过程。模型选择阶段主要是针对场景自己训练或者选择外购或开源的基础模型。如果选择训练甲方垂类模型，一般需要经过数据准备、预训练和后训练过程。预训练需要基于通用语料训练模型学习语言的结构和规律得到基础模型，后训练是基于微调(Supervised Fine-Tuning，SFT)或者强化学习(Reinforcement Learning，RL)等技术手段使得预训练得到的基础模型掌握业务场景所需要的行业知识和数据，该阶段一般会需要数据标注。训练自己的模型需要确保语料安全、数据安全等。如果是选择外部模型，一般选择方式有API调用、本地部署开源或商业模型等方式，需要考虑引入模型的安全性。

表6 各类渠道大模型风险

	风险	缓解策略	适用场景	备案要求
外部接口	输出不合规内容； 内部数据泄露； 服务稳定性；	SLA 协议； 本地部署围栏产品针对出向流量中敏感数据做规则防泄露；	普通场景、办公场景等	算法登记
本地部署开源大模型	输出不合规内容，侵犯用户权益； 供应链漏洞； 模型投毒； 算力消耗；	围栏； 及时跟进漏洞预警并进行模型升级； 对模型来源进行审核，建立内部可信源，关注huggingface 等模型源上的安全信息；	通用型场景	算法登记
本地部署外购大模型	输出不合规内容； 参数不透明，解释性不强； 算力消耗；	合同约束；	专业领域	算法备案
本地部署自训练垂类或行业大模型	输出不合规内容； 模型的稳定性； 训练泄露敏感数据； 训练数据标注偏差输出偏见； 模型逆向和窃取； 算力消耗；	正则化、对抗训练； 训练数据脱敏、差分隐私； 模型文件加密、签名和完整性检测；	专业领域	模型备案 算法备案

在完成基础模型构建后，进入智能体开发阶段。当前阶段证券业常用的AI智能体开发平台为开源或商业版本的Dify、n8n以及Coze等。该类开发平台存在漏洞风险和版权风险。开源版本的Dify目前使用的要求为不可去标识、且不支持多租户等，部分企业在使用开源版本时需要注意遵循开源协议中声明的义务，避免发生侵权行为。在构造智能体时，可以依据MAESTRO框架考虑各层级的风险，针对风险选取合适的安全防护手段。企业级智能体开发需要综合考虑集成 API 网关、操作审计、数据加密，满足 GDPR、等保合规要求等。

（三）安全评测

对客类应用应强制引入模型安全测评，模型安全测评主要包含基准评测、漏洞后门扫描及红队测试等。部分外规会对基准评测的所使用的语料测试集数量有明确要求，比如完成大模型备案需抽检1万条以上拦截关键词检测、一万条以上内容评估测试集、五千条民族信仰性别测试等。各种组织也开源了自己的测评框架可以用于测评，企业内部如自训模型可以考虑自建安全评测基准库。应用本身的安全测试可参考《AI智能体运行安全测试标准》，对大模型智能体引入了的新增攻击面进行测试。安全测试需要考虑应用整体的安全。

表7 测评方式合集

	测试内容	案例/工具	测试对象
模型测评	基准测试、基础模型扫描（漏洞扫描、后门扫描）、对抗性测试（红队测试）等	ModelScan Counterfit Garak Guardrails AI LLM-Guard BenchmarkTest（Openai 的接口或国内乙方提供的测评平台）	模型
MCP 服务	智能体调用 MCP 工具服务的安全测试	AI-Infra-Guard V2 传统黑白灰盒/开源组件扫描	MCP 服务
开发框架等其它各攻击面	第一章提到的各类场景下的各类攻击面安全测试	传统黑白灰盒/开源组件扫描 版本和配置基线检查，查看开源版本的最新安全修复记录，确认是否残存遗留安全问题 人工渗透	构造智能体引入的各模块

（四）安全运营

大模型应用上线前要做好内部AI物料登记，上线后需要保障AI运行态安全。AI智能体的防护架构也应遵循纵深防御的原则，并充分考虑生成式人工智能语言问答的风险特性。

表8 大模型防护矩阵

层次	被保护主体	拦截和检测攻击类型	措施/工具
数据层防护	数据	数据窃取攻击等	数据脱敏等
应用层防护	接口 应用	传统 web 攻击	API 网关 全流量产品 API 安全产品 智能体沙箱 应用层隔离
模型层防护	模型	模型越狱 输出不合规内容 模型功能滥用 信息窃取 提示词注入	LLM 防火墙 模型沙箱
基础设施层防护	主机 容器	针对基础侧安全对抗	主机安全产品 容器安全产品

三、总结

大模型技术的使用即给安全防护体系带来了挑战，也同时赋能安全各领域。大模型只是工具，如何使用好工具是我们需要探索的问题。在大模型领域安全治理的思路仍然是掌握资产，发现漏洞。各项技术仍不十分成熟，业务和技术都在探索中。学术和工业界对于大模型应用的研究日新月异，安全需要跟进各项最新技术的发展，建设好甲方大模型设施台账，做好基础的隔离工作。

参考文献

1.2025年中国证券业大模型应用跟踪报告https://mp.weixin.qq.com/s/_2cZLO4w_bmwU4PXD4dqwQ

2.云上LLM数据泄露风险研究系列（一）：基于向量数据库的攻击面分析https://mp.weixin.qq.com/s/5jndWjm_yMEXY0E-W369NQ

3.如何保护 RAG 应用的安全https://mp.weixin.qq.com/s/HSmyvwN4z8sScKq_ZqW9Jw

4.Agentic AI Threat Modeling Framework: MAESTROhttps://cloudsecurityalliance.org/blog/2025/02/06/agentic-ai-threat-modeling-framework-maestro

5.大模型编排框架攻防（以LangChain为例）https://xz.aliyun.com/news/18440

6.Enterprise-Grade Security for the Model Context Protocol (MCP): Frameworks and Mitigation Strategies https://arxiv.org/html/2504.08623v1

7.https://www.anthropic.com/engineering/building-effective-agents

8.AI STR系列：单AI智能体运行安全测试标准https://wdtacademy.org/publications/Agent?sessionid=1515960270

9.互联网信息服务深度合成管理规定

10.生成式人工智能服务管理暂行办法

11.当AI智能体学会“欺骗”，我们如何自保?火山引擎的MCP安全答卷https://mp.weixin.qq.com/s/GSmhUuo7msvHGHEJm4wCGg

12.AI 应用安全挑战与测评实践指南

13.国内外大模型安全技术框架汇总https://mp.weixin.qq.com/s/birTRTbNWXYmZ-ee46vlqw

05 热点解读

P79 小程序安全解决方案
张华

P85 金融行业IPv6规模化部署顶层设计与落地策略
葛锐、张绍峰、陈政、沈鑫尧

P89 金融行业勒索病毒防御评估研究
施勇、张涵

P93 2025RSAC大会解析：众声汇聚，共探全球网络安全新趋势
江爱军、王伟涛、盛浩月

小程序安全解决方案

张华 | 腾讯云计算（北京）有限公司

摘　要： 随着小程序在日常生活中的广泛应用，小程序急剧扩张、种类繁多、开发门槛低、广泛使用第三方库和插件和跨平台开发框架等，即使微信、支付宝、抖音等小程序本身的基础平台功能较为安全，80%以上的小程序都存在安全风险如：数据泄露、隐私合规、接口仿冒、薅羊毛、山寨仿冒等，国家、政府及相关行业机构开始关注小程序的安全问题。国家和地区已经出台了相关法律法规，对小程序的开发、运营和数据处理等方面进行规范和监管。

关键字： 小程序安全、隐私合规、数据泄露、山寨仿冒

一、引言

小程序作为一种无需下载安装即可使用的应用程序，具有便捷性、轻量化等特点，在移动互联网领域迅速崛起，广泛应用于社交、电商、金融、生活服务等多个领域。然而，小程序的快速发展也带来了诸多安全问题。一方面，小程序开发门槛较低，许多开发者在安全意识和技术能力上存在不足；另一方面，小程序开发过程中大量依赖第三方库和插件以及跨平台开发框架，这些外部资源可能存在安全漏洞，从而增加了小程序的安全风险。小程序安全问题不仅会导致用户个人信息泄露、财产损失，还会损害企业的声誉和利益，甚至影响整个小程序生态的健康发展。因此，研究小程序安全的现状与防护策略具有重要的现实意义。

二、小程序安全现状

（一）小程序的发展与应用

小程序自诞生以来，凭借其无需下载、即开即用的特性，受到了用户和企业的广泛青睐。截止2023年8月，中国移动互联网月活用户达到12.22亿，用户小程序使用习惯持续攀升，微信、支付宝、百度、抖音小程序整体月活规模分别达到9.25亿、6.45亿、3.67亿、2.67亿，全网用户月人均使用小程序个数达到15.9个，涵盖了电商购物、餐饮外卖、出行服务、在线教育等众多领域。支付宝、抖音等平台的小程序也呈现出快速发展的态势，为用户提供了丰富多样的服务。

（二）小程序安全风险的统计与分析

相关研究表明，80%以上的小程序存在安全风险。其中，数据泄露问题最为突出，约占安全风险总数的57%。例如，部分小程序在获取用户输入信息（如账号密码、身份证号等）后，未进行妥善的加密处理，导致数据在传输或存储过程中被窃取。隐私合规问题也较为常见，约占22%，一些小程序过度收集用户的个人信息，且未明确告知用户信息的使用目的和方式，违反了相关隐私保护法规。接口仿冒、薅羊毛、山寨仿冒等安全风险也不容忽视，接口仿冒可能导致用户数据被非法获取，薅羊毛行为会给企业带来经济损失，山寨仿冒小程序则会误导用户，损害正版小程序的权益。

（三）典型小程序安全事件案例分析

近年来，发生了多起影响较大的小程序安全事件。例如，某电商小程序因存在安全漏洞，导致大量用户订单信息和支付记录泄露，涉及用户数量超过百万，此次事件不仅给用户带来了财产损失和隐私泄露风险，也对该电商平台的声誉造成了严重损害，用户信任度大幅下降，平台业务受到了明显的冲击。又如，某互联网金融公司的小程序被发现存在接口仿冒问题，不法分子通过仿冒接口获取用户信息，进行盗刷操作，给用户造成了巨大的经济损失。这些案例充分说明了小程序安全问题的严重性和危害性。

三、相关法规

（一）国家和地区出台的相关法律法规概述

为了规范小程序的开发、运营和数据处理行为，保障用户的合法权益，国家和地区出台了一系列相关法律法规。如《中华人民共和国网络安全法》明确规定了网络运营者的安全义务和责任，要求其采取技术措施和其他必要措施，保障网络安全、稳定运行，有效应对网络安全事件，保护个人信息安全。《中华人民共和国个人信息保护法》则对个人信息的收集、存储、使用、加工、传输、提供、公开等处理活动进行

了全面规范，强调了个人信息处理者应当遵循合法、正当、必要和诚信原则，不得过度收集个人信息。此外，各小程序平台也制定了相应的安全规范和政策，如微信小程序的《微信小程序平台运营规范》、支付宝小程序的《支付宝小程序隐私政策指引》等，对小程序开发者和运营者提出了具体的安全要求。

（二）行业监管要求

1、金办发[99]号文

国家金融监督管理总局办公厅关于加强银行业保险业移动互联网应用程序，为指导银行业金融机构、保险业金融机构和金融控股公司(以下统称金融机构)进一步提升服务质量，规范移动互联网应用程序(运行在移动智能终端上向内、外部用户提供服务的应用软件，包括但不限于移动应用APP、小程序、公众号等，以下简称移动应用)管理

2、证监会

中央网信办秘书局中国证监会办公厅关于印发《非法证券活动网上信息内容治理工作方案》的通知中针对仿冒网站、仿冒APP、仿冒小程序的金融行业仿冒常态化监管，同时证监会也对移动互联网应用颁布了相关指引和规范如：《证券期货业移动互联网应用程序安全检测规范》、《证券期货业移动应用软件备案工作指引（试行）》。

四、小程序安全及解决方案

（一）小程序隐私合规

1、小程序隐私合规风险

（1）过度收集用户信息

部分小程序为了获取更多的数据资源，存在过度收集用户信息的现象。例如，一些并不需要使用用户地理位置信息的小程序，却在未经用户充分授权的情况下，获取用户的位置信息；还有些小程序在用户注册时，要求用户提供过多的个人敏感信息，如身份证号、银行卡号等，超出了正常业务所需的范围。

（2）信息共享和第三方合作中的隐私问题

小程序在与第三方进行信息共享或合作时，如果未对第三方进行严格的安全评估和监管，可能导致用户信息在共享过程中泄露。例如，小程序将用户的部分信息共享给第三方广告商，若第三方广告商的安全措施不到位，用户信息就可能被泄露。此外，一些小程序在隐私政策中未明确说明信息共享的对象和条件，用户对自己信息的流向缺乏了解，增加了隐私风险。

2、小程序隐私合规解决方案

基于相关法律法规、国家标准、行业标准，通过静态检测和动态检测技术，结合隐私合规相关实践经验，识别小程序的数据隐私合规问题。

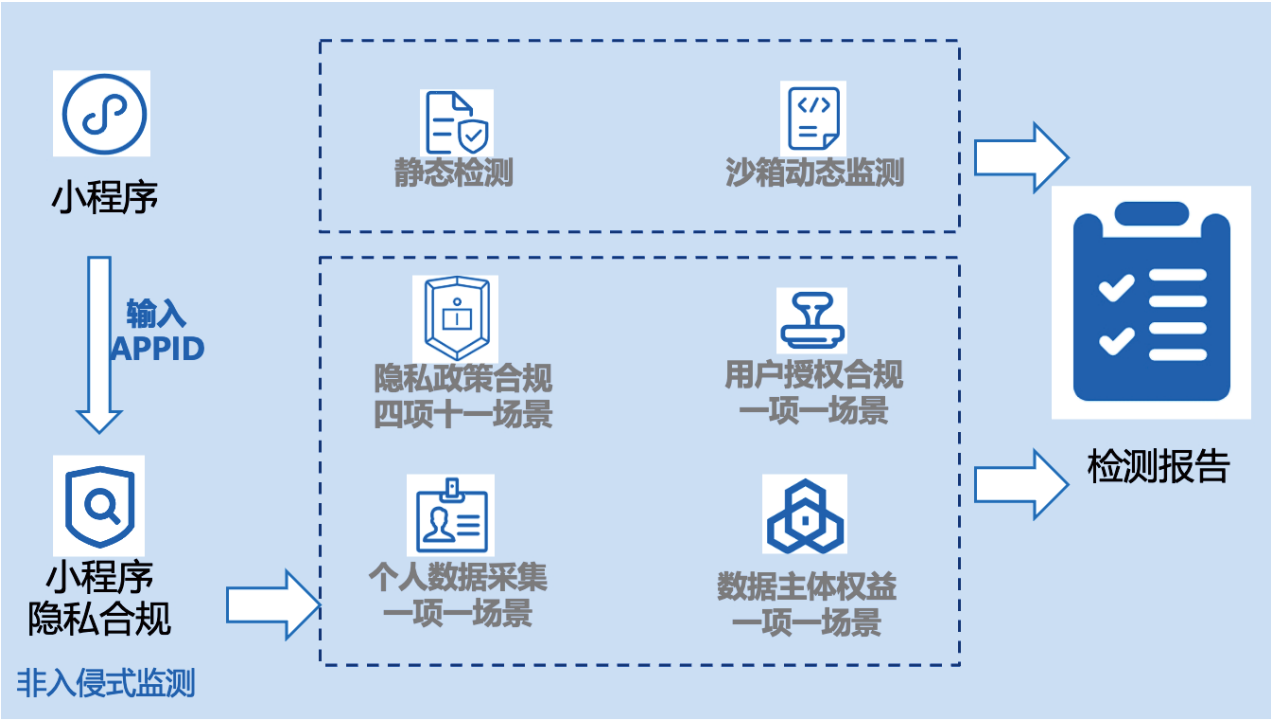


图1 小程序隐私合规检测

(1)检测技术

通过动态沙箱检测技术和DPI深度报文检测技术，覆盖产品全生命周期的隐私合规需求，多重纬度技术检测能力，得出全面精准地合规风险检测结果。

(2)检测依据

■工信部信管函〔2020〕164号文:

■工业和信息化部关于开展纵深推进APP侵害用户权益专项整治行动的通知。

■明确“APP服务提供者，即互联网信息服务提供者提供的可以下载、安装、升级的应用软件，包括快应用和小程序等新应用形态”。

(二)小程序资产排查

1、小程序资产排查

通过排查资产和相似度检测，梳理小程序资产数量、更新情况和风险详情，摸清资产底数，探查非官方、僵尸账号，提前识别资产风险，为资产有效管控提供基础。

技术方案：对接小程序平台的数据库接口（如微信、支付宝的开发者平台数据），通过爬虫或API同步全量小程序信息，建立资产库并实时更新。

■定期爬取或通过平台授权获取新增小程序的注册信息，与品牌资产库中的正版信息比对，发现未备案的相似资产。

■用定时任务（如每小时）扫描平台内小程序，触发关键词预警（如含品牌商标的新增小程序），实现动态监测。

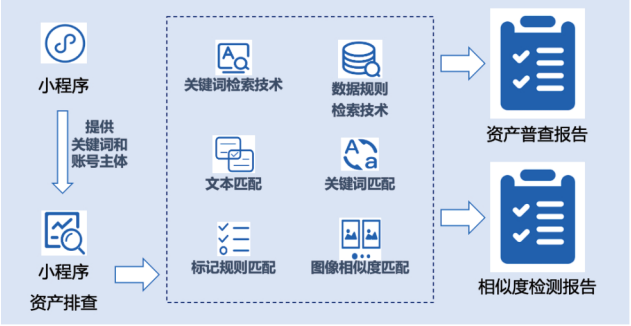


图2 小程序资产排查

2、小程序仿冒检测

(1)小程序仿冒检测

通过相似度比对技术、关键词与商标监测、品牌标识与技术认证等多种技术检索微信生态内小程序，生成相似度评分报告，快速定位仿冒者基于名称、图标、功能描述、页面结构等特征，通过文本相似度算法（如余弦相似度）、图像识别（图标/页面截图比对）、代码结构分析（若能获取代码片段），识别与正版高度相似的仿冒小程序。

(2)多维度特征提取与比对技术

技术方案：从小程序的基础信息（名称、图标、开发者）、内容（页面结构、功能描述）、代码（若可获取）中提取关键特征，通过算法比对分析差异或相似度。

■文本特征：用自然语言处理（NLP）提取名称、描述中的关键词，通过余弦相似度、编辑距离等算法，识别仿冒小程序的谐音、错字变体（如“微xin”仿“微信”）。

■图像特征：用图像识别（如SIFT算法）提取图标、页面截图的轮廓、色彩特征，比对与正版图标的相似度，识别高度模仿的图标。

■代码特征：若能获取代码片段，通过抽象语法树（AST）分析代码结构、函数命名等，判断是否抄袭正版核心逻辑。

(3)异常行为分析技术

技术方案：通过用户行为数据（访问量、跳转路径）、业务数据（交易金额、接口调用频率）识别异常，间接定位风险资产（如仿冒小程序的短期流量激增）。

■分析小程序的访问来源，若某小程序突然从非官方渠道获得大量跳转，且功能与正版相似，可能为仿冒品。

■监测接口调用异常，如仿冒小程序盗用正版API时，会出现非授权IP的高频请求，通过接口签名校验和IP黑名单拦截。

(4) 自动化合规检测技术

技术方案：将政策法规（如《网络安全法》）、平台规则（如微信小程序审核标准）转化为可量化的检测规则，通过规则引擎自动校验小程序是否合规。

预设合规关键词库（如禁止使用的违规词汇）、功能禁区（如未经许可的金融服务），对小程序内容进行文本匹配和功能扫描，生成合规评分报告。

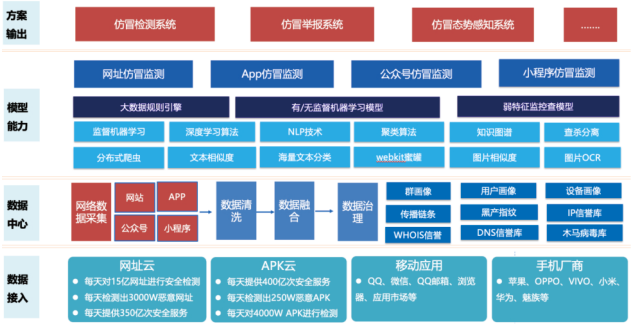


图3 仿冒小程序检测

(5)小程序仿冒防护

■快速投诉与下架

对接各平台（微信、支付宝、抖音等）的投诉通道，通过技术手段自动生成侵权证据（如相似度报告、商标证明），提交平台快速处理（下架、封号）。部分平台支持“品牌保护绿色通道”，认证企业可优先处理仿冒投诉，缩短处理周期。

■平台规则与法律防护配合

利用平台规则：向小程序平台（微信、支付宝、抖音等）提交商标注册证、著作权证明等材料，备案品牌信息，平台会对含侵权标识的小程序注册进行拦截。

法律追责：对仿冒情节严重者，通过公证固定证据（技术监测报告、用户证言），提起知识产权诉讼，要求赔偿并禁

止其运营。

（三）小程序安全诊断

日常小程序业务场景中，存在敏感信息泄漏、业务数据被篡改、恶意病毒植入、未授权或越权访问等因业务系统漏洞引起的安全风险，严重影响业务的正常运行。小程序安全诊断对小程序进行全链路的安全问题排查，并形成诊断报告和修复建议。

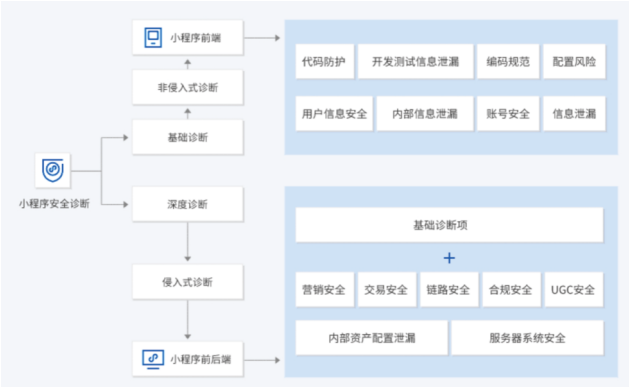


图4 小程序安全诊断

（1）基础诊断

通过小程序 AppID 即可开始诊断，覆盖代码保护检测、开发测试信息泄漏、编码规范、配置风险、账号安全、用户信息安全、内部信息泄漏、其他类安全共8大检测类型。

（2）深度诊断

是对客户通过模拟黑客攻击的形式，对小程序业务系统进行渗透测试，覆盖范围在基础诊断的范围之上又包括营销安全、交易安全、链路安全、内部资产配置泄漏、通用 Web 安全、服务器系统安全、系统安全漏洞、合规安全、UGC 安全，并在诊断报告中提供专业的修复建议，帮助企业识别代码漏洞等安全风险。

（四）小程序风险监测

在日常小程序业务场景中，存在因微信 API 调用不规范而隐藏的安全风险，及业务逻辑不严谨导致的安全隐患。同时国家对移动互联网的日益重视，不断推出安全法律法规要求企业加强对数据的监管和保护，企业亟需具备一定的安全检测能力。

结合网络安全法、等保2.0等政策要求，针对小程序前端和后台 WEB 端整体提供的自动化风险检测工具，覆盖前台代码安全和 API 使用规范，以及业务 CGI 和对 WEB 框架的安全检测，包括对 SQL 注入、XSS 跨站脚本、目录遍历、信息泄露等主流 Web 攻击方式的防范，提升小程序的安全防护能力。

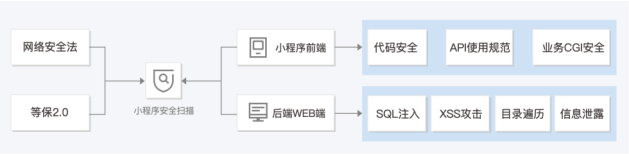


图5 小程序风险监测

（五）小程序安全防护

近些年来，随着移动互联网的井喷式发展，小程序、移动应用已经成为互联网最大流量入口，从而针对小程序、移动应用的攻击也成为攻击业务的重要突破口。

1、场景案例

以国家攻防演练中的常见攻击手段举例：

Step1：信息收集。红队通过互联网公开信息，获取目标单位的域名、小程序等资产信息 。

Step2：脆弱性分析。红队通过测试发现，目标单位没有对小程序做额外的防护措施，信息传输走公网链路，且未对数据加密，是容易实施攻击的薄弱点。

Step3：渗透攻击。抓取了用户登录小程序时的请求包，发现有一个请求包中泄露了Access Key Id / Secret Access Key；经过查询发现这是云服务器的AK/SK，填入key值后，成功接管oss云存储桶。

Step4：横向移动。下载了存储在OSS云存储桶中的所有敏感数据，包括客户数据、公司机密、财务信息；

Step5：获取权限。使用账号密码直接登录云主机控制台，完全接管所有云资源。

以上可以看出，小程序、移动应用逐步也成为攻击者又一突破企业内部安全防护体系的“纽带”，针对小程序攻击场景也越来越广泛，例如：

（1）中间人攻击，攻击者截获并可能篡改通信数据。

（2）重放攻击，攻击者重复发送已被接收的数据包。

（3）前向安全性不足，长期密钥泄露可能导致之前通信的会话密钥被破解。

（4）协议伪造，伪装 MMTLS 对业务发起请求。

（5）降级攻击，攻击者将 MMTLS 连接降级为不安全的 HTTP 连接，从而窃取用户数据。

2、对抗策略

（1）使用带认证的密钥协商，通过数字签名算法对公钥进行签名，确保通信双方计算出的密钥是一致的，防止中间人攻击。

（2）采用 0-RTT 密钥协商方法，在握手过程中安全地传递业务数据，减少握手时间。由于已由 0RTT 完成初步协商，后续密钥协商通过 1-RTT PSK 交互密钥和数据包。

（3）使用 AES-GCM 算法进行认证加密，确保数据的保密性和完整性。

（4）通过 HKDF 函数进行密钥扩展，确保生成的密钥具有足够的伪随机性一性。

(5) 0-RTT 实施基于时间序列的防重放策略，结合客户端和服务器端的时间戳来检测和防止重放攻击。

3、解决方案

(1) 安全网关接入

打通小程序、移动应用与微信网关接入链路，通过原生的高可用加速服务、防数据泄露、中间人劫持、重放攻击、防DDoS、防DNS劫持、防爬防刷等原生安全防护能力，全方位保障业务安全、高效、稳定运行。

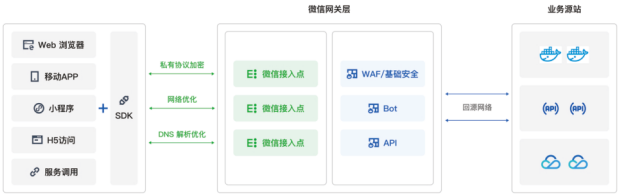


图6 小程序安全防护解决方案

■通过微信SDK前端加密，有效保护企业数据资产，防止中间人攻击、重放攻击、密钥泄漏

· 二次封装加密数据及接口，无明文数据

· 微信自研私有协议加密，破解门槛及成本极高

■在黑产「三板斧」基础上，针对小程序场景划分联合打击方案：

· 黑账号打击：联合微信账号风控体系，精准控制非法低质量羊毛党账号

· 黑设备打击：加密协议结合动态流量风控，精准识别非法设备

· 黑行为打击：结合用户行为分析模型，判断模拟行为路径，抵御群控风险

图7 微信安全网关防护示意图

■第一层防护

通过微信网关的微信私有链路，在请求进入微信网关的微信私有协议时，微信网关可识别各种协议栈、爬虫特征、模拟器攻击、黑灰产IP访问、DDoS攻击等各种异常请求，并进行及时拦截；经第一层安全防护的初步筛选后，将目前初步判定为正常的流量放入微信网关的微信服务器中进行第二层的流量筛选与清洗。

■第二层防护

通过微信网关对经过第一次初步清洗后的请求进行第二层的风控防护，针对伪造/篡改网关私有协议的请求进行清洗，同时结合用户请求频率、权限校验等风控能力筛选过滤掉非法请求、越权请求、高频请求等；经过第二层安全防护的二次清洗，已将90%以上的异常流量拦截掉，此时再将剩余的流量转发给业务源站。

■第三层防护

为防止有漏网之鱼，微信网关通过第三层自定义安全防护策略的能力，识别漏网之鱼，结合请求特征、设备信息、账号等多维的数据形成智能分析模型，对流量进行环境检测、智能验证等更高级别的处理，支持业务自行决策防护形态同时支持切换至安全性更高的验证码形态。

(2) 账号风控

依托于微信万亿级超大规模风控平台，账号风控系统通过多个维度进行风控的逐层进行风控分析和递增：

■自然人与实名认证识别

■风控识别

在自然人与实名认证识别的异常风险识别中，由于openid、unionid 对于 小程序商户为固定，因此可以快速识别当前用户的相关实名身份，避免大量小号入侵当前小程序。

