HZ BOOKS
华章IT

网络空间安全
技术丛书

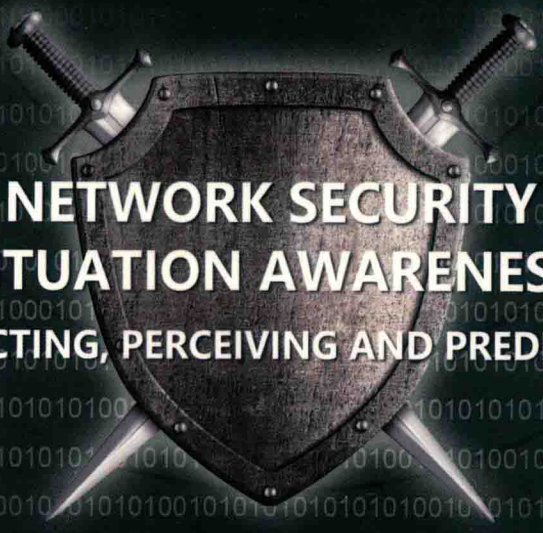
由业内安全专家倾力打造，系统讲解网络安全态势感知的技术、方法和工具。

从网络安全态势提取，到态势的理解和预测，有效识别网络入侵和潜在安全威胁。

网络安全态势感知

提取、理解和预测

杜嘉薇 周颖 郭荣华 索国伟 编著



NETWORK SECURITY
SITUATION AWARENESS
COLLECTING, PERCEIVING AND PREDICTING

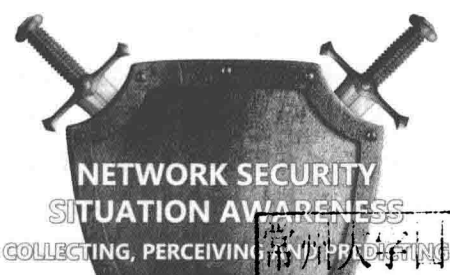
机械工业出版社
China Machine Press

网络空间安全
技术丛书

网络安全态势感知

提取、理解和预测

杜嘉薇 周颖 郭荣华 索国伟 编著



机械工业出版社
China Machine Press

图书在版编目（CIP）数据

网络安全态势感知：提取、理解和预测 / 杜嘉薇等编著 . —北京：机械工业出版社，2018.7
（网络空间安全技术丛书）

ISBN 978-7-111-60375-7

I. 网… II. 杜… III. 计算机网络 - 网络安全 - 研究 IV. TP393.08

中国版本图书馆 CIP 数据核字（2018）第 138240 号

网络安全态势感知：提取、理解和预测

出版发行：机械工业出版社（北京市西城区百万庄大街 22 号 邮政编码：100037）
责任编辑：佘 洁 责任校对：殷 虹
印 刷：中国电影出版社印刷厂 版 次：2018 年 7 月第 1 版第 1 次印刷
开 本：186mm × 240mm 1/16 印 张：18.5
书 号：ISBN 978-7-111-60375-7 定 价：69.00 元

凡购本书，如有缺页、倒页、脱页，由本社发行部调换
客服热线：（010）88379426 88361066 投稿热线：（010）88379604
购书热线：（010）68326294 88379649 68995259 读者信箱：hzit@hzbook.com

前　言

网络安全既涉及国家安全也涉及经济安全，目前世界各国每天都在进行着大量隐蔽的较量，网络安全的重要性不容忽视。5年前，我因偶然机遇进入网络空间安全领域，通过各种活动和项目实践见证了国家对网络安全重视程度的不断提升。这期间先后研究过威胁建模、信息安全风险评估与控制、流量检测和安全测试等方面，在持续学习和实践的同时，也与国内许多专家学者和企事业单位频繁接触，深感安全人员不能只关注具体点而忽略整体面。花大价钱购置一大批盒式设备“堆”在网络中的时代已经一去不复返，这只能带来虚假的安全感。安全的“短板效应”和“木桶理论”决定了我们必须以全局整体的视角去看待它，而且这种整体视角是基于动态博弈的暂时平衡。

