僵尸网络中的关键问题

王天佐[1),2)] 王怀民[2)] 刘 波[1)] 史佩昌[2)]

[1)](国防科学技术大学计算机学院网络与信息安全研究所 长沙 410073)

[2)](国防科学技术大学计算机学院并行与分布处理国家重点实验室 长沙 410073)

摘 要 僵尸网络是一种复杂、灵活、高效的网络攻击平台,在互联网中分布非常广泛.僵尸网络使攻击者具备了实施大规模恶意活动的能力,如发送垃圾邮件、发动分布式拒绝服务攻击等.由于其危害日益严重,僵尸网络已经成为网络安全研究的热点之一.但是近年来,僵尸网络新的发展、变化,突破了以往对僵尸网络的认知.文中分析僵尸网络的现有研究,对僵尸网络进行了重新定义,并从网络结构、网络独立性和信息传递方式等角度对僵尸网络的类型进行了划分;然后,梳理了僵尸网络检测技术、测量技术和反制技术等方面的工作;最后,给出了僵尸网络的演化趋势和未来研究方向.

关键词 志愿僵尸网络;自部署僵尸网络;测量技术;检测技术;反制技术;僵尸网络演化趋势

中图法分类号 TP311 **DOI号**: 10.3724/SP.J.1016.2012.01192

Some Critical Problems of Botnets

WANG Tian-Zuo[1),2)] WANG Huai-Min[2)] LIU Bo[1)] SHI Pei-Chang[2)]

[1)](*Institute of Network and Information Security, School of Computer Science, National University of Defense Technology, Changsha* 410073)

[2)](*National Laboratory for Parallel and Distributed Processing, School of Computer Science, National University of Defense Technology, Changsha* 410073)

Abstract As a complex, flexible and effective platform for network attacking, the botnet spreads widely in the Internet. Botnets can provide the botmasters with the ability to launch large-scale malicious activities such as spamming and DDoS (Distributed Denial of Service) attacks. Botnets are continuously bringing more and more severe threats, so that the study on botnets has already become one of the focuses in the field of network security. However, in recent years, some new developments of botnets are challenging the existing understanding on botnets. In this paper, according to the new conditions of botnets and the researches in existence, a definition of botnet is proposed based on the works of other researchers, taxonomies of botnets are introduced respectively from the views of network structure, dependency and delivery pattern of C&C (Command and Control) information. Then the techniques on detecting, measuring and restraining botnets are analyzed systematically. In the end, we give the evolution trends of botnets and the future research trends in this area.

Keywords volunteer botnet; self owned botnet; botnet detection; botnet measurement; botnet restrain; evolution tendency of botnets

---

收稿日期:2011-10-26;最终修改稿收到日期:2012-04-13.本课题得到国家"九七三"重点基础研究发展规划项目基金"高效可信的虚拟计算环境基础研究"(2011CB302600)、国家自然科学基金"大型分布式软件系统的行为监控与可信演化"(90818028)、国家杰出青年科学基金(60625203)资助.**王天佐**,男,1982年生,博士研究生,主要研究方向为分布式计算和信息安全.E-mail:phoenixwtz@163.com.**王怀民**,男,1962年生,博士,教授,主要研究领域为分布式计算、信息安全和计算机软件.**刘 波**,男,1973年生,博士,副研究员,主要研究方向为分布式计算和信息安全.**史佩昌**,男,1981年生,博士研究生,主要研究方向为分布式计算和信息安全.