在风控识别中，存在多种合法性校验，包括但不限于Openid、用户校验、商户属性校验，商户安全字段。安全网关在访问过程中会根据实时的流中的 appid，openid 信息聚合对应的风控业务数据标签。并根据对应风险标签提供相关反馈风险登记结果。

(3) 设备风控

在重点的风控场景下，基础协议层面对抗已经进化为终端、行为、实人上的对抗。灰黑产用户在进行对抗的场景下，通常会在设备环境完整的下进行批量大规模自动化控制，用来模拟真人的操作（点击，滑动）。

所以通过 Web 中的常用的设备可信度用来做设备异常判断信号容易出现设备碰撞导致误判的场景。针对这种场景，需要依赖各类型的传感器信号以及底层设备的架构信息来判断是否为真人在使用真实微信客户端，而不是单纯判断微信客户端的有效性。

真人操控微信客户端会包含多种方式去判断是否在机架、是否固定、是否虚拟化设备、是否猫池、是否连接池等。而 PC 微信端由于对抗场景不一致，会有更进一步风控管理策略。

同时，会通过用户的访问路径行为来判断用户的行为是否与主流用户离群，从而判断是否异常。

■用户 IP，用户访问的路径（如 Path A->B->C，缺失 B 时提高风险评级）

■用户访问某个路径的次数（如 B->B->B...，请求次数

过多时提高风险评级）

■用户访问某个路径的 Header, Body 等信息

■用户访问某个路径的开始时间，结束时间，以及大多数用户的访问模式

通过此类方式弥补了纯账号风控模式下，安全风控信息更新不及时导致的误拦截，漏拦截。

表1 针对不同种类的设备进行风控识别

设备风险	评判举例	建议处置
0	正常设备，用户正常使用设备	无风险，不做任何阻拦
1	低风险设备，有非人类操作的嫌疑，但设备本身完整	轻度可疑的风险，建议结合行为上报，综合其他维度评判是否需要进一步检测和处理
2	中风险设备，有外挂特征，但设备本身完整	建议进行简单的验证（如验证码、短信等）
3	结合设备信息与网络综合判定，设备有群控特征，或在系统层面强制阻断了一些信息的上报，以混淆判断。	建议根据业务场景采取一定措施避免伤害。例如，营销活动可降低高等级奖励的概率；打榜类活动对此类投票降低权重；登录注册要求二次验证等
4	确定为不安全设备，无法通过正常设备检测，可能为模拟器或工具	建议根据业务逻辑直接拦截。例如，红包类活动返回不中奖或最小额红包；打榜类活动不计算票数；登录/注册操作要求二次验证；高危业务可选择限制本次操作。

（六）小程序安全加固

1、小程序常见风险问题

小程序前端代码和逻辑设计存在常见风险问题如：代码易被反编译，核心业务逻辑被破译，算法易被二次打包，病毒代码和流氓广告等被恶意植入等，此类风险问题严重影响业务的正常开展，并触及安全法规的要求。

2、解决方案

在不改变应用源代码的情况下，针对小程序前端代码进行加密，实现字符串加密、属性加密、调用转换、代码混淆、反调试、防破解等多项保护措施，防止代码暴露，提高攻击者分析前端代码逻辑的难度，保护小程序安全。

图8 小程序安全加固

五、结论

小程序的安全问题已经成为制约其发展的重要因素，数据泄露、隐私合规、接口仿冒、薅羊毛、山寨仿冒等安全风险给用户和企业带来了严重的损失。国家和地区出台的相关法律法规，对小程序的安全起到了规范和监管作用，但要从根本上解决小程序安全问题，还需要小程序开发者、运营者从技术、管理和运营等多个层面采取有效的防护策略。在技术层面，要加强数据加密、用户认证、代码安全等方面的措施；在管理层面，要建立健全安全管理制度，加强人员培训，严格审核第三方资源；在运营层面，要实施实时监控、优化活动规则、加强品牌保护。只有各方共同努力，才能提高小程序的安全性，保护用户的合法权益，促进小程序生态的健康、可持续发展。未来，随着技术的不断发展和应用场景的不断拓展，小程序安全问题也将不断演变，需要持续关注和研究新的安全防护技术和策略，以应对不断变化的安全挑战。

参考文献

1.《中华人民共和国网络安全法》
2.《中华人民共和国个人信息保护法》
3.《数据安全法》
4.《金办发99号》
5.《证券期货业移动互联网应用程序安全检测规范》
6.《证券期货业移动应用软件备案工作指引（试行）》
7.《微信小程序平台运营规范》

金融行业IPv6规模化部署顶层设计与落地策略

葛锐、张绍峰、陈政、沈鑫尧 | 互联网域名系统北京市工程研究中心有限公司

摘　要： 本文围绕 IPv6 规模化部署，从政策背景出发，阐述了改造原则与策略，重点规划了 IPv6 地址管理、DNS服务器配置及流量监测等内容，提出了分阶段的改造计划，并针对网络设备、DNS 解析、地址分配和安全等常见问题给出解决方案，为 IPv6 改造工作提供全面指导。

关键字： IPv6、地址规划、改造计划、DNS配置

一、引言

贯彻落实习近平总书记关于网络强国的重要思想，为深入实施《关于加快推进互联网协议第六版(IPv6)规模部署和应用工作的通知》,扎实推动IPv6规模部署和应用向纵深发展，有力支撑网络强国、数字中国建设，中央网信办、国家发改委、国家工信化部联合印发《2025年深入推进IPv6规模部署和应用工作要点》文件，明确2025年工作目标：全面建成全球领先的IPv6技术、产业、设施、应用和安全体系，并部署深化融合应用、深化单栈规模部署、提高终端设备连通水平、强化网络安全保障等九个方面重点任务，2025年5月人行发布《IPv6技术金融应用规范》推进IPv6技术创新与金融融合应用、提升IPv6在金融行业的应用广度和深度，推进金融行业IPv6规模部署和应用创新成果标准化。

（一）改造原则

鉴于IPv6 网络演进是一个长期过程，IPv4网络仍在广泛使用，综合业务运行、改造成本等因素，改造基本原则：

（1）双栈优先：网络改造原则上采用IPv4/IPv6双栈方式，鼓励有条件的用户采用IPv6单栈方式。

（2）最小变动：网络改造应基于既有设备和网络，尽可能让网络架构最小改动及设备最少替换。

（3）安全平稳：网络改造应统一规划、因地制宜、分步实施、安全稳妥，基础设施先行，终端和应用逐步改造，保障业务平滑切换。

（二）改造策略

IPv6改造，采取基础设施先行、分批分阶段推进系统改造策略。

先总后分：以二级单位本部为改造起点，因其是网络中枢，完成后再推进分子机构改造，可形成自上而下的规范传导，避免分散建设导致的标准不统一。

先外后内：优先改造外网及相关系统，这类区域直接对接外部IPv6环境，先行打通能快速实现外部通信适配；后续再改造内网，可在外部稳定基础上保障内部网络迁移的安全性。

先新后旧：新建网络及系统本身无历史兼容负担，优先改造能快速形成IPv6应用示范；旧有区域及系统则在新架构经验基础上逐步推进，降低改造对现有业务的干扰。

二、IPv6地址规划

IPv6改造，核心在于IP地址的改造。企业在这一过程中，需要具备直观的IP地址规划管理能力、高性能的IPv6地址分配服务能力、高效的IPv6地址审计溯源能力，以及智能的IPv6终端资产探测能力。

（一）地址规划管理

1、地址申请

在IPv6网络改造过程中，IPv6地址规划将决定网络与ISP互联的规划、路由规划、AS规划、NAT规划等，IPv6地址容量应符合企业规模需求，从运营商或CINIC申请全球可单播IPv6地址，不具备条件的可使用IPv6私有地址，IPv6地址规划对网络的兼容性、可扩展性产生决定性影响，是IPv6网络改造的第一步也是最重要的一步。

(1) 模式一：向CNNIC申请地址

CNNIC（中国互联网络信息中心）是国内IPv6地址分配的权威机构。通过该模式申请，可直接获得全球唯一的IPv6公网地址段，地址归属清晰且具有官方权威性。适用于对地址自主性、长期规划有较高需求的单位，比如大型企业、科研机构等，申请需遵循CNNIC的地址分配规范和流程。

(2) 模式二：ISP申请IP地址

运营商拥有丰富的IPv6地址资源，向其申请可获得适用于自身网络环境的公网IPv6地址。这类地址能直接接入运营商的IPv6网络，适配性较强，申请流程相对简便，通常与运营商的网络服务绑定，适合依托单运营商网络的单位。

(3) 模式三：向ISP申请GUA，内部ULA

GUA（全球单播地址）是可在互联网上路由的公网地址，向运营商申请GUA能满足单位外部通信需求；ULA（唯一本地地址）是仅在单位内部使用的私有地址，无需向外部申请，可由单位自行规划使用。该模式下，外部通信依赖运营商提供的GUA，内部通信使用自主管理的ULA，即GUA为全球公网地址，ULA为内部私有地址，GUA向运营商申请，ULA自行规划。

2、规划原则

(1) IPv6地址划分应有层次性，便于简化路由表。IPv6地址分配要尽量给每个区域分配连续的IP地址空间；相同的业务和功能尽量分配连续的IP地址空间，有利于路由聚合以及安全控制。

(2) IPv6的地址分配需要有足够的灵活性和可管理性，应考虑到现有业务 、新型业务以及各种特殊的业务的需要，地址使用兼顾到近期的需求与远期的发展以及网络的扩展，预留相应的地址段。安全层面，尽可能实现IPv6地址的分配记录和历史审计。

(3) 子网网段地址要连续，便于聚合，并尽量使IPv6地址与原来的IPv4地址有一定对应关系，要预留作为Loopback地址的网段。

(4) 每个子网中最小一个IPv6地址建议作为该子网的网关地址。

(5) 各安全设备带外管理口配置IPv6地址，统一接到管理交换机上，实现带外管理。

3、分配原则

针对不同类型终端的功能特性与管理需求，建议采用差异化的IPv6地址分配策略，具体如下：

(1) 有状态分配（DHCPv6自动配置）：适用于普通PC、通用终端，以及需要进行集中IP地址管理的工业互联网终端。这类终端通常对网络参数的统一配置、地址的动态管控有需求，通过DHCPv6自动配置，可由DHCP服务器完成地址的动态分配与回收，同时同步配置DNS、网关等关键参数，满足集中化管理的要求。

(2) 无状态分配（SLAAC自动配置）：主要针对数据采集设备、传感器、视频摄像头等物联网终端。这类终端更注重接入效率和轻量化运行，可在网关部署对应地址配置策略，采用无状态SLAAC自动配置方式——终端无需依赖额外服务器，能直接基于路由通告生成地址，减少网络交互带来的延迟，适配物联网终端的特性。

(3) 手动分配：适用于服务器、存储设备、网络设备等的管理地址，以及不支持IPv6地址自动配置的设备。作为核心设备，管理地址需要保持固定以保障访问稳定性，避免动态分配可能导致的连接中断。

4、规划案例

一般我们获取的IPv6地址为48或者64位，将该/48或/64位地址段划分出多个/64位地址段，用于提供给用户终端及办公网络各类业务 ，终端通过配置自动获取对应IPv6地址；单独选择某个/64网段继续划分，划分出多个/127的段，用于提供给设备链路互联端口，对于设备链路互联地址，上联设备采用较小号IP地址，下连设备采用较大号IP地址；划分出多个/128的段，提供给设备环回接口，用于网管对设备进行管理。

假如运营商分配给企业分配的IPv6业务地址段为240E:250:2814::/48，实际使用时需要进一步将该整个大段划分成各个小段的地址分配给企业内的有线、无线终端用户，以及一些其他的IPv6业务终端。

类型	企业侧	运营商侧
边界互联接口地址	240E:250:2820::1:2814:02/126	240E:250:2820::1:2814:0 1/126
业务地址	240E:250:2814:: /48	

有线和无线用户的IPv4和IPv6网关都部署于核心交换机上， 由核心交换机统一进行IPv4和IPv6地址的分配 ； 在IPv6地址分配上，可使用DHCPv6地址分配的方式。

(1) 主要遵从四个原则：唯一性、可扩展性、连续性、实用性；

(2) 设备管理及互联地址：使用240E:250:2814:1::0/64网段；

(3) Loopback地址规划，使用整段连续IP地址，也使用240E:250:2814:1::0/64网段；

(4) IPv6业务网关统一设定为网段的第一位，即::1；

(5) 在完成IP地址规划之后，既可以配置静态IP地址，

也可以使用DHCP服务器动态分配IP地址，根据实际需求考虑。

区域	楼栋	用途	IPv6 地址	网关
设备互联	NA	设备互联	240E:250:2814: 1::0/64	NA
办公区	办公区	无线	240E:250:2814: 100::0/64	240E:250:2814 :100::1/64
		有线	240E:250:2814: 101::0/64	240E:250:2814 :101::1/64

5、地址规划

针对从运营商或上级获取IPv6地址段后，企业在规划管理中存在的缺乏统一规范、全流程监测手段，进而导致IP地址管理效能不足、难以规模化部署的问题，可通过构建系统化的IPv6地址规划管理方案解决。

方案应具备这些关键功能：需要能通过图形化操作或手动输入来划IPv6地址空间，可给不同地址空间添加标识并进行语义化说明，让地址空间更易理解和粒度更细；同时规划完成后能自动形成直观的IPv6规划图示，简化维护方式和加强二次规划能力；另外规划结果可在地址管理环节直接调用和补充；需要能贯穿地址管理全流程，形成规范的管理模式，实现对地址相关情况的有效掌握。

6、地址分配

在办公网启用IPv6地址自动分发功能，使用DHCPv6服务器中配置自动分配规则，支持终端设备通过无状态自动配置获取地址。配置中需定义分发地址段、网关及DNS信息，确保终端获取的地址与办公网络规划一致。通过实际测试验证自动分发效果，检查终端是否能够快速连接内外IPv6资源，并优化分发规则以提高效率。

7、地址管理

在IPv4与IPv6共存的双栈环境中，通过IP地址管理机制，以适配双栈的扫描方式实现两类地址全覆盖，收集包括基础状态及终端设备、所属部门等关联信息并多维度展示，形成区分地址类别、记录生命周期及关联信息的IP资产台账，可实现地址全量高效管控。

8、溯源审计

在IPv6环境中面对安全问题的事后溯源需求可依托IP地址管理相关机制实现。通过对IPv4/IPv6地址生命周期的回溯，需呈现历史事件信息像配置变更、地址冲突、地址状态等内容都能清晰展示。以地址为核心的回溯方式能帮助提升故障处理效率为故障诊断、三方鉴责提供有力支持还能精准追溯故障至具体用户从而达成一站式的IP溯源目标。

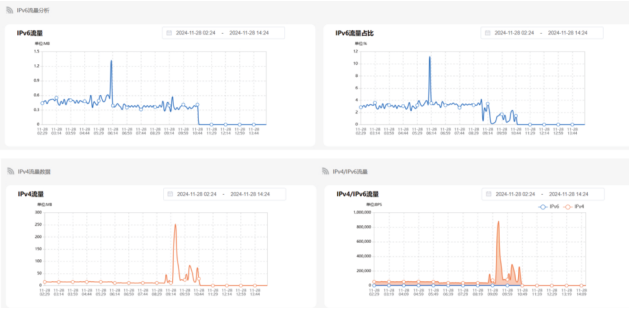
（二）DNS服务器配置

为了支持IPv6境下的域名解析功能，需要IPv6地址到相在DNS服务器上增加对AAAA记录的解析能力，确保能够解析应的域名。同时，启用提供的IPv6 DNS服务地址，并调整网络中的终端配置，使其优先使用IPv6 DNS服务。验证工作需覆盖内部和外部域名解析，确保在IPv6环境下的业务和应用均能顺利解析域名。

尤其需要注意的是，在IPv6改造过程中，对于应用间互访或应用调用数据库时，若原始采用IP地址硬编码的方式，后续必须改造为域名访问方式。这一改造是保障业务从IPv4平稳过渡到双栈模式，再到IPv6单栈模式时，最大限度降低业务影响面的关键举措。

（三）IPv6流量监测

改造后，需要建立IPv6流量监测平台，实时监测统计单位互联网出口IPv6流量使用情况，通过对比分析和占比趋势，评估IPv6的推广效果，推动网络从IPv4向IPv6的平滑过渡，确保IPv6流量占比达到要求，并逐步提升未来IPv6流量。



三、改造计划建议

序号	阶段划分	工作安排	工作内容	时间节点	实现成果	备注
1	前期准备阶段	IPv6 地址规划	1.协调开通 IPv6 链路带宽，获取分配 IPv6 内外地址信息。 2.根据上层或向运营商分配 IPv6 私网地址按规划规则进行精细化子网划分。		完成 IPv6 规划形成表格	
2		具体实施方案确认	制定详细的系统调试实施方案落实并确认。		形成具体实施方案	
3		现有环境备份	将现网所有设备的配置进行备份留存。		备份现有配置，用于失败后回滚	
4	实施阶段	硬件设备 IPv6 兼容性优化	针对内部终端、网络设备、安全设备进行兼容性优化工作，包含硬件更新迭代、软件升级、补丁修复、功能授权补全等。		软硬件设备升级到指定版本	
5		互联网出口区域改造	1.配置出口防火墙 IPv6 链路外联，建立内外网 IPv6 地址互联，配置相关地址安全策略、转换策略。 2.核心设备连接链路启用 IPv6，新建 IPv6 地址互联，配置现有 VLAN 下同时支持 IPv4/IPv6 地址转发。 3.办公网接入层 IPv6 调试，针对接入层非 PoE、POE 交换机配置 IPv6 地址协议透传，配置 IPv6 无线下发 SSID 信号。		完成办公无线网络 IPv6 接入公网互联	

序号	阶段划分	工作安排	工作内容	时间节点	实现成果	备注
6		无线网络双栈改造	4 部署 DHCPv6 服务器，支持 IPv6 地址动态下发。 5.部署 DNS 服务器兼容 IPv4/IPv6 双栈架构，配置 IPv6 业务地址解析规则与递归解析路径，解析结果优先向解析请求端返回 IPv6（AAAA 记录）地址。 6.休息日进行启用 IPv6 解析规则，进行业务割接测试。 7.检测业务是否正常运作。 8.检测 IPv6 改造是否成功。			
			2.核心设备连接链路启用 IPv6，新建 IPv6 地址互联，支持 IPv4/IPv6 地址转发。 3.周末进行割接。 4.检测业务是否正常运作。 5.检测 IPv6 改造是否成功。			
7		有线网络 IPv6 下发与业务割接	1.配置 DHCPv6 服务器办公网段动态地址分配，下发 IPv6 地址掩码、网关、DNS。 2.配置 DNS 服务器华为云业务域名 IPv6 业务地址解析规则，解析结果优先向解析请求端返回 IPv6（AAAA 记录）地址。 3.休息日进行启用 IPv6 解析规则，进行业务割接测试。 4.办公生产业务系统访问流量割接。		完成办公有线网络 IPv6 接入生产业务访问、公网业务访问互联	
8	试运行阶段	整体调优	1.网络安全设备策略调优。 2.核心互联路由配置优化。 3.DNS 服务器解析调优配置优化。		完成办公无线网络 IPv6 接入公网互联	

四、常见问题和解决方案

（一）网络设备问题

（1）出口设备IPv6流表不足导致卡顿丢包，大部分出口设备（例如防火墙）在转发IP流量时，IPv4和IPv6采用不同的表项，对应的性能容量也不同，常规网络设备的性能优化偏向于支撑IPv4的高并发流量，对于IPv6可能存在表项和容量较少。

解决方案：改流表设备容量，调高线卡IPv6流表容量或 更换高性能出口网络设备。

（2）互联网出口设备使用ADSL拨号链路，可能出现因设备不支持ADSL拨号后获取的IPv6地址，导致IPv6无法开通。

解决方案：ADSL先通过光猫进行拨号，在拨号获取IPv6地址后，再通过动态地址分配给内网设备。

（3）终端设备软件版本老旧，例如Android系统8.0、鸿蒙OS系统3.0、IOS 9及之前的版本是不支持IPv6的。

解决方案：通过各类终端设备自带的软件更新功能，将系统更新至最新版本。

（二）DNS 解析问题

（1）DNS服务器不回应IPv6的AAAA解析请求，终端等待AAAA回应超时，导致IPv4和IPv6双栈访问慢问题。

解决方案：更换支持AAAA请求响应的DNS服务器，比如61.128.128.68、114.114.114.114等。

（2）终端配置IPv6双栈后，访问互联网的流量优先走IPv6，出口负载均衡失效造成链路拥塞。

解决方案：通过DNS设置出口链路调度，配合策略路由、用户路由以及应用路由等方式进行IPv4/6流量调度实现负载均衡。

（三）地址分配问题

（1）部分安卓系统手机不支持DHCPv6，导致部分用户无法获取IPv6地址。

解决方案：对用户端开启无状态IPv6地址获取。

（2）IPv6路由选路问题，若企业存在多家运营商出口链路，可能出现流量负载不均问题，即大部分流量走向支持IPv6的出口专线。

解决方案：配置出口设备NAT66或者IPv6策略路由，结合DNS智能调度实现多链路的负载均衡效果。

（四）安全问题

（1）IPv4与IPv6均存在共性的安全问题，且由于IPv6协议头部包含一些特有的标签及扩展头部，因此IPv6还有一些特有的威胁攻击方式（如：RH0攻击、洪水攻击），在此基础上，需要一并考虑IPv6安全。

解决方案：部署IPv6后，需要及时对安全设备进行防护策略调整，避免出现安全事件。

（2）IPv6部署后，内网主机都直接暴露在互联网上，没有NAT保护，需要注意网络的安全防护。

解决方案：配置出口设备NAT66，增加IPv6安全防火墙设备和策略。

参考文献

1.RFC 3633——DHCPv6中的IPv6前缀选项
2.RFC 4192——不设定割接日期，为一个IPv6网络重新分配地址的步骤
3.RFC 6879——IPv6企业网地址重新分配的场景、注意事项和方法
4.RFC 6177——终端站点IPv6地址分配
5.RFC 4241——一种IPv6/IPv4双栈互联网接入层模型

金融行业勒索病毒防御评估研究

施勇、张涵 | 上海霞安信息科技有限公司

摘　要： 勒索病毒已成为全球最严重的网络安全威胁之一，尤其对金融行业这种高数据价值、高合规要求的领域影响尤甚。本文梳理了勒索软件的发展趋势与典型攻击流程，结合金融行业实际案例分析其攻击特征与业务影响，提出适用于金融机构的多维度勒索防御评估框架，涵盖特权账户、攻击面管理、网络与终端防护、数据加密备份、日志监控、供应链及云安全等关键领域，旨在帮助金融机构识别防御薄弱环节并提升整体安全韧性。

关键字： 勒索病毒、监管要求、防御评估

一、背景

近年来，勒索病毒已发展为全球网络安全的核心威胁。据权威报告显示[1]，2024年全球网络攻击量同比激增44%，攻击模式呈现技术升级与产业化趋势：攻击者利用生成式AI实现加密算法动态优化和攻击链智能生成，传统防御体系面临失效风险；攻击模式从单纯加密转向“数据泄露+勒索”双重勒索，医疗、教育等行业成为重灾区；勒索软件即服务的普及使攻击门槛大幅降低，而抗量子计算勒索技术的预测更凸显未来防御的严峻性。在中国，勒索攻击呈现高度本土化特征，2025年Weaxor勒索家族利用OA系统漏洞入侵十余家企业[2]，“银狐”木马伪装成税务文件定向攻击财税人员[3]，国内捕获的勒索病毒中30%已采用AI技术提升攻击效率，加之内部人员作案风险（如某医疗机构前工程师植入木马事件），形成内外交织的复杂威胁生态[4]。

金融行业因其数据高价值性与业务强连续性，成为勒索病毒的核心目标。一方面，金融机构承载着客户账户、交易记录等敏感数据，一旦泄露可能引发系统性风险[5]；另一方面，行业需满足《银行保险机构数据安全管理办法》中“零数据丢失”（RPO≈0）和“30分钟快速恢复”（RTO<30分钟）的严苛要求。

面对勒索攻击的智能化、定向化演进，传统防火墙、终端检测等防护手段已显不足。研究显示，大部分勒索攻击利用已知未修复漏洞，暴露出补丁管理机制的缺失；金融机构备份系统常被二次攻击渗透，亟需验证“真隔离、可验证”的数据保护能力。在此背景下，构建体系化勒索病毒防御评估框架成为刚需。

二、勒索攻击介绍

（一）常见攻击过程

勒索攻击通常遵循一个高度组织化的多阶段流程，旨在实现对目标系统的全面控制并最大化勒索价值。攻击者常通过钓鱼邮件、账户盗用或利用已知与零日漏洞进入目标环境，通常结合鱼叉式钓鱼与社会工程以提升初始渗透成功率。在获得初步访问权限后，攻击者会维持长期控制，例如通过植入后门、使用远程管理工具及隐蔽通信通道，以避免被检测。接着，他们会进行权限提升，利用系统漏洞、凭证转储或滥用域管理策略以获得更高权限，并进一步扩大其对环境的掌控。随后，通过横向移动技术（如Pass-the-Hash、利用远程桌面协议及漏洞横向渗透），攻击者会探索和控制更多系统，建立对网络的全局视图[6]。

在全面渗透的过程中，攻击者往往伴随数据侦察和窃取行为，提取有价值的商业数据、敏感信息及凭证，以便后续勒索或在暗网出售。数据外泄往往先于勒索行为，使攻击者即便在加密操作前已获得足够的筹码。接下来，勒索软件开始在网络内部进行自传播，常利用Windows组策略、批处理脚本或内网共享，实现大范围快速感染。最终，攻击者对关键数据进行加密，并在部分场景下重复窃取敏感信息，双重勒索成为近年来的主要手段，即威胁公开数据与恢复数据解密双重施压。整个流程高度模块化，攻击者可根据目标防御强度与业务价值灵活调整策略，以提高成功率和勒索金额[7]。

图1 常见勒索攻击流程

（二）国内外金融机构勒索攻击事件

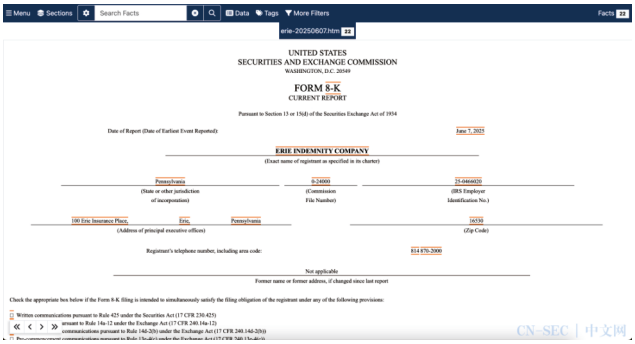
1、国内某金融机构双重勒索攻击

2023年11月8日，某金融机构全资子公司遭遇LockBit 3.0勒索软件攻击。攻击者通过未修复的Citrix Bleed漏洞（CVE-2023-4966）侵入系统，该漏洞存在于Citrix NetScaler网关设备中，允许攻击者绕过所有身份验证机制直接获取系统控制权。由于未及时修补漏洞，攻击导致部分系统中断，美国国债结算业务受阻。为维持交易连续性，被迫采用物理媒介传递结算数据，这一应急措施虽避免了交易全面停滞，但仍扰乱了美国国债市场流动性，引发证券业与金融市场协会的监管审查。事件发生后，LockBit在加密通讯平台Tox上公开宣称对攻击负责。

2024年9月，勒索组织Hunters对该金融机构伦敦分部发起新一轮攻击，宣称窃取6.6TB敏感数据，包括客户交易记录及内部通信文件。与2023年攻击不同，此次该金融机构伦敦分部通过实时威胁监测系统发现异常活动，立即启动应急隔离机制，有效限制了攻击横向扩散，未造成业务中断。在声明中强调“已采取强化安全措施并启动彻底调查”，但未披露是否涉及数据泄露实质性证据或赎金谈判。

2、国际金融机构遭遇的勒索事件

2025年6月初，美国两大保险巨头——伊利保险（Erie Insurance）和费城保险（PHLY）在72小时内接连遭勒索组织 Scattered Spider 的协同攻击。攻击者利用未公开的漏洞侵入系统，导致客服、理赔及内部系统全面瘫痪，超15TB敏感数据（含客户信息、财务记录及商业文件）被窃取并在暗网公开叫卖。伊利保险作为财富500强企业，持有600万份保单，事件直接致其股价下跌4.56%；费城保险系统瘫痪超96小时，核心业务崩溃，客户服务完全失效，紧急救援服务中断，员工无法办公，运营几近停摆[8]。


Menu Sections Search Facts Data Tags More Filters Facts 22

erie-20250607.htm 22

UNITED STATES
SECURITIES AND EXCHANGE COMMISSION
WASHINGTON, D.C. 20549

FORM 8-K
CURRENT REPORT

Pursuant to Section 13 or 15(d) of the Securities Exchange Act of 1934

Date of Report (Date of Earliest Event Reported): June 7, 2025

ERIE INDEMNITY COMPANY
(Exact name of registrant as specified in its charter)

Pennsylvania	0-24000	25-0460020
(State or other jurisdiction of incorporation)	(Commission File Number)	(IRS Employer Identification No.)

