安全+
SECURITY
2021/07 总第 049
技术版 ▶▶ 与安全人士分享技术心得 Share technique experience with security professionals

★ 本期焦点

智慧安全3.0之可信任解读
——新智慧，可信任

智慧安全3.0助力城市轨道网络安全保障升级

物联网中基于UDP的DDoS新型反射攻击研究
—— 绿盟科技刘文懋RSAC主题演讲

软件定义的原生云安全解决方案

绿盟科技官方微信

NSFOCUS

SECURITY+

安全+ 技术版
与安全人士分享技术心得
Share technique experience with security professionals

本期看点 HEADLINES

主办：绿盟科技
策划：绿盟内刊编委会
地址：北京市海淀区北洼路4号益泰大厦三层
邮编：100089
电话：(010)6843 8880-5463
传真：(010)6872 8708
网址：www.nsfocus.com

2021/07 总第 049

欢迎您扫描封面左下角的二维码，关注绿盟科技官方微信，
分享您的建议和评论，或者来信nsmagazine@nsfocus.com
与我们交流。(本刊部分图片来源于网络)

安全+ SECURITY

目录 CONTENTS

随着数字化经济的发展，跨界融合的业务需求、复杂多变的IT环境、持续创新的技术和以APT攻击为代表的威胁演进，导致网络安全攻防对抗形势愈加激烈。传统的基于边界安全防护，以及单次静态安全策略配置的安全措施已无法满足业务发展的安全需求。结合合规和攻防业务需求，从体系化角度进行安全建设，成为我们为用户赋能的主要发力点。

RSA 2021大会的主题是“RESILIENCE”，从这个主题我们也可以看到体系化建设已经成为全球数字化时代所追求的安全实效目标。在面对具备隐蔽性、非对称性、不可预见性等多种特性的网络攻击时，如何快速预测、监测并识别潜在威胁，并在持续压力下保证业务系统的连续运行，迅速恢复到被攻击前的状态，是RSA 2021大会探讨的重点内容，这与绿盟科技提出的“智慧安全3.0”理念不谋而合。

“智慧安全3.0”强调以体系化建设为指引，结合自身业务规划，实现顶层设计到不同层次的落地设计，构建“全场景、可信任、实战化”的安全运营能力，达到“全面防护，智能分析，自动响应”的防护效果。该理念继承了智慧安全2.0的精髓，并实现了内涵和外延的扩展。

“全场景”将智慧安全的外延扩展到整个网络空间，面向全部数字化应用场景；“可信任”则是内涵的扩展，安全不仅仅是攻击防护，还是信任模型的建立与保障；而“实战化”呼应当前新形势要求，同时也涵盖了智慧安全2.0的“智能、敏捷、可运营”理念，加入对运营效果的验证要求。相比智慧安全2.0运营流程的自适应闭环，“智慧安全3.0”的“体系化”表现为立体的动态演进的安全生态空间，是对智慧安全2.0的全面升华，融入了绿盟科技对安全新的理解，是对数字化经济趋势持续适应的体现。

在此基础上，绿盟科技从业务链、供应链、人员链角度梳理攻防战术，设计安全对抗的措施和场景，并总结了“1-3-4+N”模式的“智慧安全3.0”建设框架。从对抗的核心角度来讲，用户需要以建设一个“安全运营中心”为目标，遵循“全场景、可信任、实战化”三原则，梳理业务环境中的“资产、应用、数据、身份”四要素，以达到对抗的有效性。由此，绿盟科技拆解了十二项安全建设工程。用户可根据自身发展需求，逐步提升能力成熟度，构建弹性的安全能力，随时应对安全威胁和新型挑战。

围绕“智慧安全3.0”理念，绿盟科技将持续以安全能力为核心竞争力，赋能服务、产品和技术，并通过安全运营服务最终用户，向“全能力、全运营”方向进化。

叶晓虎

智慧安全3.0之可信任解读
——新智慧，可信任

绿盟科技 创新中心 刘文懋

一个好的安全体系的前提是为合法主体建立信任关系，在保证业务的前提下，通过信任降低安全成本，在运行时及时检测并消除非法主体的恶意行为，所以信任是网络安全的前提要求。

从业界近年的发展情况来看，无论是 Gartner 认为信任和弹性是自适应安全的两个原则，还是“零信任”理念成为业界热议的流行词，都说明安全行业开始反思并认识到简单堆砌的安全机制已无法抵御日渐复杂的应用场景和攻击团伙。所以回归安全的本源，思考如何构建信任体系，成为当前一种独特的现象。

绿盟科技于 2021 年推出“智慧安全 3.0”新战略，其中要素之一就是可信任，这表明，构建信任机制是新一代安全体系中非常重要的一环。本文将诠释“智慧安全 3.0”中可信任的内涵。

1. 信任与风险的平衡

在网络空间中，业务发展、环境变化或技术更新是常态，也是数字化转型的必然，因而风险是永远存在的，安全团队不可能为规避所有风险而要求业务不变，甚至让业务下线。对抗的本质是攻守成本的平衡，安全的本质是防护减少损失（收益）超过防护成本。因而安全防护最终是缓解风险，不可能完全去除风险。

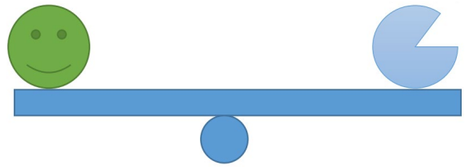


图 1 信任与风险的平衡

不信任就不可能开展新的业务，适应并拥抱变化是安全团队应信奉的哲学。为了保证业务正常运行，安全团队必须让渡部分非关键风险的处置，也就是在管理风险的前提下，主观上相信业务会安全运行。例如，边界就是信任的体现，跨边界的访问是不可信的，需要强制认证和访问控制，而内部网络的访问默认是可信的，这样就大大减少了身份认证和访问控制的部署成本，也提升了用户的使用体验。

但安全团队也应清晰地认识到，信任第三方就意味着存在风险。那么就应评估和管理风险，将业务受到的影响降到最低，保证业务的弹性（Resilience）。

总之，拥抱变化，管理风险，是可信任的内在需求。就安全厂商而言，应当构建信任管理体系，以降低整体安全投入的成本，并提升安全防护效率，帮助客户赢得安全方面投资的回报。

2. 信任管理模型

维基百科（Wikipedia）上对信任（Trust）的定义为：一方（信任方）在未来依赖另一方(被信任方）行动的意愿。假设给定三方 A、B、C，三者之间都有交互，如图 2 所示。

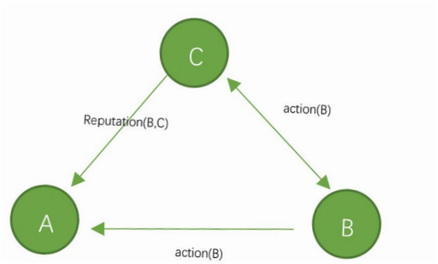


图 2 信任度模型

由此可见，主体 A 对 B 的 action(B) 行为的信任是结合了 A 对 B 的历史行为的观察 {actions(B)} 和第三方（如主体 C）对其信誉评价 Reputation(B,C) 的综合评估。事实上，信任度的度量会更复杂一些，需要考虑到观察行为（证据）的可靠度，以及信任度随着时间的推移而衰减等因素。

而信任机制在应用时，根据不同的场景和需求会有多种形态，如 IAM、访问控制、边界控制等，具体产品就更是五花八门，但从核心上看，信任管理有四个要素：

（1）主体身份属性确认，即 Identification；
（2）资源的属性确认，即 Attribute Enumeration；
（3）主体对资源操作的授权，即 Authorization；
（4）操作控制，即 Enforcement。

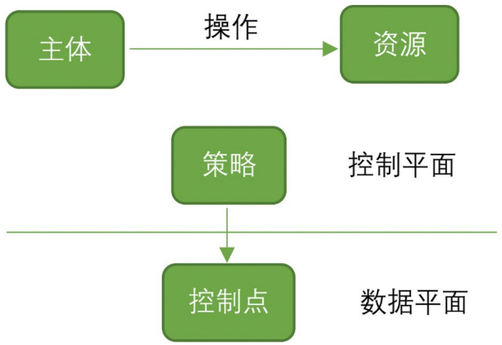


图 3 信任机制表示

当前新出现的安全模型也对信任提出了新的要求。例如，零信任(软件定义边界) 模型要求全方位的信任控制，即主体在任何时间、任何地点访问客体，均需要遵从全局的访问控制策略。而 Gartner 提出的“自适应访问控制”，则要求控制点在通过主体的访问请求后，还需要根据上下文进行调整，动态授权其访问。

而当前行业大多数的信任管理处于被动信任和静态信任的阶段，远未达到全方位、自适应的阶段。

所谓被动信任，就是无条件地信任。例如，客户在部署了某个安全产品、平台或服务后，只能黑盒地选择相信其正常工作、发挥作用；开发团队在采用某第三方软件或硬件时，只能假定其没有后门、漏洞。总之，既来之则安之，除了信任，没有其他办法。

而所谓静态信任，则是采用了确定性的信任评估方式，设置后长期不变。这也是行业内主流的信任管理机制，虽然简化了策略制定、系统运行时机制，但没考虑到上下文变化，是造成现在网络安全事件频发的根本原因之一。

事实上，信任是主观、动态、不确定的，信任管理是要以风险为基础的，安全策略需要根据主体行为等上下文进行动态调整。因而，在下一代安全体系中，信任管理应当具备可信任的安全能力、可信任的访问机制，以及可信任的供应链管理。

3. 可信任的安全能力

随着地缘政治和数字化转型的深化，网络安全将从满足合规性要求转向攻防对抗，客户对于安全厂商的要求将越来越偏向其安全能力可信任。这主要包括两个层面：安全运维可信任和安全运营可信任。

首先，网络安全设备本质上就是一类网络设备，与其他的路由器、交换机无异，都是部署在网络中。但在以往的安全运维中，除金融等重要业务场景外，客户对安全设备的稳定性要求没有网络设备那么高。但在现在的大背景下，安全设备的重要性将越发凸显，边界侧、串接型的安全设备一旦发生故障，很有可能出现断网事故，影响正常生产；而旁路侧的设备，虽然不至于影响生产，但如果出现发生故障，也可能会丢失重要的事件、告警和日志，给后续排查溯源带来不可逆的影响。因而，安全厂商的安全设备、平台和服务必须具有极高的稳定性和可靠性。

其次，安全厂商也是安全服务提供商，会参与到客户的实际安全运营中，应对各类真实威胁和攻防演练。由于这类安全运营需要实打实的安全防护、检测和响应能力，因而安全厂商的运营能力必须是可信任的。以前在安全运营中最大的挑战是告警泛洪，一台入侵检测设备单日可能会产生数万条告警，大部分告警指示的并非关键性事件，而且有不少是误报，这对于安全运营而言是灾难性的，一位安全运营人员是不会信任这样的设备产生的告警的。在海量告警中寻找真正的攻击，无异于大海捞针。

当前，安全厂商已经开始利用一些先进技术，如人工智能、剧本编排等，将顶尖的安全运营专家知识赋能智能引擎，从而大幅减少安全运营的边际成本。例如，绿盟科技于 2021 年发布的《AISec-Ops 智能安全运营技术白皮书》介绍了通过人工智能技术进行安全运营的技术路线和关键技术，如可以通过多引擎评估和智能推荐算法，将关键告警推荐给安全运营人员，从而摆脱大海捞针的困境。一旦安全运营团队通过短期调整，固化其场景下的运营需求及其推荐模型，则可在日常运营中在最短时间内发现关键安全事件，建立对安全厂商能力的信任。

最后，在对抗过程中，攻击者绕过安全设备的规则也将成为常态，这非常考验安全厂商对安全漏洞、安全事件的日常收集、研判和产品转化能力。当前的攻击手法主要是规则绕过，而随着攻防技术的进一步发展，基于人工智能的引擎也可能被攻击者绕过，这就要求人工智能算法是可解释、可信任的，不会发生类似“熊猫变长臂猿”的攻击案例。

4. 可信任的访问机制

网络安全的本质是保证网络中主体对客体的访问是合规、合法、合理的，然而太多的机构数据泄露事件表明当前的访问控制无法满足以上要求。例如，攻击者利用服务漏洞进行持续渗透，或是恶意内部用户尝试窃取非授权的数据，如果通过常规的访问控制或入侵检测是很难发现的。

当前业界一方面使用欺骗、沙箱等高级对抗技术，极大增加了检测、防护的投入成本；另一方面也会使用用户实体行为分析（User & Entity Behavior Analytics, UEBA），分析访问主体的行为模式，能从一定程度上补全行为和业务层面的安全机制空白，但也存在大量误报，同样也需要投入较多运营成本。

如果从安全对抗的本质出发，一方面，从正常业务的用户体验的角度而言，应确保大部分合法主体的访问不被安全机制所困扰；另一方面，将安全控制前移，在前期投入较少资源以获得较大收益，以避免后期的各类成本（包括投入新安全机制的投资成本 CAPEX 和缓解误报的运营成本 OPEX）。

在安全运营的闭环中，防护、检测、响应阶段的投入依次增加，安全效果却依次降低。因而，应适度将部分检测、响应阶段的投入转移到前期防护阶段，即将安全控制前置。保证访问主体对客体的访问关系是可信任的，即主体具备合法的身份，主体对客体的访问是授权的，主体的行为是合理的。

前述的“零信任”模型是可以达到这样的要求的，“持续验证、永不信任”是零信任的理念，但其本质还是信任主体的身份，需要

动态的上下文评估其行为，从而确保该访问是能够被信任的。

5. 可信任的供应链管理

可信任的供应链管理有两层含义：安全产品可信任和第三方软硬件可信任。

安全产品可信任，是客户对安全厂商信任的基石。安全产品在企业的安全运营中处于非常重要的地位，其自身安全不容忽视。无论是处于网络边界的网络安全设备，还是管理大量终端的终端安全平台，一旦被攻破，都可能导致攻击者渗透进入内网或控制大量的终端。近年的安全演练发生的安全厂商的产品因出现漏洞导致被攻破，便是明证。

因而，安全厂商应重视其产品的安全，从开发到交付过程，需建立安全开发、代码审查、上线前检查等一系列安全流程，此外，应建立漏洞悬赏计划和应急响应机制，以确保及时发现潜在的安全漏洞，并在最短时间内确认安全事件、发布安全修复计划和相应的安全更新。我们说过，攻防是成本和收益的平衡，安全产品与其他的 IT 软件一样，漏洞是不可避免的，但通过事前、事中和事后的预防和应对机制，可以将安全产品变为最不易被攻破的防线。而对于攻击者而言，当攻破安全产品的成本明显高于其他组件，而收益几乎没有时，则不太会以它为靶标。

在更多的案例中，第三方的软硬件漏洞或后门成为企业被入侵的重要原因。例如，今年的 SolarWind 旗下的 Orion 软件更新包中被黑客植入后门，波及范围极大，包括政府部门、关键基础设施以及多家全球 500 强企业，影响极其深远。此外，地缘政治冲突日益加剧，国家安全要求第三方硬软件安全可信，国内正在构建信创生态体系，覆盖非常长的硬软件供应链。无论是要避免来自第三方的安全漏洞，还是构建安全可信的体系，都需要对供应链的第三方提供商进行持续的安全评估，使客户避免使用具有风险的提供商产品，Gartner 将其称为安全评级服务（Security Rating Service, SRS）。

在实践过程中，应当根据客户的开发和运营流程，科学评估存在的供应链风险。如在开发环节，应当识别软件项目中引入的开源代码和第三方软件库，并评估其潜在的安全漏洞；在集成环节，应当对第三方的模组、系统或镜像等进行安全评估。

6. 结语

信任是人类社会快速前进的基石，也是数字化业务迅速发展的前提，在“智慧安全 3.0”的蓝图中，我们协助客户构建信任管理体系。首先，我们具备客户可信任的安全产品和安全服务能力。其次，我们提供可信任的访问机制和可信任的供应链管理，在降低整体投资和运营成本、提升用户体验的前提下，提供坚实、有效、全面的安全防护效力。

参考文献

[1] Trust (social science)_https://en.wikipedia.org/wiki/Trust_(social_science).

[2] AISecOps 智能安全运营技术白皮书 , https://www.nsfocus.com.cn/html/2020/92_1218/142.html.

[3] “熊猫”变“长臂猿”“乌龟”变“来复枪”，深度学习模型被攻击，破解之道有哪些？http://leiphone.com/news/201910/W4Wm5jfL19ZWbIbp.html.

[4] 初探 SOLARWINDS 供应链攻击事件来龙去脉 http://blog.nsfocus.net/solarwinds-unc2452-1222/.

