

# 隐蔽式网络攻击综述

曹自刚<sup>1,2</sup> 熊 刚<sup>1</sup> 赵 咏<sup>1</sup> 郭 莉<sup>1</sup>

<sup>1</sup>(中国科学院信息工程研究所 北京 100190)

<sup>2</sup>(北京邮电大学 北京 100876)

**摘 要** 近年来,随着信息化的推进,国民经济各行各业对网络的依赖性明显增强,网络信息安全问题成为关系国家和社会安全的突出问题。受经济利益驱动,加上各国之间的博弈在网络空间的体现不断加强,具有高技术性、高隐蔽性和长期持续性的网络攻击成为当前网络安全面临的主要挑战之一。文章对这种隐蔽式网络攻击进行了介绍和描述,分析其主要特点和对当前安全体系的挑战。在此基础上综述了国内外隐蔽式网络攻击检测等方面的最新研究进展。最后,对关键技术问题进行了总结,并展望了本领域未来的研究方向。

**关键词** 隐蔽; 网络攻击; 高级持续性威胁; 僵尸网络; 威胁发现

**中图分类号** TP 393 文献标志码 A

## A Survey on Evasive Network Attack

CAO Zigang<sup>1,2</sup> XIONG Gang<sup>1</sup> ZHAO Yong<sup>1</sup> GUO Li<sup>1</sup>

<sup>1</sup>(*Institute of Information Engineering, Chinese Academy of Sciences, Beijing 100190, China*)

<sup>2</sup>(*Beijing University of Posts and Telecommunications, Beijing 100876, China*)

**Abstract** In recent years, due to advances in informatization, the national economy has become more dependent on networks. As a result, the network and information security has become a prominent problem for the national security and social stability. Driven by economic interests and the game between countries reflected in the growing cyberspace confrontations, network attacks with high-tech, high concealment and long-term sustainability become one of the major challenges in the network security. In this paper, the certain kind of covert attack was referred as the evasive network attack (ENA). Firstly, the main characteristics of ENA and the challenges it brings in current security systems were analyzed, based on which the latest developments at home and abroad for ENA detection and other related studies were reviewed then. Finally, key technical issues and future research directions in this field were summarized.

**Keywords** evasive; network attack; advanced persistent threat; botnet; threat discovery

收稿日期: 2014-01-24

基金项目: 国家高技术研究发展计划(863 计划)(2011AA010703)、国家科技支撑计划(2012BAH46B02)、中国科学院战略性先导科技专项课题(XDA06030200)、国家自然科学基金项目(61070184)。

作者简介: 曹自刚, 博士研究生, 研究方向为网络测量和网络攻击检测; 熊刚(通讯作者), 高级工程师, 研究方向为网络信息安全, E-mail: xionggang@ict.ac.cn; 赵咏, 博士, 助理研究员, 研究方向为信息安全和 P2P 测量; 郭莉, 正研级高工, 研究方向为网络信息安全。

## 1 引 言

随着互联网的快速发展、工业信息化的推进以及多种网络的融合,网络信息安全问题已经成为一个突出的社会性问题。近年来,受经济利益驱动而借助僵尸网络和木马进行网络攻击和信息窃取的事件数量快速增加。网络攻击范围已经由计算机互联网扩展到工业控制系统、通信、能源、航空和交通等各个领域。根据国家互联网应急中心的网络安全态势报告<sup>[1]</sup>,针对网络基础设施的探测、渗透和攻击事件时有发生,网站被植入后门等隐蔽性攻击事件呈增长态势,网站用户信息成为黑客窃取重点;火焰病毒(Flame)、高斯病毒(Gauss)和红色十月(Red October)病毒等实施的高级持续性威胁(Advanced Persistent Threat, APT)活动频现,对国家和企业的信息安全造成严重威胁。2012年,我国境内至少有4.1万余台主机感染了具有APT特征的木马程序。与此相呼应,2009年以来,多起APT攻击事件接连被曝光,部分确认受到国家级的支持。最近美国“棱镜(PRISM)”计划告密者斯诺登于2013年11月公开的机密材料<sup>[2]</sup>显示,我国有大量用户被美国通过计算机网络入侵(Computer Network Exploitation)方式安装恶意软件操控。网络攻击和信息窃取行为,已经对公民个人财产及信息安全乃至国家信息安全都造成了严重威胁。

从网络攻击的发展趋势看,当前网络攻击具有如下特点:

(1)广泛性。攻击目标范围广,从传统的计算机互联网到各行各业,如工业控制系统、交通、能源、航空、移动互联网和物联网等。

(2)趋利性。攻击目的以信息窃取和获取经济利益为主,政治目的也逐渐凸显。窃取个人账户等隐私信息、商业机密、科技情报和利用控制的大量主机实施拒绝服务攻击、发送垃圾邮件或用于地下产业收益的攻击数量大幅度增加。

(3)隐蔽性。技术手段上,为了保持持续性控制或持续获得有用信息,攻击者采用高级隐蔽技术对抗不断增强的安全威胁检测技术,从而实现长期潜伏和信息窃取而不被发现。

具有高隐蔽性和持续性的网络攻击能长期躲避安全审查,对信息安全危害大,是目前安全防护和发现威胁的难点。为了更好地描述此类高隐蔽性的网络攻击,基于其与网络安全研究中相关概念术语在原理上的相似性,包括躲避入侵检测系统(IDS)的攻击(Evasion)、躲避被动攻击检测系统的逃逸攻击(Evasive Attacks)以及恶意软件在主机和网络层面的逃逸技术(Evasion Technique)等,我们将其称作隐蔽式网络攻击(Evasive Network Attack, ENA)。隐蔽式网络攻击的检测对于及时发现安全威胁、保护公民个人隐私和财产安全、维护公共网络安全并提高网络安全保障能力具有重要意义。

## 2 隐蔽式网络攻击的概念和特点

### 2.1 概 念

隐蔽式网络攻击是一种对抗性网络攻击,采用隐蔽的入侵手段,并将自身网络通信伪装或隐藏于合法的正常网络数据流中,以躲避主机和网络安全检测,从而长期驻留并控制受害主机,达到持续窃取信息或长期控制利用的目的。隐蔽式网络攻击的概念核心点如下:

(1)网络相关性。攻击必须有网络活动,因此单机上的恶意软件不属于此概念范畴。需要指出的是,本文讨论的网络一般限定为IP网络,但广义上讲此处所指网络的形式和协议是多种多样的。

(2)隐蔽性。入侵方式和通信行为伪装或隐藏是隐蔽式网络攻击的核心特征,因此,使用固定的非常用服务端口通信(加密或非加密)、网络通信具有明显内容签名特征或者攻击过程容易被

受害方感知的攻击(比如拒绝服务类)等不属于此概念范畴。

(3)可控性。攻击者可以通过网络远程控制受害主机执行指定操作、记录、上传指定信息,因此一般意义上的蠕虫、病毒不属于此概念范畴。

(4)目的性。以持续窃取机密信息或长期控制利用为目的,因此普通的后门程序、盗取个人信息的常规木马、恶意勒索软件等都不属于此范畴。

目前流行的网络攻击中,隐蔽型的木马(Trojan)及后门程序(Backdoor)、新型僵尸网络(Botnet)和部分APT可归为此类攻击,而拒绝服务(DOS)攻击和Web入侵渗透(SQL注入、跨站脚本攻击)等一般不属于此概念范畴(部分可归结为隐蔽式网络攻击采用的技术手段)。需要指出的是,早期的IRC僵尸网络大都采用具有明显特征的明文通信协议,而常规木马采用异常端口或明文协议,容易被发现,不具备强隐蔽性特征,显然不属于隐蔽式网络攻击。我们定义的隐蔽式网络攻击与定向网络攻击<sup>[3]</sup>、APT有较多重叠。拒绝服务攻击虽然不具有一般意义的隐蔽性,但其往往是由僵尸网络控制者(BotMaster)操纵数以万计的僵尸(Zombie)主机(俗称“肉鸡”)来实施的,而Web入侵的目的通常是窃取网站用户信息