100 Erie Insurance Place, Erie, Pennsylvania	16530
(Address of principal executive offices)	(Zip Code)

Registrant's telephone number, including area code: 814 870-2000

Not applicable
Former name or former address, if changed since last report

Check the appropriate box below if the Form 8-K filing is intended to simultaneously satisfy the filing obligation of the registrant under any of the following provisions:

☐ Written communications pursuant to Rule 425 under the Securities Act (17 CFR 230.425)
☐ Soliciting material pursuant to Rule 14a-12 under the Exchange Act (17 CFR 240.14a-12)
☐ Pre-commencement communications pursuant to Rule 14d-2(b) under the Exchange Act (17 CFR 240.14d-2(b))
☐ Pre-commencement communications pursuant to Rule 13e-4(c) under the Exchange Act (17 CFR 240.13e-4(c))

CN-SEC | 中文网

图2 伊利保险公司向美国证券交易委员会报告网络安全事件

三、勒索防御评估

（一）评估目标

通过对关键资产及合作环节开展系统性安全评估，优化现有技术架构与防护工具，确保各项勒索防御措施能够有效阻断勒索软件攻击，并在安全事件发生时实现快速响应与影响最小化。在评估实施过程中，确保对现有业务运营零干扰，避免业务中断或额外负担。通过本次评估，能够全面验证企业现有勒索防御体系的有效性，评估应急响应机制在真实情境下的可用性与效率，强化漏洞管理、数据备份与恢复等关键能力，进一步提升整体防御水平，确保在遭遇勒索事件时能够高效处置、快速恢复并最大程度降低潜在损失。

（二）评估范围

勒索防御评估覆盖数据中心、集团总部及其下属分子公司，以及与业务紧密相关的供应商体系，旨在对不同类型场景的安全防御现状开展系统化分析，识别关键风险点并提出针对性改进策略。根据对象特性，评估范围划分为数据中心、办公区域和供应商三大类别，通过文档审计、人员访谈、技术验证等方式，以确保评估结果的全面性与可操作性。

在数据中心层面，重点对数据防护、物理安全、网络防御与灾备能力进行技术性验证。通过评估核心数据加密机制与多层级备份策略，验证其在面对勒索软件加密与破坏时的抵御能力；通过检查物理访问控制、入侵报警及监控机制，评估物理安全防护的完整性与有效性；通过验证防火墙、入侵检测与防御系统等关键安全设备的部署及策略配置，分析其对外部威胁的防护能力；同时结合容灾演练与恢复策略，评估数据恢复的可行性与恢复效率，以确保在极端场景下的业务连续性。

在集团办公区域与供应商管理层面，评估重点聚焦终端防护、账户与邮件安全、网络隔离及供应链安全管理。针对办公区域，检查终端设备防病毒、补丁更新及安全配置策略，验证特权账户分配与访问控制机制的有效性，并评估邮件安全网关及反钓鱼策略在防御恶意附件与钓鱼攻击方面的实际防护水平；同时检验办公网络与核心生产系统之间的隔离策略，确保潜在威胁无法横向扩散。针对供应商，重点审查其安全管理体系及合规性，验证合同中安全责任与事件响应条款的完备性，评估远程访问权限的授权与控制机制，并核查日志留存与溯源能力，以防止通过供应链成为攻击入侵的入口。通过上述多维度、差异化的综合评估，可全面识别企业在勒索软件防御能力建设中的不足，为后续防御体系优化与技术改进提供科学依据和实践参考。

（三）评估维度

依据评估覆盖范围及国际通用的勒索防御最佳实践，勒索防御评估将聚焦于多个核心安全维度，以验证企业防范勒索软件攻击的能力，并确保在突发安全事件发生时能够快速响应和恢复业务。这些维度包括特权账户管理、攻击面管理、网络安全架构、终端防护、数据加密与备份、监控与日志管理、供应链与第三方安全、邮件安全、威胁情报能力、

云安全和安全管理制度体系建设等。每一维度均从配置有效性、策略落实和运行状态等方面进行深入检查，以形成系统化的防御能力评估。

1、在特权账户管理方面

重点检查账户权限分配是否遵循“最小权限原则”，包括特权账户是否与普通账户分离、是否定期进行权限审查、是否启用多因素认证、是否存在共享账户或默认账户未禁用的情况。如果企业未能有效控制特权账户，攻击者在初步入侵后可通过凭据窃取与权限提升，快速获得域控制权限，从而全面控制网络环境并在短时间内部署勒索软件，导致防御体系形同虚设。

2、在攻击面管理方面

重点评估企业是否定期进行外部资产扫描，全面识别互联网暴露的端口、服务和应用，验证是否及时修补高危漏洞、关闭不必要的端口和服务，并对暴露系统进行强化配置。如果此维度管理不到位，勒索软件可利用未修复的漏洞或弱口令服务（如RDP、VPN设备）作为突破口进入内部网络，成为整个攻击链的起点。

3、在网络安全架构方面

重点检查网络分区与内外网隔离策略是否合理，验证防火墙策略的有效性，评估入侵检测/防御系统与网络流量监控系统的部署与告警响应能力。如果企业的网络隔离和监测薄弱，攻击者可在获得初步访问后快速横向移动，感染更多主机并接触到核心系统与数据资源，极大增加勒索软件攻击的破坏范围。

4、在终端防护方面

评估内容包括终端设备的安全配置基线是否符合要求，操作系统与软件是否及时更新，防病毒与终端检测响应系统是否部署和启用，是否存在未授权设备接入。如果该维度存在缺陷，终端将成为勒索软件的首选入侵入口，攻击者可通过钓鱼邮件或恶意附件在终端设备落地，进而突破企业的外围防线。

5、在数据加密与备份方面

重点验证企业是否对敏感数据进行访问控制和存储加密，备份策略是否覆盖关键系统和数据，备份存储是否与生产环境隔离（离线/异地备份），以及灾难恢复计划是否经过验证性演练。如果缺乏有效的数据加密和独立备份机制，即便攻击者入侵后被发现，企业也难以通过备份快速恢复业务，只能被迫支付赎金，造成更严重的经济与声誉损失。

6、在监控与日志管理方面

重点检查企业是否建立了覆盖全域的日志采集、集中存储与分析机制，包括身份认证日志、文件访问与修改记录、网络流量日志及关键系统安全事件日志等，并验证日志留存周期、完整性保护及实时分析能力。如果缺乏有效的监控与日志管理，攻击者在入侵和实施勒索操作过程中可能长期处于隐匿状态，导致异常行为未被及时发现，企业难以及早阻断攻击链条，增加勒索攻击成功率与加密范围。

7、在供应链与第三方安全管理方面

重点检查与供应商、外包方及合作伙伴的安全协作机制，评估第三方接入企业系统的身份认证与访问控制策略，验证合同或协议中是否明确安全责任与事件响应要求，检查数据交换和接口调用的加密传输及日志审计。如果该维度管理薄弱，攻击者可通过供应链渗透或利用第三方系统漏洞获得企业内部访问权限，从而绕过外围防御直接接触核心数据和系统，成为勒索软件入侵的重要突破口。

8、在邮件安全防护方面

重点检查企业邮件系统是否部署反钓鱼网关、反垃圾邮件引擎、邮件附件沙箱及URL实时检测与阻断等多重防护措施，验证员工邮件收发安全策略及邮件访问控制。如果防护不到位，钓鱼邮件和恶意附件可能轻易绕过检测，在终端落地并触发勒索程序，成为攻击链条最常见的初始入口。

9、在威胁情报能力建设方面

重点评估企业是否具备主动收集、分析及利用威胁情报的能力，是否与行业联盟、安全厂商或政府安全机构开展情报共享，是否通过自动化情报集成与安全设备联动实现快速响应。如果缺乏有效的威胁情报机制，企业难以及时了解新兴勒索软件攻击手法与攻击指示器，导致防御策略滞后，增加被新型变种攻击成功的概率。

10、在云环境安全方面

重点检查云平台的身份与访问控制策略（如最小权限、多因素认证）、工作负载防护（如云主机防病毒、WAF与运行时防护）和数据存储加密与安全审计机制。如果云环境安全防护缺失，攻击者可通过云账户劫持、云服务漏洞利用或配置错误实现勒索程序的部署与云端数据的加密或窃取，扩大勒索攻击影响范围。

11、在安全管理制度建设方面

重点评估企业是否建立了完善的网络安全管理体系，包括定期的风险评估、勒索防御专项策略、员工安全培训及应急响应预案设计和演练。如果管理制度缺失或执行不到位，整体防御工作将缺乏系统性与持续性，使技术防护措施难以落地，企业在应对勒索攻击时可能反应迟缓、协调不力，导致事件损害扩大。

12、在实际评估过程中

将综合运用文档审计、人员访谈及技术验证等多种方法对不同维度进行深入检查，以确保评估结论的全面性与可操作性。例如，在特权账户管理维度，通过文档审计核查账户管理制度、权限审批流程及多因素认证策略是否符合合规要求，并在技术验证中抽样检查域控制器与核心系统的账户配置；在攻击面管理与网络安全架构维度，结合渗透测试与漏洞扫描，对外部暴露端口、VPN服务及网络隔离策略进行验证，并与网络管理人员访谈，了解防火墙策略更新及应急响应执行情况；在终端防护和邮件安全方面，通过现场访谈 IT 运维人员掌握补丁管理及邮件安全网关的部署现状，并在技术层面抽测终端安全基线与钓鱼邮件拦截效果；在数据加密与备份维度，结合文档审计查阅备份策略、加密规范及演练记录，并通过技术验证核实备份存储的隔离性及恢复可行性；在供应链和云安全方面，通过访谈供应商管理部门与云运维人员了解第三方安全管理与云平台访问控制措施，并抽查合同中安全条款及云资源配置合规性。通过上述多层次、交叉验证的方法，可有效识别企业在制度、人员执行与技术防御等方面的薄弱环节，形成对各维度防御能力的系统性评估。

通过对上述维度的系统化评估，可有效揭示企业在监测响应、外部协作、邮件防护、情报利用、云资源管理及制度建设等方面的薄弱环节，为勒索防御体系的全面提升提供有力的技术依据与管理支撑。

（四）安全加固建议

结合金融行业高数据价值、高合规要求及强业务连续性的特点，针对勒索防御评估各个维度可提出如下安全加固与改进建议：在特权账户管理方面，应全面落实最小权限原则，启用多因素认证并对高风险账户实施集中管控和动态审计，防止凭据被盗导致的横向渗透；在攻击面管理上，需定期开展外部资产扫描和渗透测试，及时关闭不必要端口、修补高危漏洞，尤其关注互联网暴露的金融服务系统；在网络与终端防护方面，建议优化分区隔离策略，部署行为分析型入侵防御系统和终端检测响应，确保对内部横向移动及异常行为的实时阻断；在数据加密与备份环节，应对敏感金融数据实行分级加密存储，建立异地、离线与不可变备份机制，并通过定期演练验证恢复可行性；在监控与日志管理上，构建统一的安全信息与事件管理平台，实现对核心交易系统、客户数据访问及网络流量的全量监控与智能分析；在供应链与云环境安全方面，强化第三方安全准入审查和合同约束，确保外部合作方符合金融行业安全标准，并对云资源实施精细化身份与访问管理及配置合规性检测；此外，应建立威胁情报联动机制，借助行业联盟和监管机构的情报共享提前防范新型勒索威胁，并通过制度化安全管理与应急演练提升整体快速响应与恢复能力。

通过上述多层次的加固措施，金融机构可显著提升勒索防御韧性，降低因攻击导致的业务中断、数据泄露及合规风险。

四、总结

勒索软件正向数据窃取与信息公开等复合型勒索演进，金融行业因其高价值数据与关键业务连续性需求成为主要攻击目标。本文通过分析金融行业案例，总结了金融机构勒索防御的监管要求与实践重点，并提出多维度的防御评估框架。研究表明，金融机构需在技术防护、威胁情报、合规管理及应急响应等方面协同提升，构建覆盖预防、检测与快速响应的全链条防御体系。

参考文献

1.Check Point.《2025年安全报告：网络攻击骤增44%》. 2025.

2.360.《勒索毒王Weaxor利用AI攻击告警》. 2025.

3.央视新闻.《“银狐”木马病毒变种预警》. 2025.

4.交通银行.《金融数据安全新范式构建》. 2025.

5.Dark Reading.《中国黑客后门入侵Juniper路由器》. 2025.

6.Melnick, D., et al. (2022). Ransomware Attack Vectors and Defense Strategies. Journal of Cybersecurity Research, 15(3), 45–58.

7.Coveware. (2024). Quarterly Ransomware Report Q1 2024. Coveware Inc.

8.Kapko, M. Aflac duped by social-engineering attack, marking another hit on insurance industry. CyberScoop, 2025.06.20.

2025RSAC大会解析：众声汇聚，共探全球网络安全新趋势

江爱军、王伟涛、盛浩月 | 奇安信科技集团股份有限公司

摘　要： 2025年4月，全球最具影响力的网络安全盛会——RSAC（RSAConference）如期举办。这场汇聚了42,500名与会者和700家参展商的行业巅峰盛会，以“ManyVoices,OneCommunity（众声汇聚，同心共筑）”为主题，聚焦AI技术飞速发展带来的安全挑战，通过多元化的交流与协作，为全球网络安全领域指明了新的方向。本文将解析2025年RSAC大会的核心内容，从大会概况、创新成果到技术趋势，呈现全球网络安全行业的最新动态。

关键字： RSAC2025、Agentic Ai、零信任架构、后量子密码、供应链安全、数据安全

图1 RSAC大会现场照片

一、大会概览：从“多元”到“共识”的安全协作之路

本届RSAC大会的主题“ManyVoices,OneCommunity”，直指当下网络安全领域的核心挑战：在技术快速迭代，尤其是AI的“跳脱式发展”、威胁日益复杂的背景下，单一主体的力量已难以应对全局风险。主题强调“多元”与“协作”——无论是技术路线的差异、行业需求的不同，还是地域法规的区别，都需要通过开放的交流寻求共识，最终构建一个能够共同抵御风险的“安全社区”。

这一主题并非偶然，而是RSAC大会历年主题演变的必然结果。回顾历年核心主题，可清晰看到行业重心的迁移：

历年主题	简要说明
2013年“知识安全”	聚焦基础安全认知的构建
2014年“分享。学习。安全”	强调信息互通的重要性
2015年“变化：挑战当今的安全思想”	直面技术变革带来的冲击
2016年“连接保护”	关注互联互通时代的安全边界
2017年“机会的力量”	挖掘安全领域的技术红利
2018年“现在很重要”	凸显实时响应的价值
2019年“更好”	追求安全能力的持续优化
2020年“人类要素”	回归安全中的人为因素
2021年“韧性”	强调系统对抗风险的恢复能力
2022年“转变”	聚焦安全架构的重构
2023年“一起更强”	初步提出协作理念
2024年“可能性的艺术”	探索技术创新的边界
2025年“众声汇聚，同心共筑”	将“协作”推向新高度，明确了多元化主体共建安全生态的核心方向

图2 RSAC大会历年主题演变

二、创新沙盒大赛：安全新势力的崛起与资本的聚焦

作为RSAC大会的“创新风向标”，2025年创新沙盒大赛迎来了第20个年头。这场专为网络安全初创企业设立的赛事，今年吸引了超过200家公司参赛，较2024年增长40%以上，竞争激烈程度创下历史新高。更值得关注的是，主办方首次推出“投资激励计划”——所有进入决赛的十强企业每家获得500万美元投资，这一举措不仅彰显了资本市场对安全创新的高度认可，更预示着一批突破性技术即将加速落地。

（一）冠军亮点：ProjectDiscovery的自动化攻击面监测

2025年创新沙盒大赛的冠军由初创公司ProjectDiscovery摘得，其被评为“2025年度最具创新初创企业”。该公司出身开源社区，核心产品基于开源工具“Nuclei”，能够实现企业攻击面的自动化监测与漏洞的快速修复。

Nuclei的“模板驱动”理念极具创新性，用户无需学习复杂语言，通过编写或引用YAML模板，即可灵活覆盖Web、API与多云环境中的最新威胁。社区版每月执行扫描超5000万次，免费向公众开放；企业版则在此基础上提供云端可视化管理平台，可自动同步资产、调度扫描，并生成直接用于修复的报告。更重要的是，其与Jira、Slack等协作工具的无缝对接，实现了“扫描—告警—跟进”的全流程闭环，大幅提升了安全运营效率，尤其契合DevSecOps团队的需求。

（二）十强企业：AI与安全的深度融合成主流

除冠军外，其余企业同样展现了前沿的技术视野，其中7家的核心产品与AI相关，印证了AI在安全领域的核心地位。

创新沙盒大赛十强概况		
序号	公司名称	亮点
1	Aurascape	2023 年在硅谷设立的Aurascape公司，由数位在 Palo Alto Networks、Google 与 Amazon 拥有丰富实战经验的安全与 AI 专家创立，定位于“AI 原生安全”的先行者。公司在2025年初完成了5000 万美元 A 轮融资，由 Mayfield Fund 与 Menlo Ventures 领投，并迅速与多家金融与制造业巨头达成合作试点，目标是成为企业在 Copilot、代码助手等 AI 工具使用过程中，保护源代码与敏感数据的中坚屏障。 Aurascape 的 AI Activity Control 平台摒弃传统的模块化防护思路，将可视化威胁检测、自动化修复和合规审计功能融合于同一个引擎中。该系统能够实时识别数千种多模态 AI 应用的交互行为，并在检测到异常时即时阻断或引导修复，无需人工干预就能覆盖文本、语音、图像与代码四大数据类别，为企业构建了真正意义上的“影子 AI”安全壁垒。
2	CalypsoAI	成立于 2018 年的 CalypsoAI公司，总部位于华盛顿特区，专注于 AI 推理层的安全防护，已经从 Paladin Capital、Lockheed Martin Ventures 等机构筹集了逾 3800 万美元资金。团队将自身定位为“Inference Red-Team-to-Defend-to-Observe”闭环方案的开拓者，致力于在模型实际运行时发现对抗攻击、快速修复漏洞，并将监控与告警原生化地接入企业既有的 SIEM/SOAR 平台。 从客户反馈来看，CalypsoAI 的平台能在毫秒级时间内对输入和输出进行动态审查，自动化调整模型的输出策略以规避敏感信息泄露或恶意利用。同时，其无缝兼容多家主流大模型与 Agent，使得企业在引入不同厂商技术时，不必为安全管控再额外搭建第三方中台，在“高效与安全”之间实现了近乎完美的平衡。
3	Command Zero	Command Zero 公司的创始团队汇聚了来自 Cisco、IBM 与 McAfee 的资深安全专家，瞄准 Tier-2/3 级别的威胁狩猎与调查瓶颈。他们设计了一套自治式、AI 辅助的安全调查引擎，能在海量日志与情报中自动抽取线索，并生成可执行的问答式调查流程，将传统需要数小时甚至数天的根因分析压缩到分钟级。 在实际应用中，安全团队只需将日志或告警输入 Command Zero 平台，系统便会结合内置的威胁知识编码与自然语言模型，迅速对可疑事件进行初步评级、关联历史案例，并给出下一步验证动作建议。
4	EQTY Lab AG	总部位于瑞士苏黎世的 EQTY Lab AG公司，自诞生之初就将密码学与 AI 诚信（Integrity）紧密结合。2025 年推出的 Verifiable Compute 公证系统，通过硬件级同态加密与零知识证明技术，让模型在训练与推理阶段的每一次决策都拥有不可篡改的审计凭证。该方案已获 Intel 与 NVIDIA 联合背书。 EQTY 的解决方案不仅能保证 AI 系统在高风险场景下的决策透明度，还为监管机构提供了便捷的审计报告导出接口。在生命科学、公共事业等对可追溯性要求极高的行业，Verifiable Compute 已被视为开启“可信 AI”时代的关键基石。
5	Knostic	2023 年成立的 Knostic公司，由安全大牛 Gadi Evron 与 Sounil Yu 联手打造，首创“Need-to-Know”访问控制框架，专门为大模型和 AI Agent 场景下的敏感信息访问提供动态授权。它通过企业内部知识图谱与自然语言处理技术相结合，能够在对话或请求发起时实时识别所涉数据的敏感度，并根据用户角色、业务场景与风险等级自动决定是放行、脱敏还是拒绝。 相比于传统的一刀切式脱敏，Knostic 的方案更注重上下文感知，把细粒度策略与零信任架构融合，实现对 Copilot、Glean 等常用 AI 工具的原生安全扩展。目前已有多家科技、医疗与政府机构在生产环境中部署，明显减少了因误用敏感数据引发的合规风险。
6	Metalware	Metalware公司 的团队成员多来自知名硬件安全实验室，他们将自动化模糊测试技术首次引入固件安全领域。通过自主研发的拆包引擎与 QEMU 模拟环境，结合改良版 AFL++ 智能变异算法，Metalware 平台能够对 BIOS、UEFI 及各类嵌入式设备固件进行持续、深度的模糊测试。每当发现潜在缺陷，系统便会结合实时威胁情报，自动生成补丁优先级与修复建议，帮助硬件厂商在产品出货前完成安全加固。 在 CI/CD 流水线中无缝接入后，Metalware 已帮助多家物联网与工业控制系统设备制造商将固件安全测试从手动推迟式，转变为与开发同步进行的常态化流程，有效堵住了底层固件成为攻击入口的隐患。
7	MIND	西雅图的 MIND 平台打破了传统数据丢失防护（DLP）与内部风险管理（IRM）各自为政的格局，将二者通过深度学习模型与大规模行为分析统一到同一套自动化流程中。平台能对 SaaS 应用、GenAI 工具、邮箱和终端设备上的敏感数据进行零假阳性的全量扫描，并在检测到风险倾向时即时实施阻断或提醒；同时结合用户画像，预判潜在的内部威胁，并自动触发后续合规与调查流程。 这种“一站式”的托管式服务模式，让中小团队也能以极低门槛享受大型企业级的 DLP+IRM 能力。
8	ProjectDiscovery	ProjectDiscovery 公司出身开源社区，旗下 Nuclei 扫描引擎凭借其“模板驱动”理念，在 Web、API 与多云环境的漏洞检测中大放异彩。社区版免费且月均执行扫描超 5,000 万次，企业版又在此基础上提供了云端可视化管理平台，能够自动同步资产、调度扫描并生成可直接用于修复的报告。 对于 DevSecOps 团队而言，它无需学习复杂语言，只要编写或引用 YAML 模板，就能灵活覆盖最新威胁；而云平台则将扫描结果与 Jira、Slack 等协作工具打通，实现扫描—告警—跟进的全流程闭环，大大提升了安全运营效率。
9	Smallstep	Smallstep公司 自 2016 年起专注于设备身份管理与零信任，实现了与 Google、Apple 共同制定的 ACME Device Attestation 标准无缝对接。其平台通过 TPM 与 Secure Enclave 技术，为每台设备颁发唯一、不容伪造的身份凭证，并根据网络环境、用户角色等动态因素实时调整访问策略，让 Wi-Fi、VPN、ZTNA 乃至云 API 的访问都能达成“外观简单、内核强安全”的体验。 基于开源社区积累，Smallstep 同时提供 Cert-Manager、Istio 等云原生项目的深度集成插件，使企业在部署零信任架构时，既能享受开箱即用的便捷，也能通过自定义扩展满足特殊需求。
10	Twine Security	由 Claroty 前高管发起的 Twine Security公司，主打“数字安全员工”理念，首款产品 Alex 结合 NLP 与安全知识库，能够在无需人工干预的情况下完成从账户配置到权限审计，再到合规报告的一系列 IAM 流程。随着使用反馈的不断累积，Alex 在自适应学习方面表现越发成熟，能够针对不同组织的安全政策与流程定制最优执行策略。 该方案无需二次开发，主流 IAM 系统打通，已在几家中大型企业进行大规模试点，帮助它们在 IAM 人才短缺的背景下维持高质量的权限治理和合规性。

图3 2025年创新沙盒大赛十强概况

三、展商规模与分布：新兴领域的崛起与热点迁移

2025年RSAC大会的展商规模保持稳定增长，共有约700家网络安全厂商参展，数量与2024年基本持平但略有提升。值得注意的是，大会专门设立的“EarlyStageExpo”初创企业展区汇聚了76家新创公司，较往年显著增加，反映出网络安全领域的创业活力持续高涨。

从技术焦点来看，行业呈现出明显的“新旧交替”特征，过去几年的热门领域如零信任架构，虽仍为核心议题，但热度较高峰时期有所回落。这并非因为其重要性下降，而是随着技术的成熟，企业已从“理念探讨”转向“规模化落地”，市场进入平稳发展阶段。

新兴领域如供应链安全、SaaS安全则获得了更多厂商的关注。随着数字化进程的深入，企业对第三方工具、云服务的依赖度持续提升，供应链与SaaS生态的安全风险日益凸显，成为行业新的增长点。

四、前沿技术趋势分析：AI驱动下的安全生态重构

2025年RSAC大会的技术讨论围绕五大核心领域展开，而AI作为贯穿其中的主线，正在深刻重构网络安全的技术生态。

（一）AI安全：从“辅助”到“自主”的范式跃迁

AI毫无疑问是本届大会的“绝对主角”。根据主办方统计，2025年提交的演讲主题中，AI相关内容占比高达40%，上文提到创新沙盒十强中也有7家聚焦AI安全，足以见其重要性。

1、AgenticAI：自主化的安全变革

AI正从“建议者”升级为“执行者”，即“AgenticAI（自主AI代理）”。这类系统具备多智能体协作能力（如Torq的Multi-AgentRAG架构）和实时推理能力，可在威胁检测、漏洞修复、事件响应全流程实现闭环自动化。例如，多个专门的AIAgent可协作执行半自主或完全自主的安全运营工作流（即“AgenticSOC”），大幅提升威胁响应的速度与准确性。

2025年以来，Agentic AI被Gartner列为年度十大技术趋势之首，同时，市场研究机构Markets and Markets预测，Agentic AI市场将在2025年达到138.1亿美元，并在2032年年底跃升至1408亿美元，期间复合年增长率达到39.3%。

在此背景下，全球科技巨头都在投入巨资推动Agentic AI技术演进。日前，亚马逊全球副总裁、亚马逊云科技大中华区总裁储瑞松在亚马逊云科技中国峰会期间表示，AI的发展已经来到拐点，如今我们正处在Agentic AI爆发的前夜。他表示，过去两年里，大模型与智能体的结合已经从实验室正在走入产业现实。驱动这股Agentic AI浪潮的背后，有三大驱动力。一是大模型能力的快速提升，二是MCP（模型上下文协议）和A2A（Agent-to-Agent）协同协议的出现，三是基础设施成本的大幅降低。

2、应用安全的AI赋能

Checkmarx、Semgrep等公司推出的AgenticAI应用程序安全态势管理产品，通过自主分析代码漏洞、优化检测规则，显著减少了误报率，为开发人员和安全工程师节省了大量时间。AI在应用安全中的角色，已从“被动扫描”转向“主动防御”。

（二）零信任架构：从“理念”到“规模化落地”

零信任架构已不再是理论探讨，而是成为企业安全架构的核心组成部分。大会强调，零信任的部署已从“试点”转向“规模化应用”，尤其在远程办公、混合云环境和多设备接入场景中，零信任成为默认安全模型。

1、零信任与AI的深度融合

一方面，AI赋能零信任：通过机器学习分析用户行为模式，可实时识别异常访问请求，并自动调整访问权限，让动态访问控制更精准、高效。

另一方面，零信任守护AI安全：随着AI系统的广泛应用，零信任架构被用于保护AI模型和数据，例如对AI训练数据的访问进行细粒度控制，防止模型被篡改或数据泄露。

2、身份安全的扩展：非人类身份（NHI）治理

零信任的核心是“永不信任，始终验证”，而身份验证是第一步。如今，身份的范围已从用户扩展到设备、应用和服务（即“非人类身份”）。通过多因素认证（MFA）、持续认证和基于风险的访问控制，企业可确保只有合法身份才能访问资源。例如，当用户从陌生网络登录时，系统会自动触发额外的认证步骤，实现“动态授权”。

（三）后量子密码：应对量子计算的“提前布局”

据公开资料显示，2024年8月，美国国家标准与技术研究院（NIST）正式发布了首批3项后量子加密标准，以保障互联网通信免受量子计算机攻击。另据报道，2024年9月，日本东芝数字解决方案公司Toshiba、量子通信技术公司SpeQtral与新加坡科技工程公司签订协议，合作推进包含后量子密码技术在内的量子安全通信和量子安全数据存储的发展。

随着IBM、谷歌、中国“九章”等量子计算机的快速发展，传统非对称加密算法（如RSA、ECC）面临被Shor算法快速破解的风险。更严峻的是，“先存储后解密”攻击（即攻击者现在收集加密数据，待量子计算机成熟后破解）已对长期数据安全构成现实威胁。

科技前沿

高技术前沿

科技云

空地两用无人机

后量子密码技术——

重构数字世界的安全边界

后量子密码技术应运而生

为逐步领域的数字化进程提供保护

多种技术路线构筑坚固防线

"鸟翼"无人机

"变翼"无人机

红外感知技术：人类穿透黑暗的"第三只眼"