基于智慧安全3.0，重构新时期企业安全能力

绿盟科技 行业技术中心 李成日

摘要：随着数字化的变革速度越来越快，整个场景带来的变化及国家法律法规的进一步健全，推动了整个网络安全的发展。越来越多的企业要求以合规为基准规划安全建设，并更多地从业务层面考虑安全带来的强需求，如 APT、勒索软件、数据泄露、供应链威胁等安全风险带来的威胁。因此我们要充分评估技术变革带来的新需求，如弹性、自适应的安全体系、业务连续性及快速响应的安全运营保障等。

1. 新时期数字化转型背景下网络安全形势

随着“数字中国”“网络强国”和“新基建”等国家重大战略部署的加快推进及“十四五”规划的相关要求，企业的数字化转型已是大势所趋，随着新一代信息技术的创新应用和业务运营模式的演进，网络安全工作面临新的挑战。

企业数字化转型将会改变原有生产和经营方式，信息技术与业务发展的深度融合会更加突出网络安全风险问题，影响原有业务的正常运营，进而影响生产安全、社会安全甚至国家安全。在数字化转型过程中，新技术、新应用、新模式层出不穷，新问题、新风险、新挑战不断出现，面对更加开放多元的应用场景，传统安全防护手段难以保障数字化业务的平稳、可靠、有序运营。

1.1 强化网络安全顶层设计

从国家政策层面来说，《网络安全法》《密码法》《数据安全法（草案）》、等保 2.0、《关键信息基础设施安全保护条例》、《工业互联网创新发展行动计划（2021——2023）》等法律法规、行业标准规范的陆续颁布施行，对数字化转型背景下企业的网络安全工作提出了更高要求。因此，结合政策标准展开积极应对，需要技术融合管理，加强顶层规划设计，按照“三同步”原则要求，构建新一代网络安全框架。

1.2 评估新技术新业务的深度融合带来的新问题

随着“云移大物智”的兴起，我们需充分评估不同场景下的网络安全风险，如不同网络接入带来的安全互联互通问题，网络边界模糊引起的多源异构身份互信互认问题，大量数据集中汇集带来的数据泄露问题，勒索软件、APT、供应链威胁等带来的新型网络安全问题等，传统的网络安全架构无法应对新技术带来的安全风险，网络安全面临全新挑战。

“两化”的深度融合是推动工业经济走向数字化、网络化、智能化、融合化的新动能，企业已构建深度融合的多层次数字化制造体系，如内部生产域之间、生产域与管理域之间都将形成互联。实现信息整合共享，打通智能制造的断路，是安全建设的巨大挑战，即信息化与工业化的深度融合，使越来越多的企业将网络安全的范围从传统的 IT 系统延伸到了工业系统。工业控制系统自身存在的安全漏洞加上物联网化带来的广泛安全威胁，使安全问题被视为未来实现工业互联网创新发展战略的中国制造 2025 在管理上的又一新挑战。

1.3 应对技术变革带来的客户侧新需求

新技术、新业务快速发展，培育客户网络安全新技能催生新的网络安全需求和安全理念，带来从“合规化”转变为“运营化”，从“可视化”转变为“实战化”，从“被动防御”转变为“积极防御 & 反制进攻”的一系列变化。我们需构建集防护、检测、响应、预警于一体的自适应安全防护体系，实现由静态、被动的防护向动态、主动防护的转变，打造一体化、自动化的安全防护能力，提升大数据时代防攻击、防泄露、防窃取的监测、预警、控制和应急处置能力，加强信息通报和应急演练，确保企业自身关键信息基础设施安全稳定运行。

在新时期数字化转型背景下，我们要思考企业网络安全定位问题，评估新技术、新业务的融合能力，重构安全保障框架，护航企业数字化转型。

2. 智慧安全 3.0 框架，有效解决新时期网络安全问题

通过评估网络安全形势，分析安全现状，以体系化建设为指引，坚持网络安全“三同步”原则，建设全场景、可信任、实战化的安全运营保障体系，达到全面防护、智能分析、自动响应的防护效果。

企业网络空间的应用场景应覆盖终端安全、网络安全、应用安全、数据安全等通用安全领域，加强特色安全领域，如工控安全、移动互联网安全、供应链安全等不同场景、不同流程下的安全防御，实现全部数字化应用场景下的安全保障。

在万物互联的时代，以零信任重构信任，为企业不同对象赋予唯一身份，对不同身份实体或行为（人员、设备、应用服务、数据）进行持续认证与访问控制，保证在不同场景下，访问实体身份可信、操作可控和行为可追溯。在可运行的业务环境中，通过信息技术应用创新构建自主可控的 IT 产业标准和生态，保障产品和技术安全可控，解决基础软硬件和网络安全领域面临的供应链安全问题。

在新时期数字化转型过程中，企业网络安全必须做到持续运营、平战结合。以等保合规为基准，通过安全流程与制度建设，把实战中的行为与过程规范化，构建平战结合、常态化、制度化的快速应急响应治理机制；加强对安全人才及全员安全意识的培养，引入安全专家服务，把安全专家的经验和能力融入运营体系中；构建协同指挥机制，以战领建，发现安全措施的薄弱环节，依靠威胁情报与信息共享体系，优化安全策略，形成对安全防御保障体系的反馈闭环机制；通过全天候、全方位的安全运营服务，强化实时研判响应能力，持续增强安全监测评估、主动预防、应急处置等能力，为企业提供实战化整体安全运营保障。

遵循相关法律法规要求，以技管并重，围绕全场景、可信任、实战化，基于 IPDRR 模型，提出符合新时期企业特色的网络安全保障体系整体框架，突破网络安全防护监测处置一体化、安全综合监管、多维异质数据流动安全管控等关键技术，形成主动防御网络安全保障体系。该体系包括以下部分：

安全保障框架

数字信任基础设施：通过密码密钥管理服务、数字证书管理服务，在企业异构应用场景下，为其安全应用提供加解密、签名验证、摘要计算等弹性密码应用服务及统一的证书签发应用服务；以身份为基石，实现企业不同对象的全面身份化，构筑基于身份的信任体系；通过权限管理服务，按照最小化权限原则进行细粒度授权，基于不同多源属性进行信任和风险度量，实现动态自适应访问控制。

场景化的安全保障服务：以数字信任基础设施为底座，提供统一调度管理，通过安全识别、安全防护、安全检测、安全响应、安全恢复等一体化的闭环安全能力，对终端安全、网络安全、云平台安全、应用安全等通用安全提供纵深安全防御能力，对工控安全、移动互联网安全、供应链安全等专用安全提供定制化的安全保障。在不同场景下，以数据为核心资产，强化数据安全监管能力，通过数据分类分级，有效保护敏感、重要数据，避免出现数据滥用、泄露及无法溯源等问题。

安全运营服务平台：围绕监测预警、分析研判、响应处置、追踪溯源等能力，基于数据关联分析、自学习智能网络威胁检测方法等关键技术创新研究，构建一套“技术先进、安全可靠、服务完备”的综合安全运营服务平台，通过识别、防护、检测、监测、预警和响应处置等，达到事前、事中、事后全方位覆盖，实现全天候、全方位综合安全态势感知，形成安全运营全生命周期的管理闭环，使企业完全具备“威胁预警、协同对抗、可管可控、智能防御”的安全运营保障能力。

3. 结语

在数字化转型过程中，企业应以全场景、可信任、实战化为发展目标，综合利用大数据、云计算、人工智能、物联网、威胁情报、安全运营等先进技术措施融合安全运营管理，构建自适应安全防护体系，提供一体化、自动化的安全防护能力，加强信息通报和应急演练，确保关键信息基础设施安全稳定运行，切实提升网络安全保障能力。

智慧安全3.0助力城市轨道网络安全保障升级

绿盟科技 行业技术中心 姚宇

2018 年至今，共有七次中央级会议或文件明确表示要加强新型基础设施建设，党中央和国务院对此愈加重视，相关政策路线图日趋清晰。国家持续密集部署新型基础设施，原因在于新型基础设施具备新时代的丰富内涵，既符合未来经济社会的发展趋势，又适应中国当前社会经济发展阶段和转型需求，不仅可以补短板，也将成为社会经济发展的新引擎。作为数字经济的发展基石和转型升级的重要支撑，新一代信息技术引领的新型基础设施建设已成为我国谋求高质量发展的关键要素。

习总书记指出：“城市轨道交通是现代大城市交通的发展方向。发展轨道交通是解决大城市病的有效途径，也是建设绿色城市、智能城市的有效途径”“要继续大力发展轨道交通，构建综合、绿色、安全、智能的立体化现代化城市交通系统。”习总书记还特别做出了要发展智能交通的指示，为城市轨道交通发展明确了路径指向。

作为“新基建”的重要领域之一，城际高速铁路和城市轨道交通的建设将融合吸纳我国先进信息技术，推动交通领域的数字化和智能化发展。应用云计算、大数据、物联网、人工智能、5G 等新兴信息技术，全面感知、深度互联和智能融合乘客、设施、设备、环境等实体信息，通过自主进化创新服务、运营、建设管理模式，构建安全、便捷、高效、绿色、经济的新一代中国式智慧型城市轨道。

城市轨道系统的智慧化是发展趋势，是城市轨道交通未来的发展方向，可实现各线路、各专业资源的共享，以及线网统一运维管理与安全管控，在降低资源利用效率和大数据融合等方面价值巨大。

从技术角度来看，越多线路上云，其基础资源节约能力越强，所需投资就越少。业界普遍认为：专业云是基础，线路云是提升，线网云并力争带来架构的升华是先进方案，其不仅更有益于数据打通、数据交互、多专业融合应用和数据传递快速响应，也更有益于全局感知、大数据开发、人工智能应用，为线网多专业的智慧运维、线网运营指挥中心、线网智慧建造平台等建设提供更有利的基础底座；云计算技术为实现三网融合创新应用、大数据平台、算法平台、业务平台等提供了丰厚的沃土，为实现多业务资源整合、数据快速检索响应、应急联动、预测预警等提供了加速平台，为大数据的扩展应用、数据的深度挖掘、智慧业务的融合创新提供了桥梁和纽带，云计算技术是实现城轨高质量发展的一条道路。

《交通运输部关于推动交通运输领域新型基础设施建设的指导意见》（交规划发〔2020〕75 号）文件中的助力信息基础设施建设部分，对网络安全保护提出了新的要求。“推动部署灵活、功能自适、云网端协同的新型基础设施内生安全体系建设。加快新技术交通运输场景应用的安全设施配置部署，强化统一认证和数据传输保护。加强关键信息基础设施保护。建设集态势感知、风险预警、应急处置和联动指挥为一体的网络安全支撑平台，加强信息共享、协同联动，形成多层级的纵深防御、主动防护、综合防范体系，加强威胁风险预警研判，建立风险评估体系。切实推进商用密码等技术应用，积极推广可信计算，提高系统主动免疫能力。加强数据全生命周期管理和分级分类保护，落实数据容灾备份措施。”

但在现实情况中，城市轨道系统的网络安全保障体系建设并不完善，基本停留在满足国家等级保护 2.0 的要求基础上，虽有一定的网络安全保障措施，但实际效益并不是很理想，并不能完全防范来自内外部的网络攻击。

目前在城市轨道交通行业中，网络信息安全保障基本上是单点、分散的保护模式，缺乏系统、全面、专业的安全防护规划，无行业统一信息安全标准。城市轨道企业信息安全业务的开展主要依赖市场上各类资质和能力参差不齐的信息安全服务企业，一般仅针对单独系统采取独立的信息防护；城轨企业网络安全专业技术人员普遍匮乏，技术能力不足。整个城市轨道交通行业信息安全形势严峻，不仅难以保证当前城市轨道交通网络的安全运营，更无法支撑未来智慧城轨的可持续健康发展。

绿盟科技“智慧安全 3.0”理念提出，在新时代下，需要以体系化建设为指引，构建“全场景、可信任、实战化”的安全运营能力，达到“全面防护，智能分析，自动响应”的防护效果。“智慧安全 3.0”理念提出三大核心要素：“全场景、可信任、实战化”。其中，“全场景”是指绿盟科技致力于面向全部数字化应用场景，针对全部安全要素，提供全方位的安全能力；“可信任”是指绿盟科技要支撑客户构建可信任的能力、可信任的访问与可信任的供应链；“实战化”是指绿盟科技以实战化安全运营为目标，为客户构建按需调度能力，以及响应高效的安全运营体系。

“智慧安全 3.0”理念提出实战化安全运营的思想。以战领建，通过攻防演练等方式，发现网络安全薄弱环节，对防护措施进行验证。在安全建设中融入零信任安全思想，最小化攻击暴露面，基于业务重要性和实际运营需求加强针对性能力适配。通过自动化编排结合专家研判等方式，强化攻防对抗的及时性和有效性，实现对风险的快速和自适应响应。

城市轨道系统依据国家等级保护 2.0 的要求，以“智慧安全 3.0”理念为指引，构建城市轨道网络安全运营中心。作为城市轨道网络安全的设计者、建设者、运营者、服务者，运营中心可以实现智慧轨交背景下网络空间的安全高效治理、集中统筹和集约化运营，持续满足网络政策法规要求，不断应对安全威胁的变化。在保障智慧轨交建设发展的同时，为城市轨道系统数字转型提供有力支撑。

- **智慧安全3.0提升网络威胁预警能力**

作为“智慧安全 3.0”战略的重要组成部分，威胁情报统筹威胁情报生态和能力建设，聚焦多源威胁情报管理，全面提供多源情报接入、融合存储、情报生命周期管理、情报共享输出、威胁预警、情报查询展示等能力。威胁情报支持离线、在线、云计算等不同环境下，融汇多方情报数据并整合应用到自身的安全体系中，全面提升网络威胁预警能力。

- **智慧安全3.0提升网络安全监测能力**

国内外不仅重视传统信息系统网络安全，对工控系统后门、漏洞和攻击等方面的研究也日益重视和深入，一些工控系统产品的安全漏洞信息和攻击代码在互联网上传播，导致对工控系统信息安全实施攻击的门槛逐渐降低，城市轨道系统信息安全形势日益严峻。

城市轨道网络安全运营中心在对传统的网络进行安全监测的同时，也可对城市轨道的工业控制系统进行安全监测，使城市轨道运营单位网络安全运营工作符合国家、行业各项网络安全防护规范要求，实现安全监控“动态感知、智能监控、主动响应、全景可视”的业务目标，确保网络安全事件看得见、看得准、看得深，能够对城市轨道运营单位各业务模块面临的网络安全风险进行持续管控，降低安全风险。

- **智慧安全3.0提升网络安全响应能力**

“智慧安全 3.0”理念体系下的安全防护网遍布“中枢神经与神经末梢”，兼顾有组织、有目的的恶意攻击，以及不易发觉的无意识攻击，如滥用误用等无意识行为及软件缺陷等。安全性取决于链条中最薄弱的一环，泛链式的全场景安全体系建设，可在实战化安全运营中有效提升安全效果。通过本地化的城市轨道网络安全运营，绿盟科技可提供分钟级监测频率、全天候 7x24 小时的信息安全监控、10 分钟级的安全事件响应和 24 小时内重大漏洞预警，帮助城市轨道用户完成安全事件分钟级闭环。

- **智慧安全3.0提升网络安全运营能力**

“智慧安全 3.0”最终将为城市轨道运营单位构建一套“技术先进、安全可靠、服务完备”的城市轨道网络安全运营体系，保障网络和通讯与信息系统安全、稳定、可靠地运行。

通过“基础运营保障、资产安全管理、威胁风险检测控制、脆弱性检测控制、安全风险通报处置、安全风险验证度量、安全检查与风险防范”形成城市轨道运营单位安全运营全生命周期的管理闭环，完全具备“威胁预警、协同对抗、可管可控、智能防御”的安全运营保障能力。

- **智慧安全3.0提升网络安全攻防能力**

城市轨道系统处于网络安全等级保护和关键信息基础设施保护的转型期，监管部门提出了“三化六防”新思想，以“实战化，体系化，常态化”为新理念，以“动态防御，主动防御，纵深防御，精准防护，整体防护，联防联控”为新举措，构建国家网络安全综合防控系统，深入推进等保和关保的积极实践。“以攻促防”将成为未来信息关键基础设施安全保障的常规手段。在满足攻防演练要求的前提下，实战化的运营将是每个运营者的日常工作，与生产运营并列成为轨道交通的重要业务之一。在“智慧安全 3.0”理念的指导下，建设城市轨道安全运营中心必然会整体提升网络安全能力。

智慧安全3.0助力“5G+工业互联网”安全

绿盟科技 解决方案中心 曹东

1.“5G+ 工业互联网”成为新趋势

在当下互联网发展大潮中，随着“消费互联网”的飞速发展，需求侧的数字化水平已经得到明显提升，生产供给侧相对于消费需求侧的数字化水平差异也逐步显现。因此，如何借助数字技术催生新产业、新业态、新模式，如何对传统产业进行全方位的改造以实现整个产业链的数字化转型升级，成为当前社会发展的重要课题之一。

要实现整个社会的数字化转型，首先需要推动网络化的深度发展，将人、物、设备连接起来，将上下游产业链连接起来，最终通过数字化提高全要素生产率。作为产业互联网主力军，工业互联网这一角色得到社会的广泛认同，成为实现深度网络化的重要领域。同时，工业生产中的各种数据非常庞杂，并对数据交互的运动性、低时延等提出更高的要求。而 5G 作为通信技术的最新升级和突破，是满足工业场景下特殊应用连接需求的重要支撑技术。可以预见，5G 与工业互联网的融合发展乘数效应显著，必将为各行各业的创新发展提供巨大增量。

同时，在 5G 与工业互联网融合推进供给侧数字化转型的过程中，跨界融合所带来的安全问题亟待解决。在引入全新安全风险点的同时，新技术的应用及 IT/OT 的跨界融合，还带来网络暴露面持续增大、安全场景更加复杂、管理界限不断模糊等安全问题。

因此，大力发展“5G+工业互联网”是实现社会数字化转型的关键突破点所在。而构建相应的安全体系，则是“5G+工业互联网”健康发展的重要保障。

2.“5G+ 工业互联网”的安全挑战

“5G+工业互联网”的应用和落地，打破了传统工业生产相对封闭可信的网络环境，技术的融合将大量 ICT 领域的威胁和挑战带入了工业 OT 网络。综合来看，除传统意义上的安全威胁（如边界安全、主机安全、应用安全、数据安全等）以外，“5G+工业互联网”的安全挑战主要来自以下三个方面。

2.1 新场景的安全挑战

从技术方面来看，5G 作为新一代移动通信技术，采用了众多的新技术，如 SBA、NFV/SDN、MEC、网络切片技术等；同时，在工业互联网建设过程中，工业互联网平台也以标识解析体系作为全新的应用进行部署。新技术在带来网络能力全面升级的同时，也使得网络风险点不断增加，加剧了信息泄露、数据窃取的风险，给“5G+工业互联网”安全防护体系建设带来了新的技术场景安全需求。

从业务方面来看，不同行业业务应用的网络安全需求存在巨大差异。同时，IT/OT 的跨界融合使得生产安全管理和网络安全管理的界限变得模糊，网络攻击可从 IT 层渗透到 OT 层，从而造成工业系统业务中断等风险。

2.2 信任机制的安全挑战

5G 和工业互联网的应用发展，使得万物互联和海量接入成为重要特征。越来越多的设备相互连接，越来越多的计算在云端和边缘侧进行。在实现广连接、快反应的背景下，无论从技术、业务还是行为等任一角度来看，都存在巨大的信任挑战。

5G 场景下，业务交互模式相较于前几代网络发生了本质的变化，网元功能分离、自动化编排等能力在进一步提升移动通信网络灵活度和适应性的同时，也引入了信令风暴、网元违规访问等业务可信安全问题。在工业互联网场景下，人、机、物全部上网并有了具体的身份，如何实现海量设备认证和安全接入，是万物互联面临的信任问题。

2.3 实战对抗安全挑战

近几年，全球水电、核电、制造等重要行业的企业遭受病毒攻击和感染的事件众多，大范围停电、生产线停摆等重大问题不断出现。

传统封闭的工业网络使得企业对网络安全问题的重视度不高，但随着“5G+工业互联网”的不断推进和落地，以及联网程度的不断提高，在不能快速改变工业互联网设备安全状态的现状下，安全运营和实战对抗能力的高低则成为衡量“5G+工业互联网”安全整体水平的重要因素。

3. “智慧安全 3.0”助力“5G+工业互联网”安全

“智慧安全 3.0”理念是绿盟科技在数字化经济形势下，对网络安全面临的新挑战的观察和对未来发展的深度思考，基于自身多年安全实践所提出来的创新型安全理念体系。该理念的提出旨在构建“全场景、可信任、实战化”的体系化安全能力，达到“全面防护，智能分析，自动响应”的防护效果。“智慧安全 3.0”理念指导下的绿盟科技“5G+工业互联网”安全体系，能够有效应对各类安全挑战，全面支撑用户的安全需求。