或通过植入恶意软件控制更多的主机,因此他们与隐蔽式攻击经常有着密切的联系。

典型僵尸网络的攻击流程包括利用漏洞入侵、命令与控制通信、信息窃取或实施攻击。典型木马的攻击流程包括利用漏洞、文件捆绑入侵、命令与控制通信和信息窃取。典型APT的攻击流程包括情报搜集、利用漏洞入口点突破,命令与控制通信、内部横向移动、资产/数据发现和数据隐蔽泄露。上述三者及隐蔽式网络攻击的技术特点、目的性和侧重点总结如表1所示。

可以看出,利用漏洞等隐蔽方式入侵、与命令控制服务器通信、实施信息窃取是他们的共同点。就隐蔽性而言,APT与隐蔽式网络攻击最为相近。APT与隐蔽式网络攻击的最大区别在于,前者一般被理解为定向攻击,而后者可以是非定向的。APT攻击发起者在攻击实施前首先要针对特定攻击目标进行深入调查,然后有针对性地展开全方位的入侵突破,采用手段包括高级入侵技术(常见是零日漏洞利用)或社会工程学等手段。此外,新型威胁、下一代威胁和高级网络攻击等概念实际意义与APT大同小异,此处不再赘述。

总之,隐蔽式网络攻击是对当前最具挑战性的一类高隐蔽性网络攻击的概括,其隐蔽性充分

表1 常见网络攻击与隐蔽式网络攻击比较

Table 1. Comparison between common network attacks and ENA

攻击种类	技术特点	目的性	侧重点
僵尸网络	漏洞、邮件、蠕虫传播;大规模可控命令控制通信	经济利益(DDOS,垃圾邮件,信息窃取)	恶意代码,大规模系统控制和资源利用
木马	漏洞、文件捆绑传播;小范围可控命令与控制通信;隐蔽性	经济利益(信息窃取)	控制与窃取信息
APT	零日漏洞等高级入侵技术、全程保持隐蔽性、小范围可控命令控制通信、定向攻击	经济利益、政治目的(机密信息窃取,战术用途,实施攻击)	侧重综合运用,信息窃取为主
隐蔽式网络攻击	隐蔽入侵,借助伪装变形加密等手段的可控隐蔽网络通信	持续信息窃取、长期控制利用	隐蔽性、长期控制

体现了攻击实施者与当前安全防护技术体系的对抗行为与能力。

## 2.2 特点

由于隐蔽式网络攻击是一个新的概念，其特点与目前的 APT、僵尸网络和木马存在部分重叠，因此，我们尝试从已存在的木马、僵尸网络和 APT 的典型攻击过程和技术特点中提取隐蔽式网络攻击的一般攻击流程及特点。

鉴于木马和僵尸网络是恶意软件和网络攻击领域广泛接受的概念，我们首先从最近引起关注的 APT 出发，分析典型的 APT 攻击的相关情况。根据 Chien 等<sup>[4]</sup>的研究，并结合国外 Mandiant(麦迪安)、FireEye(火眼)、McAfee(麦咖啡)和 Kaspersky(卡巴斯基)等安全公司的 APT

研究分析报告<sup>[5-10]</sup>，汇总了典型 APT 案例的具体情况，如表 2 所示。

经过对表格分析并结合部分报告内容，我们可以得到如下关于 APT 的初步经验结论：

APT 通常利用零日或未修复漏洞进行入侵突破，通过网络进行命令控制通信和信息窃取；被攻陷主机通常采用常见服务，尤其是加密服务的端口通信来躲避安全检测，最常用的端口是 443 和 80 等与 Web 相关的服务；C&C 服务器部署的 SSL/TLS 服务常采用匿名性强的自签名证书；命令控制服务器 IP 地理分布往往是相对分散的，通信内容载荷往往经过压缩、加密变形。由此可见，采用高级隐蔽技术躲避检测是 APT 的主要技术特点。

表 2 典型 APT 案例的分析

Table 2. Analysis of typical APT cases

APT 名称	目的行为	技术特点	网络通信特点
Operation Aurora (极光行动)	窃取特定 Gmail 用户的邮件内容	社会工程学；IE 漏洞利用；远程控制信息窃取	SSL 加密通信
Stuxnet (震网， 超级工厂)	长期隐蔽，悄悄破坏伊朗核设施	针对工业控制系统；多个漏洞利用；长期潜伏实施破坏	零日漏洞 U 盘感染内部主机
Night Dragon (夜龙行动)	窃取西方 5 家跨国公司包含机密信息的敏感文件	SQL 注入；内网扫描；暴力秘密破解；远程控制；数据泄露	80 端口，非 HTTP，通信有明显的固定签名特征；动态 DNS 域名解析服务
RSA SecureID 窃取	窃取 RSA 公司的 SecureID 技术及客户资料	利用零日漏洞的邮件附件；Botnet 的命令与控制 (C&C) 服务器下载指令；加密压缩文件上传	Poison Ivy 变种实施远程控制 (FireEye 发现的样本中该工具主要采用 443、80、8080 等端口通信)
暗鼠行动 (Operation Shady RAT)	渗透了 72 个全球组织	包含特定漏洞利用代码的鱼叉式钓鱼邮件；远程控制	远程控制通信
毒趣 (Duqu)	信息窃取	包含零日漏洞的 Word 文档；远程控制；自定义生命周期	支持采用 80 端口 HTTP 协议、443 端口 SSL 协议及 P2P 命令控制通信
火焰病毒 (Flame)	多种技术有选择的信息窃取	利用 Windows 漏洞，伪造证书，伪装合法软件	监听端口 22、80 和 443；C&C 采用自签名证书
高斯病毒 (Gauss)	财务信息窃取	利用 USB thumb 感染漏洞，命令控制通信	监听端口 22、80 和 443；SSL 服务器采用自签名证书

僵尸网络和木马虽然从具体特征上未必与 APT 相同, 但他们也在朝着更具隐蔽性的方向发展。主流僵尸网络已经逐步采用 HTTP 协议通信, 其通信协议采用 HTTP, 很难做到有效检测和通用性检测; 而木马已经长期采用 HTTP 协议作为突破防火墙等安全设备的方法。同时, 与 APT 类似, 采用常见服务端口如 80、443 和 8080 等的僵尸网络和木马实例也被不断发现, 比如 Zeus 和 Spyeeye 等。因此, 可以看出一个趋势是: 网络攻击由于趋利性而走向隐蔽性, 而为了实现隐蔽性, 往往伪装自身通信内容隐藏于海量的合法和正常的网络数据流中。

隐蔽式网络攻击之所以能够长期潜伏而不被发现, 正是依赖于上述躲避当前主流安全审查策略及技术的高级手段。根据对当前曝光的隐蔽式攻击实例分析, 总结出其主要特征如下:

(1) 高级入口点突破。为了保证攻击的隐蔽性, 入口点突破技术既要突破网络防护, 又要突破主机防护, 经常利用零日漏洞或未修复漏洞, 或通过社会工程学方式将恶意文件或软件传递至特定目标。零日漏洞利用(常见是电子邮件附件和水坑攻击)、SQL 注入攻击、跨站脚本攻击和特种木马等是常用手段。

(2) 隐蔽网络通信。用于受害者主机与外部命令控制服务器通信, 以及将窃取到的数据泄露, 用于躲避防火墙访问控制规则、IDS 内容检测等。通常通过出站连接外部常用的合法服务端口进行通信, 包括加密服务和非加密服务端口。最常用的端口是 80(HTTP)和 443(HTTPS), 其他常用端口包括 22(SSH)和 8080(HTTP)等, 同时也不排除采用 53(DNS)、21(FTP)、25(SMTP)、110(POP3)、465(SMTPS)和 995(POP3S)等常见服务端口。

(3) 内部网络入侵。当攻击者在内部网络找到立足点之后, 一般都需要通过内部网络入侵控制更多有经济情报价值的主机, 进而访问到高价

值的资产或信息。经常会用到的技术包括内部网络主机探测、漏洞扫描和口令破解等内网渗透技术等。由于目前的网络防护主要是边界防护, 安全设备一般部署在网关位置, 网络内部缺乏有效的攻击检测和防护, 缺乏有效的内网网络数据流的监控方法, 一旦攻击者突破网络边界的安全壁垒, 对内网的攻击就变得相对容易。