网络安全的哲学已经从“努力防住”转变为“防范终将失效”，从“发现并修补漏洞”转变为“持续过程监控”，也就是说，无论你在网络和系统中投入多少，入侵者仍可能获胜。基于这个哲学前提，我们能做的就是在入侵者实现目标前尽可能地发现、识别并做出响应，及时分析情况和通报事件的发生，并以最小代价减轻入侵者的破坏，只要入侵者的目标未能达成即是相对安全的。网络安全态势感知就是这种思路的典型体现，通过获取海量数据与事件，直观、动态、全面、细粒度地提取各类网络攻击行为，并对其进行理解、分析、预测以及可视化，从而实现态势感知。它有助于安全团队发现传统安全平台和设备未能监测到的事件，将网络上似乎无关的事件关联起来，从而更有效地排查安全事件并做出响应。

虽然网络安全态势感知的概念早在若干年前就已被提及，但由于当时的技术和认知水平，其发展是有限的。随着时间的推移，以及数据量和数据形态的改变，老问题发生了新变化，需要我们重新审视它。尤其是近几年来，随着大数据技术的迅猛发展、数据处理和分析方法的不断创新，以大数据为平台框架进行安全分析（即数据驱动安全）逐渐成为热点。新技术的驱动给网络安全带来许多新的挑战，也迫使我们重新思考。只有不断更新知识，创造性地发现问题、研究问题和解决问题，才能跟上信息化时代的脚步。正如《人类简史》中所写：“人类近500年的科学革命意义重大，它并不是‘知识的革命’，而是‘无知的革命’。真正让科学革命起步的伟大发现，就是发现‘人类对于最重要的问题其实毫无所知’……现代科学愿意承认自己的无知，就让它比所有先前的知识体系更具活力、更有弹性，也更有求知欲。这一点大幅提升了人类理解世界如何运作的能力，以及创造新科技的能力。”

读者对象

本书的读者对象主要包括：

- 网络安全领域的技术爱好者和学生
- 网络运维管理、信息安全领域的从业人员
- 网络空间安全等相关专业的本科生及研究生
- 期望在网络安全领域就业的技术人员
- 网络安全态势感知领域的研究人员

本书结构

本书分为四个部分：基础知识、态势提取、态势理解和态势预测。每章都会重点讨论相关理论、工具、技术和核心领域流程。我们将尽可能用通俗易懂的方式进行阐述，让新手和安全专家都能从中获得一些启发。本书所构建的框架和理论基于集体研究、经验以及合著者的观点，对于不同的话题和场景所给出的结论可能与他人不同。这是因为网络安全态势感知更多地是一门实践活动，不同人的认知和理解难以趋同，这也是完全正常的现象。

本书的内容框架如下：

第一部分：基础知识
　第 1 章：开启网络安全态势感知的旅程
　第 2 章：大数据平台和技术
第二部分：态势提取
　第 3 章：网络安全数据范围
　第 4 章：网络安全数据采集
　第 5 章：网络安全数据预处理
第三部分：态势理解
　第 6 章：网络安全检测与分析
　第 7 章：网络安全态势指标构建
　第 8 章：网络安全态势评估
　第 9 章：网络安全态势可视化
第四部分：态势预测
　第 10 章：典型的网络安全态势预测方法
　第 11 章：网络安全态势智能预测
　第 12 章：其他

本书涉及的网络安全态势感知主题众多，我们希望书中各章涵盖的内容能够具有一定的参考价值；同时希望读者能够获得愉悦且充沛的阅读体验，就如同我们在字斟句酌数易其稿、亲历从最开始寥寥数页的章节意向到成稿付梓形成手头这本书的过程中体会到的一样。