“全场景”把智慧安全的外延扩展到整个网络空间和全部数字化应用场景。“5G+工业互联网”使得网络安全无论在技术场景还是在业务场景方面都得到了极大丰富。作为工业互联网的重要基础设施，5G 在赋能工业互联网的过程中，也将工业网络、工业控制、工业数据、终端接入、工业应用等各个维度安全的内涵和外延进行了扩展。绿盟科技“5G+工业互联网”安全技术方案，在“全场景”方面聚焦工业互联网的联网企业、标识解析、工业互联网平台等主要场景安全以及 5G 应用场景下的 5G 核心网和 5G 企业专网场景安全问题，以态势分析作为核心分析和处理的决策点，以防护体系作为态势分析决策的执行点，同时也作为感知点为态势分析提供源源不断的数据支撑，全面打造智能、敏捷的安全闭环，最终以场景安全为切入点，实现安全能力与“5G+工业互联网”完美融合，保证新技术的安全应用和新业务的可靠运行。

“可信任”是对智慧安全内涵的扩展，强调安全不仅仅是攻击防护，还是信任模型的建立与保障。随着 5G 与工业互联网的融合，海量工业终端及 5G 终端的接入成为网络运行的重要特征，相应的安全风险也不断增加，对海量终端访问模型、认证授权模式、信令交互方式等都提出了挑战。传统安全采用边界防护方式，即在网络边界验证用户和终端身份，确定是否被信任。但随着“5G+工业互联网”场景下访问模式的变化以及攻击方式和威胁的多样化升级，传统网络接入安全架构显现出很大的局限性。绿盟科技“5G+工业互联网”安全技术方案，在“可信任”方面全面引入基于零信任的安全思想，组合终端安全、身份识别与管理、网络安全、应用和数据安全、安全分析协作与响应等模块，构建以用户信任和设备信任为基础、持续评估访问过程的行为可信、自适应访问控制的零信任安全架构，实现信任模型的升级。

“实战化”呼应当前安全新形势要求，针对当前网络安全组织化攻击的特点，构建适于实战的创新网络安全体系。“5G+工业互联网”带来的主要安全问题是暴露面的增加，以及攻击途径和攻击方式的变化，同时也造成了攻击的灵活性和不确定性。实战化思想强调提升安全运营的效果，以攻防实战的思维有效应对网络安全的变化性和不确定性。绿盟科技“5G+工业互联网”安全技术方案，在“实战化”方面强调以战领建，通过攻防演练等方式，发现“5G+工业互联网”新场景下的网络安全薄弱环节，并对防护措施进行验证。在安全建设中最小化攻击暴露面，基于业务重要性和实际运营需求加强针对性能力适配。通过自动化编排结合专家研判等方式，强化攻防对抗的及时性和有效性，实现对风险的快速和自适应响应。

4. 结语

在新基建的大背景下，5G 与工业互联网的深度融合成为现代化产业体系创新升级的重要动力，我们除看到数字化能力高效提升以外，更应该看到全面融合所带来的安全问题。“智慧安全 3.0”指引下的绿盟“5G+工业互联网”安全体系，融入了绿盟科技对“5G+工业互联网”安全需求的深入理解，强调体系化应对，全面适应“5G+工业互联网”安全建设需求。绿盟科技愿与各方携手共赢，为保障数字新基建的可持续健康发展共同努力。

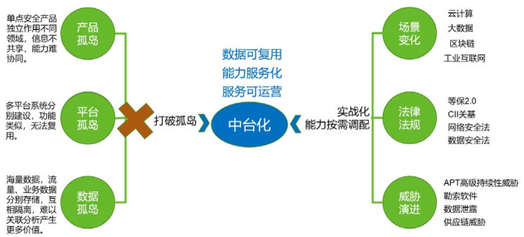
智慧安全3.0实践——中台赋能安全建设

绿盟科技 安全管理产品部 吴天昊

1. 背景

随着安全行业的飞速发展，中台越来越多的被政府和企业单位所接受，并涌现出了一批“安全中台”建设项目。绿盟科技在内部产品服务设计与客户安全运营体系建设中也进行了相关实践，本文将阐述“智慧安全 3.0”理念如何融入“安全中台”建设。

“中台”往往是业务发展到一定阶段，存在大量的冗余，无法进行快速发展后，才开始逐步推进的。在全球数字化变革的背景下，安全业务的碎片化越来越严重，碎片化速度也越来越快，如何应对碎片化的安全需求与安全场景是近年来网络安全所面临的最大挑战。通过中台建设，可以整合数据应用，提升效率，以中台为中心快速构建更宏大的上层业务能力。



数字化变革催生中台建设

2. 智慧安全 3.0 实践：中台安全建设思路

“中台”并没有一个标准的概念，一种普遍接受的观点是，中台是将系统的通用化能力进行打包整合，通过接口的形式赋能到外部系统，从而达到快速支持业务发展的目的。

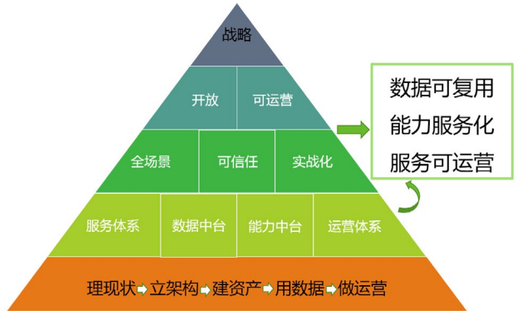
在业务发展的过程中，许多不同业务应用的数据和能力是互通的，随着大数据技术的成熟，企业开始寻求改变，将共性能力抽象出来，加大投入，建设一个完善的“中间层”，它具备强大的数据收集与处理能力，具备高度的安全能力抽象和完善的对外能力接口。它可以将各类数据源的数据集中收集，形成标注化数据资源池，供上层业务消费。又能抽象各种各样安全能力（策略管控、设备管理等），供上层业务调用。它具备完善的上层开发标准及 API 接口，通过这个“中间层”，客户或服务提供商不再需要考虑数据获取或能力对接，可以全力投入在上层应用搭建中，大大提升上层应用搭建速度。足以支撑不同业务场景下上层应用的快速搭建，这个强大的中间层，就是中台。

对于安全体系建设来说，数据和能力是至关重要的。大中型企业由于业务系统繁多，组织架构复杂，参与建设的厂商众多，很容易出现烟囱式建设，大量业务之间数据冗余，能力冗余，造成资源的浪费，建设成本大大提升。

在“智慧安全 3.0”理念指导下，绿盟科技已经有针对“安全中台”的整体方案，协助用户提升安全能力。许多大中型用户，在多年的

安全建设过程中已经积累了比较全面的安全体系，安全能力增长也会逐步乏力，基于此，用户经常会选择进行中台改造。在改造实践过程中，我们一般分为三个阶段：

1. 整体方案设计。针对已有的安全能力进行梳理，了解客户的业务需求，形成横纵结构，识别共性数据需求及能力需求，设计中台框架。

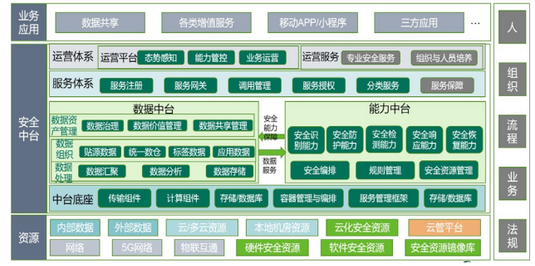


以中台为核心的安全体系规划

2. 中台能力建设。中台能力建设往往伴随着大数据平台的建设，或者在大数据平台建设完成后进行，基于大数据平台，对数据进行集中收集与治理，抽象标准化各类“原子能力”，封装对外接口，构建安全中台。

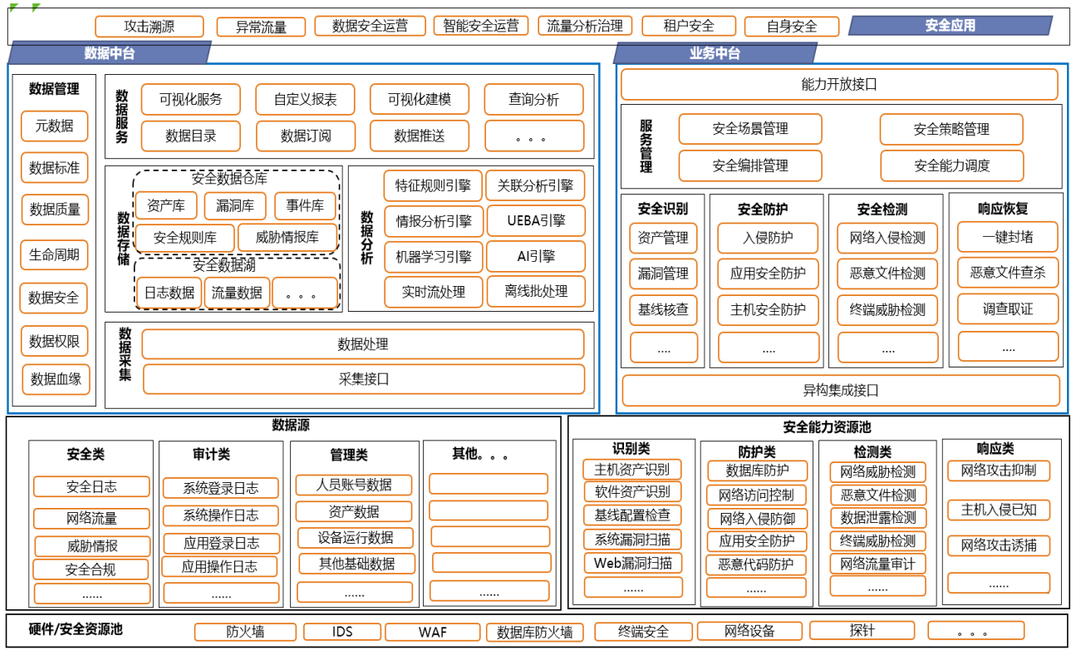
分层解耦，打造数据中台与能力中台

3. 上层业务快速孵化，经过一系列的中台改造并形成中台能力后，用户安全基底建设完成，引入各种厂商，快速构建上层安全业务，实现安全能力的高速提升。



中台拉通各类资源，安全应用高速建设

2019 年，绿盟科技参与了某客户的“安全中台”建设，该客户信息化建设程度非常高，参与厂商众多。但随着业务发展，需要安全业务上云并对外开放安全业务。安全业务建设过程中，发现安全能力由各厂商分别建设，数据很难整合，安全能力也比较分散。基于此，我们为客户量身定制了一套安全中台体系，梳理各厂商安全数据，整合底层安全能力，将数据与能力“框”在了一个系统下，并形成标准的对外接口，中台建成后，上层由绿盟科技与其他厂商快速搭建了各类安全能力，针对云、管、端、边各类安全场景进行全场景覆盖，实现底层数据及能力融合互通，随着上层业务的开展，不断强化中台能力，安全建设效率大大提升，安全运营能力快速增长。



面向全场景的安全中台建设

3. 结语

近年来，信息化建设高速发展，数字化转型势在必行。大中型政企机构要实现自身业务需求的全场景覆盖，能力与数据的集约化是必经之路。通过建设安全中台，数据上，整合数据收集能力，打破数据存储孤岛，完善数据的有效利用，实现安全数据的统一管理与融合共享；能力上，改变原有的“烟囱式”建设思路，达到安全能力的标准化、原子化，形成安全能力服务目录，实现安全能力的统一管理，高效利用。最终支撑上层业务的快速孵化，实现安全业务全场景覆盖，提高实战化运营效率。

大数据场景下的安全数据分析及威胁模型构建

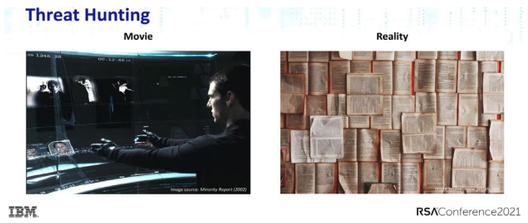
绿盟科技 能力中心&平行实验室 王津

摘要：RSA 2021 大会主题为：Resilience（弹性），强调可恢复性和健壮性。该主题在如今世界疫情导致的混乱大背景下显得非常贴切，这或许也是黑客 & 威胁及风险管理主题相关内容在本届主题中占比最大的原因之一。当然，作为世界影响力最大的信息安全大会，传统安全所关注的一系列相关主题仍是讨论热点。很多参展厂商针对安全领域持续关注的课题提出了自己的思路，其中大部分是再次强调过去实践验证有效的成功经验和方法，另一部分则是新的尝试。

大数据场景下，威胁数据分析和威胁狩猎一直是企业级和国家相关监管部门主要的安全应用场景，也都是 RSAC 一定会涉及的相关主题。内容往往涵盖宏观的威胁框架和业务流，以及具体的行之有效的算法应用和数据处理方法等。我们通过梳理本届参展厂商汇报演讲内容，对海量数据背景下部分厂商的数据处理分析和威胁模型构建思路进行总结。

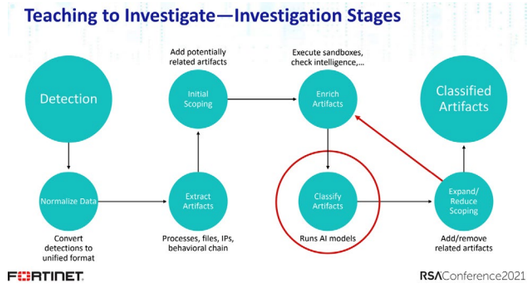
1. 海量多模态数据的处理方法

大数据场景下，威胁安全分析一开始需要面对的问题就是如何有效处理接入的海量告警。在一个典型的大数据场景下，接入的数据往往是海量且异构的。

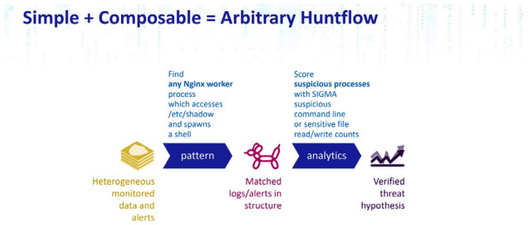


这一阶段的核心诉求在于，一方面，希望接入一切能够接入的数据以保证威胁特征的完整性，这些数据通常包括终端数据、各种网络流量探针数据、威胁情报甚至研判人员发出的相关日志等；而另一方面，又希望接入的数据能够得到有效整合和筛选，凸显出真正值得关注的少量数据，从而保证威胁特征的有效性。这两个需求在一定程度上互相矛盾，但利用行之有效的范式化方法和特征关联筛选之后，仍然可以同时被满足。

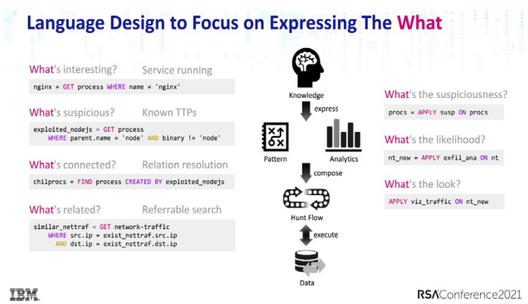
来自 Fortinet 的 Roy Katmor 和 Udi Yavo 在演讲中列出他们在数据处理阶段的一些关键步骤，包括数据范式化、特征提取、关联和富化。



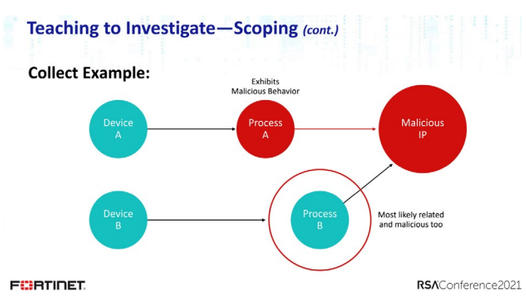
IBM 的 Xiaokui Shu 和 Jiyong Jang 在介绍他们的开源项目 Kestrel 时，将他们的威胁狩猎业务流定义为 2 个关键环节：多模态告警数据的模式化，以及基于该模式的分析模型。

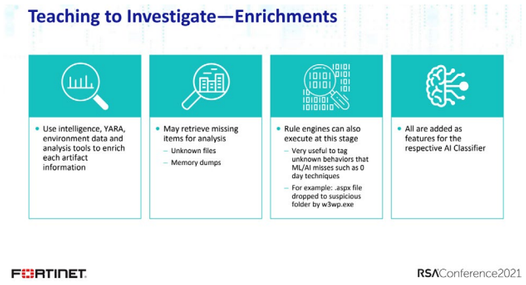


需要指出的是，Xiaokui Shu 所说的多模态数据的模式化是基于威胁特征层面的模式化，而非简单的数据（record）层面，这是他们后续进行基于实体（entity based）的威胁分析模型构建的基础。

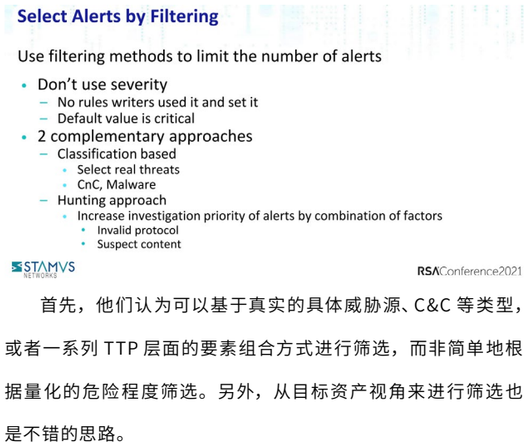


此外，基于初始数据进行有效的关联扩展（Scoping）和上下文的富化（conetxt enrich）可以有效补充更多的威胁特征，以支撑后续威胁模型的训练和推理。

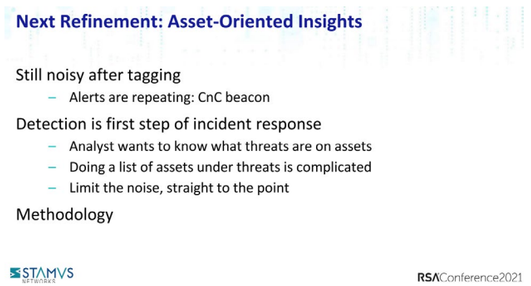




在有效聚焦和筛选数据方面，Stamus Networks 的两位专家给出了他们的思路。



首先，他们认为可以基于真实的具体威胁源、C&C 等类型，或者一系列 TTP 层面的要素组合方式进行筛选，而非简单地根据量化的危险程度筛选。另外，从目标资产视角来进行筛选也是不错的思路。