其中, 隐蔽网络通信是隐蔽式网络攻击的核心特点。

### 3 隐蔽式网络攻击对当前安全体系的挑战

自 2010 年起, 被曝光的隐蔽式网络攻击事件明显增多, 尤其是美国的一些安全公司频繁爆料此类攻击。当然, 其中存在某些安全公司为了特定目的把不具备此类攻击明显特征的网络攻击进行炒作的现象, 比如对于暗鼠行动(Operation Shady RAT), 赛门铁克、卡巴斯与麦咖啡公司就持有不同意见: 赛门铁克认为麦咖啡过分夸大了该攻击<sup>[11]</sup>, 卡巴斯则认为仅仅是僵尸网络而已。然而, 国内外的这些案例警示我们, 目前的网络安全系统已经无法应对快速发展的网络攻击技术。几十年来, 网络安全防护缺乏前瞻性研究, 核心技术思想没有实质变化, 当新型网络攻击突破当前防护体系的核心防护技术后, 整个网络安全行业似乎一下子被拉开差距, 处于疲于应付的状态。

目前的安全防护体系的主流安全防护设备包括基于主机的反病毒(Anti-virus)软件、主机防火墙和主机入侵防护系统(HIPS), 基于网络的防火墙(Firewall)、网络入侵检测/防护设备(IDS/IPS)、统一威胁管理(UTM)和 Web 应用防火墙(Web Application Firewall, WAF)等。这些设备和安全防护技术尚不能满足隐蔽式攻击检测的需求。表 3 是主流安全防护产品在应对隐蔽式网络

表3 当前安全防护手段与隐蔽式网络攻击对抗分析

Table 3. Analysis of current security protection measures fighting against ENA

当前安全防护设备及主要技术		隐蔽式攻击技术手段	防护结果	
主机层	防火墙	IP、端口、域名、出入方向等访问规则控制	合法端口（如 HTTP、HTTPS）出站连接	被突破
	反病毒软件	特征码、恶意行为分析、黑白名单	零日漏洞或未修复漏洞	被突破
	主机入侵防护系统	程序行为规则(应用、文件、注册表和网络)	零日漏洞、社会工程学（电子邮件等）	基本被突破
网络层	防火墙	IP、端口、域名、出入方向等访问规则控制	合法端口（如 HTTP、HTTPS）出站连接	被突破
	IDS/IPS/UTM	数据包载荷特征为主,连接统计信息为辅,多数包含防火墙功能	合法端口（如 HTTP、HTTPS）出站连接, 数据编码或加密	被突破
	WAF	基于攻击特征的 HTTP 请求异常检测	正常的 HTTP 连接或加密连接	被突破

攻击时的表现情况。

由表 3 可以看出, 隐蔽式网络攻击能够有效地有针对性地突破当前的安全防护体系。目前的主机防护的异常行为检测技术容易被隐蔽式攻击的正常行为绕过, 基于特征码的检测则更加容易躲避。同时, 隐蔽式攻击采用“大隐隐于市”(Hide in the Plain Sight)的方法, 主要是合法服务端口、合法行为和合法加密通道的方式突破目前的网络安全防护。

需要说明的是, 主机入侵防护系统理论上能够检测主机端隐蔽式攻击行为, 然而其总体易用性较差, 正常应用也可能被误报, 安全性很大程度上依赖于用户的个人判断能力, 因此未能广泛应用。而目前的自动化 HIPS 技术不完善, 往往因为便利性而牺牲安全性, 且 HIPS 采用的防护技术也能被绕过, 因此 HIPS 不能从单点解决隐蔽性网络攻击问题。

此外, 目前安全防护体系提出的云安全(Cloud Security)和沙箱(Sandbox)技术也不能解决隐蔽式网络攻击的问题。云安全的主要思想是

利用知识共享和统计方法。共享云中的一台主机发现恶意攻击软件后上报至云端, 云端服务下发至各个用户, 这样可以实现一处发现, 全网受益, 从而有效缩短对攻击的响应时间。而统计是指对于一个样本, 如果大部分用户都识别为安全, 那么从统计上该样本是安全可信的; 反之, 大部分人认为是攻击程序, 则其很可能是有害的。这里面可能存在很多问题, 包括对安全厂商以及公共、私有云的信任, 用户隐私泄露, 用户是否加入云, 普通用户对恶意攻击的识别度及主机端检测技术水平等。目前来看, 云安全并没有在检测技术上有根本性提升, 只是通过分布式群体协作架构在一定程度上提高安全性。因此, 云安全不能从根本上应对隐蔽式网络攻击。沙箱技术作为目前防护恶意软件的主流技术之一, 试图采用虚拟运行环境发现恶意软件的攻击性行为, 但很可能被隐蔽式网络攻击绕过或因系统、软件、环境等不能满足特定条件无法触发攻击行为。

总之, 隐蔽式网络攻击的最大特点是隐蔽

性, 从入侵技术到持续潜伏, 再到与外部通信和数据泄露, 力求做到在用户无觉察情况下突破入口点, 利用内网防护弱点进行网内移动, 从而使主机端安全软件认为其是合法程序或构成部分, 安全防护体系认为其通信行为是合法行为。这样就对目前的安全检测和应急响应带来了巨大挑战, 具体如下:

(1) 主机层面对恶意程序的检测问题。目前主机层面的主要技术有特征码、黑白名单、数字签名、沙箱和应用行为检测(包括虚拟机技术、启发式行为检测等)。简单特征码、文件哈希检测早已无法应对恶意程序数量的快速增长, 目前反病毒软件可以用一条特征码应对数十个、数百个甚至更多的恶意程序样本的检测, 但特征码是基于已有特征的抽取, 对于未知恶意程序, 尤其是针对隐蔽式攻击程序, 特征码基本无效。黑名单技术类似, 如果恶意程序不在黑名单内, 则无效。白名单则往往仅限于特定场景应用, 或是为了提高反病毒程序的效率和减少误报, 恶意程序可以采用高级技术伪装或注入到系统或常用程序内。数字签名一般用于辅助判定, 而恶意程序完全可以盗用合法签名或者自行签名, 目前已经发生了多起此类攻击<sup>[12]</sup>。行为检测技术一般误报率较高, 难以在准确度和效率之间找到平衡, 同时隐蔽式网络攻击往往有触发场景控制, 或者伪装能力很强(看起来就是“合法”正常程序), 难以从主机层面找到典型的恶意行为特征。此外, 一旦隐蔽式攻击利用零日漏洞进行攻击, 其攻击行为对于基于已知知识或行为主机的安全防护软件来讲是不可见的, 也是无法检测的。

(2) 网络层面的隐蔽攻击检测问题。防火墙的经典策略是访问控制列表, 根据网络连接的源目的地址、端口规则进行控制, 其缺点是不做内容检测, 对于合法服务的端口被恶意利用的情况无法应对。很多隐蔽式网络攻击都利用了这一显著弱点, 内网主机主动连接外部网络的主机的

HTTP、HTTPS、SSH 等常见服务, 轻易绕过端口限制。网络入侵检测系统目前主要基于内容特征签名和异常行为模式进行网络攻击检测。对于内容编码、变形及加密的应用服务, 入侵检测系统无法知道真实的传输内容, 为了不造成误报, 一般对此类服务放行。因此, 隐蔽式网络攻击可以变形伪装成加密服务躲避入侵检测系统的检查。

对于内网攻击的检测, 面临诸多问题。首先是部署位置的问题, 网关型设备很可能根本无法感知攻击流量, 在什么位置部署、部署多少设备涉及到成本和维护便利性问题。其次是防护的有效性, 不少企业或机构内部网络中主机的安全性较差, 漏洞不能及时打上, 应用程序版本陈旧, 没有安装安全防护软件或者长久无法更新, 使得他们更易受攻击, 当攻击者利用未修复的漏洞实施攻击时, 很难被发现。同时, 内网主机间往往信任度高, 伪造身份信息或通过社会工程学攻击会更容易成功, 造成的危害性也更大。