科学的历程

图4 后量子密码技术：重构数字世界的安全边界

1、混合加密：过渡期的主流策略

由于后量子密码算法（如基于格、哈希、编码的算法）在性能和兼容性上仍有改进空间，RSAC2025提出"混合加密方案"（如RSA+Kyber）将成为过渡期的核心策略——通过传统算法与后量子算法的结合，平衡安全性与可用性。

2、与隐私计算的融合

后量子密码与同态加密、多方安全计算（MPC）等隐私计算技术的结合成为趋势。在医疗、金融领域，基于后量子同态加密的方案已实现数据"可用不可见"，既满足了隐私保护需求，又抵御了量子计算的潜在威胁。

3、金融行业的迫切需求

金融行业因大量依赖传统密码算法（涉及加密传输、数字签名、敏感数据加密等场景），受量子计算的影响范围极广。大会指出，金融机构需逐步将现有密码体系迁移至抗量子攻击的安全体系，这一过程虽复杂，但已刻不容缓。

美国Gartner公司发布的《2025年十大战略技术趋势》预测，到2029年，大多数传统的非对称加密技术将不再安全，量子计算对现有传统加密技术或将造成颠覆性的威胁。从实验室到战场，从虚拟世界到现实生活，后量子密码作为一种方兴未艾的前沿技术正在重构数字世界的安全边界。

（四）供应链安全：从"企业防御"到"生态防护"

供应链安全是本届大会的另一重点议题。随着攻击者从"打企业"转向"打依赖"——通过突破上游厂商影响下游数百万用户，供应链攻击因隐蔽性强、影响范围广、回溯难度大，成为企业面临的重大挑战。

1、从SBOM到AI-BOM：清单化的安全治理

软件物料清单（SBOM）已成为行业标准，但随着AI的普及，"AI-BOM（AI物料清单）"开始出现。AIBOM用于追踪模型权重、训练数据来源、插件依赖等信息，而"ModelProvenance（模型溯源）"则聚焦模型版本迭代历史与审计信息。SBOM与AIBOM不仅是技术工具，更成为合规要求——美国行政令和欧盟CRA法案已推动其强制落地。

2、开源依赖攻击的加剧与应对

多个演讲指出，开源包注入恶意代码已成为常见攻击战术，"拼写欺骗（Typo-squatting）"和"依赖混淆（Dependencyconfusion）"再次成为案例焦点。对此，行业建议采用"软件仓库扫描+安全CI/CD插件+开源组件实时监控"的组合策略，构建全流程的开源供应链防护体系。

（五）数据安全：AI时代的全生命周期治理

当前，多家金融启动AI战略，探索大模型在金融领域的广泛应用随着技术的不断演进，尤其是自然语言处理、机器学习和数据挖掘技术的进步，证券公司通过大模型能够显著提升业务效率、优化服务质量、减少人为风险，并在日益复杂的金融市场中保持竞争力。然而，随着大模型应用的深入，数据隐私保护、合规性问题以及模型本身的安全性等挑战也成为行业亟待解决的难题。

目前已经有多起因AI大模型的不当使用导致的数据安全事件发生，三星的一名员工将机密源代码粘贴到ChatGPT中以检查错误、另一名员工与ChatGPT和"请求的代码优化"共享代码、第三个人分享了一份会议记录，并将其转换为演讲笔记，这些信息曾一度公开在 ChatGPT 上使用。2024年8月，国安部通报，部分涉密人员为节省时间，违规将涉密材料内容输入AI写作小程序，导致国家秘密泄露。通报指出，AI小程序自动采集输入信息用于学习，数据极易被境外情报机构窃取，严重危害国家安全利益。

所以在本届RSAC大会上，数据安全与AI安全、身份安全并列成为大会核心议题，其范围已从传统的加密与访问控制，扩展到生成式AI时代的数据治理、模型保护、数据主权与"AI时代的数据生命周期管理"。

1、数据安全重要性上升的三大原因

（1）AI驱动的数据依赖增强：AI模型依赖海量高质量训练数据，数据的隐私、完整性、可溯源成为AI可信的基础。

（2）法规压力陡增：欧美数据保护法、AI法案、数字主权

法规频出，企业合规已成为“先决条件”。

（3）攻击方式变化：攻击者从“攻击系统”转向“攻击数据”，如数据投毒、脱敏逆推、隐私窃取等新型攻击手段层出不穷。

2、DSPM：数据安全的“全景式管理”

多家厂商推出的数据安全态势管理（DSPM）产品，成为企业应对数据安全挑战的核心工具。这类产品可实现：自动发现企业内所有数据资产（结构化/非结构化/SaaS数据）、对数据进行分类分级（如个人信息、财务数据、源码、AI训练集等）、动态风险评估与访问控制建议。Wiz、Sentra、BiglD、Cyera等厂商展示的新一代DSPM模型，已成为云原生安全的标配能力。

五、RSAC2025的八大关键词

2025年RSAC大会的核心内容可浓缩为八大关键词，它们共同勾勒出全球网络安全的发展蓝图：

关键词	简要说明
AI Security(人工智能安全)	AI被广泛用于攻击与防御两端，生成式 AI引发 Prompt Injection、模型投毒、隐私泄露等新型风险，同时推动 AI 防御工具与红队测试平台的兴起。
Agentic AI(自主 AI代理)	新一代AI能够执行多步攻击任务，如编写钓鱼邮件、自动扫描漏洞、规避检测，RSAC强调需建立 Agent 行为识别与身份治理机制。
DSPM(数据安全态势管理)	数据安全从“静态加密”迈向“全生命周期动态治理”，DSPM工具能识别、分类、保护敏感数据，成为云原生安全标配能力之一。
Supply Chain Security（供应链安全）	供应链攻击激增推动SBOM普及，AI模型引入风险倒逼AIBOM(记录模型来源、训练集、微调过程)逐步标准化。
Post-Quantum Cryptography(后量子密码)	量子计算对现有加密算法的威胁迫使政企开始部署PQC(如 KyberDilithium)，并强调“Harvest now, decrypt later”的长期威胁。
Zero Trust(零信任架构)	零信任正从网络层、防火墙、身份扩展到 数据层、Alagent层，出现ZTDA(Zero Trust Data Access)等新模型。
Industrial Cybersecurity(工业网络安全)	电网、制造、能源等OT领域成为新重点，自动化工业攻防平台(如SpiderSim)亮相，强调数字孪生+风险模拟。
AI Governance&Compliance(AI治理与合规)	伴随 AI 立法(如《人工智能法案》)推进，RSAC呼吁企业构建模型治理机制，覆盖伦理、透明性、可追溯性、红队测试、偏见检测等。

图5 RSAC 2025八个关键词

2025年RSAC大会以“众声汇聚，同心共筑”为指引，清晰地展现了网络安全行业在AI时代的挑战与机遇。从自主AI代理的崛起，到后量子密码的布局，从供应链安全的强化到数据治理的深化，每一个趋势都指向同一个核心：只有通过多元化协作，才能构建起抵御复杂威胁的安全生态。未来的网络安全将不再是单一技术的比拼，而是整个行业在技术、治理、合规等多维度协同的综合实力较量。

六、总结：多方视角下对RSAC2025技术趋势的学习与借鉴

本届RSAC大会揭示的AI安全、零信任规模化等趋势，为网络安全领域多方参与者指明了方向。

企业作为安全需求的主体，需将技术趋势转化为可执行的安全策略，重点关注“实用性”与“适配性”，以落地适配为核心，构建动态防御体系。分阶段引进新兴技术，强化供应链与数据治理，布局后量子密码过渡方案，加强企业内部跨团队协作，并结合法律法规及时调整安全策略，避免合规风险的发生。

厂商作为技术供给方，需紧跟趋势突破产品边界，同时强化与产业链的协同能力，以“技术创新+生态协同“为导向，响应市场需求。深化AI安全技术的研发，推动零信任与新兴场景融合，布局后量子密码与供应链安全工具，以适应新的市场需求。

监管机构要以标准引导与风险共治为核心，完善技术标准与法规体系，建立跨行业风险共享机制，引导关键行业应对量子威胁。可以通过政策激励初创企业技术落地，同时规范 AI 模型的伦理审查（如要求模型决策可追溯、隐私保护合规），避免技术滥用。

RSAC2025的核心启示在于“协作”，多方协同才能将“众声汇聚”的理念转化为抵御复杂威胁的实际能力，共同构建更具韧性的网络安全生态。

参考文献

[1]中经科技前沿：Agentic AI爆发落地前夜 业界聚焦模型和成本挑战；

[2]中国军网-解放军报：后量子密码技术 重构数字世界的安全边界。

06 2023年度实验室优秀课题摘录

P99 基于可信内存指令序列检测技术的漏洞（包含未知漏洞）攻击防护能力
刘嵩、胡广跃、王磊
刘磊、刘志明

P105 挂图作战在证券行业应用实践研究
华仁杰、沈嗣贤、徐俊超、鞠叶、任思豫
张海龙、金亮成、杨云云

P111 零信任架构下的业务系统敏感数据保护实践
王洪涛、刘宏、马晓鹏、黄施宇
何艺

P118 东方证券企业终端数据安全解决方案探索
邬晓磊、甄明达
焦健、汤华晟

P124 基于漏洞情报的拟态防御技术实践
邢骁、蔡子豪
薛辛

基于可信内存指令序列检测技术的漏洞（包含未知漏洞）攻击防护能力

刘嵩、胡广跃、王磊 | 光大证券股份有限公司

刘磊、刘志明 | 奇安信科技集团股份有限公司

摘　要： 随着漏洞数量的快速增长和新型攻击手段的不断涌现，传统防御手段存在滞后性和局限性，难以应对网络空间安全挑战。基于可信内存指令序列检测技术的网络安全防护能力能够直接从内存指令层检测漏洞攻击代码的执行，不依赖已知漏洞特征和攻击代码，能够有效防护未知漏洞利用，通过不断的策略优化可以实现精准识别异常指令，降低误报率和运营成本，为增强防御能力、保护终端安全提供有力支持。

关键字： 可信内存指令序列检测技术、未知漏洞、漏洞防护能力、异常指令、终端安全、无文件攻击、攻击溯源、攻击载荷

一、概述

近年来，漏洞数量持续高速增长，漏洞利用造成的安全事件频繁发生，给国家网络空间安全、企业数字化转型以及个人信息安全带来了严重的威胁，而面对层出不穷的新型漏洞攻击，传统防御手段已力不从心，亟需更先进的安全技术来应对挑战。

本篇文章是基于可信内存指令序列检测技术，来发现和防范代码异常执行类的漏洞利用，包括操作系统漏洞、浏览器漏洞和文档漏洞，并从实战角度考量该技术在增强防守方防御能力，通过内存指令调用检测技术，关注证券行业常用系统、文件活动在内存指令级的调用情况，一旦发现指令序列的异常调用，即时向终端报出异常，保护券商业务的终端安全。通过基于内存指令调用序列检测技术，在攻击发生时，能够在内存指令层进行安全检测，从而有效抵御0day漏洞攻击。从纵深防御角度看，终端位于防护体系的最后一环，实现攻击链末端防护。当攻击者突破了边界防护，绕过身份认证/访问控制、终端杀毒软件等终端安全机制后，利用可信程序执行恶意指令时，系统将发挥防御作用，阻断程序进程并发出告警。同时，通过不断的策略优化，实现了对异常指令的精准识别，减少了误判，从而极大降低了对人工的依赖，运营成本大幅下降。

二、基于可信内存指令序列检测技术的网络安全防护能力的设计

（一）总体能力框架

在系统架构上，分为客户端软件和控制中心两个部分，在终端或服务端计算机上，安装集成了客户端组件，在服务端上安装控制中心软件，即一个服务端控制中心管理多台终端计算机。

服务端：服务端承担控制中心角色，部署在独立服务器上，通过局域网IP连接每台终端，对终端进行管理。管理员可通过浏览器访问控制中心，配置并下发终端安全策略，提供终端状态监控、防护事件感知与处置、风险事件分析溯源等功能。

客户端：客户端是部署在被保护的终端上提供安全防护的程序（Agent），可根据不同系统类型下载对应安装包进行快速部署，接收控制中心下发的策略并监控终端操作行为。客户端支持Windows全系列、主流Linux系统内核、国产麒麟\统信等主流操作系统，适配Intel x86/64、ARM、MIPS等不同系列的硬件架构体系。

系统架构图如图1所示：

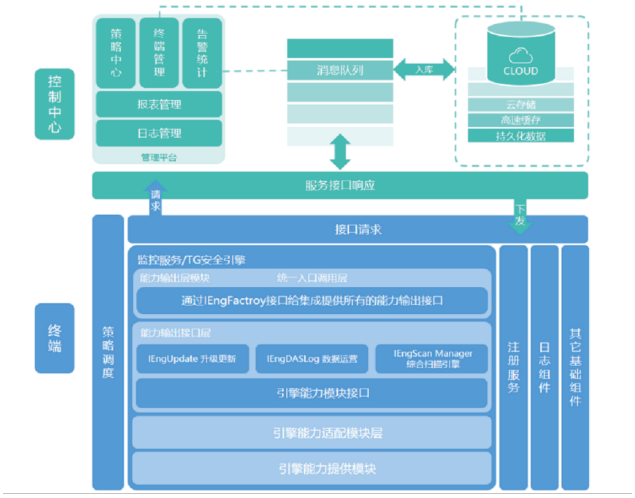


图1 系统架构图

漏洞攻击的实质，是控制“可信程序”来执行恶意指令，其继承了漏洞所在程序的一切特权，传统的安全技术是对外防护的，并未考虑到系统本身、可信的第三方程序作恶或被利用作恶的问题。基于可信内存指令序列检测技术的网络安全防护能力的设计主要针对操作系统、浏览器、文档进行可信指令的学习并构建可信指令库，来发现和防范代码异常执行类的漏洞利用，实现漏洞防护效果。系统能力架构图如图2所示：

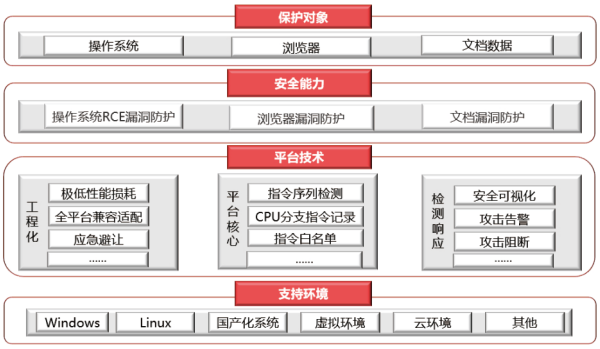


图2 能力架构图

（二）应用场景

1、有效抵御0day漏洞利用攻击

操作系统厂商一直是与漏洞攻击进行对抗的主要力量，先发现漏洞被利用，再针对已知系统漏洞开发补丁包，当系统停服不再更新或者系统存在未知漏洞时将给企业或组织带来很大的安全隐患。而在日常办公中必定会安装浏览器、文档编辑软件等常用办公工具，正版开发软件可能存在未知漏洞或后门，而非正版软件由于被破解篡改过，安全威胁更是堪忧。

基于内存指令控制流检测技术，从更底层监测漏洞攻击代码的执行，完全不依赖已知漏洞特征和已知攻击代码的特征，可第一时间发现并阻止未知漏洞利用的攻击。能够有效解决上述隐患，针对RCE远程代码执行漏洞利用攻击、浏览器漏洞攻击、办公文档程序漏洞攻击等常见的漏洞利用领域提供了全方位的漏洞攻击防护解决方案。

2、无文件攻击

高级威胁的攻击方式更加复杂而隐蔽（如主流的无文件攻击、内存攻击、内存马、ROP攻击等），导致基于文件静态检测的传统终端安全软件无法应对。采用指令序列检测的技术能够从容应对。

3、攻击溯源能力

攻击者潜入内网终端后，不会立即发起攻击，而是在傀儡机下载攻击木马或脚本，并横向扩散入侵更多的终端，而传统的终端安全软件无法全面识别威胁行为（如病毒操作、目的、影响面等），导致无法定位终端威胁的来源。针对不同的漏洞攻击领域，本次系统设计了不同的防护模块，从根源通道处解决各类漏洞攻击问题。

4、极端场景兼容

隔离网、机器性能差、软件冲突等各种场景均可兼容，不影响产品能力的发挥，摆脱了传统安全技术对文件、流量、数据、行为等特征的依赖，实时监测不滞后。

三、技术创新

漏洞攻击的实质，是控制“可信程序”来执行恶意指令，其继承了漏洞所在程序的一切特权，传统的安全技术是对外防护的，并未考虑到系统本身、可信的第三方程序恶意行为或被利用进行攻击的问题。

比如：操作系统本身的服务拥有系统的完全控制权限，如果攻击者通过漏洞控制了系统服务，通过系统进程对用户数据文件进行窃取、修改、删除等恶意操作，安全软件是难以防护的。

因此，漏洞攻击防护的核心是对可信的程序、设备的防护。

通过权限控制可以有效地遏制利用可信程序的漏洞发起攻击所带来的危险，但存在明显的缺陷，因为无论怎么限制权限，都不能杜绝当一个本应该有权限的主体存在漏洞而被利用的问题。

例如：允许Office Word程序打开和编辑Office doc文件是必然的，如果黑客通过漏洞控制了Word程序时，无论是传统安全技术的文件特征检测、行为检测，还是权限控制技术都无法有效地进行识别，这需要更底层的安全技术来支撑，即不依赖文件特征、行为特征的内存指令控制流检测技术。

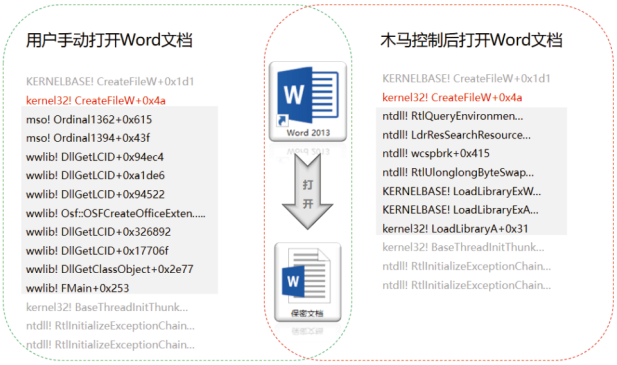


图3 内存指令控制流检测技术

通过对指令控制流的检测来发现漏洞攻击，一直是漏洞攻击发现与防护领域较前沿的安全理念，各硬件芯片厂商、操作系统厂商、安全厂商都在此领域进行着持续的技术研究与探索，其中最大的问题是指令是一个计算机系统的最小执行单元，动辄以每秒亿计的频率执行，对指令进行检测不可避免地带来性能的严重损耗，而如何在更有效地发

现攻击的同时，把对性能的影响降低到一个可接受的程度，是大家的努力方向。

通过该技术的研究，在低性能损耗的情况下，实现了有效的漏洞攻击发现能力。通过不断的策略优化，实现了对异常指令的精准识别，减少了误判，从而极大地降低了对人工的依赖，使得检测引擎的误报率大幅降低，运营成本大幅下降。

（一）内存指令采集技术

可信内存指令序列的检测技术主要用于漏洞利用攻击的防护，学习和采集系统中需要被保护的程序，或存在所有可能被利用风险的程序的二进制文件，通过反汇编引擎把二进制文件内容转换成机器指令的中间数据，再根据本系统的技术防护点及特定的算法得到指令间的调用关系，并按特定结构组织构造指令序列的白名单，形成“指令序列白名单库”；当实际在内存中执行的任意一条指令序列不在白名单中时，即认为是非原生的额外出现的异常攻击指令。

基于可信内存指令序列检测技术的网络安全防护能力研究的内存指令采集技术，从更底层来关注系统、文件活动在内存指令级的调用情况，摒弃了以往的可信程序、白名单的机制，对需要被保护的程序每一次的内存调用指令进行学习和采集。一旦发现指令序列的异常调用，即时向终端报出异常，让所有的基于漏洞的攻击行为无所遁形，从而打破了以往只能基于已知漏洞的防御手段，在面对未知漏洞、零日漏洞攻击时也能做到主动防御，确保系统的安全。

（二）内存指令攻击载荷的追踪技术

以往的病毒、木马查杀技术，或者漏洞补丁的修复方式，对用户始终是个黑匣子，漏洞在哪里，木马植入到了什么地方，漏洞的攻击方式、路径是怎样的，都是不可见的。漏洞利用攻击发生后，目前的安全技术若能发现，也都只能上报攻击事件行为数据，无法还原攻击通道，对进一步溯源和后续处置缺乏必要信息。

本系统将监测能力下探到内存指令级，具备了发现漏洞触发点的能力，因此在漏洞攻击发生后，不仅能够收缩和聚焦到漏洞发生的位置，还能捕获到攻击时的Shellcode攻击载荷，具备还原攻击通道的能力。本系统创新性推出的攻击溯源的展示技术，通过对指令序列调用的跟踪，还原出整个漏洞攻击路径，漏洞所在位置，将指令调用异常点进行可视化呈现，让运维、安全管理人员，能一目了然的看到漏洞是从哪里发起攻击，所借用的主客体是谁，整个攻击路径是怎样的，为事后补救和加固提供必要的支撑。这项攻击溯源展示技术的推出，不仅仅让漏洞防得住，更主要的是，让运维和安全人员看得懂、用得会漏洞防护产品，有效地增加了使用者的漏洞防护知识和安全防护意识。

（三）“基于内存指令执行序列”进行安全检测的技术

当应用程序中的“某一指令序列”存在一个漏洞，在被攻击者利用后，正常的指令执行序列发生改变，转而去执行攻击者的攻击指令，从而导致危害发生。

本系统通过可信内存指令序列检测技术，还原系统中被保护程序的指令调用序列，构造成“指令序列白名单库”生成程序的合法权限控制策略，当内存中执行的指令序列不在白名单中时，即认为是异常攻击的指令。完全不依赖已知漏洞特征和已知攻击代码的特征，可以发现绝大多数利用未知漏洞发起的攻击；解决未知漏洞攻击问题。

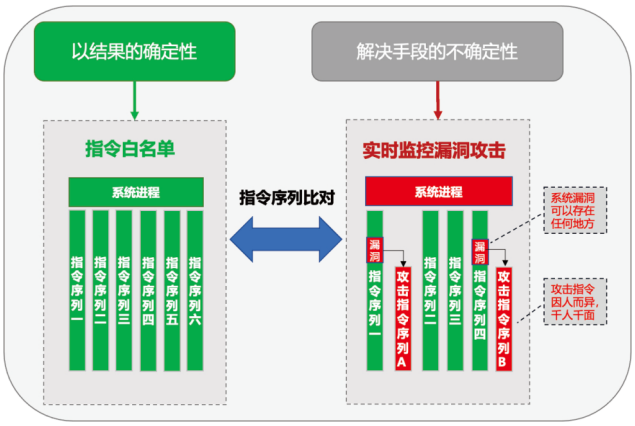


图4 基于“内存指令执行序列”进行安全检测的技术原理图

四、系统测试情况分析

（一）测试目标

本系统开发完成后，对部署了不同安全能力的终端/服务器使用相同的POC用例进行对比测试，以达到以下目标：

（1）验证在没有使用补丁以及漏洞缓解技术的前提下（模拟未知漏洞攻击），内存指令序列检测技术对漏洞利用的防护能力；

（2）通过对防护结果的数据比对与分析，体现内存指令序列检测技术与传统安全检测技术之间的差异；

（3）通过防护结果的数据比对与分析，判断内存指令序列检测技术是否可替代传统安全检测技术。

（二）测试环境和拓扑

1、环境说明

本次测试共部署三套靶机和一台攻击机，在三套靶机内分别部署，杀毒软件、EDR和本系统开发的防护软件，进行防护效果对比。由于测试需要利用特定漏洞，因此靶机系统版本及所安装的应用版本有特定要求。本次测试靶机环境具体如下：

测试环境1：准备Windows 10操作系统一台，安装WPS10700.12012.2019和浏览器chrome89.0.4389.114、准备Windows server2016操作系统一台启用DNS服务器，并在两台系统中分别部署杀毒软件进行防护。

测试环境2：准备Windows 10操作系统一台，安装WPS10700.12012.2019和浏览器chrome89.0.4389.114、准备Windows server2016操作系统一台启用DNS服务器，并在两台系统中分别部署内存指令序列检测技术进行防护。

测试环境3：准备Windows 10操作系统一台，安装WPS10700.12012.2019和浏览器chrome89.0.4389.114、准备Windows server2016操作系统一台启用DNS服务器，并在两台系统中分别部署EDR进行防护。

攻击机：准备1台kail。

2、测试拓扑

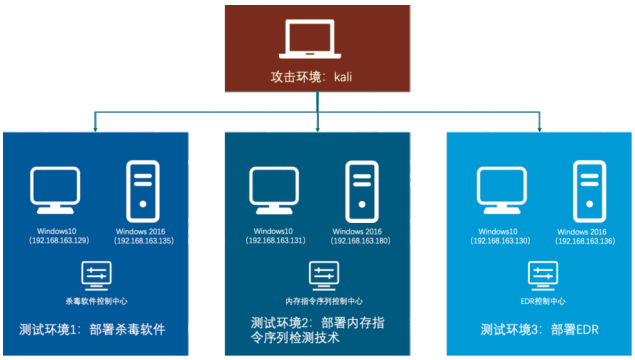


图5 测试拓扑图

（三）测试用例

1、CVE-2020-1350

漏洞类型：整型溢出

漏洞危害：远程代码执行

漏洞介绍：2020年7月14日，微软修复了一个Window DNS Server中的严重远程代码执行漏洞，该漏洞由于Windows DNS Server未能正确处理特定畸形数据交互，从而导致远程无需验证的攻击者通过利用在本地系统账户下执行任意代码。该漏洞在微软的通告中被定级为可蠕虫化级别，此类型的漏洞在被利用后，往往会导致严重的安全威胁，类似之前被利用来传播WannaCry蠕虫的永恒之蓝漏洞。

预期效果：模拟攻击者利用该漏洞进行远程攻击，最终拿到目标靶机的SYSTEM权限。

图6 CVE-2020-1350漏洞测试图

2、CVE-2021-30551

漏洞类型：类型混淆

漏洞危害：远程代码执行

漏洞介绍：2020年6月9日，谷歌修复了一个Chrome 浏览器中的严重远程代码执行漏洞，在91.0.4472.101 之前的Google Chrome 浏览器的 V8 中存在类型混淆，远程攻击者可通过制作的恶意HTML 页面利用此漏洞实现RCE。

预期效果：模拟攻击者利用该漏洞制作恶意页面，诱使用户访问该页面，当用户访问页面时，会触发漏洞利用，实现对用户电脑上的文档类文件进行加密勒索，被加密的文件的后缀为encrypt。

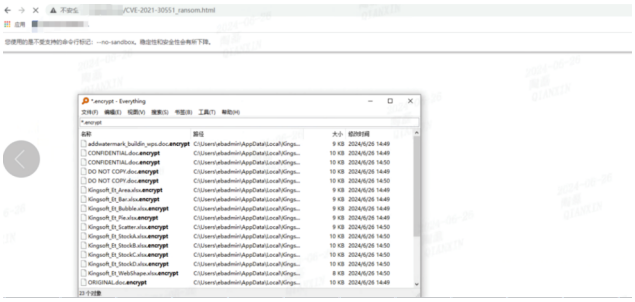


图7 CVE-2021-30551漏洞测试图

3、WPS 远程代码执行漏洞

漏洞类型：逻辑漏洞

漏洞危害：远程代码执行

漏洞介绍：2021年，金山软件修复了一个WPS中的远程代码执行漏洞。WPS Office的内置浏览器存在漏洞，攻击者可以利用该漏洞专门构造出恶意文档，受害者打开该文档并点击文档中的 URL 链接或包含了超级链接的图片时，存在漏洞版本的 WPS office 会通过浏览器加载该链接从而执行恶意代码导致RCE。

预期效果：模拟攻击者利用漏洞制作恶意文档，进行传播。当用户收到恶意文档后，使用WPS打开文档，在文档中单击一次，即可触发漏洞利用，实现在用户电脑上创建任意进程，本次攻击时创建了一个计算器进程。

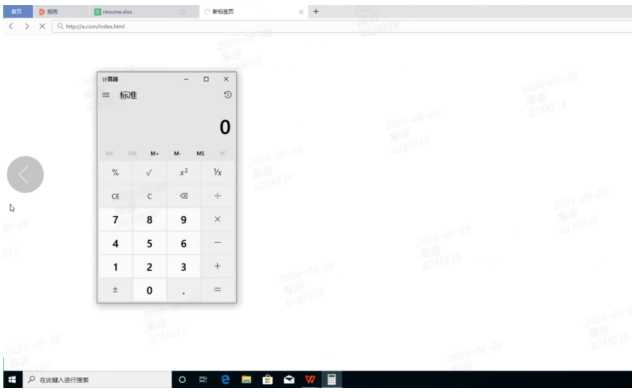


图8 WPS远程代码执行漏洞测试图

（四）测试结论

1、测试结果

（1）对测试环境1，安装了杀毒软件的测试结果如下：

依次执行完三个测试用例，在测试终端并无弹窗提示，到杀毒软件的控制平台分别搜索Windows10和Windows server2016的IP地址192.168.163.129和192.168.163.135，也无告警信息。