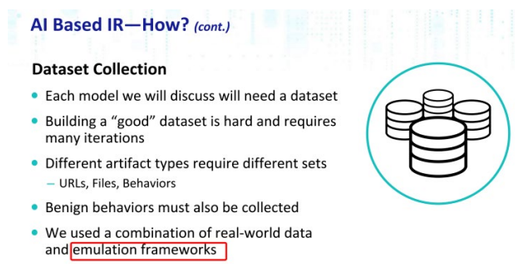
综上所述，大数据场景下的海量多模态数据处理思路可以总结为几个关键环节：多源数据的采集、数据的范式化、数据的特征富化以及基于特征的筛选。每一个关键环节的具体做法往往依赖于具体安全业务场景和需求，更取决于后续威胁模型的具体数据要求。

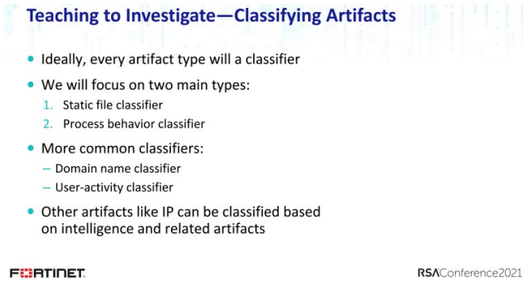
2. 威胁模型构建方法

大数据场景下的威胁模型构建往往绕不开各种人工智能算法的参与，但与当年机器学习（尤其是深度学习）刚取得突破性进展时“机器学习无所不能”的氛围不同，近年来，包括信息安全在内的各个行业对于人工智能，特别是机器学习的局限性等问题的认识越来越清晰，Fortinet 的两位专家在他们的《Applying Artificial Intelligence to the Incident Response Function》中就指出，在事件响应方面，AI 不能完全取代人工。

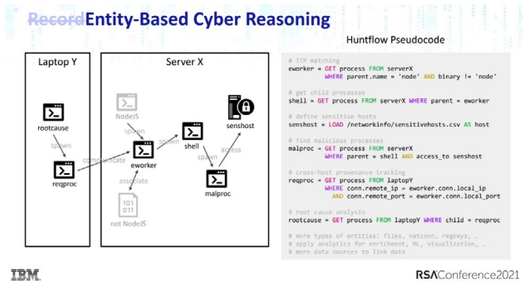
因此，目前绝大多数研究人员不再盲目相信智能算法，而是转而寻求人工深度参与的“半智能”方法，将专家知识和智能算法进行结合，从而提升算法的可控性和可解释性。

Fortinet 的两位专家通过在分类模型的训练数据中引入模拟攻击数据来进一步加强对分类模型的人工干预，并基于细化的威胁特征场景来进一步构建不同的分类模型，降低对分类模型的过度依赖，提升分类模型的可控性。



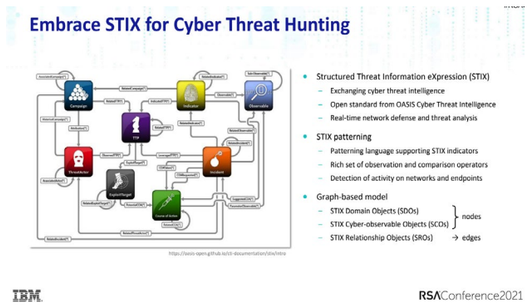


而 IBM 的 Xiaokui Shu 和 Jiyong Jang 则提出了另一个相对较为新颖的思路：对基于行为特征构建的威胁实体，结合专家构建分析模型进行推理。



这个方法有一个前提，就是前文提到的利用行为特征模式化多模态数据，将海量多源异构数据转化为威胁实体，从而从接近行为的层面进行关联推理。

他们还进一步提出，可以参照 STIX 的框架进行对应，将 STIX 中的域对象和关系对象对应威胁实体和推理生成的关联边，从而极大地提升模型的共享和导入能力。

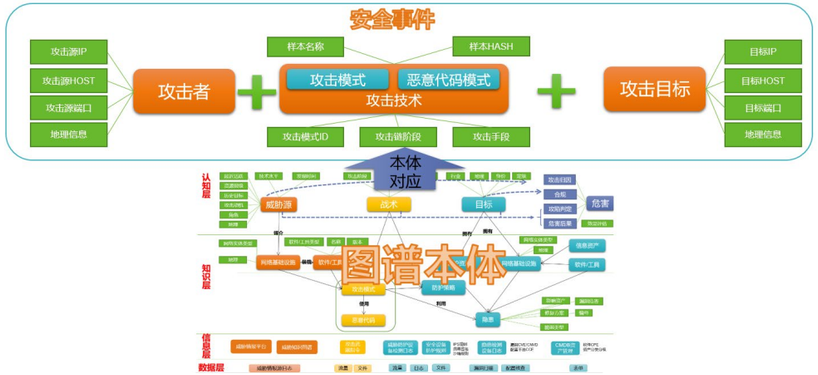


3. 绿盟科技相关研究

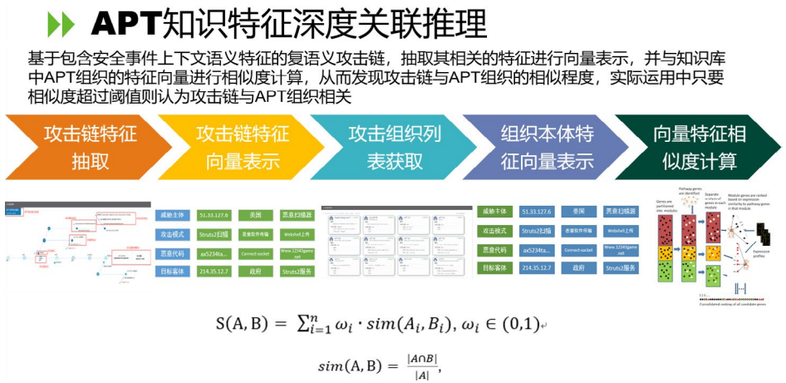
绿盟科技平行实验室一直持续关注大数据场景下多模态数据的感知理解和威胁模型构建方面的研究。与上文所介绍的几个厂商研究者的方法类似，我们基于对威胁安全的认知，构建了绿盟科技威胁安全知识图谱，并基于图谱本体，将大数据平台接入的多模态数据范式化理解为威胁实体，依托知识图谱保存的威胁语义知识，在实体层面扩充并关联事件语义，结合专家知识和攻击链等模型对事件进行整合及筛选。

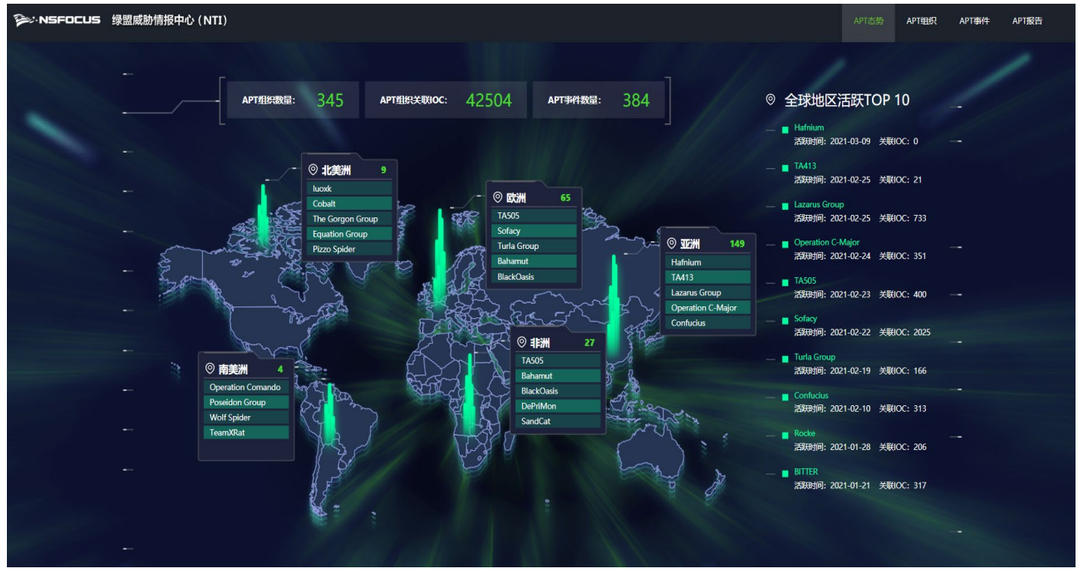
特征融合-范式化安全事件定义

基于威胁知识图谱的本体结构，从**攻击者、攻击技术、攻击目标**三个维度定义安全事件



在威胁模型构建方面，我们通过抽取事件及相关上下文语义特征，与图谱中的 APT 组织、恶意代码家族进行特征向量相似度计算，对威胁事件的攻击团伙进行归因。





此外，绿盟科技构建了以威胁源为核心的特征图模型，并利用图计算进行多次迭代的聚类，从而发现隐藏于海量事件中的团伙活动。团伙特征也会在简单研判之后被保存至图谱团伙知识库中，团伙知识库同样支持 STIX 格式数据的导入和导出。

4. 结语

通过梳理 RSA 2021 大会中大数据场景下的数据分析和威胁模型构建相关方面的研究汇报，我们发现一些传统的思路没有改变，如尽可能接入可能包含威胁特征的多源数据，在保留威胁特征的前提下进行数据的范式化和筛选等，威胁情报引入、上下文语义的富化等处理方法也逐渐被更多厂商提及。

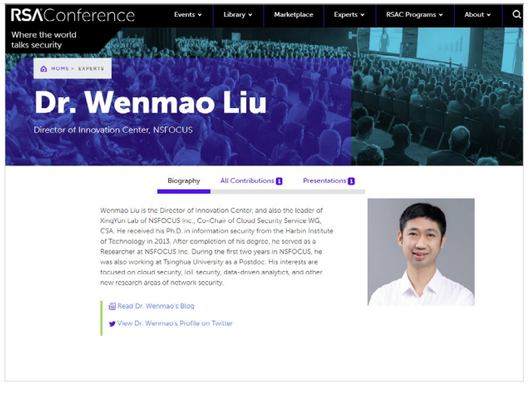
另外，值得一提的是，随着业界对于包括机器在内的人工智能算法的理解逐渐趋于理性，几乎未再看到单纯依靠人工智能算法支撑安全业务的情况，更多安全研究者正在考虑进一步分解安全业务，并加强专家知识的主动干预，从而在有效利用人工智能算法高效处理能力的基础上，提升算法的可控性和可解释性。这个思路也是绿盟科技近年来秉承的基本思路。

物联网中基于UDP的DDoS新型反射攻击研究

——绿盟科技刘文懋RSAC主题演讲

绿盟科技 物联网安全产品部 刘军

RSA 作为全球规模最大的网络安全行业会议，迄今为止已举办 30 届，一直着眼于推动全球网络安全界的共享、创新与进步。绿盟科技在 2021 年脱颖而出，发表物联网安全领域主题演讲《物联网中基于 UDP 的 DDos 新型反射攻击研究》（Research on New Vectors of UDP-Based DDoS Amplification Attacks of IoT，[SAT-M19]）。绿盟科技创新中心高级总监刘文懋博士，代表绿盟科技的物联网安全研究团队进行了主题演讲。

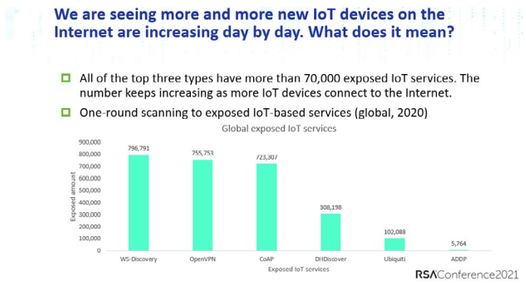


下面与大家分享绿盟科技在 RSA 2021 大会上的演讲精华。

1. 全球物联网资产暴露情况

随着越来越多的物联网设备接入互联网，物联网联网数每天都在增加。通过对互联网上的设备进行扫描，我们发现全球开放 WS-Discovery、OpenVPN 和 CoAP 协议的物联网服务超过 70000 个。

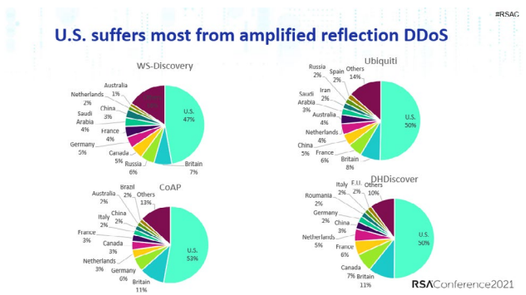
不仅安全厂商可以发现这些物联网暴露资产，攻击者也能够发现这些资产。通过扫描器、僵尸网络或任何可以使用的工具发现物联网资产后，这些资产容易遭到攻击，以及被利用发起攻击。



2. 美国是遭受放大反射 DDoS 攻击影响最大的国家

通过物联网蜜罐捕获和收集受害者的 IP 地址，在计算目标 IP 地址的地理位置后，可以看到美国受放大反射 DDoS 攻击的影响最大。

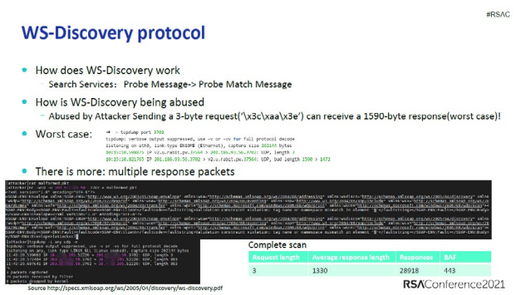
无论是勒索软件还是 DDoS 攻击，都可以作为一种黑产服务，此前在暗网上已经出现明码标价提供的租用 DDoS 服务，而暴露的脆弱物联网设备成为黑产潜在的攻击武器。美国拥有庞大的商业和 IT 产业，因而也成为网络犯罪的最大目标。



3. WS-Discovery 协议介绍

WS-Discovery 是基于 UDP 的、用于 Web 服务发现的单播协议。其工作原理是客户端发送 UDP 探测消息搜索服务，然后等待应答。该协议具体被滥用的情况是：攻击者发送一个 3 字节的请求：3c、aa、3e，并携带一个欺骗的源地址，服务会回复一个 1590 字节的响应。

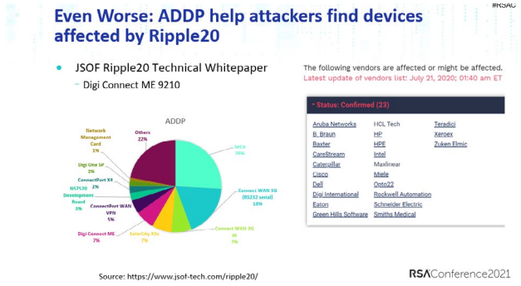
这里使用了 BAF（Bandwidth Amplification Factor），带宽放大系统的概念，这个概念最早在 2014 年 NDSS 的论文《Amplification Hell: Revisiting Network Protocols for DDoS Abuse》中提到。为了计算 BAF，可以将有效负载发送给使用真实源地址公开的所有服务，验证获得的响应结果数据。经过测试，发送的请求的长度 3 字节时，收到的响应的平均长度是 1330 字节，计算出 BAF 数值是 443。利用该协议漏洞，通过 WS-Discovery 可以产生比请求流量大 400 多倍以上的恶意流量。



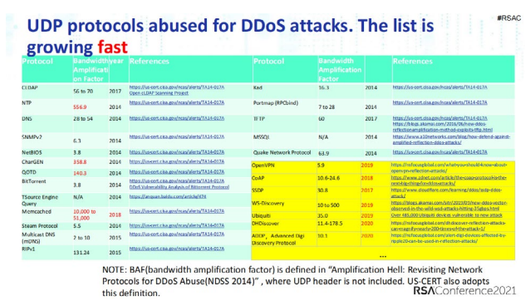
4. ADDP 帮助攻击者找到存在 Ripple20 漏洞的设备

ADDP 是高级设备发现协议（Advanced Device Discovery Protocol），是 Digi International 公司开发的基于 UDP 的多播协议。借助 ADDP，无论网络如何配置，设备都可以在本地网络上发现其他设备。经过全网扫描测试，我们发送的请求的长度是 14 字节，而收到的响应的平均长度是 141.7 字节，对应的 BAF 是 10.1。

事实上，ADDP 被许多数码网络设备使用，这些设备中的大部分可能存在 Ripple20 漏洞。因而，攻击者通过发现暴露的 ADDP 服务，再验证 Ripple20 漏洞，则可以发起一些攻击。



除 WS-Discovery、ADDP 之外，报告还分析了 OpenVPN 协议的脆弱性，此外绿盟科技在 2018 年、2019 年和 2020 年发布的《物联网安全年报》中分析了 SSDP、DHDiscover、Ubiquiti 等协议，这些物联网协议都存在相似的脆弱性：基于 UDP、支持单播、响应远大于请求长度，因而容易被利用发动 DDoS 攻击。事实上，在 2017 年后，利用物联网协议发动 DDoS 攻击俨然成了攻击者的重要选择。



5. 一些建议和观点

给物联网厂商的建议：

首先，设置首席安全官，组建安全团队。其次，在设计环节，默认禁用服务 / 设备发现功能，非多播不响应，非内网不响应。最后，在运营环节，建立应急响应流程并及时发布安全补丁。

给最终用户和机构的建议：

识别自有的物联网设备，检查配置、访问控制策略；持续地使用网络空间测绘技术监控暴露资产；构建识别—评估—治理的安全运营闭环，将物联网安全融入统一的安全运营体系内。

给物联网客户的解决方案：

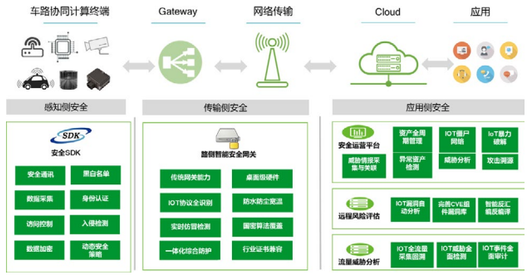
关注城市、企业物联网安全隐患，综合展示物联网各垂直领域风险态势，以及各地区、部门威胁情况，使用绿盟物联网保护伞解决方案，通过终端 SDK、固件检测、准入网关、物联卡分析、物联网安全测评等多个系统的数据，支撑物联网安全态势。

一平台，五系统支撑

关注城市、企业物联网安全隐患，综合展示物联网各垂直领域
风险态势，以及各地区、部门威胁情况

未来一段时间内，更多物联网协议和设备的漏洞会不断出现，绿盟科技通过创新中心、格物实验室、物联网安全产品部的联合，将“研、产、用”进行结合，通过物联网保护伞解决方案，为物联网场景及客户安全保驾护航。