1 引　言

僵尸网络通常指攻击者通过一对多的命令与控制信道，控制大量主机所组成的恶意网络. 这种恶意网络在互联网中的渗透范围非常广. 赛门铁克公司的数据表明①，2010 年仅 Rustock[1] 僵尸网络的结点规模就超过了 100 万. 中国国家互联网应急中心年度报告②也指出，2011 年中国境内约有 890 万个 IP 地址感染了木马和僵尸病毒，较 2010 年同比增长 78.5%. 凭借大量网络资源，僵尸网络在攻击者的控制下，对网络安全构成巨大威胁，主要体现为发送垃圾邮件[2-3]、窃取用户隐私[4-5]、进行点击欺诈[6]、发动分布式拒绝服务攻击(Distributed Denial of Service，DDoS)[7]、为恶意网站提供 Fast-Flux 保护[8]等. 例如，2010 年全球 47.5% 的垃圾邮件来自 Rustock，其日发送量约为 441 亿封①，这些垃圾邮件被用于投送海量广告、传播恶意代码和散播不实股市消息. 另据分析③，到 2011 年初点击欺诈也成为了 Rustock 的主要功能. 通过 DDoS 攻击 BGP(Border Gateway Protocol)路由器来瘫痪互联网[9]，更凸显了僵尸网络的巨大破坏力.

目前僵尸网络已经得到网络安全学术界的持续重点关注. 德国蜜网项目组的 Holz 等人[7,10-11]将蜜网技术用于僵尸网络检测，做出了突出贡献；美国乔治理工学院的 Gu 等人设计了 BotHunter[12] 等多个僵尸网络检测系统；美国约翰霍普金斯大学的 Rajab 等人[13]在僵尸网络规模测量方面研究较早，提出了一种根据 DNS(Domain Name System)缓存记录进行测量的方法；德国曼海姆大学的 Holz、加拿大麦吉尔大学的 Davis 等人[14-15]分别研究了 P2P 僵尸网络反制方法；美国加州大学圣芭芭拉分校的 Stone-Gross 等人[16]的工作证明了劫持僵尸网络的可行性. 2008 年以来，CCS(ACM Conference on Computer and Communications Security)、SIGCOMM(ACM annual conference of the Special Interest Group on Data Communication)、USENIX Security、NDSS(ISOC Network&Distributed System Security Conference)等网络安全重要学术会议都对僵尸网络给予高度重视，而由 WORM(ACM Workshop on Recurring/Rapid Malcode)和 HotBots(USENIX Workshop on Hot Topics in Understanding Botnets)合并而来的 LEET(USENIX Workshop on Larger-scale Exploits and Emergent Threats)会议更将僵尸网络作为重点关注方向.

最近几年僵尸网络的发展出现了很多不同以往的新情况. 在网络形态上，随着智能手机的迅速普及，不仅僵尸程序的载体由相对固定的主机扩展到了移动平台[17-18]，控制信道的构建方式也由基于 Internet 延伸到了基于短信[19-20]、蓝牙[21]等技术；随着匿名黑客社区的逐渐成熟[22]，僵尸网络的传播不再纯粹依靠恶意攻击，用户通过自愿安装他所信任的 LOIC④等开源程序加入僵尸网络已经是不争的事实[23]；随着云计算技术的快速发展，攻击者基于云服务平台(如亚马逊的 EC2⑤等)提供的廉价资源构建高效能、高可控的僵尸网络也变得方便可行[24]. 在网络功能上，Stuxnet⑥和 Duqu⑦的相继出现标志着僵尸网络的攻击技术走向专业化，开始具备向工控系统等非通用平台渗透攻击的能力. 在威胁水平上，僵尸网络的目标已经逐渐由单纯谋取暴利和攻击某个站点转向更高层次. 如随着社交网络(如 Twitter、Facebook 等)的广泛使用及对社会舆论引导能力的增强⑧，僵尸网络不仅能够借助社交链进行社交渗透和扩大感染规模(如 Koobface[25])，还可以借助社交网络对大众心理产生重要影响[26]. Ratkiewicz 等人[27]研究发现在 2010 年美国中期选举中有部分受控社交结点在 Twitter 上发布大量误导信息，试图对选情造成影响. 据称，目前地下市场中一个这种结点的售价高达 29 美元⑨.