致谢

写书的过程是漫长而艰辛的！我有个习惯，就是无论在何种环境下都会不断告诫自己："人不能贪图安逸，总要有一个目标，时不时给自己一些挑战，做一些有难度的事情。"从设定这样一个目标到谋划这件事情，再到搭建书的整体框架，对每个章节进行布局，完成初稿、中间修改和定稿的整个过程中，是心中的信念让我克服了人固有的惰性并坚持了下来。

当然，本书之所以能完成，离不开许多朋友直接或间接的帮助，我也想借此机会感谢他们。

感谢父母，在你们的影响下成长，使我成为一个独特的人。作为子女所能做的是，传承他们赋予的性格并分享他们给予的爱。

感谢我的家庭和可爱的女儿，家人给我的爱是浓厚的，对我非常重要，他们在生活中对我的关心和支持，让我有动力完成各种艰难的挑战。

感谢单位的领导和同事，他们给予我充分的信任和支持以开展这方面的研究和实践，并对我的研究工作提出了诸多宝贵意见和建议。

杜嘉薇

目　录

第二部分 态势提取

第一部分

基础知识

第 1 章

开启网络安全态势感知的旅程

我们的生活样式就像一幅油画，从近看，看不出所以然来，要欣赏它的美，就非站远一点不可。

——亚瑟·叔本华，德国哲学家

1.1 引言

网络空间实在太宽泛，包罗万象，而且已经发展得异常复杂，远超人类直觉所能感知的范围，每天流经网络上的比特数比全世界所有海滩上的沙子还要多。过去的十年间，我们目睹了计算能力的指数级增长和各种计算设备的爆炸式应用，IT 基础设施正在发生深刻变化，虚拟化技术、软件定义网络、移动互联网技术逐渐从概念走向实际应用，云计算的兴起、BYOD[⊖]的普及改变了传统的数据中心架构和人们的工作方式，使得传统的网络边界变得模糊甚至消失，这给传统的、以安全边界为核心的防护思想和安全产品带来了巨大的挑战。与此同时，非法利用和破坏信息系统也发展成为有组织的犯罪行为和敌对国家的活动。网络攻击的实施者不再是个人，而是有着明确政治、经济利益目的的“黑产”组织、国家机构等，攻击的手段和工具也日新月异（“零日漏洞”已成为网络空间地下黑市的抢手货）。网络空间威胁已经呈现出集团化、工具化、流程化的趋势，这给传统的以检测为核心的防御手段带来了巨大挑战。

过去，人们更多地依靠安全分析员的经验和安全工具来感知和分析网络的安全状态，然而，安全分析员所拥有的知识量有限，各种安全工具也都有短板。现如今，面对网络空间安全形势所带来的挑战，在强大的计算机和数据分析平台的支持下，我们希望网络安全态势感知能改变这一局面。借助新型网络安全态势感知技术，可更全面地了解当前网络安

⊖ BYOD（Bring Your Own Device）指不受时间、地点、人员、网络环境等限制，携带自己的设备办公，这些设备包括个人计算机、手机、平板电脑等。

全状态，预测其发展趋势并做出有效规划和响应，更高效、更科学地检验和支撑人的直觉观点，保护如今日益庞大和复杂的基础设施系统。

网络安全态势感知本质上就是获取并理解大量网络安全数据，判断当前整体安全状态并预测短期未来趋势。总体而言，其可分为三个阶段：态势提取、态势理解和态势预测。其中，态势提取至少包含通过收集相关的信息素材，对当前状态进行识别和确认；态势理解至少包含了解攻击造成的影响、攻击者的行为意图以及当前态势发生的原因和方式；态势预测则包含跟踪态势的演化方式，以及评估当前态势的发展趋势，预测攻击者将来可能采取的行动路径。虽然我们的理想状况是可以在没有人工干预的情况下进行自动化感知和防御，但目前的技术发展还未能达到如此智能化的水平。也许，随着新技术的发展和人工智能的革新，未来有一天真的能够实现这个愿景。但在目前，我们所研究的网络安全态势感知系统仍是硬件设备、计算软件和人类思维决策的共同组成体。