绿盟科技车路协同网络安全技术方案，着眼于规模化车路协同应用，采用了车载可信级“安全 SDK+ 路侧智能安全网关 + 安全运营平台端到端”的安全联动架构模式，构建监测、检测、预警、防御、响应与应急处置安全能力，全面覆盖感知侧、传输侧、平台 / 应用侧防护场景，为智能交通领域的网络安全保驾护航。



通过车联网终端及平台探针部署、威胁情报采集等，收集车路协同通信网内安全数据信息，并基于大数据关联分析处理，形成了主动探测、被动诱捕、流量分析、僵木蠕、DDoS 攻击、APT 检测等安全监测、检测、预警、防御、响应与应急处置安全能力，结合安全咨询、渗透测试、全生命周期安全风控等安全服务，构建车路协同安全运营体系。从方案价值来看，能够满足车路协同安全合规及新基建网络安全建设安全的迫切需求，进一步保障整个车路协同应用安全、可控、健康发展。

网络威胁狩猎：回归“乐趣”

绿盟科技 伏影实验室

威胁狩猎的过程就像网络世界中的“科学探索”，狩猎成功前需要进行大量繁杂、乏味的分析任务，如复杂的数据查询、大量手工数据关联等。这些烦琐的操作不仅影响效率，也降低了安全人员的战斗力。在 RSA 2021 大会中，来自 IBM 的研究员提出了全新的威胁狩猎工具，该工具可以有效减少狩猎过程中的烦琐工作，使狩猎重新聚焦到其创造性和令人兴奋的部分，实现乐趣回归。

1. 威胁狩猎的痛点

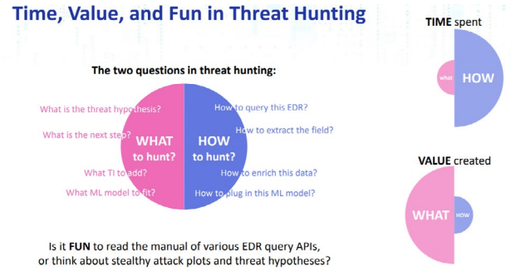
“威胁狩猎”指采用人工分析和机器辅助的方法，针对网络和数据进行主动搜索、关联和分析，从而检测出逃避现有安全防御措施的高级持续性威胁（APT）。威胁狩猎过程中，安全人员主要思考和解决两大类问题：

- 如何狩猎？

例如，“如何从 EDR 查询数据”“怎么提取数据字段”“如何补充线索的上下文信息”“如何使用机器学习模型”等。

- 狩猎什么？

例如，“如何建立假设”“如何基于假设进行分析”“需要哪些威胁情报数据参与分析”“哪种机器学习模型适合分析”等。

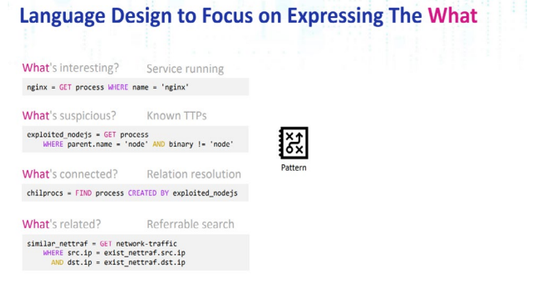


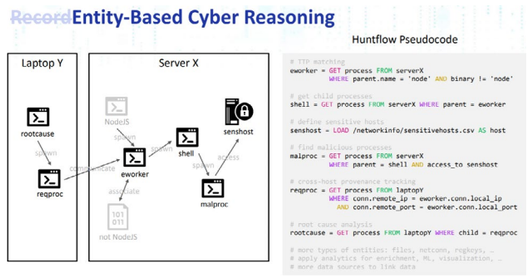
当前威胁狩猎的痛点在于：“如何狩猎”所花费的时间远超“狩猎什么”，但后者才是创造价值的重点。安全人员将太多时间浪费在阅读各类 EDR 的 API 查询接口上，而无法聚焦到对威胁狩猎更有价值的威胁假设和攻击情节分析部分。

2. 推荐“狩猎编程语言”——Kestrel

研究人员推荐了一个全新的开源项目“Kestrel”，它可以降低安全人员在“如何狩猎”上的投入，将精力聚焦到威胁狩猎上。为了实现这个目标，Kestrel 对面向威胁狩猎的编程语言进行了定义，可帮助使用者更高效地进行狩猎。

为帮助使用者将目标聚焦在狩猎上，Kestrel 编程语言在设计理念上采用了方便狩猎目标描述的语法，有助于对狩猎目标实体进行表示。

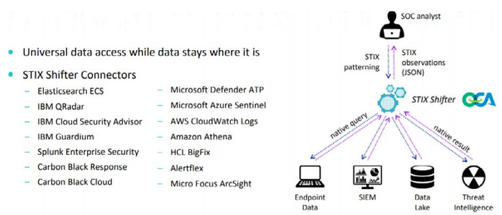




3. 推荐规范化数据格式——STIX

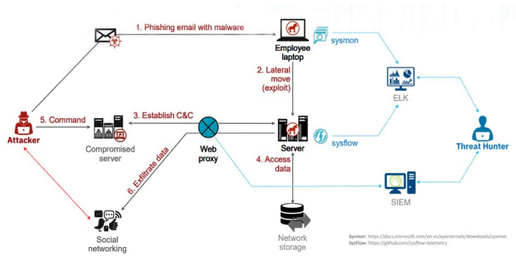
数据的标准化、规范化是威胁狩猎有效进行的关键，只有采用规范的数据格式，才能从不同系统收集到统一的数据进行狩猎。来自 IBM 的研究人员推荐采用国际威胁情报标准 STIX 进行数据表示和共享。

为实现安全数据在不同系统之间直接的安全共享，IBM 提出了 Security Connect 的概念，其核心技术之一是 STIX-Shifter 工程。该工程使用结构化威胁信息表达（Structured Threat Information eXpression，STIX）进行安全共享威胁信息。在该工程中，STIX-Shifter 能保持从所有 IBM Security 产品和各第三方产品中收到的数据的一致性。通过提供同一个通用 API 的通用服务，IBM 如今可以跨任意数据源查询数据，无论数据源是产品还是存储仓库，并收获同样的数据投入分析和搜索中。



4. 如何构建威胁狩猎场景

研究员分享了如何基于免费工具进行威胁狩猎环境搭建的案例。案例中结合黑客攻击的路径，使用 sysmon、sysflow、SIEM 等软件工具平台，分别在攻击者的钓鱼邮件投递、横向移动、C&C 联络通信、数据访问、数据渗出等过程中进行数据监测和收集。收集到的数据经过规范化处理后，提供给威胁狩猎平台。最终安全人员可以在威胁狩猎平台上使用 Kestrel 进行威胁狩猎分析。



5. 结语

威胁狩猎进入实战应用阶段后，如何在实际操作过程中提升安全人员的狩猎效率是业内关心的问题。IBM 的研究人员推出的方案可以成为企业进行威胁狩猎的选项之一，该方案有很多亮点：

- 狩猎方案采用开源/免费软件进行搭建，搭建成本降低；
- 狩猎场景基于黑客攻击过程进行狩猎点设计和部署，在体系化建设威胁狩猎思路方面值得参考；
- 新推出的狩猎工具/语言Kestrel可以大幅提升溯源效率，让安全人员脱离苦海。

Kestrel 可以帮助安全人员降低异构数据信息获取、烦琐基础操作等方面的开销，但仍是人工狩猎操作。接下来，如何基于攻击线索的自动关联尽可能实现威胁狩猎的自动化有望成为新的技术方向。

深度社会工程学攻击，你了解多少？

绿盟科技 伏影实验室

摘要：被誉为全球网络安全发展风向标的 RSA 2021 已闭幕，大会分享了最新的安全动态、前沿的技术理念及未来的安全方向，其中与社会工程学攻击相关的议题被广泛探讨。在信息安全这个链条中，人的因素是最薄弱的一环节。社会工程学是一种针对被攻击者的心理弱点、本能反应、好奇心、仁慈、信任、贪婪等心理陷阱，采取的诸如欺骗、伤害、信息盗取等对社会及人类带来危害的行为。而社会工程攻击，就是一种利用“社会工程学”来实施的网络攻击行为。

1. 深度社会学攻击特点解读

BioCath 公司在 RSA 2021 会议上带来了一场精彩的演讲。演讲人从福尔摩斯红发会的故事讲起，引发了对社会工程学（Social Engineering）的讨论，并进一步探讨了关于深度社会工程学攻击的课题。

2. 网络安全与社会工程学

结合网络安全来定义社会工程学：

从心理学的角度出发，密谋一场精心的骗局，诱使目标人物泄露机密信息，以达到收集信息，欺诈或访问用户系统等目的。而社会工程学的运用通常是复杂骗局中必不可少的步骤之一。

目前讲，在社会工程学的范畴下，网络安全可能会遭受的攻击类型分为以下四种。

2.1 静态的机密信息收集（Static Credentials Harvesting）

钓鱼攻击 / 语音钓鱼攻击 / 短信钓鱼攻击：这类攻击会诱骗受害者主动地泄露机密信息，如个人信息、银行信息等敏感内容。

2.2 RAT 陷阱（RAT Traps）

在攻击前，攻击者会诱导受害者在其个人电脑或移动手机上安装远程控制工具（RAT）。

2.3 OTP 的收集与用户的分心

通过电话诈骗收集 OTP 以供立即使用。例如，木马 MITB，就

旨在分散用户注意力和收集 OTP。

2.4 深度社会工程学攻击

与传统的社会工程学攻击不同，深度社会工程学攻击，看起来是一个更完美的骗局，让受害者浑然不知，从心理上认识不到自己已深陷骗局之中。其真正的目的是使受害者将资金直接转给欺诈者。

3. 标准的社会工程学攻击

随着网络的发展，以及电子金融的普及，越来越多的网络骗局已经转向电子银行，其最终的目的都是使用各种手段来骗取受害者的财产。但是，对于前三种社会工程学攻击，从一些细节上是可以识别正常和非正常操作的。例如，下图中的两个例子是在登录页面，正常操作和不正常操作所带来的差异，具体表现为鼠标移动的轨迹和付款流畅程度等。

正常操作下登录网银：

鼠标移动的轨迹流畅且连续，键盘用于输入 OTP，所有的行为与账户的基准是匹配的。

非正常操作下登录网银：

鼠标的移动轨迹出现跳跃、卡顿或中断，付款存在异常（因为远程控制你的攻击者可能在不同的国家）。

尽管这些离线的社会工程学攻击是无法直接检测的，但是我们可以通过用户级别和总体级别的异常来检测欺诈，如从鼠标移动的轨迹、滚轮滚动方式与时间、键盘删除信息的方式、选择国家的方式等。

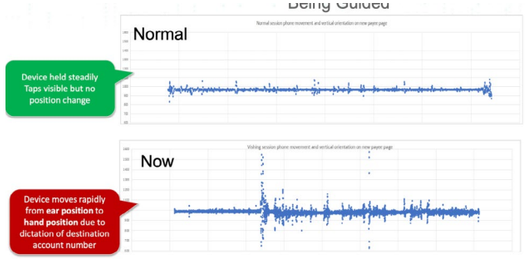
4. 深度社会工程学攻击

深度社会工程学攻击，是一种全新的骗局。2019 年首次在英国出现，后来逐渐蔓延到欧洲、澳洲以及北美洲各地。

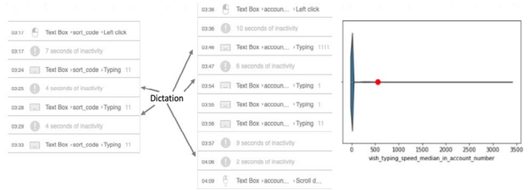
我们可以从一些细微的痕迹上来观察用户在遭受这种攻击时，所表现出的一些不寻常行为。

例如，被攻击者停留在银行的页面上时间过长，鼠标会有过长的时间来回移动，且从行为上疑似用户不知道要干什么 。因为在整个过程中，攻击者不断地利用语言去营造一个故事，让被攻击者去相信自身并不是在一个骗局中，所以被攻击者的行为看起来是分心的，并不是专注在银行的页面上。并且，在最终点击提交的按钮上，被攻击者的行为显得极为犹豫。

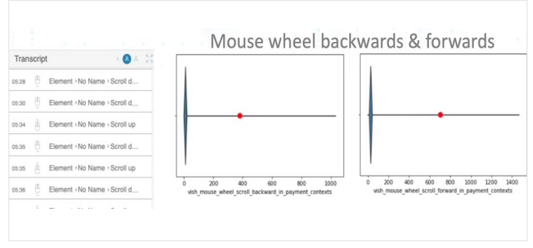
攻击者采用电话语音时，在正常的情况下，语音的总体音量是平稳的，不会出现大的改变。但是在深度社会工程学攻击下，语音会出现波动。根据 BioCath 公司的分析，因为用户要记录目的账号，所以手机经常会从耳边移动到不同位置。如下图所示：



从用户输入信息(如账号等) 的时间上,我们可以发现一些不同。总体时间上，用户的输入所用时间比正常情况下的要短，因为用户需要听写对方给予的新账号。



同样，还有一些细微的行为可以看出用户的犹豫和不安。例如，点击提交按钮的时间，以及在交易后，用户频繁地滚动鼠标滚轮。如图所示，滚动滚轮的频率远高于正常情况下的次数。



最后，整合多个微弱信号中的不同，通过机器学习，最终去判定是否受到攻击。

5. 结语

标准的社会工程学攻击 ——目前的趋势

（1）说服用户在 PC / 移动设备上安装远程访问工具。

（2）假装自己是银行，诱骗用户通过电话提供 OTP 码。

（3）在特洛伊木马攻击中使用，以分散用户的注意力和收集 OTP。

深度社会工程学攻击——一种全新的犯罪类型

（1）引导用户向犯罪分子汇款。

（2）受信任的设备，没有恶意软件 / RAT，没有犯罪的行为。

（3）由于这是完全授权的交易过程,因此不是真正的欺诈行为，但监管机构要求采取行动。

如何建立基于情报和威胁狩猎能力的实战化运营体系

绿盟科技 安全运维保障部 邵子扬 李子奇

图 1 RSAC 2021 主题“弹性”

1. 安全弹性

享誉全球的安全大会 RSA 2021 已经落下帷幕，RSA 公司每年都会根据网络形式为 RSAC 确定一个主题，今年的主题是“弹性”(Resilience)。

NIST.SP 800-160（卷 2）将网络弹性定义为“预防、抵御、恢复、适应那些施加于含有网络资源的系统的不利条件、压力、攻击或损害的能力”。基于“对手成功突破防御，并且在组织中长期存在”这一假定，弹性网络安全要求企业更加快速地发现攻击痕迹、降低事件响应时间、抑制攻击者对企业的损害并进行恢复。

网络弹性实施方法中要求企业进行分析监控，包括监控并评估损害、数据和情报关联分析、取证并进行行为分析。威胁狩猎则是其中一种方法——同样基于“对手成功突破防御”，其旨在通过关联情报、数据和攻击假设，主动发现关键信息基础设施中是否存在隐蔽恶意行为并及时进行应急响应，能够帮助企业快速发现入侵、降低威胁风险并恢复业务正常运行。

NIST.SP 800-160（卷 2）还指出，使现有技术更具有网络弹性，需要用威胁情报信息补充现有监控服务，更好地发现与寻找入侵。组织需要参与威胁情报消费、共享、协同，应对千变万化的网络威胁。

在 SANS 的网络安全滑动标尺中，威胁狩猎处于滑动标尺的主动防御阶段，是在组织安全架构、被动防御、主动监控较为成熟的条件下进行的主动发现潜藏威胁的行为，通常被描述为无事件的事件响应。同时，企业构建自身网络威胁情报、参与外部情报协同（例如作为威胁情报的使用者、双向威胁信息共享、协同合作以应对威胁），形成更加成熟的网络威胁情报体系，能补充现有监测服务的监测能力，为威胁狩猎增加更多可能及行动价值。

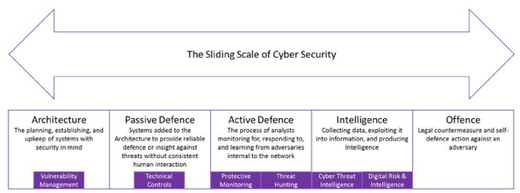


图 2 SANS 网络安全滑动标尺

2. 威胁狩猎

威胁狩猎分为结构化狩猎和非结构化狩猎。结构化狩猎主要基于 IoA（攻击信标）以及 TTPs 进行，而非结构化狩猎则是基于情报的数据驱动狩猎，与结构化狩猎相比更加宽松，即使是低置信度的 IoC 也可以作为一个狩猎规则。

Tim Bandos 在“Hunt and Gather: Developing Effective Threat Hunting Techniques”的分享中提到结构化狩猎的流程，是一个基于 ATT&CK 威胁模型的假设型搜寻（当然猎人也可以混合情报，甚至结合行业态势以及基线异常进行狩猎），包括模型的选择、攻击特征识别、狩猎规则建立，然后再进行规则部署并从基线数据中发现威胁。

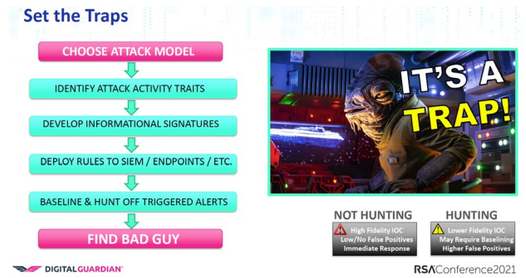


图 3 威胁狩猎流程

其整个步骤基本匹配了建立假设、印证假设、调查事件、事件分析四个主要流程。图 4 是一个扩展型的狩猎流程，更好地展现

四个基础流程的附加细节，包括建立假设所需情报、警告输入、事件调查中进行漏洞管理、SoC 数据分析的操作、输出到应急响应团队（CIRT）的流程。

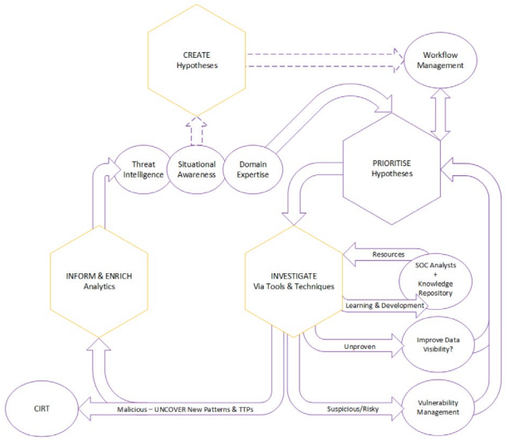


图 4 扩展狩猎流程

面对海量网络威胁情报、安全日志 / 设备信息以及重复性的威胁狩猎流程工作，IBM 的研究员在“The Game of Cyber Threat Hunting: The Return of the Fun” 中提出一种开源的威胁狩猎语言——Kestrel-lang，将资深威胁猎人描述的狩猎步骤转化成为 Huntflow（猎流），将 Huntflow 适配多种语言，同时进行共享以便于重新执行，最后提供逻辑数据表示。看到这里，Huntflow 可能会让读者想起 SOAR 编排脚本，其实它们的目的都是减少人力，提高效率，一个针对发现，而另一个针对处置。