(3) 多来源数据的关联融合分析问题。隐蔽式网络攻击是多阶段的, 具有较长时间跨度。入侵阶段的入口点是动态的, 方式是多变的, 内部攻击的方法和网络通信的时间、网络出口、数据泄露的方法等也是无法确定的, 因此势必需要多方面的全方位检测体系。就目前的防护体系来看, 虽然对攻击准确识别不太可能, 但攻击流程的任何一个环节都有可能触发低安全等级的警报或产生日志, 比如反垃圾邮件日志、IDS 的可疑远程控制通信行为日志、WAF 的可疑连接日志等。在极端情况下, 需要记录所有的网络行为, 包括正常行为。如何从多个来源的海量日志进行融合分析, 通过关联分析和挖掘, 从低安全等级日志和正常日志中根据时间和空间关系发现隐蔽性网络攻击, 以及采用什么数据分析模型或方法, 是具有很强挑战性的问题。

总之, 目前以已知应对未知, 异常应对“正常”, 单一层面检测应对隐蔽式的复杂攻击, 以

实时短时检测应对长期持续性攻击,这种网络安全防护手段对抗网络攻击尚未能取得较好的效果。隐蔽式网络攻击提高了攻击的技术水平,突破了传统的成本约束规则,打破了当前的安全信任与分工体系,并通过拉长攻击周期使实时检测和应急响应几乎成为空谈。隐蔽式网络攻击的攻击行为不是单个突出的,而更可能是多个、持续性、有序的和隐蔽的看似正常的或低威胁性的行为组合。

对于隐蔽式网络攻击的检测,需要主机和网络检测相结合,在提升多个点各自能力的同时,更需要从多个维度开展,从长时间尺度上进行关联分析,从而发现隐蔽的安全威胁。

## 4 隐蔽式网络攻击的攻与防

对隐蔽式网络攻击的研究目前处于起步阶段,主要是通过实例分析找到此类攻击的主要技术手段,并相应采取新老交替的方法进行应对,同时借助于一些新的技术进行对抗,比如零日漏洞检测技术、隐蔽网络通信行为的检测以及多源数据的关联融合分析等。下面将从其相关技术和检测两个方面分别进行阐述。

### 4.1 隐蔽式网络攻击

隐蔽式网络攻击之所以采用零日漏洞等攻击手段,无非是为了突破当前的网络安全防护体系。而当今网络安全防护系统中,防火墙和IDS仍占据主要地位,但较以往的粒度更细,功能亦更多。下面就从入侵检测系统逃避和新型隐蔽攻击方法探索两个方面进行简单总结。

在躲避入侵检测系统方面,相关研究工作至少可以追溯到上个世纪末,且持续至今。Ptacek等<sup>[13]</sup>指出了IDS系统被动协议分析可靠性的两个根本问题,即信息不全和被动方式,并定义基于上述问题的针对IDS系统的三类攻击:插入、逃避和拒绝服务,针对市场上四款IDS产品的测

试表明上述攻击是有效的。Wagner等<sup>[14]</sup>引入了模仿攻击(Mimicry Attacks)的概念,并开发了一个评估IDS对抗模仿攻击安全性的理论框架,针对一个典型主机IDS采用6种思路实施攻击,包括在避免改变应用可观察到的行为、耐心等待适时插入攻击序列、寻找最佳执行路径、替换系统调用参数、插入无用操作和生成等效攻击。Kayacik等<sup>[15]</sup>对模仿攻击进行了揭秘,将缓冲区溢出攻击分为前奏和利用两个组成部分,通过对四个有漏洞的UNIX应用程序进行监控研究前奏和利用部分的异常行为,结果表明虽然利用部分可以逃避异常检测,前奏部分却难以完全逃避。Kolesnikov等<sup>[16]</sup>探讨基于正常流量变异的多态蠕虫(每个实例具有不同形态)的概念,研究当蠕虫已经进入目标系统内时如何躲避本地IDS检测的问题。针对IDS实例的实验表明,简单的多态蠕虫可以躲避基于特征签名的IDS,而高级多态蠕虫通过特定方法将自身混入正常HTTP流量中并保持流量属性的统计分布基本不变,可以躲避基于异常的IDS的检测。在Rajab等<sup>[17]</sup>的研究中,他们指出随着被动网络监视器司空见惯,恶意软件可以主动探测他们并躲避他们,并提出一种简单有效的动态躲避网络监视器的技术。该方法采用轻量级的采样技术探测活动网络的IP前缀,隔离不活动的IP前缀对应的网络(可能是未被使用或作为被动监测),从而避免被监视器所捕获。Marpaung等<sup>[18]</sup>对恶意软件的逃避技术进行了总结,包括混淆(Obfuscation)、分片和会话拼接(Fragmentation and Session Splicing),应用特定的违规行为,协议违规行为,在IDS处插入流量,拒绝服务和代码重用攻击。

在新型隐蔽攻击的探索方面,最近较为火热的是高级逃逸技术(Advanced Evasion Technique)<sup>[19]</sup>,其技术大部分都是上述逃避IDS的方法中提及的技术。2010年10月,芬兰公司STONESOFT发现,很多高级逃逸技术可以穿透很多当时国际上



著名的大公司网络。这一方面说明当前安全设备存在诸多不完善之处, 同时也为攻击者提供了躲避安全检测的思路, 即利用攻击目标安全防护系统的自身漏洞。刘超等<sup>[20]</sup>设计实现了一种基于 TLS 加密通信协议 HTTPS 隧道木马, 可以有效绕过防火墙和入侵检测系统。此外, 零日漏洞的挖掘和利用对于隐蔽攻击往往是十分便利的, 相关研究此处并未包含。

需要指出的是, 对于安全行业, 最好的技术、最新的方法未必体现在学术研究成果上。受经济和政治利益驱动, 黑客或研究人员一旦找到可以突破安全防护的新方法(比如可利用的系统或常用软件零日漏洞、新的攻击技术), 往往不是将相关信息提供给安全公司并提醒用户防护, 而是用来谋取利益。比如 Vupen 公司<sup>[21]</sup>就将其团队发现的漏洞出售, 这种情况在黑客社区并不少见。对于国家来说, 新的隐蔽攻击方法可能被作为网络对抗中的有力武器, 很可能属于国家秘密, 并不轻易示人。这样以来, 学术研究往往会滞后于实践, 甚至在一定程度上缺乏实用性。因此, 我们的总结必然是不全面的。但至少可以看出, 安全本身就是一个蕴含着攻防双方对抗的动态概念。攻击方对于躲避安全防护系统的探索从未停止过, 也不可能停止。隐蔽式网络攻击的凸显是防护方的方法和技术较明显落后于对方的表现。

#### 4.2 隐蔽式网络攻击的检测

由于隐蔽性网络攻击不是具有精确定义的称呼, 根据其内涵与外延, 相关的检测与防护研究工作可以从僵尸网络、隐蔽木马和 APT 三个相关的方向进行总结。

僵尸网络的检测研究方面。目前基于 HTTP、P2P 的隐蔽性和生存性强的僵尸网络呈增多趋势, 对应地, 基于群体行为特征和行为关联的检测方法成为主流, 基于日志数据的学习挖掘方法为事后分析提供了支持。僵尸网络的检测从根本上来讲是异常检测, 即区分出僵尸

网络通信区别于一般通信的特点。目前有以下检测方法<sup>[22-24]</sup>: 基于内容签名的方法(简单准确, 对未知的无效)、基于行为特征的方法(流量的群体相似性和时间关联性)、基于关联分析的方法(分布式拒绝服务攻击、扫描、垃圾邮件、二进制文件下载和漏洞利用等行为活动)和基于数据挖掘的方法(动态 DNS、DNS 集中请求, NXDOMAIN 和攻击事件日志分析)。Gu 等<sup>[25]</sup>提出的 BotMiner 通过将可疑活动聚类 and 相似通信流聚类, 以及对两者结果进一步关联分析找出 Botnet 中的主机。该方法不依赖协议、网络架构和先验知识, 具有可扩展性。Gu 等<sup>[26]</sup>提出了 BotProbe 通过主动探测来识别 IRC 僵尸网络的隐蔽或混淆通信。Yadav 等<sup>[27]</sup>借助 DNS 请求失败和域名(成功和失败)的熵(归一化的编辑距离)关联分析加速 Botnet 检测。Zhao 等<sup>[28]</sup>利用机器学习方法借助于流统计属性通过对流量行为进行分类的方法对僵尸网络进行识别。