本书试图提供当今网络安全态势感知中重要主题的概览。每一章都将着重阐述网络安全态势感知的某个方面，并讨论网络安全人员进行态势感知所采用的理论、方法和技术。尽管每个主题都可以作为一个方向扩展出丰富的内容，甚至写出一本书，但我们仍然只是对它们进行概述，这样做的目的是想为读者进一步深造提供一个良好的起点，希望本书能够激起读者进一步探究网络安全态势感知领域的兴趣。

本章将带领读者开启通往网络安全态势感知世界的旅程，我们将从网络安全简史谈起，然后引入核心术语和相关模型，从宏观上介绍网络安全态势感知的产生背景和基础知识。同样不能忽视的是各国在网络安全领域进行的认知和实践活动，因此本章还会谈到我国网络安全态势感知相关的政策和发展历程，以及以美国为代表的先进国家在网络安全方面的实践经验，最后给出一些系统建设方面的意见和建议。

1.2　网络安全简史

不了解过去就难以理解未来。在探讨网络安全态势感知这一主题之前，先来看一下网络安全的简缩版发展史。

1.2.1　计算机网络

计算机网络从 20 世纪 60 年代发展至今，已经形成从小型的办公局域网络到全球性广域网的规模，对现代人类的生产、经济、生活等方方面面都产生了巨大的影响。1962 年，由美国国防部（DOD）资助、国防部高级研究计划局（ARPA）主持研究建立了数据包交换计算机网络 ARPANET。ARPANET 利用租用的通信线路，将美国加州大学洛杉矶分校、加州大学圣巴巴拉分校、斯坦福大学和犹他大学四个节点的计算机连接起来，构成了专门完

成主机间通信任务的通信子网。该网络采用分组交换技术传送信息，这种技术能够保证四所大学间的网络不会因为某条线路被切断，而影响其他线路间的通信。当时的人们根本想不到，20 年后计算机网络在现代信息社会中会扮演如此重要的角色。ARPANET 已从最初四个节点发展成为横跨全世界一百多个国家和地区，挂接数万个网络、数千万台计算机、数亿用户的互联网（Internet）[⊖]。由 ARPANET 发展而来的 Internet，是目前全世界最大的国际型计算机互联网络，目前仍在快速发展中。

1.2.2　恶意代码

网络的发展给人们带来了诸多便利和好处，然而也带来了恶意行为的肆虐，这就是下面要讲的恶意代码。北京邮电大学的杨义先教授说："如果说普通代码是佛，那么恶意代码就是魔。佛与魔除了分别代表正义与邪恶之外，其他本领其实都不相上下。"从恶意代码的作恶能力来看，只有你想不到的，没有它做不到的。

对于 2010 年席卷全球工业界的"震网"病毒事件，相信读者多少有所耳闻。这是一款专门定向攻击真实世界中基础（能源）设施的"蠕虫"病毒，该病毒具有超强的破坏性和自我复制能力，已感染全球超过 45000 个网络，曾造成伊朗核电站 1/5 的离心机报废、约 3 万终端被感染、监控录像被篡改、放射性物质被泄漏，危害不亚于切尔诺贝利核电站事故。这款病毒以其强大的破坏性，使得伊朗被迫关闭核电站，让美国不废一兵一卒就将其工业控制系统摧毁。该恶意代码能够成功的关键是它同时调用了几个所谓的"零日漏洞"，即新发现的还未被人恶意利用过的软件缺陷，这几个漏洞分别是 RPC 远程执行漏洞、快捷方式文件解析漏洞、打印机后台程序服务漏洞、内核模式驱动程序漏洞和任务计划程序漏洞[⊖]。然而，更加可怕的是，"震网"病毒的历史使命并未结束，它还能够进入多种工业控制软件并夺取一系列核心生产设备的控制权，攻击电力、运输、石油、化工、汽车等重要工业和民用基础设施。这还只是影响力较大的病毒之一，事实上，全球每天都在上演着各式各样的恶意代码进行破坏的"戏码"。