（2）对测试环境2，安装了内存指令序列检测技术软件的测试结果如下：

1）启动微软Windows DNS Server 中的远程代码执行漏洞（CVE-2020-1350）的用例，被成功拦截，界面中除了显示了引发告警的主体对象、客体对象和命令参数外，还展示了栈信息和异常指令信息。

2）启动Chrome浏览器中的远程代码执行漏洞（CVE-2021-30551）的用例，被成功拦截，界面中除了显示了引发告警的主体对象、客体对象和命令参数外，还展示了栈信息、Shellcode和异常指令信息。

3）启动WPS远程代码执行漏洞的用例，被成功拦截，界面中除了显示了引发告警的主体对象、客体对象和命令参数外，还展示了栈信息、Shellcode和异常指令信息。

（3）对测试环境3，安装了EDR的测试结果如下：

依次执行完三个测试用例，在测试终端并无弹窗提示，到EDR的控制平台分别搜索Windows10和Windows server2016的IP地址192.168.163.130和192.168.163.136，也无告警信息。

2、测试结论

（1）杀毒软件和EDR对于本次系统使用的漏洞利用样本无法检出，该结果表明基于知识和基于行为的防护方式对于漏洞利用的攻击，存在防护无效情况。这种情况，导致了防护软件出现漏报现象，即存在漏洞利用手段会绕过安全检测的可能。该情况在真实场景下必然存在。在安全防护手段无法检测、或没有安全告警记录的时候，安全人员难以通过现有手段知晓是否存在漏报，会导致对当前安全状况产生误判，影响安全运营的结果；

（2）杀毒软件和EDR作为终端安全重要的防护能力，当出现漏报时，会让安全人员将原本“控制失效的”状况误认为是“安全可控的”。结合三个用例可以看到，攻击者无论是获取系统SYSTEM权限、还是创建任意进程、亦或是对系统内的文件进行加密，都是控制失效结果，而安全人员对于这些结果难以在第一时间知晓，这便给了攻击者扩大战果的机会，同时也给安全运营增加了难度与挑战。

（3）可信内存指令检测技术在不依赖知识库与行为特征的前提下对未知漏洞攻击拥有防护能力，在没有使用补丁以及漏洞缓解技术的前提下，可信内存指令序列检测技术对于本系统使用的漏洞利用的样本百分百检出，并拥有拦截能力。该技术在不依赖知识库与行为特征的前提下，不仅检出了攻击，还提供了指令异常信息、栈信息、Shellcode信息等，为定位漏洞所在位置提供价值数据。由此表明可信内存指令检测技术对未知漏洞攻击同样有防护能力。

（4）与杀毒软件和EDR的告警方式不同，可信内存指令检测技术提供了一种全新的告警方式，不仅仅是简单的以主客体信息作为告警中的数据，还提供了更精细的内存数据（指令异常信息、栈信息和Shellcode信息）供安全人员进行分析。例如：通过对指令异常信息和栈信息的查看，可以知晓内存指令跳转发生异常时的具体节点；通过查看Shellcode信息可以知晓攻击者申请内存空间的过程。这些信息都有助于安全人员了解攻击过程的全貌，即便不使用可信内存指令检测技术的拦截能力，也可以从中提取知识与行为特征，用于其他安全能力的预防和检测规则，从而提升整体的安全运营能力。

（5）可信内存指令检测技术，通过检测识别被保护主体的非可信指令，判断其风险，不仅能够防护漏洞（包含未知漏洞）攻击，更能避免漏报产生；

（6）基于知识（如杀毒软件）和基于行为（如EDR）的防护方式，因其特性必然存在漏报的情况的，而基于可信内存指令检测技术的防护方式，其本质是采用白名单理念，将需要被保护的主体对象的指令进行采集，形成可信指令序列白名单。任何对被保护的主体发起的漏洞攻击，想要攻击成功，必然会出现与白名单不符的指令。而通过与白名单比对，可以发现任何不在白名单内的指令以及不符合白名单序列的指令调用顺序。因此使用基于可信内存指令检测技术的防护方式可以避免漏报，也能防护未知漏洞攻击。

（7）可信内存指令检测技术的原理及实现方式和当前使用的安全能力（基于知识、基于行为）截然不同，可以组成异构体系，实现对纵深防护能力的提升；

基于知识（如杀毒软件）和基于行为（如EDR）的防护方式虽然是两种检测方式，但实质并没有脱离历史经验，即需要先掌握/发现已有的知识和行为，才能进行检测。将这两种方式同时使用组成纵深防护时，并未形成真正的异构，仅是采用了不同的检测手段。而基于可信内存指令序列检测技术并不依赖于历史经验，其依赖的是被保护主体本身的可信指令序列。因此可信内存指令序列检测技术能够与当前安全防护方式组成真正的异构体系，将检测能力下沉到内存指令层，弥补当前防护领域的空缺，实现对纵深防护能力的提升。

五、总结

本课题主要研究“基于内存指令执行序列”的检测技术，旨在探索证券行业常用系统、文件活动基于可信内存指令序列检测技术的网络安全防护，通过实战的角度发现和防范代码异常执行类的漏洞利用，从而增强防守方防御能力，保护券商业务的终端安全。

本课题在实践中通过在办公区、营业部终端上安装系统客户端组件，在服务端上安装控制中心软件，实现了一个服务端控制中心统一管理多台终端的功能。通过内存指令调用检测技术，采集证券行业常用系统、文件活动在内存指令级的调用情况，一旦发现指令序列的异常调用，即时向终端报出异常，确保系统的安全。通过本次研究实践，解决了行业未知漏洞防护问题，能够为证券行业其他单位完善企业网络安全防护体系，加强未知漏洞安全防护能力建设提供案例参考，同时，也可以促进行业内安全领域的交流和合作，推动整个行业漏洞安全防护的发展。我们相信通过后续系统的不断优化，将进一步实现各类终端安全可信的目标。

参考文献

1.刘建亮.计算机网络安全漏洞分析研究[J].中国新技术新产品,2019, (15): 31-32. DOI:10.13612/j.cnki.cntp.2019.15.018.

2.潘然.软件系统漏洞发现及倾向性预测关键技术研究[D].桂林电子科技大学, 2019.

3.刘鹏毅,何嘉栋.APT防御与网络协议缺陷探析[J].保密科学技术,2023,(12):32-39.

4.张耕源.论网络空间安全与攻防技术[J].信息系统工程,2023,(12):137-140.

挂图作战在证券行业应用实践研究

华仁杰、沈嗣贤、徐俊超、鞠叶、任思豫 | 东吴证券股份有限公司

张海龙、金亮成、杨云云 | 金证金融科技（北京）有限公司

摘　要： 随着证券行业数字化、网络化的不断深入，网络攻击手段也日益多元化，从数据泄露、系统入侵到恶意软件的传播，行业传统安全运营体系面临着前所未有的安全挑战。挂图作战能力体系从网络安全业务场景的实际需求出发，依托平台基础设施，通过面向分类业务的图层映射、要素提取与场景绘图，以丰富的图形展现网络安全保护平台信息数据，赋予作战“指挥官”判断、决策、指挥能力。本次课题完成“资产安全知识图谱构建”“资产全生命周期安全管理”“安全风险级别标准化”三大核心能力以及“资产风险监测”与“攻击路径”两大可视化挂图建设，夯实挂图作战体系基础。

关键字： 挂图作战体系、安全业务场景、资产安全图谱、资产全生命周期安全管理、安全风险级别标准化

一、研究背景

金融行业作为国家经济的核心，近年来却成为网络安全事件的高发区。随着数字化、网络化的不断深入，网络攻击手段也日益多元化。从数据泄露、系统入侵到恶意软件的传播，金融机构面临着前所未有的安全挑战。

随着金融科技快速发展，公司近几年也逐步加大信息化建设的投入，与此同时，安全运营也面临着更大的挑战。基于实践角度，主要存在以下问题：

1、难以有效应对新型安全威胁

证券公司因业务类型多、应用系统数量多、业务场景复杂等问题，即使获取到供应链/应用组件漏洞情报，也很难高效、准确地判断是否涉及、影响范围，从而难以高效展开事前防御工作。

2、安全分析治标不治本

攻击者在进行单点突破后，往往会进行横向突破以尽可能获取更多权限。因此需要快速、准确地确定横向范围和影响范围，避免遗漏，从根本上遏制此次事件。

3、应急响应优先级缺乏数据支撑

当前安全告警更多是以风险级别作为主要依据，如漏洞级别等作为依据，缺乏从业务、资产、风险可利用性视角进行整体的考量。存在应急响应计划优先级制定不合理、与业务诉求脱节等问题。

4、安全管理任务繁重

当前安全运营体系尚未完全打通组织架构、IT资产、业务流程等要素，很多安全管理的流程类事务仍依赖于OA工单、线下联系等方式进行，难以进行统一的指挥协同。

5、难以直观掌握攻防过程

作为安全运营人员，希望可以结合网络拓扑、IT资产、安全覆盖度等防守者视角，更直观掌握攻击者的攻击路径信息，以便开展监测与响应工作。

挂图作战本质是一套完整的安全运营体系理念。挂图就是网络空间可视化表达，目标在于以网络空间地图形式全面展示网络信息，实现网络空间的具体化和数字化，从而为决策者提供直观、有价值的信息，以降低决策的不确定性。

东吴证券为进一步落实网络安全保障责任，支撑公司数字化转型战略，有效应对网络安全严峻威胁、挑战和网络安全管理工作需求的扩展，通过网络安全挂图作战体系建设与落地，夯实IT资产管理基础、重塑安全事件定级标准、规范安全管理工作流程，以达到如下安全运营工作成效：

（1）平时——常态化、体系化、有数据支撑的安全运营工作；

（2）战时——实战化、智慧化、一体化的指挥协同安全运营工作。

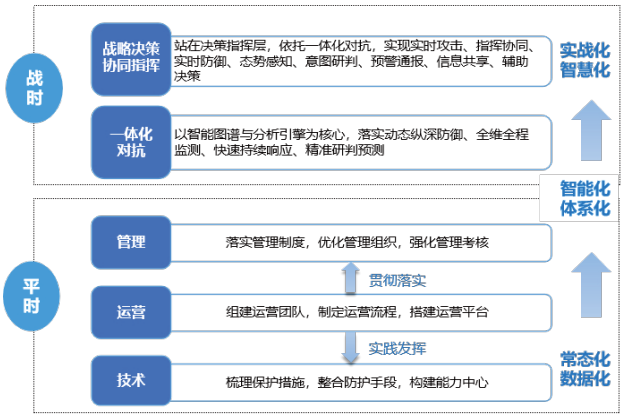


图1 挂图作战体系建设的目标

二、挂图作战实践方法理论

（一）挂图作战体系设计

实施网络安全挂图作战，必须从网络安全业务场景的实际需求出发，依托平台基础设施，通过面向分类业务的图层映射、要素提取与场景绘图，以丰富的图形展现网络安全保护平台信息数据，赋予作战“指挥官”判断、决策、指挥能力。

具体在挂图作战体系设计中，我们首先结合当前安全现状进行需求分析，完成对挂图作战场景的需求梳理，详细了解不同场景下的业务运作流程、业务关注要素，得出“挂图作战场景列表”及对应需求文档。接着分析挂图作战不同场景下的表达维度、配套模型、图层要素和图层要素之间的关联关系，进而确定挂图作战总体架构和图层要素间关联关系。最后围绕图层要素，基于公司现有网络安全、内部组织架构、基础设施等相关数据资源，抽取必要数据，进行清洗、处理和映射等处理工作。

至此，东吴证券完成“数据资源－图层要素设计提取－挂图作战场景构建”三步走的挂图作战体系设计思路。

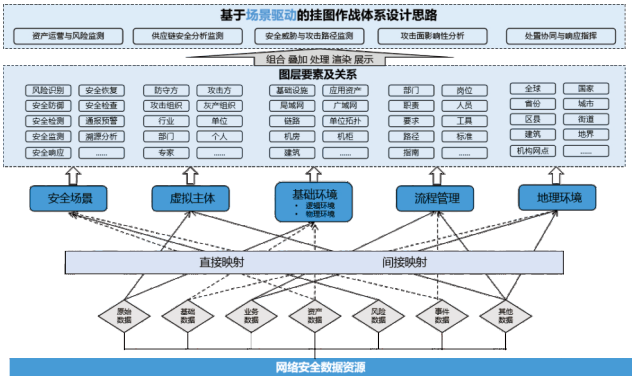


图2 场景驱动的网络安全挂图作战体系设计思路

（二）东吴挂图作战实践架构

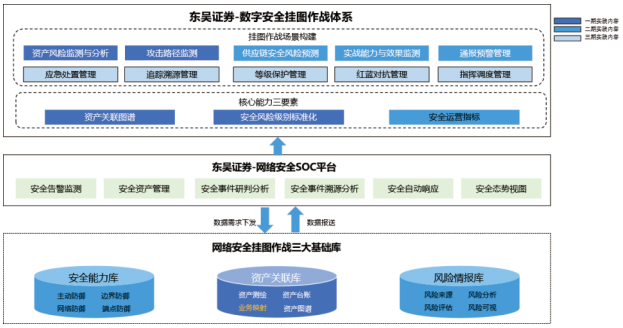


图3 东吴证券挂图作战实践架构

依据挂图作战体系设计思路，结合当前安全现状、安全业务需求、业内最佳实践，提出一套东吴证券挂图作战体系实践架构，实践架构主要分为：

1、网络安全挂图作战三大基础库

以安全能力库、资产关联库和风险情报库作为挂图作战的基础底座，建立三大基础库运行体系，确保资产可知、风险可控、安全能力可用。

2、安全集中运营中心SOC平台

基于现有SOC平台作为挂图作战的支撑，当前SOC可支撑非场景化的网络安全管理和运营工作，也已对接了大量安全数据，这为挂图作战平台提供基础可信的安全源数据。

3、核心能力三要素：

(1)以业务线为视角串联公司IT资产，并将“安全”视作资产的一种属性形成资产关联图谱，完善全维全域监测，实现精准研判预测和智能分析能力。

(2)重构安全风险级别，不再仅依赖漏洞评分等基础安全维度，而是综合资产重要性、影响范围、部署区域、漏洞可利用性、漏洞评分等维度对安全事件告警级别进行重构，提高应急响应效率。

(3)围绕资产管理、终端安全、应用安全、开发安全、网络安全、安全建设、安全运维、风险管理等19个能力域，结合攻击者视角出发的杀伤链2.0模型框架，形成东吴证券网络安全运营指标体系，为挂图作战提供数据支撑。

4、数字安全挂图作战平台

数字安全挂图作战平台是挂图作战的呈现形式，也是挂图作战的场景。以实际安全业务需求场景出发，依托于三大基础库的能力、SOC平台的源数据，基于资产图谱、安全风险级别标准化、安全运营指标进行挂图作战场景图层设计，以更直观的方式实现安全业务场景化。

（三）东吴挂图作战实践路径

为切实保障挂图作战体系的建设成效，计划分三期开展实践应用工作：

1、第一期重点实践内容

一期重点工作目标是夯实挂图作战体系基础，对当前安全能力进行优化升级。基于此目的，一期重点打造3项能力+2项场景。

核心能力建设部分：主要围绕“资产安全知识图谱构建”“资产全生命周期安全管理”“安全风险级别标准化”三项内容开展研究应用。

安全场景设计部分：构建“资产风险监测”与“攻击路径”两张可视化挂图。

2、第二期重点实践内容

二期工作是在一期基础工作上重点对一期内容进行有效性验证，并以降低安全运营成本、提高安全运营效率为目

标展开。

核心能力建设部分：主要围绕“安全风险情报体系建设”“安全运营指标构建”“安全能力迭代”三项内容开展研究。

安全场景设计部分：构建“供应链安全风险监测”“实战能力与效果监测”“通报预警管理”三张网络安全场景地图。

3、第三期重点实践内容

通过一期、二期的建设，挂图作战体系的核心能力已构建完毕，第三期内容更多围绕安全管理、围绕常见的安全业务场景开展，构建“应急处置管理”“追踪溯源管理”“等级保护管理”“红蓝对抗管理”“指挥调度”等场景地图，真正将网络安全从“量”实现到“质”的飞跃。

三、挂图作战一期内容实践

（一）资产全生命周期安全管理实践

挂图作战一期的核心能力建设围绕“资产安全知识图谱构建”“资产全生命周期安全管理”“安全风险级别标准化”三项内容开展研究应用，以及“资产风险监测”与“攻击路径”两项安全场景。资产管理作为挂图作战一期的核心能力建设的一部分，既完成资产全生命周期安全管理，同时也为知识图谱的构建提供数据支撑。

资产的全生命周期管理覆盖资产上线、资产运行、资产下线，在这过程中要对资产进行分类分级、对资产进行风险评估、对资产的台账进行维护。资产管理存在很多维度，资产的安全管理则在常规资产管理的基础上叠加安全维度，从安全运营管理角度看待资产，通过对资产属性进行安全管理，消除资产管理中的安全隐患，保障IT资产的安全性。

1、资产全生命周期管理实践

IT资产的复杂性，决定了很难依靠单一的方式将其管理起来，需要创建一条“人+工具+流程”的资产安全运营体系，从运营团队、制度流程、技术工具的全盘考虑，整合团队、技术、规范、流程、平台等全要素，打通资产管理全流程，才能形成完善的资产安全管理体系。

资产安全全生命周期管理涵盖资产上线、资产运行以及资产下线全过程。

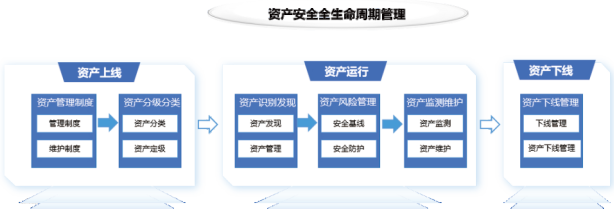


图4 资产安全全生命周期管理

2、资产安全分类分级体系

资产分类是企业资产保护工作的重要前提，是建立统一、标准化资产管理体系的基本条件，是资产安全管理的基础。东吴证券参照《证券期货业证券业务标准规划（2022-2025）》《证券期货业信息系统分类与代码（征求意见稿）》《证券期货业网络安全等级保护基本要求》，结合自身业务系统建设情况，建立资产分类分级指南。

通过对东吴证券业务的梳理，将业务系统划分为经纪业务、资产管理、信用业务等一级、二级业务系统。按照业务系统重要性分为办公系统、一般业务系统、核心业务系统。

根据资产安全属性（保密性、完整性、可用性）遭到破坏后的影响程度大小，直接资金损失等，进行资产级别的分类。

（二）资产安全图谱实践

从公司角度来看，网络安全建设主要是保障公司各业务活动能够安全、有序地开展；从IT角度来看，网络安全保护对象主要包括网络系统软/硬件、基础网络、业务应用及应用中的数据。

基于以上两点，本次研究应用了“以业务线为视角，以IT基础资产为对象”的资产安全图谱，基于图谱实现三项目的：

（1）资产安全图谱中任意“节点”监测到安全告警，可快速定位到所属业务线，将安全告警转化为业务风险告警，提高应急响应效率。

（2）资产图谱中任意“节点”监测到安全风险，通过图谱能直观反映风险影响范围、影响程度，提前进行布防工作，避免事态进一步扩大。

（3）资产安全图谱作为挂图作战体系的基础能力之一，通过构建图谱，进一步夯实安全基本功，将安全运营体系从“以事件为中心”到“以资产为中心”转变，提升安全运营效率。

1、资产安全图谱实体节点

图5 资产安全知识图谱关键要素

基于业务视角、安全视角出发梳理IT资产间逻辑关系，自上而下得出“业务条线依赖于业务应用、业务应用使用应用程序接口（API）进行交互、业务应用通过软件包集成、软件包安装依赖于操作系统环境、操作系统需部署于硬件服务器，服务器之间通信依赖于网络区域划分”的逻辑链路，每一个模块都是安全知识图谱中的一个实体节点。

2、知识图谱构建流程

在确定了知识图谱的关键实体节点后，通过资产测绘、知识收集、知识抽取、知识融合、知识计算推理、知识图谱可视化完成资产安全图谱实践。

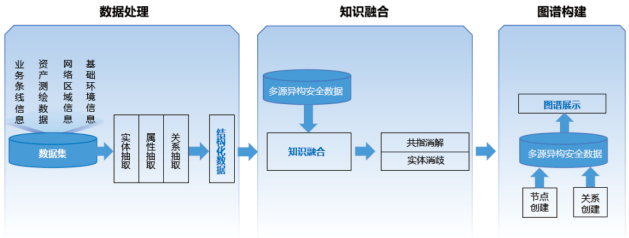


图6 知识图谱构建流程

3、应用成效

通过资产测绘、流量监听手段，结合现有CMDB数据库与手工填报的方式，将资产从业务视角进行逻辑串联，形成了层层递进的资产逻辑链条，赋能整个IT基础设施团队。从安全角度看，图谱中任意节点出现安全告警，可以快捷、高效地串联出风险链、横向连接区域等要素，为后续安全分析、应急处置夯实基础。

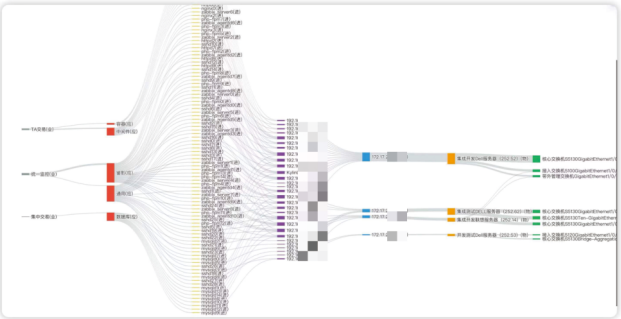


图7 资产知识图谱

在此基础上，结合并结合威胁情报、流量监测、漏扫、HIDS等现有安全能力进行知识融合，将安全视为资产的一类属性，形成完整的资产安全图谱。真正落地“安全作为业务属性”这一概念，可以直观了解到业务线及所属资产的安全防护现状、安全风险、安全事件。

图8 资产安全管理列表

（三）安全风险级别标准化实践

挂图作战是带有一种军事色彩的网络安全综合运营体系。从军事角度看，不同的对象具有不同军事价值，高价值对象往往面临着被更饱攻击的风险，当多个对象被攻击后，高价值对象一定需抢险修复。同理，站在安全运营角度来看，一方面高价值的业务线更易成为黑灰产团伙的目标，需要提高防范与监测能力；另一方面，在资源有限的情况下，应急响应时也应当遵循“抓大放小”原则，按照影响对象、影响范围、影响程度进行综合考量。

1、漏洞标准归一化实践

当前公司内使用的漏洞标准包括CVE、CNVD、CNNVD、AVD（阿里云漏洞标准）五种类型。结合内部使用率、易用性、可维护性等因素综合考量，以CVE作为东吴主漏洞标准，建设和维护东吴证券漏洞库。

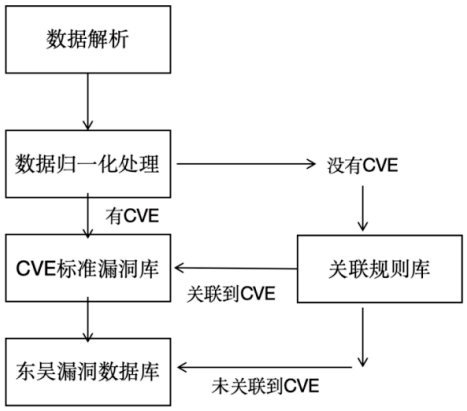


图9 数据归一化处理逻辑

数据解析：对异构的漏洞数据进行数据解析，得到统一的数据字段、数据格式，作为后续归一化处理的基础数据输入。

数据归一化处理：对基础数据输入进行漏洞标准字段提取。如CVE编号、漏洞特征、影响对象等。

CVE标准漏洞库：即MITRE公司运营的CVE公开漏洞库，通过CVE漏洞编号作为唯一标识符进行关联匹配；

关联规则库：用来存储异构漏洞标准编号与CVE标准漏洞数据的关联关系，用于不具备CVE标准编号的漏洞数据进

行归一化处理。

东吴漏洞数据库：归一化后形成的漏洞知识库，作为公司整体安全的“漏洞数据底座”赋能其他产品和安全工作。

漏洞名称	CVE编号（主键）	CNNVD编号	AVD编号	漏洞类型	披露时间	漏洞状态	危害等级
Splunk Enterprise Windows平台 modules/messaging 目录遍历漏洞	CVE-2024-36991	CNNVD-202407-071	AVD-2024-36991	未定义	2024-07-02	PoC	高危
GeoServer property 表达式注入代码执行漏洞	CVE-2024-36401	CNNVD-202407-085	AVD-2024-36401	CWE-94	2024-07-02	EXP	超危
OpenSSH Server 远程代码执行漏洞	CVE-2024-6387	CNNVD-202407-017	AVD-2024-6387	CWE-362	2024-07-01	PoC	高危
Ollama 目录遍历致代码执行漏洞	CVE-2024-37032	CNNVD-202405-4994	AVD-2024-37032	未定义	2024-06-25	PoC	中危
UI for Apache Kafka 后台 jmx jndi 代码执行漏洞	CVE-2024-32030	CNNVD-202406-2233	AVD-2024-32030	未定义	2024-06-20	PoC	高危
Magento estimate-shipping-methods XXE漏洞	CVE-2024-34102	CNNVD-202406-1362	AVD-2024-34102	CWE-611	2024-06-13	PoC	超危
PHP CGI Windows平台远程代码执行漏洞	CVE-2024-4577	CNNVD-202406-852	AVD-2024-4577	CWE-78	2024-06-07	PoC	超危
SolarWinds Serv-U FTP 目录遍历文件读取漏洞	CVE-2024-28995	CNNVD-202406-444	AVD-2024-28995	CWE-22	2024-06-06	EXP	高危
Triton Inference Server 远程代码执行漏洞	CVE-2024-0087	CNNVD-202403-2522	AVD-2024-0087	未定义	2024-06-06	EXP	超危
Apache Wicket XSLT 代码执行漏洞	CVE-2024-36522	CNNVD-202407-1346	AVD-2024-36522	CWE-74	2024-06-05	PoC	超危
Apache OFBiz 目录遍历致代码执行漏洞	CVE-2024-36104	CNNVD-202406-159	AVD-2024-36104	CWE-22	2024-06-03	PoC	超危
Ivanti EPM EventHandler SQL注入致命令执行漏洞	CVE-2024-29824	CNNVD-202403-5051	AVD-2024-29824	未定义	2024-06-01	PoC	超危
Rejetto HFS 2.x 远程代码执行漏洞	CVE-2024-23692	CNNVD-202403-5001	AVD-2024-23692	CWE-1336	2024-05-31	EXP	超危
CheckPoint Gateway 文件读取漏洞	CVE-2024-24919	CNNVD-202403-4696	AVD-2024-24919	CWE-22	2024-05-29	PoC	高危
Atlassian Confluence 登陆后代码执行漏洞	CVE-2024-21683	CNNVD-202403-4060	AVD-2024-21683	未定义	2024-05-22	PoC	中危
Zabbix Server Audit Log SQL注入漏洞	CVE-2024-22120	CNNVD-202403-3146	AVD-2024-22120	未定义	2024-05-17	EXP	中危
Nexus Repository 3 目录遍历与文件读取漏洞	CVE-2024-4956	CNNVD-202405-2896	AVD-2024-4956	未定义	2024-05-17	PoC	低危
Mura CMS processAsyncObject SQL注入漏洞	CVE-2024-32640		AVD-2024-32640	CWE-89	2024-05-09		低危
Apache OFBiz 目录遍历致代码执行漏洞	CVE-2024-26026	CNNVD-202403-1736	AVD-2024-32113	CWE-22	2024-05-08	PoC	中危
F5 BIG-IP Next Central Manager SQL注入漏洞	CVE-2024-26026	CNNVD-202403-4722	AVD-2024-26026	CWE-89	2024-05-08	PoC	低危
瑞友天翼应用虚拟化系统 uploave SQL注入漏洞		CNNVD-202403-021	AVD-2024-1723189	未定义	2024-05-07	CVE	低危
瑞友天翼应用虚拟化系统 bmsvcs.php SQL注入致代码执行漏洞			AVD-2024-1723009	CWE-89	2024-05-06	CVE	中危
Apache ActiveMQ Jolokia REST API 未授权访问漏洞	CVE-2024-32114	CNNVD-202405-256	AVD-2024-32114	CWE-1188	2024-05-02	PoC	中危
ZenTao PMS 项目管理系统身份认证绕过漏洞		CNNVD-202401-1132	AVD-2024-1720490	未定义	2024-04-25	CVE	低危
Pomocon EOS Platform jmx 远程代码执行漏洞		CNNVD-202405-12	AVD-2024-1720287	未定义	2024-04-24	CVE	低危
CrushFTP 认证绕过模板注入漏洞	CVE-2024-4040		AVD-2024-4040	CWE-94	2024-04-23	EXP	低危
Apache HugeGraph-Server v1.3.0 Gremlin命令执行漏洞	CVE-2024-27348		AVD-2024-27348	未定义	2024-04-22	PoC	低危
ioKiloView 任意文件上传致远程代码执行漏洞			AVD-2024-1710112	未定义	2024-04-17	CVE	PoC	低危
IP-guard WebServer 权限绕过漏洞		CNNVD-202406-6271	AVD-2024-1718774	未定义	2024-04-16	CVE	
		CNNVD-202405-5061					高危
		CNNVD-202405-5054					
Palo Alto Networks PAN-OS GlobalProtect 命令注入漏洞	CVE-2024-3400	CNNVD-202405-5053	AVD-2024-3400	CWE-77	2024-04-12	PoC	