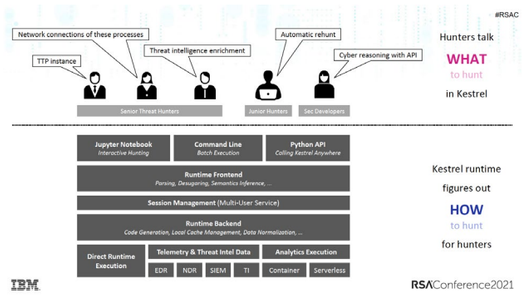


图 5 自动化狩猎语言 Kestrel

威胁猎人使用 Kestrel-lang 进行狩猎的过程，就是将其狩猎思路表达成 Kestrel-lang 中的匹配语句（展示猎人希望看到的元信息或者关系内容）和分析逻辑（引入外部的情报或者其他计算逻辑进行数据分析），匹配语句以及分析逻辑聚合成为 Huntflow。形成的 Huntflow 将能被用于运行，完成自动狩猎。

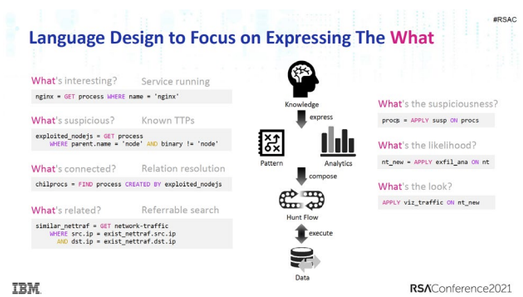


图 6 Kestrel 进行狩猎

自动化能够切实解决如狩猎复用、多平台、人力依赖等问题，这也是组织迈向更高成熟度威胁狩猎的必要条件。自动化不应该单用在狩猎脚本上，也可以用在情报整合与消费、情报生成与共享以及自动化处置上，并且逐渐向人工智能的威胁狩猎过渡。

3. 威胁情报

前面提到完整的威胁情报体系包括构建自身网络威胁情报、参与外部情报协同，其能够补充监测能力，为威胁狩猎提供更多可能。但是美国国土安全部针对 2017 年和 2018 年信息共享效果的一份发布报告中指出，在美国政府组织的信息共享活动中：

信息并未包含充分缓解潜在威胁的足够细节：共享内容不包含充足的上下文（IP、协议、域名等）或者背景资料；

共享网络威胁信息的参与者数量有限：2017 年，88 名参与者中只有 2 人（1%）共享网络指标。而在 2018 年，252 名参与者中只有 9 人（3%）共享指标。

针对普遍存在的威胁情报共享不理想等情况，非盈利情报共享机构 Cyber Threat Alliance 的 CEO——Michael Daniel 在“Faulty Assumptions: Why Intelligence Sharing Fails”中提出了三个有关信息共享时候的错误假设：

(1) 认为所有的威胁情报（CTI）都是技术数据；

(2) 所有的组织都应该进行相关技术数据分享；

(3) 一旦进行共享渠道，那么分享就是一件非常容易的事情。

图 7 情报共享的误区

他随后纠正了大家的三个错误认知：

【误解一：认为所有的 CTI 都是技术数据】

Michael 将网络威胁情报（CTI）分为四类，包括：技术、战术、运营、战略，涵盖了技术以及非技术的 CTI。根据 Threat Intelligence Sharing: State of the Art and Requirements，这是英国国家基础设施保护中心提出的 CTI 分类，区别于 Gartner 的三层模型，其添加了技术层分类旨在归纳更加短期使用的哈希、IP 地址等信息，其他分类的情报内容也进行了一定调整。

#RSAC | 5

Categories of Cyber Threat Information

Category of Cyber Threat Information	Examples of Information Conveyed	Intended Audience	Decision Example	Timeframe of Use
Technical	Indicators of malicious activity (e.g., malware hashes or IP addresses)	Cyber security vendors and network providers	Should the network security tool allow this packet through?	Immediate
Tactical	Details related to a specific/impending cyber attack	Network defenders (i.e., relevant staff and decision-makers)	Do we need to change a security setting today?	Short term
Operational	Malware types; Attacker tactics, tools and procedures (TTPs)	Senior-level security personnel/ managers	How often should we patch our networks?	Medium Term
Strategic	High-level information on changing cyber risk	Executives/ senior decision-makers	Should we change our risk calculation because a new adversary is targeting our industry?	Long Term

CYBER THREAT ALLIANCE

RSAConference2021

图 8 情报共享分类

每个类别 CTI 在情报内容、情报来源、消费人员、消费途径、情报时效上都有所不同，这样的分类有助于解决更多问题或者进行决策——领导者应该更专注于战略情报（如某政府被认为入侵了拥有直接竞争关系的外国公司）以调整安全资金投入预算，而威胁猎人则应更加关注于战术内容（如 TTPs）以发现恶意攻击行为。

【误解二：所有的组织都应该进行相关技术数据分享】

不同类型的大量 CTI 分享带来一个问题——" 如果每个组织都分享 CTI，那么大家都会被 CTI 淹没”。Michael 认为需要了解组织情报分享两个驱动力：业务模式相关和比较优势。

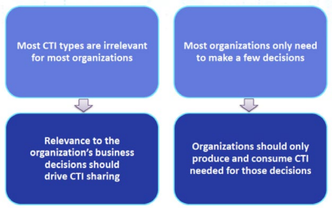
业务相关性表现为获取的情报以及产生的情报，例如教育行业就无法产出和消费制造业产生的工控相关情报，工控相关情报对于教育组织来说就是业务不相关情报。由于各个组织产出的 CTI 不一致，因此需要想清楚是否需要进行协同共享。同时每个组织都应该清楚自己需要什么情报，什么情报有利于你的业务或

者防守策略，如演练期间大家都会需要各种攻击方 IP、工具哈希等 IoC。

只有关注相关威胁信息，才能够有效实现威胁信息降噪，并且看到威胁信息带来的明显价值。

#RSAC | 9

Relevance Drives Utility



CYBER THREAT ALLIANCE

RSAConference2021

图 9 相关性驱动情报分享

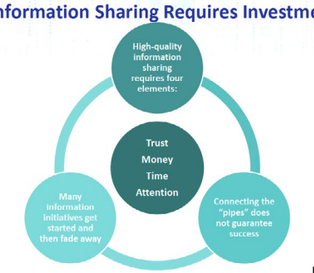
比较优势驱动则是强调需要符合企业自身资金以及技术能力，例如作为一个软件供应商可能没法像互联网大厂一样能够招聘到专门的人员进行 CTI 的运营，更多是消费大厂的 CTI。另外就是需要符合自己属性能力，例如政府就适合发布国家范围的威胁情报。

【误解三：一旦进行共享渠道，那么分享就是一件非常容易的事情】

Michael 认为，高效高质量情报共享需要四要素：信任——建立充分信任才能不保留地共享情报；金钱——用于购买更多情报；时间——投入人员进行情报运营；关注——企业组织由上而下对于情报的关注让共享可持续。

#RSAC | 12

Effective Information Sharing Requires Investment



CYBER THREAT ALLIANCE

RSAConference2021

图 10 有效情报共享投入

除了建立信任以及持续投入，Michael 提出了一些提高效率的假定以及落地措施，例如：

- 减少共享的组织以降低CTI的噪音；
- 标准的分享形式可以提高情报分享的效率，往往能够提供更好的情报。

要让威胁情报体系落地，需要组织中不同消费者的共同努力。组织领导者可以统计事件之间的时间或者检测事件的时间，然后检查他们随着时间变化的情况，可以适当地安排情报分享行为活动，以将这些活动变得更合理；而网络团队则可以决定决策相关 CTI 需求，根据需求进行 CTI 收集和应用；对于第三方 CTI 提供者，供应商应该更加聚焦，而政府则需要更广的 CTI（如上下文）以及激励带领 CTI 共享项目。

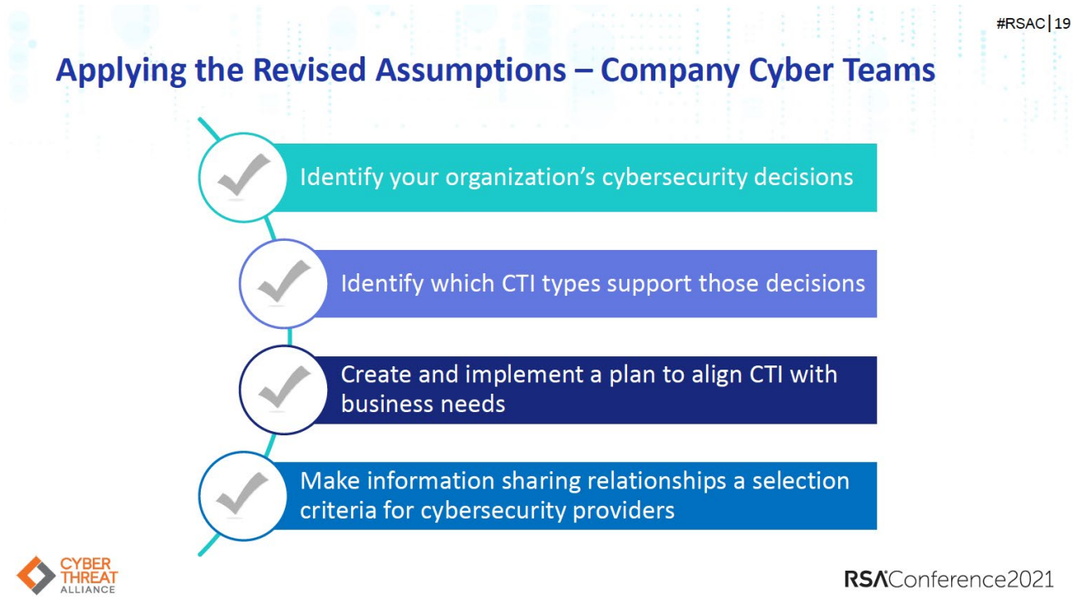


图 11 网络组的行动

4. 实战运营

随着国际形势的变化，在网络安全攻防实战演习的推动下，国内攻防对抗态势逐步升级，激烈强度越发贴近真实对抗。

攻击组织在 0day 数量上、储备上、针对性上优势明显，并随着开源攻击工具的成熟和完善，武器化、自动化成本大幅度下降，钓鱼攻击专业化、专职化，结合 0day 储备进行利用，攻击更加难以防护。另外，随着免杀、伪装、隐匿技术的发展，相关技术应用更加广泛，攻击组织的活动，更加难以监测及捕获。

为应对当前的现状，作为企业和关键信息基础设施的运营单位，“孤军奋战”已不再适应当前的攻防趋势，“情报驱动，主动狩猎，基于实战，归于常态”更能适应当下，甚至未来的态势。

【情报驱动，主动狩猎】

“实战对抗瞬息万变，基于情报的联防联控机制，主动狩猎，提前防控，是应对规模化、武器化、自动化的攻击集团及境外势力的有效手段。”

面对国内当前的攻防态势，在境外攻击组织频繁活动，以及趋向常态化、实战化的各项网络攻防演习及检查之下，各关键信息基础设施运营单位及各大企业以往常规的应对方案已无法彻底解决当前所面临的威胁及挑战，并常年处于疲于奔命的状态下，更难以推动新方案的实践和落地，陷入了恶性循环。

绿盟科技通过多年对安全产品的研发投入、在威胁情报领域的持续积累与沉淀，实战攻防中积累的经验以及培养的专家人才，为威胁狩猎能力的落地夯实了基础。

根据情报来源，我们通过互联网情报 / 流量 / 样本监控、溯源前置方案、SaaS 云监控方案、本地安全事件运营等方式驱动威胁狩猎实战体系：

- **互联网威胁捕获驱动**

关键在于三点：一是建立完善的威胁情报捕获方式与可靠的情报渠道；二是以智能平台为基础结合专家经验对威胁情报进行分级分类与分析；三是基于行业信息、企业资产信息进行威胁情报降噪。

- **溯源前置诱捕驱动**

关键在于两点：一是溯源前置技术创新避免攻击者识别并形成安全产品侧可落地的方案；二是攻击诱捕中结合企业实际环境，实现诱饵合理部署及日常运营增加诱饵的成功诱捕概率。

- **本地安全事件驱动**

关键在于两点：一是建立合理的运营流程，关注内部安全事件的监测与处置；二是合纵连横，建立企业本身威胁情报体系以及情报运营能力，深度利用情报发挥情报价值。

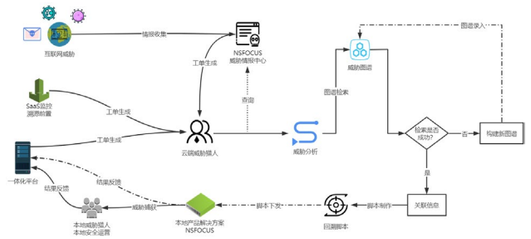


图 12 威胁狩猎能力运转流程

以云威胁猎人与威胁情报库、威胁图谱为云端大脑，进行威胁情报筛选整合与分级分类，并关联监控诱捕数据、项目安全事件数据进行分析、溯源反制，最终形成更完整全面的威胁情报、威胁狩猎脚本进行下发；本地专业威胁猎人以及运营人员作为桥梁，根据资产信息等组织特征调整本地防御策略、依托攻防类安全产品进行本地威胁狩猎，共同完成实现威胁狩猎常态化运转。

【基于实战，归于常态】

“通过运营实现对抗能力的常态化，在实战当中不断推动能力优化。”

多年来，绿盟科技在众多的重大网络安全保障实战以及网络安全攻防演习对抗中，积累了丰富的对抗经验。在每一次的大型网

络安全保障及对抗当中，我们需要同时保障全国数以千计单位的网络安全，涉及关键信息基础设施以及重要资产不计其数。

随着对抗的逐步升级，我们孵化并落地了“常态化保障中台”，基于情报驱动联防联控机制，主动进行威胁狩猎，并实现提前防控，以确保在对抗中占据更有利地位。

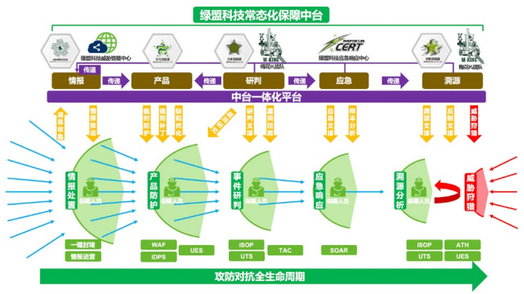


图 13 绿盟科技常态化保障中台

绿盟科技常态化保障中台，以情报驱动为核心，协同研判分析、应急响应、溯源反制、威胁狩猎、产品支持五大能力模块，通过中台一体化平台实现一体化作战体系，以点带面，实现“一点发现，全国闭环”的联防联控。

保障中台通过研判、应急、溯源，从事件中提取攻击手法、特征、行为等，结合产品规则、补丁、方案全国统一下发，完成防护能力闭环。

与此同时，对非法攻击者或攻击组织进行画像，并关联威胁情报库及图谱，通过中台一体化平台向全国下发回溯脚本，以流量回溯、特征回溯、行为回溯等多种方式，完成威胁狩猎，实现全面防护和精准打击。

另外，绿盟科技常态化保障中台通过分析提炼对抗中所产生的大数据，可形成适应实战的解决方案，结合中台所提供的能力，全面推动企业安全建设工作。

参考资料

[1] Hunt and Gather Developing Effective Threat Hunting Techniques.

[2] The Game of Cyber Threat Hunting The Return of the Fun.

[3] Faulty Assumptions Why Intelligence Sharing Fails.

[4] Hunting Threat Actors with Attack Surface Management.

[5] DHS Made Limited Progress to Improve Information. Sharing under the Cybersecurity Act in Calendar Years 2017 and 2018.

[6] Detecting the Unknown: A Guide to Threat Hunting.

[7] Threat Intelligence Sharing: State of the Art and Requirements.

[8] Threat Intelligence: Collecting, Analysing, Evaluating.

[9] NIST Special Publication 800-160 Volume 2.

[10] https://en.wikipedia.org/wiki/Cyber_resilience .