为什么恶意代码有如此大的破坏力呢？因为计算机由硬件和软件两部分组成，前者决定了它的"体力"，后者决定了它的"智力"。软件不过就是指令和数据的集合，表现形式是代码。人类把做"好事"的代码称为善意代码，把做"坏事"的代码称为恶意代码，但从计算机角度看，它们都是代码，没有任何区别。在信息时代到处都有计算机的身影，它可以大至一间屋子（如超算中心），小到一个微型器件（如嵌入式芯片），并与人们的生活密切相关。凡是计算机能做的事情，即（善意）代码可做的"善事"，也都可以由恶意代码转

⊖ 该词的出现始于 1982 年美国国防部信息系统局和高级研究计划局发布的传输控制协议 / 互联网协议（TCP/IP）。

⊖ 如果你想了解这些漏洞是怎样协同分工以完成目标的，详见百度百科的"震网病毒"。

换成“恶事”，从而影响人们的工作生活，这就是恶意代码“邪恶无边”的原因。而且从制造难度上看，恶意代码比善意代码更容易编写，因为普遍的规律是“败事容易成事难”。恶意代码以其制造的容易性、破坏的强大性，已经形成了一个个“黑色部落”，演化出了一个庞大的“家族”。这个“家族”里比较典型的有四种：病毒、僵尸网络、木马、蠕虫（简称为“毒僵木蠕”）。此外，还有后门、下载器、间谍软件、内核套件、勒索软件等其他类型。

1. 病毒

病毒（这里指计算机病毒），从其名字就可以感受到它的威力，就像生物病毒一样，感染者非病即死。世界上第一款病毒雏形出现在 20 世纪 60 年代初的美国贝尔实验室里，三个年轻的程序员编写了一个名为“磁芯大战”的游戏，游戏中可通过复制自身来摆脱对方的控制。20 世纪 70 年代，美国作家雷恩在其出版的《 P1 的青春》一书中构思了一种能够自我复制的计算机程序，并第一次称之为计算机病毒。1983 年 11 月，在国际计算机安全学术研讨会上美国计算机专家首次将病毒程序在 vax/750 计算机上进行了实验，世界上第一个计算机病毒就这样出现了。但真正意义上在世界上流行的第一个病毒出现在 20 世纪 80 年代后期，在巴基斯坦，两个以编程为生的兄弟为了打击那些盗版软件的使用者，设计出了一款名为“巴基斯坦智囊”的病毒，该病毒在全世界广为传播。虽然其形状不像生物病毒，但在行为特征方面，计算机病毒比生物病毒有过之而无不及，都具有感染性、传播性、隐蔽性、可激发性等破坏性，而且还能自我繁殖、互相传染和激活再生。按照感染策略，病毒可分为非常驻型病毒和常驻型病毒。顾名思义，前者短暂停留，一旦摸清被攻击者的情况就快速展开感染、复制、繁殖；后者则长期隐藏在受害者体内，一旦时机成熟就会像癌细胞一样不断分裂，复制自身，消耗系统资源，不断作恶。当然，与生物分类的多样性类似，病毒还有其他许多分类方法，在此不一一列举，你只需要知道它的基本行为特征和破坏力就够了。