基于 HTTP 协议的僵尸网络很容易隐藏在普通 HTTP 流量中, 难以发现; 同时不少僵尸网络开始采用加密手段或 P2P 模式增强隐蔽性和可生存性。很多隐蔽式网络攻击是定向攻击, 很可能不具备僵尸网络的协同性和时空相似性等特征, 且往往采用加密信道, 难以找到特征。隐蔽性网络攻击与一般意义上的僵尸网络存在较大差异, 不一定具有分布式拒绝服务、垃圾邮件和扫描等相关活动特征。因此, 对僵尸网络的检测方法不能广泛应用于隐蔽式网络攻击。

木马的检测方面。主要检测方法可以归纳为基于特征匹配的检测技术和基于网络行为的检测, 其中后者涵盖基于心跳检测(时间间隔、数据包大小和数量)和交互操作行为(人的行为, 两个方向数据包大小、间隔、比例等)。Chen 等<sup>[29]</sup>提出基于应用层内容检测已知远程控制木马的框架, 并通过会话关联检测简单加密(XOR)的木马。李世淙等<sup>[30]</sup>等提出了一种分层聚类的木马

通信行为检测方法,通过对IP通信的双向字节数、包数、持续时间和包间隔进行异常检测发现木马通信行为;类似的工作还有张晓晨等<sup>[31]</sup>针对窃密木马检测的研究。基于特征匹配的方法属于事后检测,需不断进行特征库更新,难以应对封装,编码变形;对基于心跳等行为检测方法而言,网络视频、网络游戏、聊天软件也有类似“心跳”行为,易误报;基于聚类等学习方法在真实场景中的实用性难以达到要求。

APT检测是目前安全产业界和学术界的焦点之一。产业界有FireEye、趋势科技(TrendMicro)、RSA等业内网络安全公司针对APT提出了检测方案,主要包括三方面的检测:恶意代码检测、可疑通信检测和攻击行为检测,分别针对包含漏洞利用代码的恶意文件、命令控制通信和下载、传播、数据泄露等。具体来说:

第一是主机端防范入侵突破,解决的问题是恶意代码检测和主机应用防护。典型代表有国外FireEye的恶意代码防御系统MPS(Malware Protection System)和Bit9的可信安全平台,国内南京瀚海源的产品“星云”和安天的“追影”高级威胁鉴定器。

第二是网络通信,包括命令与控制通信信道发现,数据泄露检测以及内网威胁检测。国外趋势科技的Deep Discovery专门检测APT攻击的命令控制通道,其入侵检测系统引擎上部署了恶意代码检测沙箱来弥补传统特征攻击检测的不足。国内尚未见到对应产品,但网络IDS/IPS产品与其原理类似,技术细节可能存在一定差异。

第三是综合检测、大数据分析方面,国外产品代表有RSA的NetWitness,国内有启明星辰天阗威胁检测与智能分析系统等。

产业界各类产品采用的主要技术包括黑白名单、虚拟执行、智能沙箱、特征匹配和自动提取、软件信誉库和智能上下文分析等。此外,为了提高准确度和减少漏报,多个厂商之间的产品

出现了相互集成和共享信息的情况。

学术方面,由于零日漏洞常出现在APT攻击前期的入口突破阶段,对于整个攻击过程来说往往十分重要,因此研究者尝试从零日漏洞攻击检测着手发现APT攻击。Alazab等<sup>[32]</sup>针对恶意软件的混淆技术,根据Windows系统API调用频率借助有监督机器学习算法对零日未知恶意软件进行识别。赛门铁克的Bilge等<sup>[33]</sup>对2008年到2011年间的零日攻击进行了系统性的实证研究,利用1100万真实主机的历史场景数据进行分析,通过将可执行程序与利用的漏洞关联,并借助程序最早出现日期对比,识别出一定数量的零日漏洞利用程序,并分析了零日漏洞生命周期。Aleroud等<sup>[34]</sup>借助线性数据变换和异常检测技术基于已知攻击签名上下文属性检测零日攻击,其中线性数据转换技术依赖于几个判别函数,用于借助分析网络连接特征计算零日攻击的估计概率。异常检测技术识别使用一类最近邻算法,并使用奇异值分解技术来实现降维。Dai等<sup>[35]</sup>通过分析全网主机的系统调用构建系统对象依赖图,并通过前向和后向追踪识别零日攻击路径。此外还有人利用蜜罐来检测零日攻击。

Binde等<sup>[36]</sup>针对APT攻击远程控制环节,提出了以下几个方面的检测方法:规则集合(钓鱼、PI-RAT和注册表项)、统计和关联方法(Fast-flux域名和注册表检查和Snort规则检查PI-RAT关联)、人工方法(DNS日志、异常流量、奇怪的出站流量)和自动化数据泄漏阻断(检测和阻止RAR文件、限制出站连接、OSSEC主动响应)。Tankard<sup>[37]</sup>指出对抗APT需要多种网络监控措施(日志分析、文件完整性检查、注册表监控和rootkit检测),而且分析出站流量非常重要,对443端口流量进行严格的访问控制;Giura等<sup>[38]</sup>提出了一种利用MapReduce分布式计算来揭露APT攻击的框架,通过事件来源、事件存储、多平面关联和时间关联进行APT检

测; Virvilis 等<sup>[39]</sup>对 stuxnet、duqu、flame 和 Red October 进行了技术分析, 并针对性地提出了对抗 APT 的方法: 补丁管理、强网络访问控制和监控, 严格的互联网访问策略和内容检查, 电子邮件附件异常检查, DNS 查询检查, 蜜罐和主机入侵检测系统。

从上面可以看出, 对于隐蔽式网络攻击, 主流检测方法不再依赖于具体特征本身, 而是利用攻击行为与正常行为在上下文属性、对象依赖关系和网络连接特征等方面的差异来识别判断。而且, 采用多种方法进行综合检测是应对隐蔽式网络攻击的主流思路。

## 5 隐蔽式网络攻击检测的未来

虽然隐蔽式网络攻击检测尚处于探索阶段, 目前甚至短期内找不到较好的解决方法, 但至少需要做好以下三个方面: 隐蔽式网络攻击的主机层面检测、隐蔽式网络攻击的网络检测(包括出入口和网络内部)和多源海量数据的融合关联分析。从做好各个关键点防护出发, 减轻此类攻击的损害, 然后构筑一个新的安全体系, 有效预防、检测和阻止此类攻击。具体来说至少需要在以下关键研究点上取得突破。

### 5.1 隐蔽式网络攻击的主机层面检测

研究隐蔽式网络攻击能够逃避主机端检测的原因、技术手段和特点, 并探索新的有效检测方法。基于现有知识的方法无法应对未知攻击, 新的隐蔽性攻击在已知特征库里面找不到特征标识, 因此被判定为合法、安全。这种理论对于隐蔽式网络攻击的安全防护是不可行的。特征库类似于黑名单, 黑名单无法检测出没有历史记录的非违法行为。而采用白名单对于用户来说限制过多, 频繁扩充或变更白名单可能出现配置不一致等问题, 而且如果过分依赖白名单, 一旦其出现问题, 后果将十分严重。

既然目前的检测方法能被突破, 那么研究新的主机层次检测方法很有必要。主机检测的优点是可以知道行为的产生者是哪个进程、哪个线程, 缺点是一旦对手比你的技术水平更高、更底层, 那么你就“有眼无珠”, 很难检测对方的行为。这个时候, 可以从主机外来看, 即从其网络连接发现蛛丝马迹。因此, 通过将主机检测和网络检测融合起来, 将是很有意义的。在主机软件判定方面, 采用白名单机制或者在白名单、黑名单中间加入“灰名单”, 借助于行为特征、历史知识等来进行综合判定是比较可行的方法。

### 5.2 隐蔽式网络攻击的网络检测

包括网络出入口隐蔽通道检测和内网攻击行为检测。具体来说, 至少包含以下内容:

#### 5.2.1 加密服务中的隐蔽攻击通道检测

研究如何将识别不同加密服务以及将常规加密应用服务(P2P 和 VoIP 等)与隐蔽式攻击加密通道进行粗粒度区分的挑战性问题<sup>[40]</sup>, 便于在此基础上对可疑攻击流量进行进一步分析。

近年来, 互联网飞速发展, 普通个人用户的可用网络带宽成倍增加, 网络用户数量呈现高速增长趋势。根据中国互联网信息中心(CNNIC)的统计, 截至 2013 年 6 月底, 我国网民规模达 5.91 亿<sup>[41]</sup>。网络用户增加促进了网络经济的繁荣, 也引起网络流量的大幅度增加。流量的增长给用户网络安全管理的有限计算资源带来极大挑战。尤其是最近几年, 随着 P2P 应用借助加密混淆技术逃避运营商的封锁以及网络安全问题的日益突出, 采用 HTTPS 进行 Web 访问, 采用加密 FTP、电子邮件、以及加密通道进行远程访问和办公的服务越来越多。根据 CNCERT 统计, TCP 443 端口(HTTPS)流量长期在 TCP 端口流量排行中位列第 3 位<sup>[42]</sup>。加密网络服务(加密 P2P 等)的大量增长和服务内部的复杂多变性<sup>[43]</sup>使得恶意行为更容易隐藏其中, 增加了隐蔽式攻击的检测难度。

目前观察到的隐蔽式攻击较多借助加密方式

躲避检查。因此,对于企业、机构的安全管理员来说,需要关注所有网络通信行为,并且尽可能准确地排除正常访问行为,从而将可疑隐蔽式攻击流量所在的大集合经过初筛形成一个较小的集合,便于后面进一步分析。

隐蔽式网络攻击的网络检测问题归根结底是流量中隐蔽攻击行为模式的识别问题,且基本上是一个异常流量的发现问题。但由于这些恶意攻击流量隐藏较深,且伪装较好,看起来与正常“无异”,所以能够欺骗当前的检测设备,突破网络安全防御。同时,流量变形(Traffic Morph)<sup>[44]</sup>和混淆技术给恶意攻击者提供了网络流量层面躲避检测的新途径。因此,如何有效检测隐蔽式攻击的网络通道,是当前亟待解决的技术难题。当网络攻击流量模式不具有明显异常时,可以运用逆向思维,通过从加密服务整体中精确识别合法服务来筛选出可疑隐蔽攻击通道。

### 5.2.2 伪装及利用合法加密服务的隐蔽攻击通道流量识别

面对安全设备的网络检查,隐蔽式网络攻击中受控主机上的恶意软件(或泄密者)和命令与控制(C&C)服务器(或接收者)通信或数据泄露逐渐转变为采用加密控制信道。目前,部分僵尸网络、后门已经开始采用 HTTPS、SSH 等加密服务端口或伪装成相应流量躲避检测。最近发现的 APT 多采用 HTTPS 和 SSH 协议。根据根据安全厂商的分析报告<sup>[6,7]</sup>, Gauss、Duqu 和 Flame 等

APT 均支持采用 SSL 或 SSH 协议与命令控制服务器通信。

需要指出,采用或伪装加密不是偶然个例,而是必然趋势。2013年4月,国内厂商发现利用波士顿爆炸事件的 APT 样本采用 HTTPS 与 C&C 服务器进行网络通信<sup>[45]</sup>;2013年5月,赛门铁克发现一个 Linux 后门程序 Fokirtor 将其流量伪装成正常 SSH 流量躲避检测<sup>[46]</sup>;知名僵尸网络 Spyeeye 采用 80(HTTP)和 443(HTTPS)端口躲避安全检查。图1是2013年4月捕获的 Spyeeye 样本的网络数据包示例(客户端到服务器方向),服务器端口为 443,但通信内容不是标准 SSL/TLS 协议格式。网络安全公司 Mandiant 的报告(M-Trends)<sup>[47]</sup>指出,APT 攻击样本 100% 仅采用出站连接,且 83% 采用 TCP 的 80 或 443 端口。

隐蔽式网络攻击之所以采用 SSL 等加密服务或服务端口进行通信,有以下原因:

- (1) 不能用明文,因为控制通道通信、上传文件等很容易被安全设备检测到;
- (2) 不适宜用简单编码或变形,可能被安全设备解密或被异常检测发现;
- (3) 不适宜在非加密服务端口采用强加密手段,容易被异常检测发现;
- (4) 只有采用合法加密服务端口,才容易绕过检测,从而长期保持隐蔽性。

因此,采用合法加密服务端口的对应加密应用类型或伪装成其流量进行通信是隐蔽性网络攻

Transmission Control Protocol, Src Port: 11286 (11286), Dst Port: https (443)		Secure Sockets Layer	
00	00 30 48 23 98 a1 44 37 e6 4d 81 0c 08 00 45 00	.0H#..D7 .M....E.	
10	01 26 7d 85 40 00 40 06 83 e8 ac 10 1e dd c1 6b	.&}.@.@. ....k	
20	ac 0b 2c 16 01 bb f5 a4 5e 05 3b dd 45 c8 50 18	..... A.;.E.P.	
30	3e bc 90 6a 00 00 f0 10 60 59 4d 4c 4c 4c 4c 4b	>..j..j.. YMLLLLK	
40	48 48 48 48 4a 4a 4a 4a 47 47 47 47 33 5c 7c 08	HHHHJJJJ GGGG\ .	
50	69 0b 67 02 02 70 15 65 54 24 26 26 3d 17 17 17	i.g..p.e T\$&=...	
60	17 75 1a 6e 31 56 23 4a 2e 2e 1b 35 04 2a 18 2e	.u.n1V#J...5.*.	
70	1e 2e 0f 4e 0a 47 0e 40 6d 5d 6e 57 60 52 6b 5e	..N.G.@ m]rW Rk^	
80	69 5f 7e 4c 7c 3d 0e 4b 0f 38 bd bb aa 6e 68 62	i.-L =K .8...nhb	
90	14 71 03 70 19 76 18 18 29 19 2a 19 20 49 40 62	.q.p.v...).*. I@b	
1a0	3e 35 35 25 49 26 45 24 48 17 63 0a 67 02 02 30	>55%i&E\$ H.c.g..0	
1b0	00 31 02 2c 1c 2f 01 32 03 23 12 25 1f 2f 1c 26	.1.../.2 .#.%./.&	
1c0	16 25 0b 32 06 34 51 54 7b 3f 3a 56 52 52 56 2c	%.2.4QT {?:VRRV.	

图1 僵尸网络 Spyeeye 通过 TCP443 端口躲避检查

Fig.1. Botnet Spyeeye exploits TCP port 443 to escape detection

击的必然选择。如何对合法加密服务端口的流量进行检测, 识别出其中极小比例的隐蔽式网络攻击的通道, 是一个亟待解决的问题。

### 5.2.3 HTTP 等常见非加密协议中的隐蔽攻击通道的识别

HTTP 是目前网络流量的最主要来源之一, 根据国家互联网应急中心的 TCP 端口流量排行统计, 80 端口的流量始终排名第一位, 且遥遥领先于其他端口<sup>[42]</sup>。根据 Mandiant<sup>[47]</sup>, 大部分 APT 攻击采用了 80 和 443 端口进行命令控制通信。主流的僵尸网络也主要采用 HTTP 协议隐藏受控端和控制端的通信。而对木马来说, 反弹式木马很早就采用了 HTTP 协议来躲避主机反病毒软件和网络防火道的检查。隐蔽式网络攻击之所以一致选择了 HTTP 服务端口, 首先是考虑到 HTTP 协议应用的广泛性; 其次, 相对于 HTTPS 的加密流, 明文 HTTP 协议的内容通过编码、变形或加密更容易逃避安全设备的检查。一旦命令和数据经过编码, 又作为 HTTP 的 URL 或数据的组成部分, 就非常难以做到有效检测, 一般情况下会被当作正常数据放行。