软件定义的原生云安全解决方案

绿盟科技 解决方案中心 杨长茂

摘要：在近三十年的发展历程中，网络安全行业由最初的计算机安全时代，经过了信息安全时代，来到了现如今的网络空间安全时代。而未来，由云计算、5G 通信、人工智能和大数据构建的万物互联的数字世界即将来临。云时代的安全，将是融合云原生环境的安全和具有云属性的安全。

我们处于一个软件和数字化世界：导航软件优化交通出行、外卖改变用餐方式、网约车提升用车体验。借助云计算、大数据和 AI 等技术，各种创新应用正渗入各行各业，促进组织和企业数字化转型，带来新的思维方式和商业模式。同时，网络安全作为数字化基础底座，面临诸多挑战，需要新模式来适应和促进数字化转型。

1. 安全挑战

1.1 云计算应用多样化

经过多年推广，云计算应用已经深入各行业和多种场景中，成为企业数字化转型的基础设施，应用形式多样。在部署方面，数据显示，超过 90% 采用混合云或多云方式。在视频渲染、游戏、工业制造和自动驾驶等场景下，以云计算为主要载体的边缘计算开始兴起。在计算资源方面，DevOps 已向传统行业应用加速，企业会采用虚拟机、容器和容器即服务等，这些计算资源的生存时间更短、自动化程度更高。在多样化的环境下，为信息系统提供一致的安全策略、自动化和敏捷的防护，成为最大挑战。

1.2 网络环境仍在恶化

随着企业将更多的信息系统和数据上云，数据安全成为关注焦点。然而，这些系统和数据不再是孤立的，更多员工和合作伙伴将在不同地方、利用多种设备、以不同方式使用它们，势必会遭受更多攻击。据统计，近五年来，我国零日漏洞收录数量持续走高，年均增长率高达 47.5%；蔓灵花和海莲花等 APT 组织已经将通信、能源和金融等重要领域作为重点攻击对象；而云平台也已成为网络攻击的重灾区。云平台仅靠边界的网络安全设备难以有效应对，需要整体防护和纵深防护。

1.3 安全合规要求增强

随着一系列网络安全相关法律法规和标准规范的发布和实施，

我国安全合规要求逐步增强。除防护能力建设外，企业将需要履行更多职责，如安全监测、监测预警和应急响应。安全已不再是企业个体行为，不同单位之间、与主管单位之间应加强联动联防，实现协同防护。然而，企业内部各部门之间、各系统和各设备之间的联动和协同仍在建设中，如云平台和云上信息系统，信息系统使用云服务时也产生了许多安全数据，如操作云服务、云服务间通信和云服务配置等，并未与安全设备及安全设备的日志数据相结合，未能提升安全防护效果。

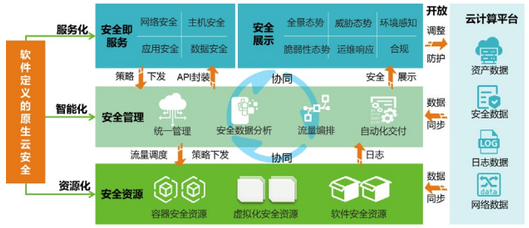
1.4 安全交付模式演变

云计算应用初期，安全设备通过虚拟化快速满足了信息系统的防护需求。当下，企业更加注重云效率，即更好地使用云计算。虚拟化安全设备在订购后部署和登录设备使用的模式，已无法满足云计算的服务模式和使用体验。与此同时，业务需求变化、网络环境恶化和 DevOps 应用，使得安全建设不能仅是一次性的，更应是动态演进的。我们不能仅关注安全建设，更应关注如何发挥好已建安全（人员、技术和流程）的作用，审视环境变化，及时做出调整和安全事件处置，充分发挥云安全价值，保护好云计算环境。

2. 软件定义的原生云安全

面对诸多挑战，云安全应利用新技术赋能安全，建设适合云计算环境的原生云安全。软件定义安全将传统安全设备的安全能力和安全管理分离，实现安全管理集中和安全能力分布执行，与云计算实现异曲同工，有助于新技术应用，提升安全实现效率和防护效果。

绿盟科技将应用软件定义安全理念，把安全能力整合为统一安全资源池，将管理能力汇集到统一平台，通过统一平台与云计算平台融合，实现安全能力的统一管理和扩展、安全数据的统一收集和分析、流量与能力的自动编排和安全事件的自动化响应。最终实现安全的服务化交付、安全能力之间的协同，安全事件的快速发现、快速处置和自动化交付，为云计算平台构建具备弹性、正面、协同和开发特点的原生安全防护。

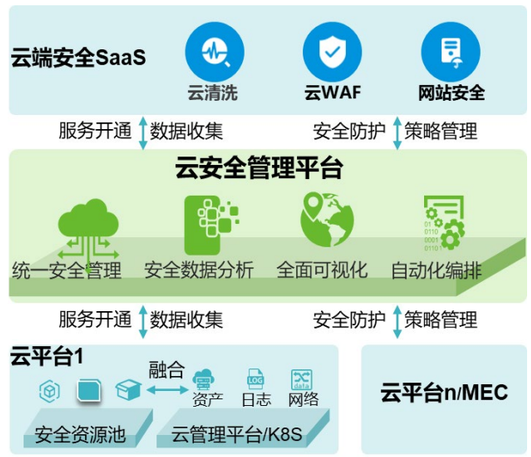


- **软件定义的资源和服务**

借助网络功能虚拟化，构建统一的安全资源池，实现安全的弹性、可扩展和服务化。一方面，将安全能力与安全设备解耦，实现传统安全能力的资源化和服务化，选择即可简便使用。围绕 PPDR，资源池提供覆盖网络、工作负载、应用和数据等关键对象的安全能力，并可随着需求变化，动态添加新能力（如容器安全和零信任等）和弹性扩展处理性能。这些安全能力包括软件、虚拟化和容器等形态，可适用于云计算和边缘计算等多样化环境。另一方面，将各种安全能力作为云安全的安全触角，感知和收集云中安全信息，为进一步安全处理和响应创造可能。

- **软件定义的安全管理**

利用 MANO、大数据安全分析和 SOAR，建立智能的云安全管理平台，实现安全的融合、可见和开放。一方面，云安全管理平台提供统一安全管理界面，实现多云环境和边缘计算，甚至是云端安全 SaaS 的统一订购、统一策略配置和统一交付。通过大数据安全分析能力，统一收集和分析各种安全能力或服务所产生的安全数据，云平台所产生的安全告警，甚至云环境中生成的大量日志和配置数据，建立统一和更全面的安全可见性。根据安全状态，安全管理人员自主选择安全服务和调整安全策略，为多云环境和边缘计算等环境提供统一的安全防护。另一方面，云安全管理平台采用开放架构，可与云平台、DevOps 进行集成和自动化交付，融入云计算环境，成为云或 DevOps 的一部分，也可将安全数据和安全事件提供给第三方平台。



- **软件定义的持续和协同防护**

依托云安全管理平台，开展安全运营，实现持续和协同的安全防护。一方面，利用云安全管理平台的安全分析能力和安全编排与自动化响应能力，可以自动发现和处置安全事件，实现安全持续、安全监控、安全平台与安全能力之间的协同。另一方面，当出现安全事件无法自动处置，安全专家团队（本地和云端）进行研判和分析，确定安全事件、调整安全策略，交由安全设备自动化执行。最终实现安全设备与安全平台之间、本地和云端安全团队之间、安全设备与安全专家团队之间的协同防护，对安全体系进行持续优化和调整。

- 一种模式：各方安全责任共担。在云服务模式下，责任共担已经成为事实标准。云安全建设将从管理、技术和运营等不同维度，提供涵盖方案设计、安全即服务、态势感知、监测预警、应急响应和检测评估等不同类型的安全能力和服务，保护云平台和云上系统/数据。并借助相关流程和机制，帮助企业的云平台建设者和使用者、企业主管单位之间共同承担安全责任和履行安全义务，为云平台构建责任清晰和明确的安全体系，消除安全隐患。

3. 云安全整体方案

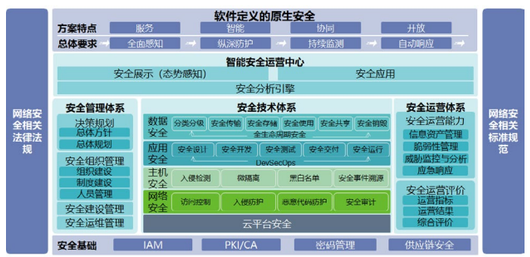
3.1 总体思路

为更好地应对网络安全风险，云安全建设将进一步融合人员、流程和服务等因素，打造体系化和常态化的云安全解决方案。本方案将采用“一个模型、一种模式”的总体设计思路，为云计算构建自适应防护和责任共担的安全防护体系，保障企业数字化转型。

- 一个模型：自适应安全防护模型。为平衡转型机遇和风险，云安全建设将在纵深防护的基础上，依托云安全管理平台的大数据分析技术，建立全面的风险可视化和持续的安全风险评估机制。使用智能研判和SOAR等自动化技术，快速实现响应处置，从而增强风险检测和响应能力，提升检测和响应效率。并利用相关组织建设、流程和评价机制，持续进行提升，最终建立自适应的安全防护体系。

3.2 整体架构

按照我国网络安全相关标准规范和法律法规要求，依据总体设计思路，本方案将采用软件定义的原生云安全理念，以数据为核心，建设“一个基础、三个体系和一个中心”的整体安全架构，构建具有全面感知、纵深防护、持续监测和自动响应的自适应安全防护体系，满足企业用云的各类安全需求，保障数字化转型顺利推进。



- 一个基础：安全基础，为企业云计算安全提供包括身份管理、证书管理和密码管理等在内的基础能力，以及采用满足供应链安全的IT基础设施，共同组成云计算的安全基础。

- 三个体系：安全管理、安全技术和安全运营三个体系。

安全管理体系：根据网络安全相关标准规范和法律法规要求，建立安全管理体系。它为企业提供安全方针和规划，将安全前置到云平台和云上系统的规划和开发等阶段；建立云平台建设部门、运营部门和使用部门，以及主管单位之间相互协作、监督的安全组织架构、制度和流程，推动建设部门和运营部门安全建云和维护云、各使用部门的信息系统和数据安全上云和安全用云。

安全技术体系：依据安全规划，利用安全资源池内的各种安全能力和服务，以数据安全为中心，围绕数据的流转和使用等全流程，构建 PPDR 的安全能力闭环，保护云平台、网络、主机和应用等关键对象，形成安全的通信网络、区域边界和计算环境。最终为云平台、云上系统和数据构建纵深防护，以应对更复杂的网络环境。

安全运营体系：围绕人、流程、安全技术和服务，构建安全运营体系。借助智能安全运营中心和安全技术体系，开展信息资产管理、脆弱性管理、威胁与事件管理和应急响应等安全运营服务，帮助企业开展安全预警、安全评估和响应处置等工作。采用各类安全指标，对安全运营工作进行考核，以及评估和分析安全管理体系和安全技术体系，提出具体的优化建议，实现云安全的持续优化和提升。

- 一个中心：智能安全运营中心。利用云安全管理平台，建立本地智能安全运营中心，统一收集和处理各类安全数据，进行持续的安全检测和分析，可视化展示安全态势。根据安全态势，安全管理人员主动并持续地调整和优化安全防护，为纵深防护赋予主动和自适应防护能力，更好地应对复杂的网络环境。

通过开放接口，云安全管理平台可将监测到的安全事件上报给主管单位或同步给其他企业，以及接收主管单位的预警通告，进行集中的展示和预警，实现安全风险的提前预防。最终形成主管单位与企业、企业与企业之间的安全联动和联防。

4. 结语

作为数字化转型的基础，云计算应用将会更加深入和多样化，其面对的网络环境也更加复杂。云安全建设应该与云计算相适应和融合，采用软件定义的原生云安全既可以应对云计算带来的安全风险，又可以更好地保护云计算环境，为企业数字化转型打下良好的安全基础。现在，绿盟科技仍在开展云安全研究，如服务网格和 Serverless 等，未来将会应用到云安全解决方案中，持续为云计算安全带来新能力，促进云计算安全应用。

参考文献

[1] 2019 年我国互联网网络安全态势综述 .
[2] The State of Cloud Native Security 2020.
[3] 云计算发展白皮书（2020 年）.
[4] 软件定义安全白皮书（2016）.

基于HTTP 响应信息降维可视化的资产特征分析

绿盟科技 合规安全技术部 张卓 张迎苹 吴磊

1. 背景简介

网络资产探测指的是对网络资产情况进行追踪、掌握的过程，其通常包括主机发现、操作系统识别以及服务应用识别等内容。从网络安全的角度来讲，网络资产探测能够为统一软硬件版本、更新升级软件和设备等工作提供基础信息，是进行网络安全监控、漏洞扫描上报、威胁态势感知等网络安全管理活动的重要前提。

从技术起源上讲，目前广泛使用的新型网络资产探测技术可以分为主动探测、被动探测以及基于搜索引擎的非入侵探测三种技术，其中被动探测方法是指采集目标网络的流量，对流量中应用层 HTTP、FTP、SMTP 等协议数据包中的特殊字段 Banner 或 IP、TCP 三次握手、DHCP 等协议数据包的指纹特征进行分析，从而实现对网络资产信息的被动探测。

在被动探测技术当中，基于 HTTP 协议进行 Web 指纹构建，对软件服务、操作系统进行识别是比较常用的一种资产探测与识别技术。而基于 HTTP 协议进行网络资产探测与识别的方法，又分为基于服务标识（Banner）的识别方法、基于头部字段顺序差异与语法差异的识别方法以及基于处理方式差异的识别方法等。其中基于 HTTP Banner 信息的识别方法是目前使用较为广泛，且简单有效的一种识别方法。通过 Banner 信息，按照 HTTP 协议规定，提取诸如 Server、User-Agent、Authorizatio 等 Response 头部字段，依据这些字段，我们能够轻易探测获得目标 Web 服务器的软件类型，甚至可以精确获取到版本信息及操作系统信息。但是由于 Banner 可以进行人为修改、伪装或模糊，这使得单纯基于 Banner 的字段信息进行资产探测与识别的准确率无法得到保证。

本文基于 HTTP 的 Response 信息，从 Response Header 与 Response 正文中提取了 Server 字段、Title 字段、Authenticate 字段、Content-length 字段和 Status 字段五个特征字段，然后对 HTTP 的 Response 信息利用 BOW 模型与 TFIDF 模型进行文本向量化建模，利用 TSNE 算法对高维文本向量进行可视化降维。依据 TSNE 降维的可视化结果，对比分析提取的五个特征字段在表征资产类别、刻画资产特征方面的有效性。最后，基于 BOW 向量使用文本聚类方法验证聚类结果与特征字段所刻画的簇类差异之间的一致性。

通过本次的验证分析，一方面验证了 Banner 特征字段所包含资产信息的有效性，另一方面验证了相应文本建模技术在资产特征刻画方面的有效性，为后续进一步利用文本挖掘、NLP 等技术对基于 HTTP 协议进行文本建模的网络资产自动探测识别技术研究奠定了基础，验证了相关技术落地的可行性。

2. 相关技术介绍

2.1 文本向量化技术选择

文本向量化即将一个文本表示为一个向量的方法，其可以使用一个统一的向量空间模型（Vector Space Model，VSM）进行表示。所谓 VSM 模型，基本思想是将文本表示为向量空间当中的一个点，在进行文本的 VSM 构建时，有两个要点：

（1） 如何确定向量的维度？

也就是说，将一个文档表示为一个向量时，向量的每个维度表示什么？目前主要有如下定义：

- 在VSM的原始定义中，并未限制每个维度的含义。
- 如果将一个维度定义为一个词、一个短语，或者特定的关键词等，在这种情况下每个维度是有具体含义的，比较典型的是BOW模型、TF-IDF模型；
- 不对每个维度进行具体的定义，进行模型自适应训练获得一些词、句的向量化表示，比较典型的是word2vec词嵌入模型。

（2） 如何确定每个维度的取值？

即在确定了每个维度的意义后，每个维度的取值如何确定？实际上取值方式的不同也导致模型的不同，目前主要使用如下方式：

- 以0/1取值表示相应的维度（词、短语等）是否在该文本中存在，即BOW模型与Bit Vector的组合；
- 对每个维度代表的词项进行词频统计，以词频作为维度取值，如最常用的BOW模型；
- 不仅考虑单个文本中的词频，而且考虑整个文本集合当中词项出现的频率大小，通过对每个词项进行重要性评分，获得每个维度的取值，如TF-IDF模型；
- 上述表达方式都忽略了词序、词义对文本向量化的差异表达，因此忽略了每个维度的具体含义，基于词的上下文语义关系为每个词自适应训练一个词向量，进而构成文本的词向量，也就是word2vec的训练方式。

在 VSM 框架下，BOW 不会对文本中的任何词进行编码，同时会忽略文本中的单词的次序，忽略所有语法，是对文本的整体建模，以单词出现的一致性作为文本相似性的判断基准。而基于 BOW 的 TF-IDF 则对单词与文档的相关度进行了编码，引入了逆文档频率来识别每个文本当中所具有的独特的、对本文重要的单词。而基于主题（Topic）分布假设的 LSA（Latent Semantic Analysis，潜在语义分析）、PLSA（Probabilistic Latent Semantic Analysis，概率潜在语义分析）以及 LDA（Latent Dirichlet Allocation，潜在狄利克雷分配）模型，则通过基于共现矩阵的矩阵分解或者基于贝叶斯的变分推断方法，获得了文本在主题空间的向量表示。进一步，以 word2Vec 为代表的基于浅层神经网络的词向量，在充分考虑词的上下文环境的基础上，表征了词的语义相似性，比上述模型更适合复杂的文本建模任务。

为简化分析过程，本文主要选择了 BOW 模型和 TF-IDF 模型作为文本向量化的分析模型。

2.2 特征降维技术选择

利用 BOW 模型对大量文本进行建模时，向量维度很容易就能达到上千维度乃至上万维度，对此可以使用 LSA 或者 LDA 模型、使用较低维度的主题空间对文本进行刻画，但这样会丢失一定的

信息。因此本文选择了一些特征降维技术，对原始的 BOW 文本向量进行降维，力求在保证减少信息损失的前提下获得文本向量在较低维度的表达，以二维可视化的方法直观观测文本在向量空间中的分布。

目前使用比较广泛的特征降维技术有 PCA、EFA、ICA、SVD、LDA（线性判别分析）等，但这些相对主流的方法无法维持在低维空间文本向量之间的分布关系，仅仅从方差解释或者特征分解的角度，找出变换特征。

而为了保证高维空间中文本向量之间的分布关系在低维空间中尽可能保留，本文选择使用 TSNE（t-Distributed Stochastic Neighbor Embedding）算法，尽最大可能保留了原始高维空间中的文本局部的相似性不被破坏，在二维空间原样呈现出文本向量之间的距离分布关系。这对后续采用的基于距离尺度的文本聚类分析，具有十分好的验证效果。

2.3 聚类算法选择

聚类分析是无监督机器学习算法当中的主体，是基于样本之间具有的聚类关系对原始样本进行无监督簇类划分的方法。目前广泛使用的聚类算法主要有基于距离的层次聚类、Kmean 聚类，基于密度的 DBSCAN 聚类算法、OPTICS 算法以及基于网格的聚类算法等。

其中，Kmean 算法是一种简单且有效的算法，但其聚类结果受初始聚类中心的选择影响较大，且对于分布复杂的簇类情况很难进行识别；DBSCAN 算法则基于样本的密度可达及样本点之间的联通关系，能够有效发现复杂的簇类分布状况，但对于样本密度不同的簇类其发现能力较差；基于 DBSCAN 改进的 OPTICS 算法依照数据点的密度可达聚类对所有样本进行排序，进而能够发现不同样本密度的簇类，但其计算复杂度相对较高，在本次实验中算法计算始终很难收敛。GMM 算法基于样本的高斯分布假设，通过 EM 算法能够很好发现样本中的簇类分布信息，但其聚类效果同样受初始簇类参数影响较大。

本文主要选择 Kmean 算法、DBSCAN 算法与 GMM 算法对文本进行聚类分析，并结合 TSNE 算法进行 2D 可视化，与相应特征所刻画的簇类情况进行对比验证。

3. 实验

3.1 技术处理路线

本次实验选取了 NTI 平台的 10 万条 HTTP 协议数据，受限于数据质量，主要从 HTTP Banner 信息中抽取了 Response Header 中的 Server 字段、Authenticate 字段、Content-length 字段、Status_code 字段，以及 HTML 文本中的 Title 字段、手动构建的全部 Banner 文本的单词长度 Banner_len。