图10 关联规则库数据示例

2、安全风险级别标准化研究

常态化安全运营中，每天面对海量告警，人员资源一定是不够的，难免顾此失彼。一方面需提高“降噪”能力，减少误报；另一方面面对海量告警，需能制定合理的关注优先级。当前安全运营平台针对安全事件提供“优先级”级别标签，但更多只是基于单点事件的风险级别给出该标签，缺乏整体视角考量。

通过资产安全图谱的实践应用，可通过IP、MAC等信息快速串联出资产所属业务线、应用系统、硬件设备甚至网络可达区域，但仍是碎片化的。故开展安全风险级别标准化研究工作，提出一套方法论指导开展安全事件定级策略优化工作。

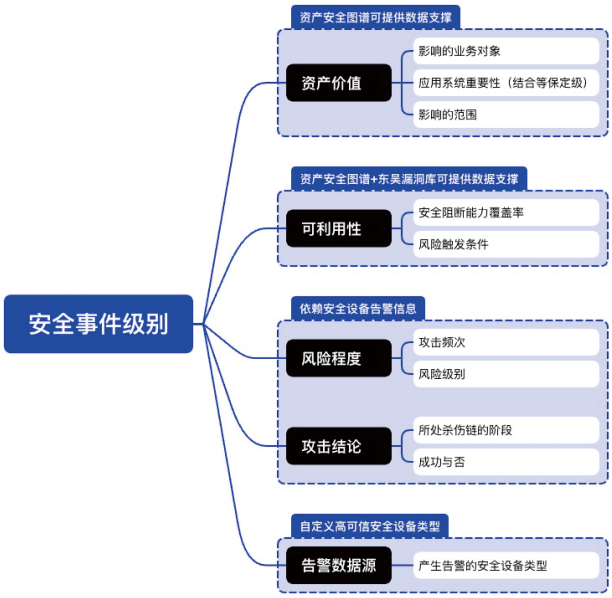


图11 安全事件定级关键要素

安全事件的告警级别主要依赖五个关键要素：

资产价值：需要给资产赋予价值属性得到资产权值，通过资产安全图谱可以获得资产完整链条信息。

可利用性：通过资产安全图谱可以知悉资产当前的安全防护情况，若处于阻断类设备的“安全保护区”中，则可利用性较低；其次漏洞类风险，触发条件较为苛刻的漏洞相应可利用性较低。

风险程度：依据安全设备的告警日志，安全设备均具备该能力。

攻击结论：依赖安全设备的告警日志，安全设备均具备该能力。

告警数据源：由于安全设备检测原理不同、部署位置不同等客观因素，对于其告警的信任度也各不相同，因此针对告警数据源也进行“高可信－可信－一般信任”的区分。

图12 安全事件告警等级调整实践

图13 安全事件告警等级调整规则示例

（四）资产风险监测挂图实践

在本期课题实践中，以资产安全图谱为资产数据底座、以安全事件定级标准为指标，结合安全运营平台和外部攻击面管理平台重构了风险监测数字化视图，将以安全事件为导向的风险监测升级为以资产（链）为导向的资产风险监测视图，提升可视化视图在安全运营中的应用价值。

图14 外部资产风险预警挂图

外部资产类型包括互联网资产、数字资产与数据资产，对这三类外部资产从漏洞、安全基线、攻击行为、诈骗仿冒、数据泄漏这五个安全角度进行监测预警从而可以快速了解当前外部资产整体安全风险，获取概览信息。基于资产安全图谱的实践，每个外部资产也有完整的资产链条，支持层层

下钻开展分析工作。

图15 内部资产安全态势

在内部资产风险监测挂图实践时，以网络拓扑为核心重构的专属内部资产安全态势，将网络拓扑关系映射到网络空间，实现网络关系、资产关系全方位可视化；借助资产安全图谱，可以直观看到该资产风险潜在的影响范围、影响路径，更有利于开展资产安全告警分析与处置工作。

（五）攻击路径监测挂图实践

安全的本质在于对抗，常说“未知攻，焉知防”，一方面对攻击手法有深入了解，另一方面也希望能直观、实时地掌握攻击者的攻击路径。基于这个诉求，开展了攻击路径监测挂图实践。

基于网络安全资产图谱的资产链条信息、网络拓扑信息，结合网络安全杀伤链2.0模型进行挂图设计，将攻击者视角与防御者视角叠加，形成一张完整的攻击路径地图。

通过外部入侵路径图可以直观、清晰地知悉攻击者在互联网侧的攻击行为，包含所使用的攻击、攻击流程、攻击所涉及的外部资产等信息。

通过内部横向路径图可以直观、清晰地获知攻击者在内网的攻击行为、攻击对象、攻击流程、攻击所处阶段和影响的内网资产范围。

结合资产安全图谱，可以更细颗粒度筛选出影响的资产范围，如API资产、业务应用资产；结合外部入侵路径图，可以更快研判出攻击行为是否奏效，如外部入侵路径图显示有攻击事件，但内部横向路径图没有该事件，该攻击事件很大可能在边界已被阻断，攻击未奏效。

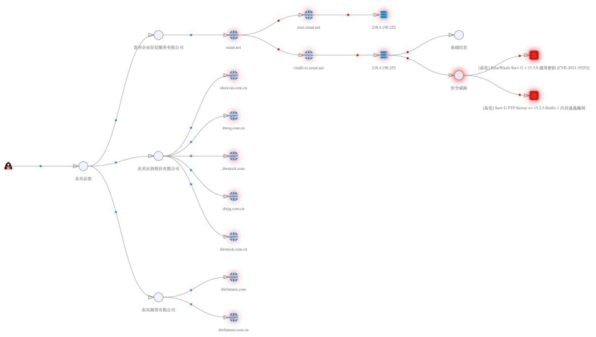


图16 外部入侵路径图示例

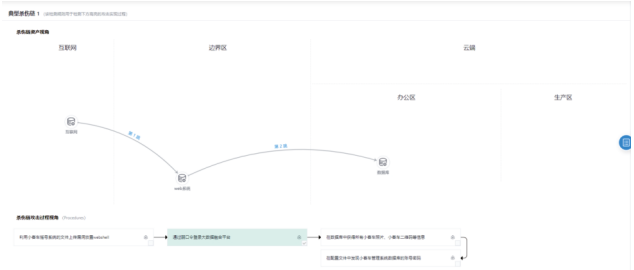


图17 内部横向路径图示例

四、总结与展望

通过本期课题，东吴证券进行了资产安全图谱实践、安全实践级别标准化实践、资产全生命安全管理实践、资产风险监测挂图实践、攻击路径监测挂图实践五项内容，将公司网络安全从“量”实现到“质”的提升，有效提升了安全运营工作效率，为网络安全决策提供数据支撑，真正实现了挂图作战的落地，而不仅仅是一张图。

未来，东吴证券也将在现有的基础上进一步强化挂图作战能力，以安全保障业务为目标，从切实场景需求着手，倒逼技术创新，引导优势技术与资源整合，不断提高自身网络安全综合运营水平，为行业总体安全能力建设作出贡献。

零信任架构下的业务系统敏感数据保护实践

王洪涛、刘宏、马晓鹏、黄施宇 | 国金证券股份有限公司

何艺 | 北京持安科技有限公司

摘　要： 本文聚焦于构建零信任数据安全访问体系，旨在为证券行业提供更安全可控的敏感数据保护方案。核心包括建立基于业务身份的敏感数据识别和资产分级分类机制，实施最小化权限管理，以及全面监控和审计数据访问。

零信任策略要求每次访问都经过严格的身份验证和授权，通过应用层代理实现安全访问。这种方法不仅能确保敏感数据的可知性、可察性和可控性，还能有效应对内外部威胁。通过整合身份认证、访问控制、数据分类、风险评估等技术，本课题旨在建立一个全面的数据保护体系，在确保合规的同时，平衡数据安全与业务需求，为证券行业的数字化转型提供坚实的安全基础。

关键字： 零信任、数据安全、动态访问控制、敏感数据

一、概述

近年来，互联网、大数据、云计算、人工智能和区块链等技术的创新融入经济社会各领域，金融服务业广泛利用这些技术进行决策和服务。随着证券业务数字化的快速推进，业务场景愈加复杂，科技架构更加智能化和隐蔽化，数据共享和交互成为高频应用。

《2018-2019年度金融科技安全分析报告》指出，大数据发展背景下，数据治理面临多重安全威胁和挑战。2023年《证券期货业信息安全运营管理指南》提出了严格的数据安全要求，包括身份认证、访问控制和审计等。

传统数据安全防护模式存在诸多不足，如部署成本高、影响用户体验、缺乏全面审计等。证券机构面临客户隐私和重要交易数据泄露的风险，尤其是员工可能受利益引诱售卖数据或不当分享。在此背景下，基于业务视角融合零信任理念的数据保护方案为证券行业带来了新的解决思路。

随着科技与金融的深度融合，数字化彻底改变了证券业务的运作模式，网络安全风险与业务风险同等重要。业务间的频繁交互和数据快速流动虽带来便利，但也加剧了数据安全问题的复杂性。

金融业数字化转型加速，新技术和业务模式不断涌现，带来了新的安全挑战：

1、业务敏捷发展下的安全风险：

■新业务可能引入数据安全漏洞

■需确保旧系统兼容性

■监管、审计和溯源难度增加

■可能导致数据丢失、隐私泄露和资金安全问题

2、传统数据安全方案不足：

■边界防御不足

■访问控制缺乏灵活性

■缺乏实时监控和响应能力

3、应用系统自身防护面临挑战：

■开放环境增加攻击风险

■业务系统漏洞可能导致数据泄露

■应对0day漏洞的能力亟需提升

本项目旨在通过零信任理念构建更安全可控的数据访问模式。零信任策略要求每次访问都经过严格的身份验证和授权，并通过应用层代理实现安全访问。在证券行业，零信任架构的数据保护具有以下意义：

1.保障业务发展与风险防护：增强身份认证和访问控制，确保系统安全稳定，细粒度控制敏感数据，预防泄露。

2.基于风险实施安全管控：根据业务需求和风险评估，实施精细化数据访问控制，降低泄露和误操作风险。

3.提供免开发的数据安全能力：标准化安全策略和控制措施满足合规要求，统一访问控制和审计确保遵守行业标准。

4.创新提升行业安全水平：零信任架构作为新型安全理念和技术，促进金融证券行业数据安全创新，增强抵御内外部威胁能力。

二、零信任体系下数据治理目标

（一）建立零信任数据安全访问体系

为实现敏感数据的全面掌控，建立零信任数据安全访

问体系至关重要。该体系包括以下关键内容：

1.识别和分类敏感数据

2.监控和审计数据访问

3.实施最小化权限管理

通过这些措施，可确保敏感数据的可知性、可察性和可控性，有效应对内外部威胁。

基于业务身份的敏感数据识别和资产分级分类是整个体系的基础，通常包含以下步骤：

1.全面审查组织内部数据，识别敏感信息

2.将敏感数据与具体业务相关联

3.根据业务身份和数据关联性进行分类分级

4.根据分类分级结果，制定并实施相应的安全控制措施

这种方法可以更有效地保护敏感数据，确保数据安全与业务需求的平衡。

（二）实现应用可信访问具备0day防护

基于零信任架构，建立完善的可信访问机制。采用多因素认证验证用户身份，并检测终端设备安全性，确保只有合法可信的用户和设备能访问应用。利用零信任网关实现细粒度资源访问控制，严格管理可能受0day漏洞影响的应用资源。对含0day漏洞的资源或接口进行精准阻断和隔离，同时保证无漏洞资源的正常访问，实现有效的0day防护。

（三）敏感信息访问控制、脱敏和溯源

零信任动态策略是一种基于最小化特权原则，在访问请求时动态评估用户身份、设备信息和行为情况，从而决定是否授予访问权限的安全策略。通过零信任动态策略可以有效实现敏感信息访问控制、脱敏和溯源，从而提高数据安全性和保护隐私。

以下是如何使用零信任动态策略来实现这些目标的步骤：

1.动态访问控制：利用零信任动态策略，在每次访问请求时对用户身份、设备信息和行为进行综合评估，根据评估结果动态决定是否授予访问权限。

2.敏感信息脱敏：访问敏感信息时根据动态策略对敏感数据进行脱敏或掩码处理，只有授权的用户才能查看完整的敏感信息，以减少敏感信息泄露风险。

3.访问溯源：记录审计用户访问敏感数据，即可一键溯源。

三、零信任体系下数据治理实践

（一）零信任数据安全管控平台

图1 零信任数据安全管控平台

1、可信代理应用研究与实践

可信应用代理如图1所示是企业网络中用于保护和控制应用程序访问的安全技术，为内外部用户提供安全可控的访问通道，同时保护企业敏感数据和资源。通常部署在业务区域边界，作为应用与用户间的中间层，主要功能包括：身份认证和授权，业务隐身，加密传输，细粒度访问控制，威胁防护，应用访问可视化。

2、零信任实时动态决策引擎研究与实践

在零信任安全架构中，动态决策引擎是安全控制中心的核心，负责评估和授权访问请求。它持续收集并分析多方面信息，如账号状态、设备安全状态、网络环境、应用安全信息、应用敏感数据等。

引擎根据预定义策略实时评估请求，动态调整权限。发现异常或风险时，可采取拒绝访问、要求额外认证或限制权限等措施。动态决策引擎在零信任架构中起关键作用，通过实时信息和预设策略做出智能决策，帮助企业建立更安全、灵活的应用访问控制体系。

（二）基于可信访问的信任链机制

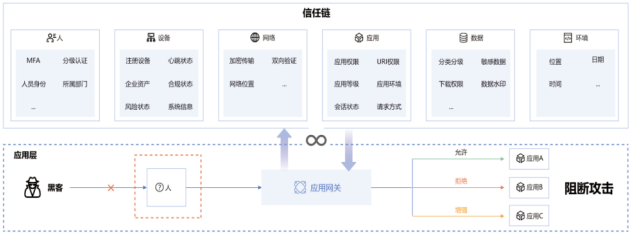


图2 基于可信访问的信任链机制

基于信任链的可信访问机制如图2所示，通过多维度实时动态信任评估来分析用户访问请求。评估考虑用户历史行为、地理位置、设备特征和上下文信息等因素，构建动态信任模型。每次访问都会重新计算信任值，并根据用户风险

等级决定是否允许、限制或拒绝访问。该机制确保多方面的可信性，如用户可信、设备可信、环境可信、传输可信、行为可信、权限可信等。

信任链机制不仅提高系统安全性，还能灵活应对不断演变的安全威胁。通过持续监控和实时反馈，系统能快速识别风险并作出响应。同时，这种方法在安全性和便利性之间取得平衡，避免过度验证影响用户体验。

（三）零信任动态决策引擎控制敏感数据访问和数据处置

图3 零信任动态决策引擎

零信任决策引擎如图3所示是零信任网络安全框架的核心，实时评估用户、设备和应用程序的可信度，动态调整访问权限和控制策略。它实现了敏感数据访问和处置的精细化控制，帮助组织建立智能灵活的数据安全管理体系，有效应对内外威胁。

利用零信任决策引擎控制敏感数据访问和处置的主要步骤包括：

1.实时评估用户身份和行为：对用户身份、设备状况和网络环境进行评估，确定可信度水平，动态调整访问权限。

2.动态控制访问权限：根据用户身份、角色和设备健康状况设置差异化访问权限，监控访问行为，发现异常时自动触发警报或阻止访问。

3.数据处置和监管：通过设定访问和操作规则，执行数据审计、脱敏处理等操作，确保敏感数据不被泄露或滥用，提高数据安全性和隐私保护。

（四）创新点

1.业务系统无缝融合：将零信任策略与证券业务系统整合，无需改造即可实现动态访问控制、数据可见性和追踪，满足金融行业特殊安全需求。

2.动态可信访问控制：通过应用层零信任策略，实现可信验证和细粒度访问控制，最小化攻击面，提高数据安全性。

3.敏感数据可见性：从人员角度分析数据访问关系，精确定位访问时间、人员、系统和数据类型，突破传统IP分析局限，形成完整访问上下文。

4.敏感数据动态处置：基于实时风险策略，在数据返回过程中进行动态处理，包括访问控制、下载限制、脱敏、水印和追踪，提供创新的数据保护解决方案。

四、零信任体系下的数据治理应用成果

国金零信任平台数据访问控制技术基于最小权限原则和持续授权的思想，实现了对数据访问权限的细粒度控制和动态调整。通过访问控制策略、安全策略引擎和审计机制等组件，对用户和设备的访问请求进行实时监控和决策，并根据实际访问情况和风险评估动态更新授权策略，从而确保敏感数据的安全访问。

目前两地三中心架构实施搭建两套零信任平台、使用15台服务器，8台网关。数据安全防控，业务场景覆盖等方面已初见成效，覆盖了400名用户、上线403个应用系统以及各种业务请求数据。涵盖了终端信息检索、攻击检测、文件外发、合规检测、应急响应、敏感数据防护等七种数据检测类型。同时，我们还内置了金融证券行业模板和通用模板，共有162条敏感信息条例和154条敏感规则，以及275条敏感词库。此外，我们还实现了数据安全API接口分析、数据识别、附件水印识别等三种管控审计策略模型，业务覆盖率已经达到了90%。鉴于在复杂业务场景下卓越的零信任数据安全能力、面对未知威胁的安全防护能力、显著提升企业业务效率等优势。

（一）平台建设情况说明

1、系统部署架构

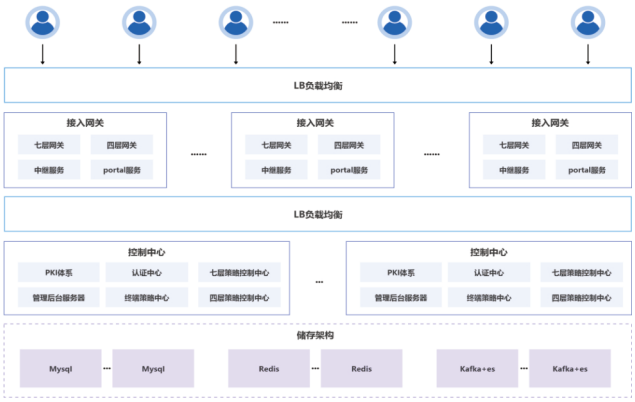


图4 平台分层架构图

如图4所示零信任平台由接入层、控制层和存储层组成，每层均具备负载均衡和高可用能力，确保服务持续可

用。采用双数据中心分布式+高可用部署架构，两个数据中心部署相同组件，本地采用高可用架构，异地采用分布式模式。整个系统在各层级均有高可用设计，最大程度满足国金证券业务连续性要求。

平台采用"两地三中心"部署架构，15台服务器分布在上海和成都，包括8台网关、4台控制中心和3台日志中心。网关分为4个区域（一办公三交易），区域内两两互为主备；控制中心在两地各两台互为主备。此架构实现了网关和控制中心在区域内外的多重高可用性。结合实际业务需求、突发状况切换和主动灾备演练，该架构符合设计预期，具有极高的可用性价值。

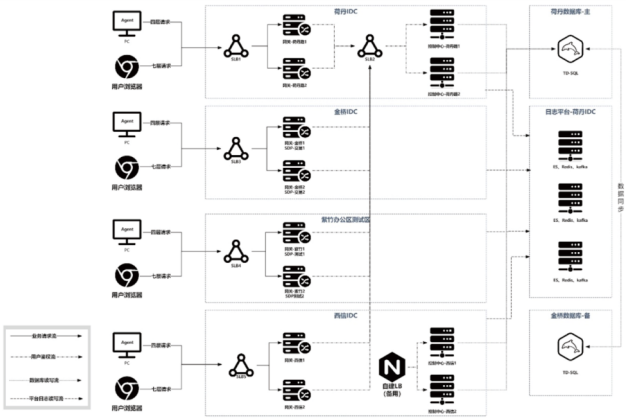


图5 部署架构图

如图5所示，网络层高可用设计采用分布式+高可用模式。零信任网关在各数据中心采用双主设计，确保转发链路可用性。决策中心采用分布式集群部署，保证故障时的快速切换。

应用层高可用采用独立组件设计，避免网络层异常影响应用层网关。

数据层高可用涉及平台配置、身份信息和数据日志存储。配置信息采用分布式高可用架构，每个数据中心部署两个数据库节点，实现实时同步和故障接管。

业务层高可用强化了应用层Bypass设计，包括三个维度：单应用Bypass、全网关Bypass和网关节点快速摘除。这确保了业务连续性和灵活的故障处理。

基于国金证券的需求和实际业务场景，本方案以零信任框架为基础，完善现有防御体系，构建以业务为核心的动态访问控制能力，在用户和业务间建立持续的安全访问信任链，实现精细化访问控制。

2、敏感数据分类分级

如图6所示系统自动识别敏感数据信息，并将企业内部敏感数据分级和分类，自定义/导入敏感数据识别规则，敏感数据访问审计。满足数据分类分级的合规要求，实现办公网敏感数据可视可知可溯源。支持金融行业的敏感数据分类分级，例如对个人敏信息分类分级：姓名、银行卡号、密码、住址、联系方式、身份证号、交易信息等信息。

图6 敏感数据分类

3、敏感数据识别与管控

如图7所示，系统通过基于应用层数据分析，可以准确还原业务系统数据，利用敏感数据标签，可以帮助安全团队及时发现系统中存在的客户敏感信息，并对访问行为实现可视化跟踪审计和数据管控。

图7 敏感数据识别流程

4、数据安全动态策略模型

数据安全动态策略模型是一种基于实时分析和智能决策的方法，用于动态调整数据安全策略，以应对不断变化的威胁和风险。该模型整合了实时监测、数据挖掘和机器学习等技术，有助于组织及时发现并应对安全事件，提高数据保护效果。

如图8所示模型的主要特点包括：

1.实时监测与分析：采集和监测数据访问行为、安全事件和上下文信息，实时评估数据使用情况。通过监控关键指标，及时识别潜在风险。

2.动态策略调整：根据实时分析结果，灵活调整安全策略和控制措施。针对不同场景和风险等级，修改访问权限、审计设置和加密策略等。

3.自适应安全控制：建立自动响应机制，根据实时监测结果和安全策略调整，采取相应的控制措施。如自动阻止异常访问、禁止非法操作和通知相关人员等应急响应。

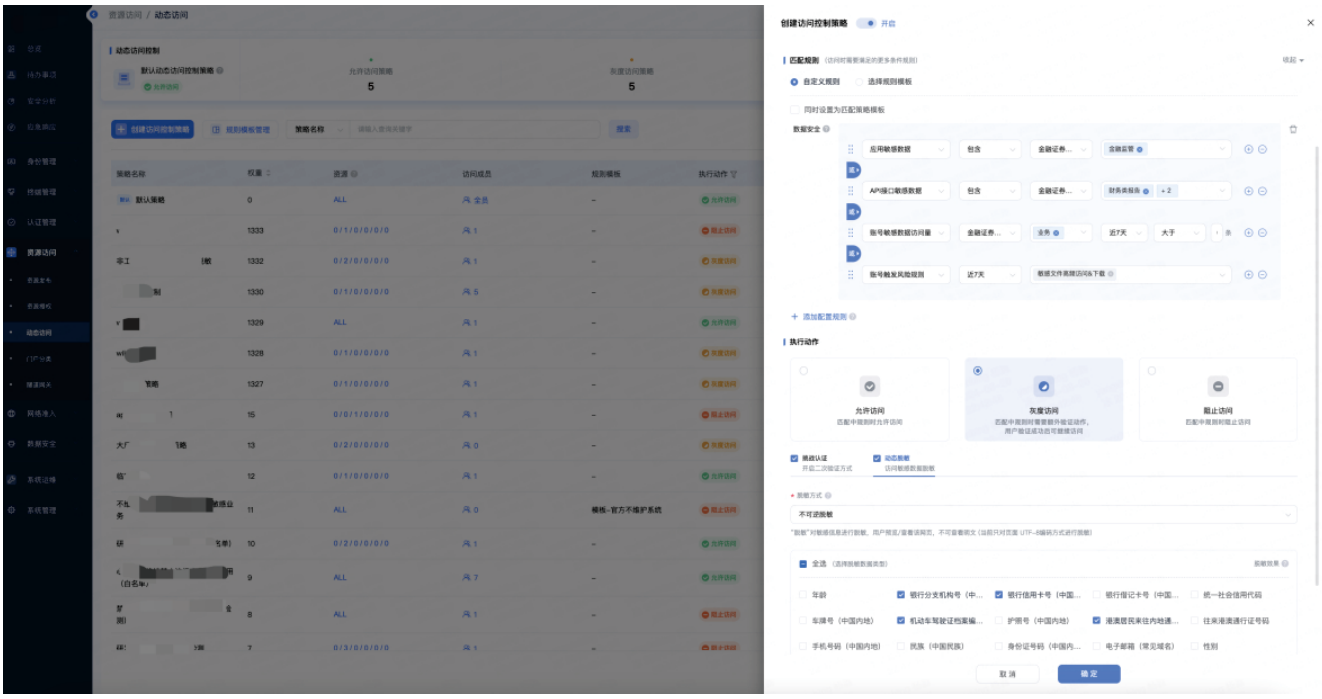


图8 数据安全动态决策

5、应用漏洞热补丁防护

零信任应用网关采用"先认证，再访问"的方式，可有效防御未知漏洞的匿名攻击。对于已知漏洞（如log4j漏洞和办公应用系统在攻防演练中发现的0DAY漏洞），由于原系统厂商修复周期较长，如图9所示，零信任应用网关能迅速实施URI级别的精准防护策略。此外，网关的身份识别能力可对特定用户授权访问或下发安全策略，既保证业务正常运行，又能提升安全防护效果。

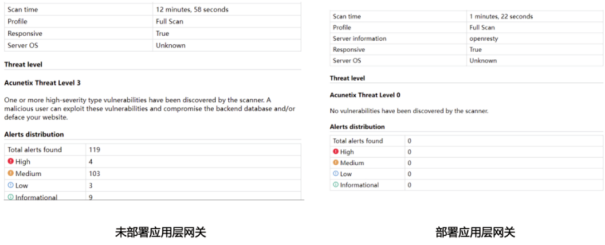


图9 应用漏洞防护

6、应用动态水印防护

如图10、11所示，无需改造业务即可实现动态水印配置。支持明水印、暗水印及自定义内容，便于信息泄露时快速定位源头。鉴于敏感数据常通过非正规API接口泄露，企业应识别和监控关键应用系统的API接口，发现异常和滥用情况。同时支持对指定应用附件实施访问和下载管控。水印功能适用于PDF、DOCX、XLSX、PPTX等文件格式。

图10 应用水印防护--明水印

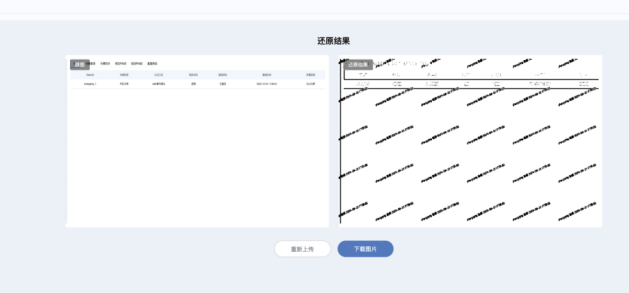


图11 应用水印防护--暗水印还原

（二）、成果应用

零信任敏感数据安全在多种场景中得到成功应用，包括内网办公、业务处理、远程办公、开发测试、外包运维和数据治理。通过实施零信任模型，企业可有效防范内外部威

胁，提升整体安全水平。

1.内网办公：防止内部人员跨权限访问，减少内部威胁和数据泄露风险。

2.业务处理：保护客户数据和业务信息，实现全链路审计防护，确保敏感数据安全和可溯源。

3.远程办公：通过多因素身份验证和终端设备检测，安全管理远程员工对敏感数据的访问。

4.开发测试：限制开发人员和测试团队仅能访问特定测试数据，监控活动防止不当操作和数据泄露。

5.外包运维：确保外包方仅在临时授权下访问必要数据，监控活动保障数据安全。

（三）能力提升

1、应用安全和数据安全一体化防护，未知风险防护能力

图12 一体化防护

如图12所示，零信任网关无需改造应用即可实现HTTPS全面覆盖，提升应用访问安全性并将业务暴露面集中到网关。通过严格控制用户身份和访问权限，阻止未经授权用户访问，防止应用暴露，并实现应用访问流量的身份化，追踪用户行为，精准定位安全问题。同时，自动防御未知风险，减少漏洞修复次数，并提供无需特征升级的保护能力。

此外，零信任网关的数据安全功能可监控敏感文件，支持灵活的数据脱敏规则配置，并提供数据流转地图，帮助管理员全面掌握企业敏感数据访问情况，有效防止数据滥用、泄露和合规风险。通过整合应用安全和数据安全，零信任网关形成了全方位、多层次的安全防护体系，这种一体化方法有效应对复杂安全威胁，提升整体安全防护能力，降低安全管理复杂性，提高安全防护效率和效果。

2、全面提升基于精准业务身份的数据访问可见性

通过识别敏感数据并根据其重要性和敏感程度对其进行分类，将敏感数据与具体的业务相关联；并严格遵守零信任最小权限授权管控理念，同时确保数据访问的可见性，帮助企业可以更好地监控和管理员工、合作伙伴、供应商等各方对敏感数据的访问行为，及时发现异常操作和潜在威胁，并采取相应的安全措施来保护数据安全。

3、免改造下的敏感数据合规管控能力

在保持现有系统架构不变的情况下，提升敏感数据的合规管控能力可通过以下方式实现：

（1）数据分类标记，实施差异化管理
（2）加密敏感数据，确保传输和存储安全
（3）严格访问控制，限制用户权限
（4）数据脱敏和匿名化处理
（5）定期备份，确保数据可快速恢复
（6）实时监控和审计数据访问
（7）定期进行风险评估和漏洞修复
（8）遵循并更新相关法规和行业标准

综合应用这些方法可有效提高敏感数据的合规管控能力，保障数据安全，降低泄露风险，同时符合合规要求。

4、基于场景的动态风险控制能力

基于场景的动态风险控制是一种实时监测、评估和应对不同环境下潜在风险的方法。实现这一能力的关键措施包括：

场景化动态风险控制是一种实时监测、评估和应对多样环境风险的方法。其关键措施包括：

（1）实时数据监测分析，为决策提供支持
（2）运用智能算法构建风险评估模型
（3）根据情况灵活调整控制策略
（4）建立可视化决策支持系统
（5）加强跨部门协作和应急响应
（6）持续总结经验，优化控制策略

通过综合实施这些措施，组织可有效应对复杂多变的风险，确保业务安全稳定发展。

5、业务敏捷发展能力

无需业务改造，如图13所示通过内置脱敏算法与动态脱敏配置，对高敏数据自定义选择脱敏类型，如：身份证号、银行卡号、手机号、邮箱等；脱敏效果如图14所示。

图13 动态脱敏策略

图14 脱敏效果

参考文献

[1]乔梦梦,李彦彪,王嘉源.数字化转型背景下金融数据安全面临的风险及对策建议[J].金融科技时代,2023,31(12):76-80.