因此, 如何从各种 HTTP 请求中挖掘出这些恶意攻击的通道, 将这些深度伪装的攻击流量或主机识别出来, 是一个非常具有挑战的问题, 甚至难以有效解决。在目前的情况下, 采用基于 IP、域名的声誉, 请求的行为特征并结合学习方法是可能的缓解方案之一。

需要指出的是, 除了 HTTP 服务的 80 端口外, 53 (DNS 服务)、110 (POP3)、25 (SMTP)、21 (FTP)、23 (TELNET)、554 (RTSP) 和 161 (SNMP) 等常用服务端口, 以及 8080 (HTTP 代理)、1080 (Socks 代理) 和 8000 (互联网广播、iRDMI) 等端口都有可能被隐蔽式网络攻击用于隐蔽通信。因此, 该研究应扩展为常用服务端口的异常攻击行为模式的发现。这就要求安全设备不能简单的基于端口, 而要深入理解常用协议的

细节, 这样才有可能发现伪装行为或异常模式。

### 5.2.4 内网攻击行为检测

内网攻击行为是指从内网发起的针对网络内部主机的攻击行为。内网攻击对于入侵者获取有价值的信息和资产是很重要的一步。对于安全防护级别较高的企业或机构来说, 包含重要信息的主机一般是不会直接暴露给攻击者的。因此, 攻击者在突破外层防护后, 一般需要在内网进行多次跳转或移动, 以控制更有价值的主机, 从而实施进一步的攻击行为。目前安全设备比如防火墙、IDS 和 WAF 等基本属于边界防护, 一般部署在网关, 无法对网络内部攻击进行审核, 一旦攻击者侵入内部网络, 就能够较容易地进行数据嗅探、网络主机扫描、漏洞扫描, 并利用漏洞进行技术性或非技术性的攻击。

内网攻击行为检测的主要难点在于如何在主机之间部署安全防护设备, 若每台主机之间都部署, 势必造成成本太高, 故此方法不可行。一种较为可行的方案是将内网产生的流量复制后转发到特定的处理设备上进行监控并分析处理, 但势必会增加网络复杂度和经济成本。

## 5.3 多来源海量数据的融合关联分析

就目前的研究成果来看, 对隐蔽式网络攻击的检测研究还处在过渡阶段, 大部分方法都是在原有基础上的改进, 并没有质的突破。如果要构建一个新的检测体系, 那么很多问题亟待解决, 其中很重要的就是多来源数据的存储、管理与综合关联分析。

目前大数据的思路<sup>[48]</sup>在应对新型安全威胁发现方面较为突出。对于日益增长的网络流量, 我们如何存储对日后分析有用的数据, 仅依赖安全设备日志肯定是不够的。那么是否全流量存储, 还是提取关键信息存储, 或者两者结合? 对于主机端是否要存储信息来弥补网络层面的不足? 面对海量数据如何设计存储架构来节约空间并满足日后检索分析的效率需求? 但无论如何, 数据来

源必定包含各种安全设备的日志信息，及大量主机的不同类型的数据信息。

然而，数据不会告诉我们攻击是如何发生的，需要我们运用智能去挖掘。智能上下文分析技术，时序事件关联有助于多来源数据的融合分析。关联分析的结果很可能是一个不确定的结果，因此隐蔽式网络攻击的识别决策问题也是未来的研究内容之一。

具体来说，对于大数据，需要解决的问题有：如何进行智能分析，如何运用机器学习的方法实现去伪存真、去粗取精，如何把事件顺序理清，如何从横向和纵向进行关联分析，如何排除干扰并最终追踪溯源等。如何从数据中提取有用知识并将知识融合形成能力是未来安全系统解决数据分析问题的关键。

#### 5.4 总结和展望

从当前的主流检测技术来看，采用常见服务端口的隐蔽式网络攻击通信行为能容易地够绕过安全防护。对于隐蔽式攻击加密信道的检测，研究者在运用多种检测手段并进行综合关联分析方面达成了一致的看法，但缺乏对关键环节关键问题的深入研究，更多的是综合性的检测方案。反病毒软件和网络入侵检测设备分别作为主机和网络层面的主要防护手段，如何在技术层面提升他们应对隐蔽式攻击的能力是需要解决的难题。从隐蔽式攻击的生命周期看，在命令控制通信和数据隐蔽泄漏阶段实施检测是基于网络检测较为合适的选择，而主机端检测技术对于早期发现并阻止隐蔽式攻击入侵具有关键性意义。需要指出的是，隐蔽式网络攻击的检测是整个安全行业的挑战性问题，需要花费大量精力进行相关技术突破，从关键点出发，由点及面，构建综合立体防御，最终实现从分散的、异构的海量数据流中发现隐藏的威胁。

当前针对隐蔽式网络攻击的检测方法可归结为以下几方面：

(1) 入口点恶意代码检测：在互联网入口点和主机层面对 Web、邮件、文件共享等可能携带的恶意代码进行检测，在早期及时发现 APT 攻击的蛛丝马迹。

(2) 出口点数据防泄密：APT 攻击目标是有价值的信息，在主机上部署数据泄漏防护产品，防止敏感信息的外传也是防御 APT 攻击的方法之一。

(3) 攻击过程中网络通信检测：在网络层面对 APT 攻击的行为进行检测。

(4) 大数据分析：全面采集网络中的各种数据(原始的网络数据包、业务运行日志和设备安全日志)，采用大数据分析技术和智能分析算法来检测 APT，可以覆盖 APT 攻击的各个阶段。

我们认为，要想做好隐蔽式网络攻击防御，需要几大转变，即：从被动防御到主动感知；从异常检测到威胁发现；从单点决策到融合分析；从攻击特征到行为路径；从实时检测到长期追踪；从黑白名单到安全信誉；从企业责任到全民协作。同时，人为因素应该受到充分重视，包括了内部人员威胁、社交网络、社会工程和安全意识等。

隐蔽式网络攻击的检测不是一朝一夕的事情。既然信息安全问题已经成为各行各业的共同问题，那么解决这一问题就不仅仅要靠少数安全行业的企业来完成，而是需要大家的共同参与，不仅是安全行业内部的协作，更需要企业之间、政府多个部门之间以及政府与民间的紧密合作。随着隐蔽式网络攻击实例的发现，经过不断总结和积极探索，新的方法和思路将越来越成熟，我们对其响应速度将会不断提升，最终其隐蔽性将被淡化直至消失，而其对整个行业的安全威胁也将得以消除。

#### 参考文献

- [1] 国家互联网应急中心. 2012 年我国互联网网络安全态势综述 [EB/OL]. (2014-2-14).