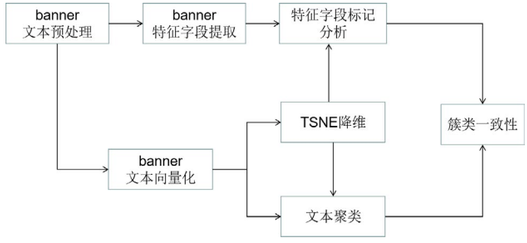


图 1 技术处理路线

如图 1 所示，本实验按照如下步骤进行分析。

（1）文本预处理：对原始的 Banner 文本进行分词、停用词删除、特殊字符移除、大小写转换等文本预处理；

（2）特征字段提取：从 Banner 文本中提取构建 Server、Authenticate、Content-length、Status_code、Title、Banner_len 等特征；

（3）文本向量化：使用 BOW 模型与 TF-IDF 模型对 Banner 文本进行向量化处理；

（4）降维可视化：使用 TSNE 模型对文本 - 单词矩阵进行降维，在 2D 空间对文本向量进行可视化；

（5）特征字段分析：将 Banner 中提取获得的 Server、Authenticate、Title、Status 字段分布作为分类响应，在 TSNE 刻画的 2D 空间中对其进行不同取值的可视化分析，与整体可视化进行对比，验证特征字段对 Banner 信息簇类刻画的有效性；

（6）文本聚类：分别使用 Kmean 算法、DBSCAN 算法与 GMM 算法对原始的文本向量进行聚类，获取每个文本的聚类标签；

（7）特征字段与文本聚类的簇类一致性分析：在 2D 降维空间内对不同簇类进行可视化，对比聚类算法的聚类结果与特征字段所刻画簇类之间的一致性，验证聚类算法在该问题场景下的可行性。

3.2 基于 TSNE 的降维可视化

10 万条样本在经过 BOW 及 TF-IDF 模型向量化后，每个文本都被转换为一个 2000 维的向量。显然，当样本达到百万级后，转换后的向量维度会更高，因此使用 TSNE 算法对样本进行降维后，两种模型所刻画的文本空间分布如图 2 所示。

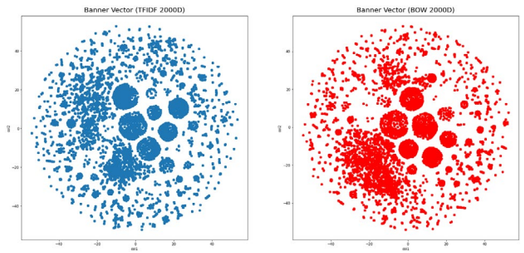


图 2 文本向量化降维可视图

对比来看，基于 BOW 与 TF-IDF 的向量化文本在 2D 空间中均具有明显的簇类分布特征，其差异如图 3 所示。

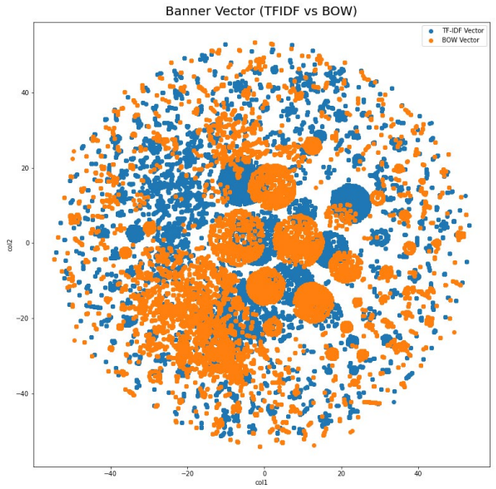


图 3 两种模型的向量化差异

总体来讲，两种模型的文本向量化差异不大，后续的特征字段分析将采用 BOW 向量化文本数据进行处理。

3.3 特征字段的可视化对比

不同的特征字段当中可能还有丰富的网络资产信息，具有相同取值的特征字段直观上讲被作为同一类资产。

而为了探究含有相同取值的字段在实际文本向量空间中的分布是否具有明显的簇类分布规律，我们主要对 Server、Authenticate、Title、Status 等字段进行了分析。

（1） Server 字段

在 10 万条数据当中，含有 Server 字段的约为 5 万条，其字段取值分布频次统计如图 4 所示（取频次大于 100 的取值）。

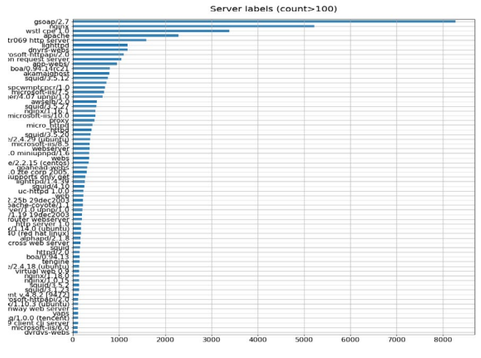


图 4 Server 字段取值分布

我们选取前 9 个簇类进行可视化，其与全部样本的簇类分布对比如图 5 所示。

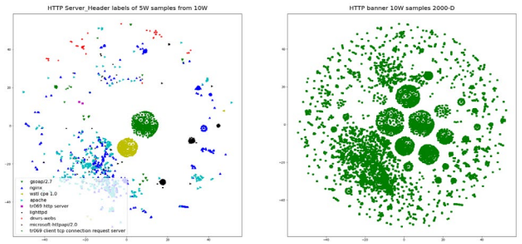


图 5 Server 字段簇类分布对比

显然，对于取值为 gsoap/2.7 与 wstl cpe 1.0 的资产，其分布簇类明显，在原始分布中具有明显的对应簇类。图 6 是包含与不包含 Server 字段之间的簇类分布对比图，显然两者分布具有较大差异。

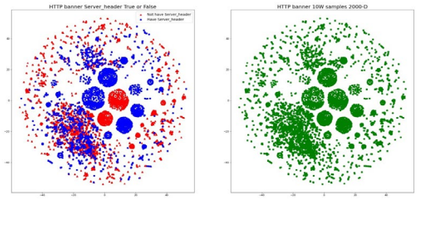


图 6 包含 Server 字段与不包含该字段的差异对比

（2）Status 字段

即对 HTTP 响应的状态码进行分析，发现其也具有明显的簇类分布特性，这里选取了 HTTP/1.1 协议的 200、401、404 状态码以及所有的 400 特征码进行分析，其簇类分布结果如图 7 所示。

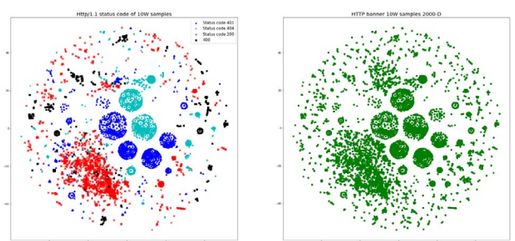


图 7 状态码特征簇类分布对比

（3）Title 字段

Title 字段不属于 HTTP 的响应头字段，是响应正文中 HTML 格式文本的 Title 属性字段，在一些情况下，该字段也包含了丰富的资产信息。就当前样本集分析来看，不存在 Title 字段的为 47418 条样本，而在大于 1000 条统计取值的 34440 条样本中绝大多数信息也是与状态码相关的信息。但这也给了我们一个方向，即关注取值频次较少的字段取值，这也可能含有别于其他资产的特异性文本信息。

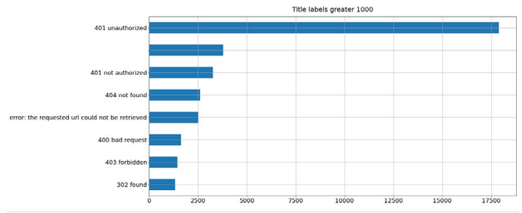


图 8 Title 字段取值分布

对 Title 字段当中的信息按照包含信息、不包含信息以及包含信息但取值为空的类别进行可视化簇类对比，其结果如图 9 所示。

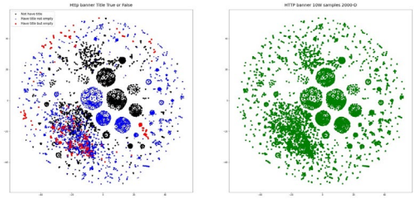


图 9 Title 字段取值簇类分布对比

（4）Authenticate 字段

Authenticate 包含 Web 访问的认证信息，分为 WWW-Authenticate 与 Proxy-Authenticate 两种字段。在 10 万条数据当中，存在该字段信息的有 34142 条样本，对其中又有 15011 条为空字符，对其剩余字段取值进行簇类分析，取值分布状况如图 10 所示，其含有丰富的资产信息。

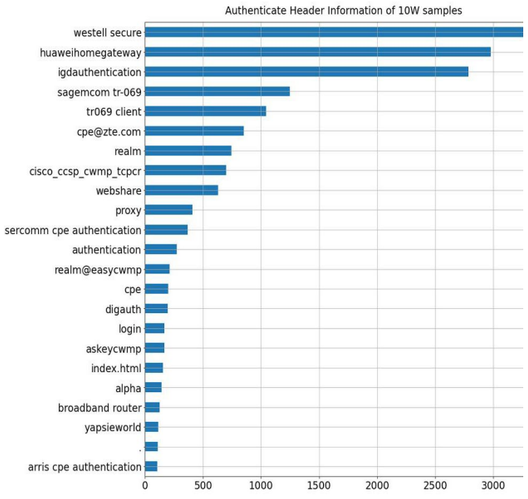


图 10 Authenticate 字段取值分布

其与原始簇类的分布对比如图 11 所示，显然取值具有明显的簇类分布特征。

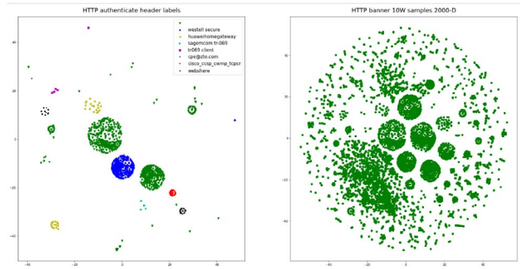


图 11 Authenticate 字段簇类分布对比图

3.4 聚类结果验证

通过上述特征字段的簇类分布分析发现，一些特征字段对网络资产特征的刻画具有明显的差异，基于 BOW 模型的文本向量化可以有效表征不同字段取值的差异。为此，我们采用了基于距离尺度的 Kmean 聚类、DBSCAN 聚类算法，并使用 GMM（混合高斯聚类）模型作为对比，其结果如图 12、图 13 所示。

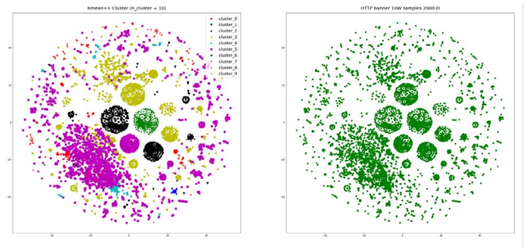


图 12 Kmean 聚类结果与原始簇类分布对比

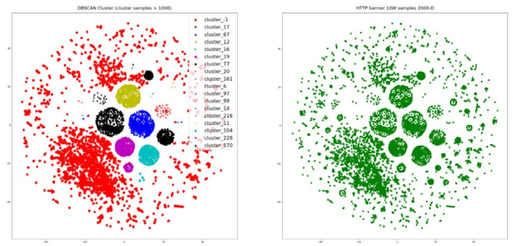


图 13 DBSCAN 聚类结果与原始簇类分布对比

对比发现，两种聚类算法都能发现较为明显的簇类样本，但在细节上 DBSCAN 趋向于发现具有同一密度的簇类，对于密度较低的簇类，较容易判为噪声点。相比之下，GMM 算法可以将低维空间紧邻的两个簇类识别标识为一个簇类，其簇类分布与之前特征字段的簇类分布具有较高的一致性，这是由于 TSNE 算法下，低维空间中数据的 T 分布近似高维空间中的样本的高斯分布距离，低维中较近的两个簇类对应到高维当中将使簇类中心很容易受到边界点的干扰。

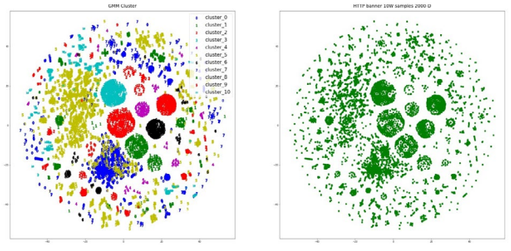


图 14 GMM 聚类结果与原始簇类分布对比

不过这些缺陷对目前的粗略分析影响不大，我们依旧可以通过多种算法聚类结果与特征字段分布簇类的对比，得到一些技术分析的指导信息。

4. 结语及下一步工作方向

总体来讲，目前的工作主要有如下结论：

（1）基于 BOW 模型及 TF-IDF 变换的 HTTP 文本向量化，能够有效地捕捉到文本当中包含的资产特征，通过几种聚类算法的自动聚类结果分析，可以验证聚类结果与不同特征字段取值在向量空间中簇类分布的一致性。

（2）侧面验证了特征字段本身包含的资产信息的有效性，可以将基于 HTTP 协议字段提取的资产信息作为 HTTP 文本算法分析的辅助标注信息，进一步提升有监督算法训练的有效性。

当然，当前从特征字段当中获取的信息还太过粗糙，基于聚类获得的不同簇类结果还需要进行大量的人工分析与筛选之后，才能确定是否发现了新的资产指纹。这其中存在以下问题。

（1）进行有监督算法学习时，如何确保标注信息的有效性：即便使用最粗糙的二分类判断，确保标签信息本身与文本内容的信息一致性依旧是十分大的挑战。

（2）如果只采用有确定标注信息的数据进行训练，目前可考虑的技术路线有两种：一是只基于有效标注信息训

练异常检查算法，对无效文本进行过滤识别，但这种方法很难发现新的资产信息；二是为每类资产指纹训练单独的识别模型，但由于不同的资产指纹可能覆盖了多种不同的乃至相互冲突的文本向量模式，因此模型的构建变得十分复杂。

以上这些问题需要我们探索更多的解决方法，以提高我们在资产发现与识别方面的能力。下一步，我们将首先使用 word2vec、textCNN 等相对更复杂一些的方法进行基于 HTTP 文本的资产探测与自动发现的技术研究。

5. 致谢

本次实验分析参考了桑鸿庆、张胜军等同事关于物联网设备的资产聚类发现经验，相关工作得到了系统架构部李瀛的支持，在此一并表示感谢。

参考文献

[1] 王宸东 , 郭渊博 , 甄帅辉 , 等 . 网络资产探测技术研究 [J]. 计算机科学 , 2018, 45(12):31-38.

[2] 易运晖 , 刘海峰 , 朱振显 . 基于决策树的被动操作系统识别技术研究 [J]. 计算机科学 , 2016, 43(08):79-83.

[3] 刘翔元 . 基于网络流量分析的网络设备类型识别关键技术研究 [D]. 南京邮电大学，2020.

[4] 赵建军 . 网络空间终端设备识别技术研究 [D]. 兰州理工大学 , 2016.

[5] 曹来成 , 赵建军 , 崔翔 , 等 . 基于余弦测度下 K-means 的网络空间终端设备识别 [J]. 中国科学院大学学报 , 2016, 33(04):562-569.

[6]Shah S. An introduction to HTTP fingerprinting[J]. Net-Square Solutions ,2004: 1-21.

[7]AHMED ABDOAZIZ AHMED ABDULLA. 击败 HTTP 指纹识别技术 [D]. 吉林大学，2012.

[8]https://blog.csdn.net/flying_all/article/details/77152409.

[9]http://lvdmaaten.github.io/publications/papers/JMLR_2008.pdf.

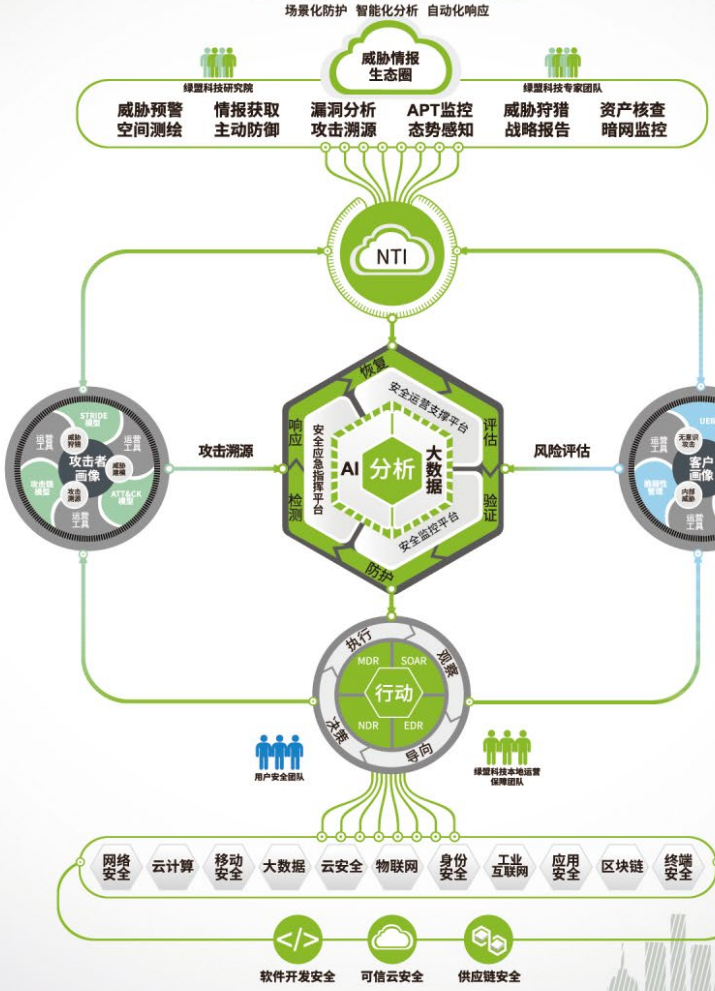
[10]https://mp.weixin.qq.com/s?__biz=MzIyODYzNTU2OA==&mid=2247488749&idx=1&sn=48f410d9cdd71d9db5d43d47c9be0184&chksm=e84fb232df383b2487332bbeec14a505ef53aa91679ca0d36ffa64670571c8def0f169bfa07&token=1483816502&lang=zh_CN&version=3.1.0.3004&platform=win&scene=21#wechat_redirect.

智慧安全3.0

NSFOCUS 绿盟科技

全场景 · 可信任 · 实战化

场景化防护 智能化分析 自动化响应



THE EXPERT
BEHIND GIANTS
巨人背后的专家

多年以来，绿盟科技致力于安全攻防的研究，
为政府、金融、运营商、能源、交通、科教文卫等行业用户和各类型企业用户，
提供具有核心竞争力的安全产品及解决方案，帮助客户实现业务的安全顺畅运行。
在这些巨人的后面，他们是备受信赖的专家。

客户支持热线：400-818-6868

NSFOCUS 绿盟科技