[2]黄莉群,官心果,钟宇.数字经济时代的数据安全研究——以金融行业为例[J].商业经济,2024(02):174-179.DOI:10.19905/j.cnki.syjj1982.2024.02.046.

[3]刘人杰.数字时代金融数据安全保护的思考[J].科技与金融,2022(05):82-86+95.

[4]CSO Online. (2023). Zero Trust Security Model Explained. CSO Online.

[5]钟红,马天娇.金融数据安全风险及监管研究[J].清华金融评论,2021(10):96-98.DOI:10.19409/j.cnki.thf-review.2021.10.026.

[6]李松涛,谢宗晓.数据资产化时代的金融数据安全[J].中国信息安全,2021(05):37-38.

[7]王京晶.新形势下金融数据安全面临的挑战与思考[J].金融科技时代,2020,28(08):53-54+58.

[8]National Institute of Standards and Technology (NIST). (2020). Special Publication 800-207: Zero Trust Architecture. National Institute of Standards and Technology.

[9]李振魁.零信任下的数据防护研究[J].网络安全技术与应用,2023(01):60-62.

[10]周潮洋,谢琴.零信任理念下的企业新型安全技术防护体系研究[J].网络安全技术与应用,2022(02):99-101.

[11]Forrester Research. (2023). The Zero Trust eXtended Ecosystem. Forrester Research.

[12]Google LLC. (2024). Zero Trust Security with BeyondCorp. Google Cloud.

[13]Microsoft Corporation. (2023). Zero Trust Security Model. Microsoft Azure.

[14]柏东明,曾丽花,董之光.企业网络零信任架构应用研究[J].信息系统工程,2022(12):15-18.

[15]周岳亮.基于零信任安全模型的数据中心安全防护研究[J].网络安全技术与应用,2020(10):88-89.

[16]SANS Institute. (2023). The Zero Trust Journey. SANS Institute. Retrieved from https://www.sans.org/white-papers/

[17]Cittadini L, Spear B, Beyer B, etal. BeyondCorp: The Access Proxy.[J]. ;login:, 2016(Vol. 41, No. 4):28-34.

[18]Rose, S. , Borchers, O. , Mitchell, S. and Connelly, S. Zero Trust Architecture, Special Publication (NIST SP), National Institute of Standards and Technology, Gaithersburg, MD. 2020. https://doi.org/10.6028/NIST.SP.800-207

[19]Ward R, Beyer B. BeyondCorp: A New Approach to Enterprise Security.[J]. ;login:, 2014(Vol. 39, No. 6):6-11.s

东方证券企业终端数据安全解决方案探索

邬晓磊、甄明达 | 东方证券股份有限公司
焦健、汤华晟 | 数篷科技（深圳）有限公司

摘　要： 当前，数据已成为全球产业升级和数字化转型的核心要素，为经济发展提供强大动力。然而，数据安全作为这一过程的基石，面临着数据流通量激增带来的挑战，尤其在金融机构数字化转型中更为凸显。国内外政策环境对此积极响应，我国通过《十四五数字经济发展规划》等多部法律法规，构建数据安全治理体系。全球范围内，数据安全与个人信息保护也受到高度重视，相关法规不断完善。东方证券在此新形式下，通过研究和实践，采用零信任沙箱、文件流转网关等技术强化数据安全，实现数据的分级分类管理和全生命周期保护。

关键字： 零信任沙箱、非结构化数据分类分级、流转控制、文件溯源

一、课题背景和研究内容

（一）课题背景

当前数据正成为全球产业升级，数字化转型的“血液”，为经济发展带来源源不断的“养分”，助力产业发展。而数据安全正是这个过程的基石，数据流通环节和数据量的显著增加，正在推动数据安全需求的大幅增长。同时，金融机构数字化转型是当下金融发展的重要趋势，伴随着大数据、人工智能、云计算等技术的广泛运用，大量金融数据快速频繁地交互流转，数据属主和管理边界愈发模糊，数据泄露、滥用、窃取等安全威胁日益加大。

在此背景下，国内外政策环境相继做出响应，我国《十四五数字经济发展规划》中重点提出了建设数据安全治理体系，完善行业数据安全管理政策的要求。2017年颁布的《网络安全法》已经对数据安全做出了相关的基础规定，2021年9月《数据安全法》正式颁布实施，同年11月我国《个人信息保护法》也正式实施，立足于数据产业发展实践和个人信息保护的迫切需求，更全面的保护了个人权利。

在金融领域，2018年5月银保监会发布《银行业金融机构数据治理指引》引导银行金融机构加强数据治理，提高数据质量，充分发挥数据价值。2020年9月中国人民银行正式发布《金融数据安全　数据安全分级指南》，给出了金融数据安全分级的目标、原则和范围，明确了数据安全定级的要素、规则和定级过程。2021年4月人行进一步发布《金融数据安全　数据生命周期安全规范》，在数据安全分级基础上，结合金融数据特点，梳理数据安全保护要求，形成覆盖数据生命周期全过程的、差异化的金融数据安全保护要求，并以此为核心构建金融数据安全管理框架。

随着信创设备、移动办公逐渐成为东方证券办公场景中的常态，传统的安全系统无法覆盖信创终端以及员工个人设备，如何在信创环境中应用一套适配信创操作系统、传统windows、MacOS等多种操作系统终端的数据安全解决方案是一项挑战。因此东方证券正在探索一条可行的信创数据安全能力建设之路。

（二）研究内容

本课题计划通过零信任接入、终端内核级数据隔离、文件控制网关等新型理念、技术的整合打造新一代企业终端数据安全平台实现员工企业终端数据安全隔离、权限可控。

本课题研究内容包括：

1.零信任沙箱—在各型终端建立可信计算区域，可信区域将从零信任隧道获取的企业数据进行隔离保护；

2.文件流转网关—在同一员工的各类终端上进行数据安全同步，对于虚拟云桌面终端提供专用客户端，非虚拟云桌面终端环境文件流转网关和零信任沙箱进行强耦合；

3.零信任安全隧道—用以控制非内网终端访问企业应用，零信任隧道和零信任沙箱强耦合使用；

4.非结构化数据分级分类—对流转网关管理的数据进行分级分类处理赋能东方证券审批人员在数据流转审批时快速处理。

（三）研究意义

本次课题研究探索在东方证券目前“自有设备和配发设备混合”、“实体终端和虚拟云桌面混合”、“传统和信创混合”等多类型终端混合办公场景下，通过终端零信任沙箱、文件流转网关、零信任安全隧道、非结构化数据分级分类等新技术的整合、探索构建东方证券办公终端数据使用过程中的安全保障能力。通过本课题的研究和推广，未来东方证券员工可使用各种类型终端安全获取、安全存储、安全流转企业文档类数据，实现精确分级分类和访问控制。

二、企业终端安全基础设施

（一）零信任沙箱

零信任沙箱技术主要是通过可信终端沙箱容器、软件定义网络、软件定义边界、后端文件存储、零信任安全引擎等关键技术构建的。

1、可信终端沙箱容器

该模块提供终端数据安全隔离环节，构筑安全工作空间。

加密与隔离——在安全空间内数据存储高强度加密，密钥由零信任安全大脑统一管控；数据可在安全空间内灵活使用，还原操作系统原生使用体验，流畅兼容；安全空间提供深度隔离环境，安全空间内数据不能通过任意途径非法提取到安全空间之外。

数据转储——安全空间内的重要数据可选择同步至后端文件存储，特别是针对外发导出或共享给其他人员的数据文件则必须同步至后端文件存储。这里一方面提升了重要文件的可用性，即使设备丢失或损坏，数据依然可用；另一方面，后端存储上的数据可依托后端服务器强大算力进行进一步的细粒度处理。

为了保障数据在任何环境下、任何终端上灵活使用，该系统在终端设备上构筑软件定义的安全计算环境。安全计算环境有以下特点：第一，数据可在环境内灵活存储和计算；第二，安全计算环境内所有计算过程被全程监控等。

2、软件定义边界

在当前的数据应用场景中，数据的流动性、分布的广泛性大大增强。基于对整个网络环境的不信任前提，平台在安全计算环境之间建立安全的数据流通通道，保障数据的流转安全。

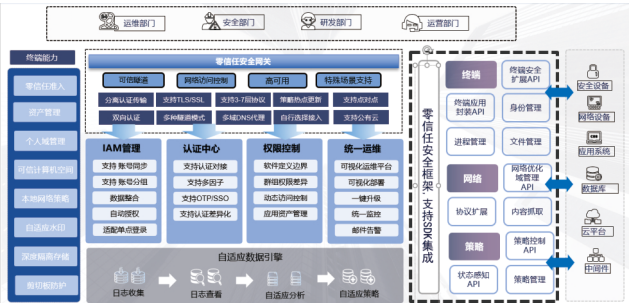


图1 零信任沙箱能力

通过零信任沙箱的建设完成后，员工配发设备、个人设备只能通过零信任沙箱访问企业应用获取企业数据，对于原分布于配发设备、个人设备上的企业数据进行第一次归集。

3、后端文件存储

该模块基于分布式文件存储系统，具有高可用、高性能、高扩展架构，可充分保证承载数据的可靠性和访问体验，也支持使用量的灵活弹性伸缩。

数据转储——后端文件存储承载来自终端侧安全空间内的重要数据文件的转储，通过多副本的方式保证数据可靠性，也同时增强了数据的访问性能。

后端文件存储上的文件生命周期可配置，可根据需要配置不同文件存储的有效期限，以达到资源高效利用的效果。

4、零信任安全引擎

该模块是整个平台的中枢，具有统一协调管控，流程支撑，智能分析，运营运维的复合能力。

策略与配置管理——零信任安全引擎提供整个系统平台各种安全策略与流程的配置与管理。管理各类主体包括：人员、设备、数据、应用，以及主体间的关系；以及各类变更和配置过程的流程以及审批流的支持。支持策略配置、策略审批、策略审计的三权分立管理。

全流程行为审计——本系统平台上发生的事件和行为会全量被零信任安全引擎记录，并结构化组织和关联；支持行为日志和事件的查询搜索，以及关联信息的快速定位。

智能文件溯源——零信任安全引擎将根据两类技术进行文件溯源。首先，上述全流程行为审计流水经过关联分析，可有效恢复安全空间内部数据流转路径；第二，通过基于人工智能的暗水印技术可针对离开安全空间，外发导出的数据甚至数据片段进行跟踪溯源。

异常检测与持续验证——零信任安全引擎能够基于上述全流程行为结合时间、空间、环境等状态信息进行异常事件检测，并基于异常事件序列进行行为意图分析，达到持续验证的效果；即会对系统平台上发生的所有行为进行持续的判断，对不合规或非法的行为进行告警以及响应处理。

运营与运维——零信任安全引擎全面支持该系统平台的运营运维功能，包括：运营数据的统计分析与报表展示，基于三权分立的系统权限管理，系统后台与客户端平滑升级，客户端热补丁管理，系统容量管理，可用性监控与告警等功能。

（二）文件流转网关

文件流转网关设计和零信任沙箱强耦合用以对沙箱归集的企业数据进行统一上传至存储并且开发了VDI专用客户端用以归集VDI上的企业数据。

网关对于不同员工零信任沙箱内数据交换以及同一员工间不同终端设备上的数据流转进行权限控制、安全策略执行、自动同步至服务端保存，文件流转网关的建设后续为文件级分级分类提供数据底座。

文件流转网关能力如下：

1、文件统一汇聚

文件流转网关作为非结构化数据分级分类的数据底座将原先分散在VDI实例、公司配发设备以及个人设备上的企业数据进行统一归集/汇聚以便后续分类分级。

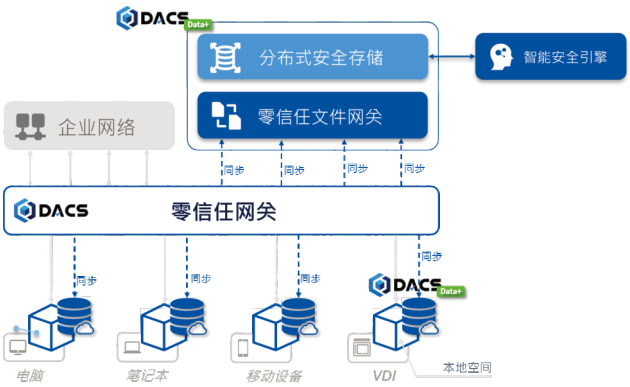


图2 文件网关运作机制示意图

2、文件安全漫游

为了解决员工远程办公时利用各种智能设备安全、高效访问员工个人数据，采用新一代文件流转技技术打造安全文件共享网关—Data+。Data+和安全工作空间客户端整合后分别为VDI、PC和智能移动设备提供专用客户端大大提高员工使用体验。

虚拟云桌面所有计算和数据均位于内网区域，安全文件流转网关客户端在云桌面自动登陆后客户端自动在VDI桌面挂载对应盘符的同时接管VDI常用的文件存储目录，例如Windows桌面的文档、下载、桌面；员工日常使用中对于文件的保存均通过文件网关同步至公司存储中，实现了类似Apple iCloud和微软OneDrive的效果。

信创设备、存量PC设备以及个人设备使用的安全工作空间客户端均整合Data+客户端保障员工从任意设备上登录均可以按需获取个人办公数据实现员工办公数据全终端安全漫游。

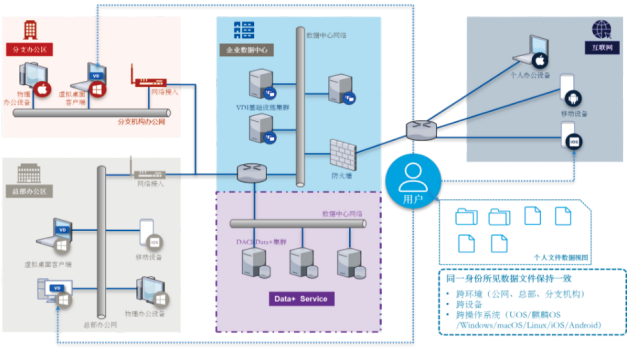


图3 文件流转网关运行示意图

3、文件分级分类

随着自然语言处理、图像识别、模式识别和视频处理等技术的成熟，在数据安全领域深度应用这些技术，可将数据的识别能力提升到新的高度。另一方面，结合高性能网络技术，可以将终端文件发送至服务端利用服务端更丰富的计算资源完成更加精准的识别。

东方证券采用文件网关的分类分级组件对归集的数据全盘扫描后进行标签化分类，分类后的结果服务于整个平台文件流转控制、外发审批辅助等文件流通行为。

针对日益增长办公终端数据应用多样性需求，我司与可信沙箱供应商联合构建面向数据流动的零信任安全解决方案，引入安全即基础设施理念，一体化解决了企业内部数据的静态访问控制，与数据流动的动态安全问题；具有技术领先，应用创新，安全高效，使用便利，体验自然，适应性、鲁棒性、可扩展性好等特点。

（三）零信任安全隧道

零信任概念的提出，推动企业安全从单纯的网络安全，转向资源访问安全，从而实现零信任无边界安全。《中国网络安全产业白皮书》中也指出在国内零信任正从“概念”走向落地。工信部在《关于促进网络安全产业发展的指导意见（征求意见稿）》中也将可信计算、零信任安全定义为“网络安全关键技术”。将零信任理念与相关技术应用于非结构化数据的访问控制保护，可以在保障数据安全性的同时释放数据的流动性，持续的精准的控制数据的访问过程，从而适应行业发展中日趋灵活多样的数据应用场景。

该模块部署于企业内网入口处，负责与终端网络边界管理模块对接，提供企业内网对终端安全计算环境的接入功能，是端到端的加密链路进入企业内网的入口。同时，基于链路对端的终端安全计算环境的上下文信息，该模块可依据安全策略决策结果，对该链路限制可达网络范围，保障数据安全计算域边界的落地管控，实现零信任网关功能。

三、课题亮点

（一）建立终端安全工作空间

为了解决企业数据在终端上存储的问题，采用操作系统内核级隔离技术打造了企业安全工作空间（可信沙箱容器）用来在终端上安全存储企业数据。操作系统提供了大量的驱动、服务、注册表（Windows）供应用系统使用，通过内核级沙箱对于这些启动、服务进行统一的虚拟、加固、隔离从而在终端上打造安全的数据隔离区和可信计算环境配合高性能SDP隧道，从而实现企业数据通过安全隧道传输至终端后存储于安全工作空间的加密虚拟磁盘内和本地文件系统完全隔离。

通过内核级隔离，存储于可信沙箱容器中的数据无法被容器外的app读取到、可信沙箱容器内部的应用如对磁盘进行写操作也会被可信沙箱容器重定向至可信沙箱容器专用的加密虚拟磁盘内进行保存通过整体的安全隔离保证企

业数据在任意终端上不可被可信沙箱外的非授权应用读取到。通过在终端上严格的数据隔离，传统的复制/黏贴、截屏/录屏、蓝牙以及较新型的NFC近场通信、进程间通信、共享内存等手段均无法获取到可信沙箱容器内的企业数据保证企业数据在终端上的存储安全。

通过企业安全空间的打造在终端数据层保障了员工在配发的信创设备、个人设备上均可以安全存储企业数据，通过和SDP高性能安全隧道的融合，东方证券打造了支持远程办公的安全基础设施平台，支持员工在任意网络、任意终端、任意时间安全远程办公。

（二）通过AI建立“文件-标签”映射关系

通过对东方证券归集的结构化数据进行扫描后进行标签化处理后服务于数据流转管控。根据规划，首期对下表用户进行分类分级扫描:

表1-分类分级用户表

用户姓名	用户账号	文件数量（个）	主要数据
用户 1	dxxa	675269	金融数据查询导出、测试
用户 2	lxxi	667501	系统运维、测试
用户 3	wxxo	530479	研发设计、测试
用户 4	xxxu	329131	测试
用户 5	gxxg	308466	研发设计
用户 6	xxxi	132695	对账单
用户 7	zx1	99446	服务、项目管理
用户 8	cxy	40797	基建
用户 9	zxxn	26061	服务、合同协议
用户 10	yxxg	24202	系统运维、项目管理
用户 11	wxi	11333	综合
用户 12	wxi	10234	开发

通过对上表用户目录的扫描，共计2855614个文件其中可被正常解析的占39.29%，不可被正常解析的占比60.71%。经过对不可被解析数据的审查其中87.2%为图片、音频、视频等暂时不支持解析的内容，该类不可被解析的内容均通过文件后缀识别的方式被证券归入“图片”、“音频”、“视频”标签，其他可备正常解析的文件根据《金融数据安全 数据安全分级指南》的标准进行标签化处理；剩余的12.8%不可被解析的非结构化数据经过人工审查均为系统文件、脚本及开发工具自动生成的资源类文件，该类不可被分类的数据采用人工审核、手动标签的形式进行补足。

（三）建立零信任高性能接入隧道

为了应对传统VPN权限与账号强关联无法判断根据终端、网络、应用等环境状态控制访问权限的现状，基于软件定义边界的模型打造高性能接入隧道，实现了对远程接入时安全隧道的双向认证以及基于用户身份、设备风险情况、网络环境等多种客观环境构建自适应安全体系，对接入员工设备、网络风险进行持续性评估适时调整用户可访问的系统资源保证应用的访问安全。

通过高性能隧道的建设在网络接入层保障企业应用的安全访问，结合终端层企业安全工作空间、原数据中心安全基础设施构建了完整的“云-管-端”一体化远程安全办公平台在符合ZTNA（Zero Trust Network Access）标准的前提下额外构建了终端企业数据保护能力，保障员工远程办公场景下的企业数据安全。

安全工作空间集成双向认证的零信任网络隧道实现用户终端的网络隔离，只有安全工作空间内程序具有隧道使用权限结合安全工作空间的隔离能力保障访问企业数据安全。每个安全工作空间单独配置网络访问权限（ACL）能力，保障员工网络资源最小化访问权限。

（四）建立员工数据多终端安全漫游能力

为了解决员工远程办公时利用各种智能设备安全、高效访问员工个人数据，采用新一代文件流转技技术打造安全文件共享网关—Data+。Data+和安全工作空间客户端整合后分别为VDI、PC和智能移动设备提供专用客户端大大提高员工使用体验。

虚拟云桌面所有计算和数据均位于内网区域，安全文件流转网关客户端在云桌面自动登陆后客户端自动在VDI桌面挂载对应盘符的同时接管VDI常用的文件存储目录，例如Windows桌面的文档、下载、桌面；员工日常使用中对于文件的保存均通过文件网关同步至公司存储中，实现了类似Apple iCloud和微软OneDrive的效果。

信创设备、存量PC设备以及个人设备使用的安全工作空间客户端均整合Data+客户端保障员工从任意设备上登录均可以按需获取个人办公数据实现员工办公数据全终端安全漫游。

（五）建立员工企业数据数据流转管控能力

员工通过管理存储于Data+的企业文件/数据的读写权限、分发范围、生存周期等权限；文件所有人配置分发范围后文件的接收人也不可超范围共享该文件。通过对源文件分发范围的管理，保障公司内的数据均可进行源头控制，保障数据不可被超范围使用。

通过Data+文件流转网关的建设，建立了完全私有化的“OneDrive”保障员工在任意时间任意地点安全、便捷地获取、使用、存储办公数据。

四、课题收益

（一）建立企业数据边界

终端安全工作空间旨在建立柔性的数据安全边界，其由终端侧的安全工作空间和零信任网络接入体系构成。终端侧的安全工作空间为企业数据在终端设备上开辟隔离出一块安全可信的数据存储与使用环境，企业数据在其中可以被受控使用。既控制了数据泄露风险，又实现了“数据可用，不可拿”的效果，在数据安全性、可用性方面给出了更好的技术实现方案。

而另一方面，数据孤立在单个安全工作空间内是无法充分发挥其价值的。为了让数据安全地流动起来，东方证券将终端安全工作空间与零信任接入网络有机融合起来，让数据通过零信任网关，可以在基于身份、持续验证的零信任访问控制监管下，与后端的企业内部应用系统进行交换，也可以实现不同设备上安全工作空间之间的数据安全流转。

随着终端安全工作空间的应用推广，在这一层数据安全基础设施上逐渐承载了企业绝大多数的敏感数据，这为建立标准化的非结构化数据处理过程形成了天然的感知和控制抓手。

依托安全工作空间的数据安全基础设施，东方证券将面向数据业务与场景构建数据安全流转中间件，向下连接各种形式的存储设施，向上连接各种异构终端、应用服务、跨地域组织机构，将企业所有数据都集中于分布式的数据安全流转中间件中，进而构建起一个综合性数据安全治理平台。在此平台上，融合多项数据安全关键技术，如：基于AI的分类分级技术、基于ABAC的访问控制技术、基于水印的数据流转追踪技术等，并利用云端的强大算力，实现对非结构化数据的规范化、统一的综合治理。

类似云计算对资源的虚拟化，数据安全流转网关可以实现数据的虚拟化。数据虚拟化技术将数据实体存储于中间件平台上，而与其对应的虚拟体则可以以任何方式进行流转，即符合用户习惯的任何方式，如：IM工具、邮件、蓝牙、短信、U盘等等；而流转中的数据虚拟体被打开时则必须回到底层的终端安全工作空间之内。这样一来，数据实体仅会存在于中间件平台和安全工作空间这两者之内，而流转则可以通过任意方式进行。这样就形成了“数据有界，流转无界”的数据安全流转能力，安全与高效兼得。

（二）分级分类辅助外发审批

通过对用户非结构化数据的归集、标签化处理后，该用户在零信任沙箱内的企业数据在进行数据流转审批时审批管理员审批界面提供该文件的标签、属性，审批管理员可以通过对标签的识别快速获取需要流转的文件属性帮助审批管理员判断该员工的数据跨域流转行为是否存在风险，如遇风险文件流转行为可以及时阻断。

通过对员工非结构化数据的深入建模，我司计划在现有标签分类的基础上增加“分级”标签；分级采用《东方证券有限公司数据分类分级管理指引》中定义的4级体系。在员工文件流转的过程中，平台识别到需要流转的文件是“1级”密级时采用自动审批、服务端归档审计模式减轻管理员审批负担。流转文件级别大于“1级”时，审批自动提醒决策人含有对应等级的敏感信息为审批决策人提供判断依据，判断该文件外发时需要添加“明水印”、“暗水印”或是该文件外发后仅有只读权限。

（三）建设文件跟踪溯源能力

本系统支持两种文件跟踪溯源方式，即基于文件流转日志关联分析的文件跟踪溯源和基于文档内容的文档跟踪溯源两种方式。两种方式同时作用，叠加生效，保障跟踪溯源效果。

1、基于文件流转日志关联分析的文件跟踪溯源

这种方式将安全空间内文件全生命周期流转过程中的重点使用行为上报入库，如文件来源（导入、业务系统下载、新建、共享接收）、文件定级与级别变更、文件重命名、文件路径移动、设备移动、文件共享事件触发的等流水信息。之后对这些信息进行关联分析，流转链路复原，最终可形成沿时间（事件）线索关联的文件变换流转信息链，用以跟踪溯源。

整合上述文件编号流转信息链，可进行进一步统计分析，对系统平台内数据流转情况生成运营报表，展现数据流转情况。可支持报表有：各安全级别文件分布状况，各安全级别数据流转活跃度，热点数据文件，明文使用行为分析，人员组织文件流动关系等。

2、基于文档内容的文件跟踪溯源

这种方式基于文档外发和流转过程中对文档进行水印植入操作，水印伴随文档全生命周期，依附于文档本体，不易被去除和篡改。水印信息与安全空间内原始文档一对一唯一关联。当获取到泄露的数据片段后即可对数据片段中的水印进行提取并匹配到对应的原文件。亦或者对数据片段进行特征提取，包括sha1/MD5，关键字，自然语言特征等，进而与原始文件进行匹配，以进一步补充追踪溯源效果。

综上所述，零信任数据安全平台将新一代网络与沙盒技术相结合，形成了企业从网络到终端的新一代软件定义安全架构。同时解决了企业内部数据的静态访问控制，与数据流动的动态安全问题。具体讲，这两种技术的结合，可以在网络和终端上划分出软件定义的企业边界，数据可以在边界以内的网络和终端上的可信沙盒内自由流动，但仅能在企业边界内流转。从而能够达到“数据给得出，收得回”、“阅后即焚”的效果。

五、课题展望

通过课题的研究，东方证券联合数篷科技旨在打造新一代办公终端企业数据安全控制体系，服务于东方证券各类业务、办公场景。

（一）客户经理现场办理业务

证券公司客户经理常常面临一种困境，一方面如股权质押等业务通常需要质押人亲自办理，而此类客户大部分

是日理万机的公司领导，通常要求业务办理人员到客户现场办理；另一方面根据《网络安全审查办法》和《证券法》的规定，证券公司在办理业务过程中要保护好客户隐私，防止客户信息泄露。这种情况下就可以利用零信任网络的特点，在客户现场启动安全域，将办理过程在可信任的计算环境中完成，通过加密存储和加密传输机制确保客户信息不泄露。另外由于安全域使用本地资源进行计算，办理业务在进行音视频录制的过程中也能保证流畅度和清晰的效果。