2. 僵尸网络

僵尸网络，听上去也是一个让人犯怵的名字。攻击者通过各种途径传播僵尸程序（虽然本质上是病毒，但它只是充当了一个攻击平台的角色）以感染互联网上的大量主机，而被感染的主机通过一个控制信道接收攻击者的指令，组成一个受控的“僵尸网络”，众多计算机就在不知不觉中成为被人利用的一种工具。“僵尸网络”是一种由引擎驱动的恶意因特网行为，常与之一起出现的词还有 DDoS（Distributed Denial of Service，分布式拒绝服务攻击），后者是利用服务请求来耗尽被攻击网络的系统资源，从而使被攻击网络无法处理合法用户的请求。DDoS 形式多样，但最常见的是流量溢出，它可以消耗大量带宽，却不消耗应用程序资源。正是“僵尸网络”的兴起，使得 DDoS 迅速壮大和普及，因为“僵尸网络”为 DDoS 提供了所需的“火力”带宽和计算机以及管理攻击所需的基础架构。发现“僵尸网络”是非常困难的，因为黑客通常远程、隐蔽地控制分散在网络上的“僵尸主机”，这些

主机的用户往往并不知情。因此，“僵尸网络”是目前互联网上黑客最青睐的作案工具之一。而对于上网用户来说，感染“僵尸病毒”则十分容易，因为网络上各种有趣的小游戏、小广告都在吸引着网友。

3. 木马

木马也称木马病毒，名字来源于古希腊传说（《荷马史诗》中“木马计”的故事），但它与一般的病毒不同，它不会自我繁殖，也不会刻意感染其他文件，而是通过将自身伪装起来以吸引用户下载执行。正如它的全名“特洛伊木马”，意思是“害人的礼物”，比喻在敌方阵营里埋下伏兵，等待命令开始行动。攻击者通过特定的木马程序控制另一台计算机，等到合适的时机，攻击者在控制端发出命令，于是隐藏的木马程序就开始进行破坏性行动了，比如窃取文件、修改注册表和计算机配置、复制、移动、删除等。从行为模式上看，木马程序与普通的远程控制软件相似，但后者进行维护、升级或遥控等正当行动，而木马程序则从事着非法活动，且因有着很好的隐蔽性而不容易被发现。也正因为木马程序的隐蔽性，普通杀毒软件难以发现它的行踪，而且它一旦启动就很难被阻止。它会将自己加载到核心软件中，当系统启用时就自动运行。木马的种类也是异常繁多，挂载在不同应用上就表现出不同的功能，不一而足。

4. 蠕虫

蠕虫也是一种病毒，其利用网络进行复制和传播。最初的蠕虫病毒定义源于在 DOS 环境下，该病毒发作时会在屏幕上出现一条类似虫子的东西，胡乱吞吃屏幕上的字母并将其改形。蠕虫病毒是自包含的程序，它能将自身功能的拷贝或自身的某些部分传播到其他计算机系统中。与普通病毒不同的是，蠕虫不需要将其自身附着在宿主程序上就能干坏事。蠕虫主要包括主机蠕虫和网络蠕虫，前者完全包含在其运行的主机中，并且通过网络将自身拷贝到其他计算机终端。一旦完成拷贝动作就会自毁，而让其“克隆物”继续作恶，因此在任何时刻都只有一个“蠕虫拷贝”在运行。蠕虫会“游荡”在互联网中，尝试一个又一个漏洞，直到找到合适的漏洞进而损害计算机，假如成功的话，它会将自己写入计算机，然后开始再次复制。比如近几年来危害很大的“尼姆亚”病毒就是蠕虫病毒的一种，感染该病毒的邮件即使在不手工打开附件的情况下，也会激活病毒。著名的“红色代码”也是蠕虫病毒，其利用微软 IIS 服务器软件的远程缓存区溢出漏洞来传播。SQL 蠕虫王病毒则是利用微软数据库的一个漏洞进行大肆攻击。与传统病毒不同的是，许多新的蠕虫病毒是利用当前最新的编程语言与编程技术实现的，易于修改以产生新的变种，从而逃避反病毒软件的搜索。

1.2.3　漏洞利用

漏洞利用是采用一组恶意软件的集合进行攻击的技术，这些恶意程序中包含数据或可