- [http://202.108.212.119/publish/main/upload/File/2012CNCERTreport\(1\).pdf](http://202.108.212.119/publish/main/upload/File/2012CNCERTreport(1).pdf).
- [2] Boon F, Derix S, Modderkolk H. NSA infected 50,000 computer networks with malicious software [EB/OL]. <http://www.nrc.nl/nieuws/2013/11/23/nsa-infected-50000-computer-networks-with-malicious-software/>.
- [3] Sood A, Enbody R. Targeted cyberattacks: a superset of advanced persistent threats [J]. *IEEE Security and Privacy*, 2013, 11(1): 54-61.
- [4] Chien E, OMurchu L, Falliere N. W32.Duqu: the precursor to the next stuxnet [C] // *The 5th USENIX Workshop on Large-Scale Exploits and Emergent Threats*, 2012.
- [5] McAfee Foundstone Professional Services and McAfee Labs. Global energy cyberattacks: "Night Dragon" [EB/OL]. (2014-2-15). <http://www.mcafee.com/in/resources/white-papers/wp-global-energy-cyberattacks-night-dragon.pdf>.
- [6] Kaspersky Lab Global Research and Analysis Team. Gauss: abnormal distribution [EB/OL]. (2014-2-10). <http://www.securelist.com/en/downloads/vlpdfs/kaspersky-lab-gauss.pdf>.
- [7] Bencsáth B, Pék G, Buttyán L, et al. Duqu: a stuxnet-like malware found in the wild [EB/OL]. (2011-10-14). <http://free.ebooks6.com/Duqu-A-Stuxnet-like-malware-found-in-the-wild--Technical-Report-by-pdf-e7121.html>.
- [8] Litan A. RSA SecurID attack details unveiled—lessons learned [EB/OL]. (2014-02-10). <http://blogs.gartner.com/avivah-litan/2011/04/01/rsa-securid-attack-details-unveiled-they-should-have-known-better/>.
- [9] Alperovitch D. Revealed: operation shady RAT [EB/OL]. (2014-02-10). <http://www.mcafee.com/us/resources/white-papers/wp-operation-shady-rat.pdf>.
- [10] Bennett JT, Moran N, Villeneuve N. Poison ivy: assessing damage and extracting intelligence [EB/OL]. (2014-02-10). <http://www.FireEye.com/resources/pdfs/FireEye-poison-ivy-report.pdf>.
- [11] Lau H. The truth behind the shady RAT [EB/OL]. (2014-02-15). <http://www.symantec.com/connect/blogs/truth-behind-shady-rat>.
- [12] Amann B, Sommer R, Vallentin M, et al. No attack necessary: the surprising dynamics of SSL trust relationships [C] // *The 29th Annual Computer Security Applications Conference*, 2013: 179-188.
- [13] Ptacek TH, Newsham TN. Insertion, Evasion, and Denial of Service: Eluding Network Intrusion Detection [Z]. Calgary: Secure Networks Incorporated, 1998.
- [14] Wagner D, Soto P. Mimicry attacks on host-based intrusion detection systems [C] // *Proceedings of the 9th ACM Conference on Computer and Communications Security*, 2002: 255-264.
- [15] Kayacik HG, Zincir-Heywood AN. Mimicry attacks demystified: what can attackers do to evade detection? [C] // *The 6th Annual Conference on Privacy, Security and Trust*, 2008: 213-223.
- [16] Kolesnikov O, Lee W. Advanced polymorphic worms: evading IDS by blending in with normal traffic [R]. GIT-CC-05-09, Atlanta: Georgia Technology College of Computing, 2005.
- [17] Rajab MA, Monroe F, Terzis A. Fast and evasive attacks: highlighting the challenges ahead [C] // *Proceedings of 9th International Symposium, Lecture Notes in Computer Science*, 2006: 206-225.
- [18] Marpaung JAP, Sain M, Lee HJ. Survey on malware evasion techniques: state of the art and challenges [C] // *The 14th International Conference on Advanced Communication Technology*, 2012: 744-749.
- [19] Gold S. Advanced evasion techniques [J]. *Network Security*, 2011, 1: 16-19.
- [20] 刘超, 王轶骏, 施勇, 等. 匿名 HTTPS 隧道木马的研究 [J]. *信息安全与通信保密*, 2011, 9(12): 78-80.
- [21] VUPEN Security—The leading provider of defensive and offensive cyber security intelligence [EB/OL]. (2014-02-10). <http://www.vupen.com/english/>.
- [22] Feily M, Shahrestani A, Ramadass S. A survey of botnet and botnet detection [C] // *The 3rd International Conference on Emerging Security Information, Systems and Technologies*, 2009: 268-273.
- [23] 方滨兴, 崔翔, 王威. 僵尸网络综述 [J]. *计算机研究与发展*, 2011, 48(8): 1315-1331.
- [24] 江健, 诸葛建伟, 段海新, 等. 僵尸网络机理与防御技术 [J]. *软件学报*, 2012, 23(1): 82-96.
- [25] Gu GF, Perdisci R, Zhang JJ, et al. Botminer: clustering analysis of network traffic for protocol

- and structure independent botnet detection [C] // Proceedings of the 17th Conference on Security Symposium, 2008: 139-154.
- [26] Gu GF, Yegneswaran V, Porras P, et al. Active botnet probing to identify obscure command and control channels [C] // Annual Computer Security Applications Conference, 2009: 241-253.
- [27] Yadav S, Reddy ALN. Winning with dns failures: strategies for faster botnet detection [C] // The 7th International ICST Conference on Security and Privacy in Communication Networks, Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering, 2012: 446-459.
- [28] Zhao D, Traore I, Sayed B, et al. Botnet detection based on traffic behavior analysis and flow intervals [J]. Computers & Security, 2013, 39: 2-16.
- [29] Chen ZQ, Wei P, Delis A. Catching remote administration trojans (RATs) [J]. Software: Practice and Experience, 2008, 38(7): 667-703.
- [30] 李世淙, 云晓春, 张永铮. 一种基于分层聚类方法的木马通信行为检测模型 [J]. 计算机研究与发展, 2012, 49(增刊): 9-16.
- [31] 张晓晨, 刘胜利, 刘龙. 网络窃密木马的自适应检测模型研究 [J]. 计算机应用研究, 2013, 30(11): 3434-3437.
- [32] Alazab M, Venkatraman S, Watters P, et al. Zero-day malware detection based on supervised learning algorithms of api call signatures [C] // Proceedings of the Ninth Australasian Data Mining Conference, 2011: 171-182.
- [33] Bilge L, Dumitras T. Before we knew it: an empirical study of zero-day attacks in the real world [C] // Proceedings of the 2012 ACM Conference on Computer and communications security, 2012: 833-844.
- [34] Aleroud A, Karabatis G. Toward zero-day attack identification using linear data transformation techniques [C] // IEEE 7th International Conference on Software Security and Reliability, 2013: 159-168.
- [35] Dai J, Sun XY, Liu P. Patrol: revealing zero-day attack paths through network-wide system object dependencies [C] // Proceedings of the 18th European Symposium on Reseaech in Computer Security, Lecture Notes in Computer Science, 2013: 536-555.
- [36] Binde BE, McRee R, O'Connor TJ. Assessing Outbound Traffic to Uncover Advanced Persistent Threat [Z]. Maryland : Sans Technology Institute, 2011.
- [37] Tankard C. Advanced persistent threats and how to monitor and deter them [J]. Network security, 2011, 8: 16-19.
- [38] Giura P, Wang W. Using large scale distributed computing to unveil advanced persistent threats [J]. Science, 2012, 1(3): 93-105.
- [39] Virvilis N, Gritzalis D. The big four-what we did wrong in advanced persistent threat detection? [C] // The 8th International Conference on Availability, Reliability and Security, 2013: 248-254.
- [40] 熊刚, 孟姣, 曹自刚, 等. 网络流量分类研究进展与展望 [J]. 集成技术, 2012, 1(1): 32-42.
- [41] 中国互联网信息中心. 第 32 次中国互联网络发展状况统计报告 [EB/OL]. (2014-02-10). [http://www.cnnic.net.cn/hlwfzyj/hlwzxbg/hlwtjbg/201307/t20130717\\_40664.htm](http://www.cnnic.net.cn/hlwfzyj/hlwzxbg/hlwtjbg/201307/t20130717_40664.htm).
- [42] 国家互联网应急中心. 协议流量排名 [EB/OL]. (2014-02-10). <http://www.cert.org.cn/publish/main/index.html>.
- [43] 曹自刚, 熊刚, 赵咏. 基于 X.509 证书测量的隐私泄露分析 [J]. 计算机学报, 2014, 37(1): 151-164.
- [44] Wright CV, Coulls SE, Monroe F. Traffic morphing: an efficient defense against statistical traffic analysis [C] // Proceedings of the 16th Network and Distributed Systems Symposium, 2009: 237-250.
- [45] 南京翰海源. 利用波士顿马拉松爆炸案热点的新 APT 攻击 [EB/OL]. (2014-02-18). [http://blog.vulnhunt.com/index.php/2013/04/19/new\\_apt\\_attack\\_exploit\\_theboston\\_marathon\\_boom\\_event/](http://blog.vulnhunt.com/index.php/2013/04/19/new_apt_attack_exploit_theboston_marathon_boom_event/).
- [46] Symantec. Linux back door uses covert communication protocol [EB/OL]. (2014-02-18). <http://www.symantec.com/connect/blogs/linux-back-door-uses-covert-communication-protocol>.
- [47] Mandiant. M-Trends: advanced persistent threat malware [EB/OL]. (2014-02-10). <https://www.mandiant.com/blog/m-trends-advanced-persistent-threat-malware/>.
- [48] 周涛. 大数据与 APT 攻击检测 [J]. 信息安全与通信保密, 2012, 7: 23.