（二）安全远程办公

2020年疫情前期，各大公司都实行了AB岗办公制度，这就导致在家远程办公的需求剧增。传统的解决方案可以通过开通企业网访问控制、申请外网虚拟桌面等进行远程办公。在实际的操作中遇到很多问题，比如某些系统开通到公网的访问控制前要进行漏洞扫描、防网络攻击检测等；使用虚拟桌面技术就算不考虑大量的虚拟机资源扩容，在进行音视频审核等业务时也会时常出现跳帧、卡顿情况，影响工作效率。

零信任企业网络解决方案可以较好的适应该应用场景。首先，该方案有本地安全计算的特点，利用员工终端设备的计算资源，不需要服务器资源的大量扩容。其次，安全域严格的网络权限控制、安全传输和软件定义边界的特点可以有效屏蔽网络攻击和数据窃取，许多业务和办公系统在安全域中访问本质上并没有改变其内网访问的属性，可以迅速部署并提供服务。另外，由于安全域可以直接使用本地外设资源的特点，对于处理音频和视频的办公场景可以提供高品质的用户体验。

（三）证券研究团队的安全高效数据协同

以对证券研究团队的服务为例，证券研究团队作为券商核心团队，负责对目标企业进行调研分析、出具企业分析报告等业务。研究团队出具的企业调研报告在很大程度上会对标的企业的股价构成影响，具有极高的机密性和重要性。

研究团队在工作保密的基础上需要进行大量的协同工作，其产出的报告、材料等成果均需多人协同工作，对于工作数据在各个相关员工工作终端上的高效安全流转有着极高要求。同时，研究团队经过多部门协同工作生成的调研报告、指导意见即使对于客户也属于保密程度极高的重要文件，同样存在接收人指定、内部分发控制、读写权限控制等要求。

应用平台后，一方面研究团队员工个人可以实现多个设备间的数据同步与协同，无论使用企业的标装设备，虚拟桌面设备，还是个人设备，甚至移动设备均可对指定数据进行合规的处理。数据始终在平台构建出的软件定义的柔性边界之内，在网络上始终处于零信任网络边界以内，在终端设备上始终处于终端安全工作空间内。保障核心数据安全的同时，达成了个人办公协同效率的大幅提升，做到了随时随地放心使用数据。

另一方面，对于研究团队的项目组，办公安全平台可以构建围绕项目组的数据协同群组，可实现多人多设备对同一组数据的实时协同。项目组每个成员均在各自的终端安全工作空间内进行数据操作，原数据通过数据虚拟化技术统一汇聚归并至中间件平台上。协同过程安全灵活，数据可用性高，整个协同过程可追溯，可回滚。从而实现了多人统一的高效安全协同。

（四）合作伙伴安全数据流转

通过本课题研究、实践东方证券打造了新一代的终端企业数据安全治理体系实现了数据“给得出”、“收的回”，终端用户“可见”、“可用”、“不可泄”的数据边界。通过相关安全能力的建设，东方证券规划向合作伙伴、供应商赋能实现东方证券企业数据在多类型用户、多种类终端上的安全存储、有序使用、合规流转。

参考文献

规划文件

[1] 国务院.《十四五数字经济发展规划》. 2022-01-12

法律法规

[1] 全国人民代表大会.《中华人民共和国网络安全法》. 2016-11-07

[2] 全国人民代表大会.《中华人民共和国数据安全法》. 2021-09-01

[3] 全国人民代表大会.《中华人民共和国个人信息保护安全法》. 2021-11-01

[4] 国家互联网信息办公室.《网络安全审查办法》. 2020-06-01

[5] 全国人民代表大会.《中华人民共和国证券法》. 2029-12-28

行业规范

[1]银保监.《银行金融机构数据治理指引》. 银保监发[2018]22号 . 2018-05-21

[2]人行.《金融数据安全 数据安全分类分级指南》. JR/T0197-2020 . 2020-09-01

[3]人行.《金融数据安全 数据生命周期安全规范》. JR/T0223-2021 . 2021-04-01

[4]工信部.《关于促进网络安全产业发展的指导意见（征求意见稿）》. 2019-09-27

基于漏洞情报的拟态防御技术实践

邢骁、蔡子豪 | 西部证券股份有限公司
薛辛 | 北京长亭科技有限公司

摘　要： 目前国内网络安全防护手段多采用被动防护措施，在主动防御方面还处于不断探索中；在国外，拟态防护作为一种新兴的网络安全主动防御措施，已经得到了广泛的关注与应用；在国内，多所高校和科研机构也在拟态技术领域取得了初步的成果，但实际应用于生产防护体系中较少。本次研究以伪装欺骗技术与拟态防护为理论基础，结合漏洞情报，以场景实验研究为重点，对Web拟态防护、主机拟态防护效果做了深入的实验测试研究，总结了一些适合应用于生产防护体系的主动防御拟态防护手段，通过这些拟态防护措施，可以打破网络空间“易攻难守”的被动防护状态，改变网络安全游戏规则，从而提升网络安全建设中的主动防御能力。

关键字： 被动防护、主动防御、伪装欺骗、Web拟态防护、主机拟态防护

一、概述

当下网络安全防护多处于被动模式，主动防御虽引入伪装欺骗技术，但应用效果欠佳。高校与科研机构对拟态技术领域研究初有成果，然而实际落地于生产防护体系的案例较少。网络安全局势日益严峻，攻击手段复杂多样，传统静态防御策略亟需提升。一方面，被动防护易陷入被动挨打局面；另一方面，主动防御中的伪装欺骗技术在高效仿真、持久吸引牵制攻击者等方面面临挑战，需要创新突破。拟态防御技术作为一种新兴的网络安全防护理念，为突破传统静态防御的瓶颈提供了新思路。其核心思想源于自然界的拟态现象，通过构建具有动态性、异构性和冗余性的执行环境，使攻击者难以掌握系统的真实状态和行为模式，从而大幅提高攻击的难度和成本。

二、基于漏洞情报的拟态防御系统建设方案

（一）构建漏洞情报分析检测系统

当前Web应用服务面临的风险多种多样，随着技术的发展和攻击手段的不断进化，这些风险也在不断变化，提前获取漏洞情报信息对我们进一步加强Web防护极其重要，当前各种威胁情报信息众多，在情报综合能力整合与自动化收集方面缺少相关措施手段，本次研究通过多种漏洞情报收集技术搜集大量漏洞情报，并构建漏洞情报分析检测系统，提升对Web网站的安全防护能力，并且相关能力实现进行了验证，实验资源环境如下表1：

表1 实验资源列表

序号	资源项	用途说明	备注
1	测试 PC	访问 web 应用系统，并尝试攻击	
2	测试靶站	模拟 web 应用系统	
3	漏洞情报分析检测系统	提供 web 访问入口，威胁风险分析，并将请求进行定向转发	Linux 系统
4	漏洞情报收集系统	情报收集，并将结果提供给漏洞情报分析检测系统	
5	二层交换机	网络互联互通	

整体的实验拓扑如下图所示：

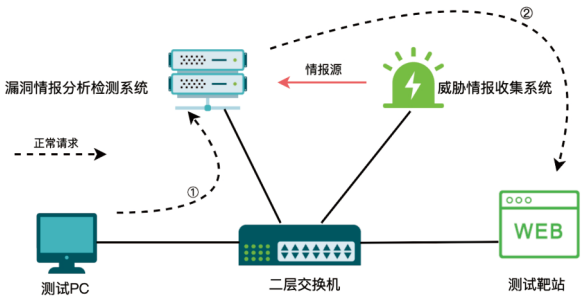


图1 情报分析检测系统拓扑图

实验环境如上图，经过多次验证测试达到以下三个效果。首先，正常的Web请求通过漏洞情报分析检测系统代理转发模块直接到测试靶站；其次，攻击者通过测试PC发起的Web攻击请求，经过漏洞情报分析检测系统的威胁分析模块判断后，会将请求进行拒绝，从而对测试靶站达到web防护的效果；再者，漏洞情报分析检测系统的自动化情报分析模块与威胁情报收集系统相结合，可以针对Nday漏洞进行提前分析形成防护规则措施，从而降低web攻击的安全风险。

漏洞情报分析检测系统是本次研究的基础，为后续其他内容的研究做好铺垫。

（二）构建主动防御伪装欺骗系统

在前一个章节中我们构建了漏洞情报分析检测系统，可以阻止或减少攻击的成功率，随着互联网的普及和技术的进步，网络攻击手段变得越来越复杂和多样化，被动的防护策略往往难以应对新型的、未知的攻击。伪装欺骗技术在主动防御中应用较多，其核心思想是通过创建和部署虚假的、看似有价值的目标来吸引和误导潜在的攻击者，从而间接保护真正的网络资源，但是如何能够高效地仿真虚假服务或者系统，并且能够长期吸引攻击者，是当前伪装欺骗技术应用于主动防御的一个挑战，我们主要从如何高效地构建极度仿真的目标诱饵，通过自动化响应、智能学习来驱动威胁检测和智能诱饵生成，从而提升主动防御的效果，并且对相关能力实现进行了验证，实验资源环境见表2：

表2 实验资源列表

序号	资源项	用途说明	备注
1	测试 PC	访问伪装欺骗系统、测试靶站等	
2	测试靶站	模拟 WEB 应用系统	
3	智学习仿真测试靶站	提供各类伪装欺骗服务，智能学习仿真测试靶站，管理监听探针	Linux 系统
4	伪装欺骗系统监听探针	监听端口，并对流量进行转发到智学习仿真测试靶站	
5	漏洞情报收集系统	广泛收集漏洞情报	
6	二层交换机	网络互联互通	
7	漏洞情报分析检测系统	解析规则、转发、阻拦	

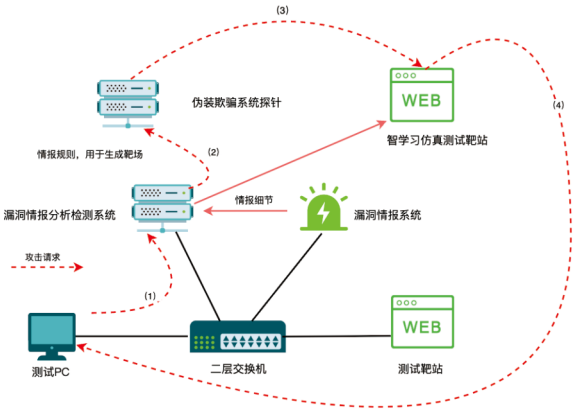


图2 伪装欺骗系统拓扑图

蜜罐的高度仿真性在主动防御技术中起着至关重要的作用，为了实现更好地迷惑攻击者，我们在蜜罐生成的时候对仿真的蜜罐服务实时更新并预制漏洞缺陷，给攻击者提供一个更趋近于真实场景的带有漏洞的业务系统，首先通过漏洞情报收集解析攻击者POC中的请求方法、请求URI、请求参数、漏洞判断参数等信息，将获取到上述信息的相关判断规则提前预制在蜜罐中，并且蜜罐将对该POC中的漏洞判断条件进行模拟，通过解析计算，构造出漏洞存在时给攻击者的返回数据包。当攻击者采用大量漏洞探测脚本探测系统是否存在漏洞时，都会收到漏洞存在的返回数据包，这增加了攻击者对漏洞判断识别工作，以此达到干扰攻击者的目的。即使系统本身存在漏洞，攻击者也会因为收到大量误报信息，无法快速判断哪个漏洞是真实存在的，从而为安全防护策略调整赢得宝贵的时间。

（三）拟态防护场景实践探索

无论是漏洞情报分析检测系统，还是主动防御伪装欺骗系统都是单方面从被动或者主动的方式去构建安全防御体系，尽管被动防御和主动防御各有侧重，但是两者之间缺少互补联动机制，在攻防对抗的两端能力较量中，如何能够有效吸引攻击者，并且能够动态地分析攻击者，甚至对抗攻击者，对整个安全防御体系的动态响应与灵活性是一个挑战，也是一个趋势，通过把被动与主动防御的深度有效融合，能够更有效地应对不断演进的网络威胁，拟态防御是一个很好的实现路径。

拟态防御采用了先进的网络安全防御对抗策略，其核心思想是通过不断变换系统的行为特征和表面形态，使攻击者难以掌握和利用系统的固有模式，从而提高系统的抗攻击能力和生存性。在本次研究中我们将拟态防御技术与被动、主动防御相结合，在漏洞情报分析检测系统与主动防御伪装欺骗系统中我们深度融合了拟态防护，并且从多个防护场景探索在网络安全建设中的实践应用，并且对相关能力实现进行了验证，实验资源环境如下表3：

表3 实验资源列表

序号	资源项	用途说明	备注
1	测试 PC	访问伪装欺骗系统、测试靶站等	
2	测试靶站	模拟 web 应用系统	
3	漏洞情报分析检测系统	提供 web 访问入口，威胁风险分析，并将请求进行定向转发	Linux 系统
4	漏洞情报收集系统	情报收集，并将结果提供给漏洞情报分析检测系统	
5	伪装欺骗系统管理平台	提供各类伪装欺骗服务，智能学习仿真测试靶站，管理监听探针	Linux 系统
6	伪装欺骗系统监听探针	监听端口，并对流量进行转发到伪装欺骗系统管理平台	
7	主机安全管理平台	主机资产风险入侵、主机探针管理	Linux 系统
8	主机安全探针	实时感知主机入侵风险，接受管理平台策略	
9	二层交换机	网络互联互通	

整体的实验拓扑如下图所示：

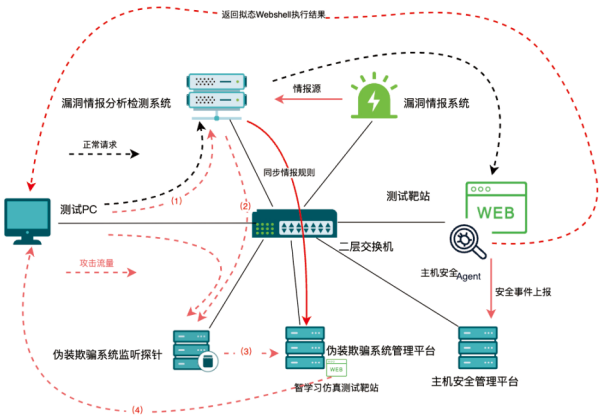


图3 拟态防护系统拓扑图

拟态防护的场景实践中，我们综合了漏洞情报分析检测系统与主动防御伪装欺骗系统，同时为了探索前端web防护被绕过直接进行主机入侵，在主机层面的拟态防护效果，我们增加了主机安全管理系统包括了主机安全管理平台与主机安全探针。如上图所示正常的请求通过漏洞情报分析检测系统后代理转发到后端web站点，攻击流量直接转发到伪装欺骗系统的蜜罐中，并且动态响应攻击者形成拟态防护效果，在拟态防护效果上从web防护与主机防护分别达到不同的应用场景效果。

安全总是相对的，存在攻击者绕过了漏洞情报分析检测系统，攻击行为直接到达站点服务器，在主机层面如果进行webshell上传攻击，主机拟态防护整体思路与流程效果如下图所示。①攻击者对站点主机服务器上传webshell，②触发主机安全探针的检测并上报给主机安全管理平台，③主机安全管理平台与伪装欺骗系统管理平台联动，将攻击者上传的webshell同步给拟态蜜罐，④拟态蜜罐给主机安全管理平台返回代理程序，⑤主机安全管理平台将返回的代理程序通过主机探针放置在站点主机服务器上，⑥攻击者以为webshell已经上传成功，发起webshell连接，⑦主机安全探针检测到刚刚上传的webshell连接时，将流量转发到主机安全管理平台，⑧主机安全管理平台将流量直接代理转发到伪装欺骗系统监听探针，⑨伪装欺骗探针将监听流量转发至拟态蜜罐中，通过拟态蜜罐再依次反馈给攻击者连接成功。

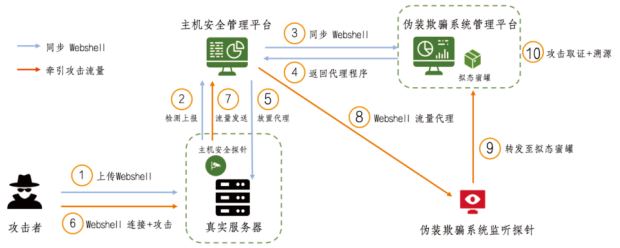


图4 主机拟态防护系统拓扑图

在实网攻击中，会有不同的攻击场景，下面我们主要从OWASP Top 10漏洞拟态响应、web攻击拟态防护与主机拟态防护三个场景进行了拟态防护的实践验证，并到达了拟态防护的效果。

三、多场景拟态防御实践

（一）OWASPTop10漏洞

实践资源见表4

表4 实践资源列表

序号	资源项	用途说明	备注
1	测试 PC	访问伪装欺骗系统、测试靶站等	
2	测试靶站	模拟 web 应用系统	10.1.9.3:80
3	漏洞情报分析检测系统	提供 web 访问入口，威胁风险分析，并将请求进行定向转发	10.1.9.10:80
4	伪装欺骗系统监听探针	监听端口，并对流量进行转发到伪装欺骗系统管理平台	10.1.9.10:81

实践场景：信息泄露攻击Web拟态防护

在实验环境中，我们的靶站不存在文件信息泄露漏洞，如下图所示，直接访问靶站返回404页面无法找到，没有相关的文件信息泄露

图5 信息泄露场景

我们将靶站通过漏洞情报分析检测系统代理访问，漏洞情报分析检测系统检测到信息泄露攻击行为，与伪装欺骗系统联动，通过拟态防护将攻击流量引入拟态蜜罐中，给攻击者反馈出/etc/group与/etc/passwd文件信息泄露，如下图所示。

图6 信息泄露拟态防护场景

（二）Web拟态综合实践

实践资源见表5

表5 实践资源列表

序号	资源项	用途说明	备注
1	测试 PC	访问伪装欺骗系统、测试靶站等	
2	测试靶站	模拟 web 应用系统	10.1.3.233:8080
3	漏洞情报分析检测系统	提供 web 访问入口，威胁风险分析，并将请求进行定向转发	10.1.9.10:80
4	伪装欺骗系统监听探针	监听端口，并对流量进行转发到伪装欺骗系统管理平台	10.1.9.10:81

实践场景：Web拟态防御综合实践

漏洞情报系统通过多方渠道采集漏洞信息包括不限于漏洞POC、EXP、防护措施等，并将漏洞POC通过调用API方式同步给漏洞情报分析检测系统。在漏洞情报分析检测系统添加识别规则。同时同步漏洞POC中的请求方法、请求URI、请求参数、漏洞判断参数等信息给伪装欺骗系统，并新建对应的漏洞回显数据包提前预制在相关蜜罐中，构建有缺陷性蜜罐。

当攻击者通过通用漏洞进行攻击时，漏洞情报分析检测系统会通过内置规则及情报收集系统规则合集判断是不是攻击行为，如果发现是攻击行为，则将其转发到智学习蜜罐中，智学习蜜罐会对攻击脚本进行虚假响应，干扰攻击者漏洞探测结果。当识别是正常流量时，转发到对应的业务系统，通过拟态蜜罐与业务系统的深度融合，让攻击者的漏洞探测技术无法发挥价值。在实验验证中我们采用nuclei和xray工具的POC进行验证，POC格式如下：

```
id: CVE-2020-14750

info:
  name: Weblogic Auth Bypass
  author: medbsq
  severity: critical
# link: https://www.oracle.com/security-alerts/alert-cve-2020-14750.html
requests:
  - method: GET
    path:
      - "{{BaseURL}}//console/images/%252e./console.portal"
    matchers-condition: and
    matchers:
      - type: word
        words:
          - "common/NoJMX.jsp"
          - "ADMINCONSOLESESSION"
        condition: and
        part: body
      - type: status
        status:
          - 200
```

请求路径

响应体命中规则

图7 漏洞检测POC实例

对蜜罐请求访问，智学习蜜罐对上述POC的返回结果如下。

图8 漏洞存在虚假信息

通过智学习蜜罐，可以构造虚假响应，让xray、nuclei等工具误以为漏洞存在。当攻击者用批量漏洞利用工具进行扫描时，会给攻击者返回大量扫描结果，攻击者无法在短时间内分辨哪些是误报。严重干扰此类漏洞扫描探测结果，给安全运营人员提供时间窗口进行应急响应与溯源反制。如下图所示，攻击者访问漏洞情报分析检测系统，正常访问时，流量都转发到业务系统。

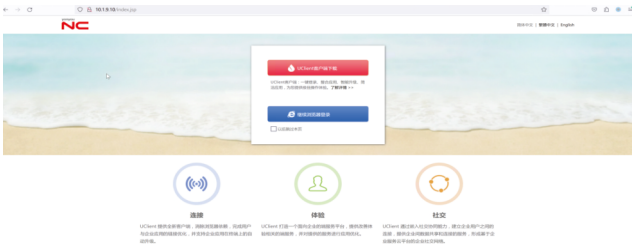


图9 正常访问跳转业务系统

攻击者用大量攻击payload探测系统是否存在漏洞时，由于智学习蜜罐对攻击数据包进行伪造响应，使得攻击者得到的扫描结果中，发现了93个漏洞。实际系统并没有这么多漏洞如下图所示。我们选取几个场景详细分析解释如下：

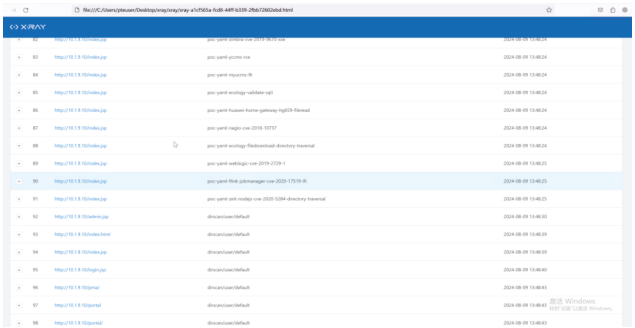


图10 漏洞扫描报告

(1) 动态计算payload响应

智学习蜜罐对POC中的漏洞存在匹配规则进行动态处理，如下图，攻击者验证脚本是打印862208410+932341588，并验证结果是否为1794549998，智学习蜜罐通过对POC脚本中的检测逻辑进行判断，将1794549998打印在响应体中，攻击者的探测脚本在响应包中识别到1794549998存在，认为漏洞存在，实际上此漏洞并不存在于当前系统。

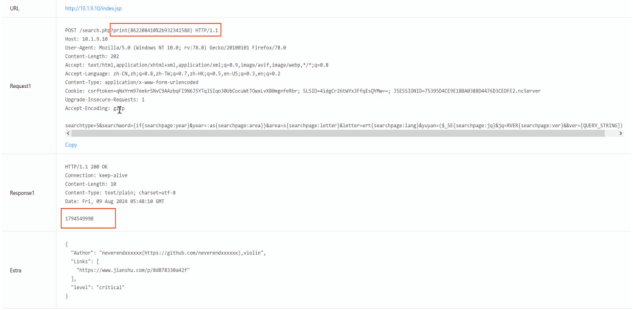


图11 动态计算payload

(2) 真实业务正常访问

如下图，正常请求业务网站，业务服务器返回正常响应。

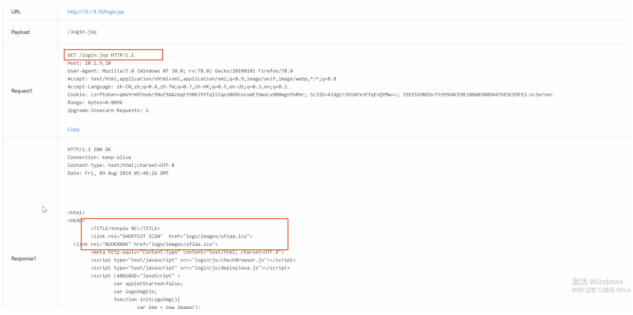


图12 真实业务正常访问

（三）主机拟态防护

实践场景：Webshell攻击主机拟态防护

主机防护是业务系统安全防护的最后一道防线，攻击者总是有概率突破前端的各种安全防守手段，对主机直接入侵，如下图所示，攻击者通过哥斯拉生成一个webshell并上传在web站点服务器的www目录下。

Web站点测试服务上有安装主机安全探针，立即发现了主机入侵的webshell文件，但是并没有直接对文件进行查杀或者隔离处理，而是自动启动了拟态防护措施，如下图所示：

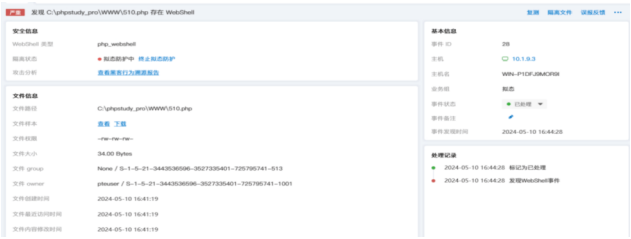


图13 主机拟态防护

攻击者通过哥斯拉连接刚刚上传的webshell，并执行命令写入文件等操作均可以成功进行，以为已经成功拿下服务器，实际上其所有的攻击行为均在拟态防护的蜜罐中进行，并且所有的操作均在拟态防护系统的监测记录中，从而了解攻击者的攻击手法与真实动向。如下图所示，攻击者的操作行为与拟态蜜罐捕获的攻击数据包。

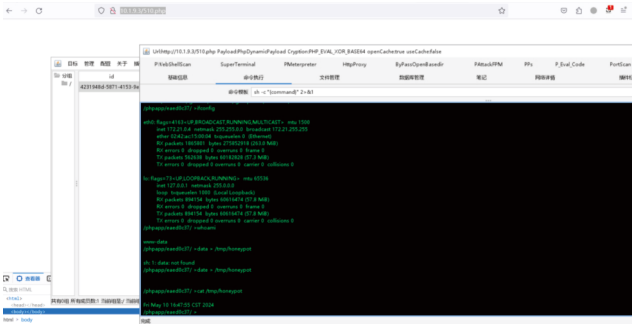


图13 攻击者连接蜜罐webshell

四、总结

无论是被动防御还是主动防御技术主要都采用了静态的防御策略，随着网络攻击技术的日益精进，越来越难以应对新型攻击。本次研究基于漏洞情报的拟态防御技术实践，通过拟态防御动态改变系统状态，使得攻击者难以捕捉到固定的攻击窗口。结合漏洞情报与拟态技术的融合，避免已知漏洞被利用，显著提升防御效果。

研究采用了理论与实践相结合的方式，首先结合情报收集系统通过收集、分析和整理漏洞情报，了解攻击者可能利用的漏洞和弱点，构建了被动防御的漏洞情报分析检测系统；其次，通过构建伪装欺骗系统故意暴露虚假的目标系统并融合漏洞脆弱性，吸引攻击者攻击虚假目标，通过精确的判断和模拟，以确保在拟态环境中看起来与真实系统一致；再者，通过拟态技术，将web攻击行为以及主机层面的webshell攻击与拟态蜜罐相结合，正常的请求转发到业务站点服务器，真实的攻击直接进入拟态蜜罐，让攻击者看似攻击成功实则在虚假目标中，并且使攻击行为无所遁形；最后，通过多个场景实践，验证了拟态防护在多个攻击场景的防护应用效果，为后期在真实的生产业务环境应用奠定了良好基础。

总的来说，基于漏洞情报的拟态防御技术实践对于提高网络安全建设具有重要意义。根据攻击者的行为，结合拟态技术动态自适应调整防御系统，有效提升了网络防御系统的抗攻击能力。它可以增加攻击的复杂性和不确定性，为防御团队提供更多的机会来检测、响应、处置攻击。未来的研究方向将着重考虑实际应用中的挑战和限制，提高情报处理的智能化水平，增强防御策略的自适应性和精细化，同时探索与大模型新技术的融合，构建更为智能、弹性的网络安全防御体系。

参考文献

1.王禛鹏,扈红超,程国振,张传浩. 软件定义网络下的拟态防御实现架构[J]. 网络与信息安全学报, 2017, 3(10): 52-61.

2.任权,贺磊,邬江兴. 基于离散马尔可夫链的不同抗干扰系统模型分析[J]. 网络与信息安全学报, 2018, 4(4): 30-37.

3.李建华. 网络空间威胁情报感知、共享与分析技术综述[J]. 网络与信息安全学报, 2016, 2(2): 16-29.

投稿邮箱：
wsil@sse.com.cn

ITRDC

ITRDC证券信息技术研究发展中心（上海）

上海證券交易所
SHANGHAI STOCK EXCHANGE

中国上海市杨高南路388号
邮编：200127
公众咨询服务热线：4008888400
网址：https://www.sse.com.cn/

内部资料 免费交流
本资料仅为内部交流使用，本期印200册，编印单位为上海证券交易所，面向证券期货行业发送，印刷时间为2025年11月，印刷单位为上海町麦广告有限公司。
部分图片或文字来源于互联网等公开渠道，其版权归属原作者所有，如有版权相关事宜，请发送邮件至wsil@sse.com.cn。