形态的多元化、功能的专业化、威胁的复杂化，不仅对当前僵尸网络防护研究提出了严峻挑战，而且说明僵尸网络的概念、外延已经扩展，迫切需要对僵尸网络从基本涵义到研究状况进行全方位的重新分析，认清威胁本质，抓住问题关键，预测发展趋势，

① Symantec. MessageLabs Intelligence：2010 Annual Security Report.
② CNCERT. 2011 年我国互联网网络安全态势综述. http://www.cert.org.cn/UserFiles/File/201203192011annualreport(1).pdf
③ World's Largest Spam Botnet Switched to Click Fraud. http://news.softpedia.com/news/World-s-Largest-Spam-Botnet-Switched-to-Click-Fraud-176510.shtml
④ Low Orbit Ion Cannon. http://en.wikipedia.org/wiki/Low_Orbit_Ion_Cannon
⑤ Amazon Elastic Compute Cloud. http://aws.amazon.com/ec2/
⑥ Symantec，W32. stuxnet dossier. http://securityresponse.symantec.com/en/id/content/en/us/enterprise/media/security_response/whitepapers/w32_stuxnet_dossier.pdf
⑦ Symantec，W32. Duqu The precursor to the next Stuxnet. http://www.symantec.com/content/en/us/enterprise/media/security_response/whitepapers/w32_duqu_the_precursor_to_the_next_stuxnet.pdf
⑧ Morozov E. Swine flu：Twitter's power to misinform. http://neteffect.foreignpolicy.com/posts/2009/04/25/swine_flu_twitters_power_to_misinform
⑨ Jet bots. http://allbots.info

找出研究盲区,明确努力方向.

本文深入剖析了僵尸网络的基本概念、类型、关键技术和演化及研究趋势.文章内容组织如下:第2节在分析僵尸网络概念的基础上,提出了僵尸网络的新定义;第3节从网络结构、网络独立性和信息传递方式等角度对僵尸网络的类型进行了划分;第4节从僵尸网络的检测、测量和反制等方面对僵尸网络研究的关键技术进行了系统总结和分析比较;第5节对僵尸网络自身演化趋势和未来研究方向进行了分析和展望;最后总结全文.

当前对僵尸网络的研究主要分为3类:(1)认识型研究.主要研究僵尸网络本身的特点和规律,包括僵尸网络的概念、功能结构和工作机制等;(2)防护型研究.主要研究如何应对僵尸网络的危害,包括僵尸网络的检测、测量和反制技术等;(3)攻击型研究.主要研究如何提高僵尸网络的功能和性能,包括增强僵尸网络隐蔽性和鲁棒性等技术.本文重点关注前两类研究.为便于表述,本文将用户自愿安装僵尸程序而形成的僵尸网络称为志愿僵尸网络,如LOIC网络;将攻击者使用自有网络结点构建的僵尸网络称为自部署僵尸网络,如基于云的僵尸网络[24].

2 基本概念

2.1 僵尸网络定义

关于僵尸网络的概念,Honeynet项目组①认为僵尸网络是由受害主机组成并被攻击者远程控制的网络;Gu等人[28]指出僵尸网络由感染了特定恶意代码的大量受害主机组成,能够被单个操纵者控制;而诸葛建伟等人[29]认为僵尸网络是攻击者出于恶意目的,传播僵尸程序控制大量主机,并通过一对多的命令与控制信道所组成的网络.然而新型僵尸网络的发展不仅超出了这些定义的范畴(如僵尸程序不再必须运行于计算机主机,不再必须构建于Internet之上,也不再必须以破坏的方式进行传播),而且已经对网络安全造成了重大现实威胁(如2011年4月,来自LOIC僵尸网络的分布式拒绝服务攻击使索尼PSN网站无法正常提供服务,该公司遭受严重损失).为了正视威胁,认清本质,有必要对僵尸网络的定义进行重新修正.

定义1. 僵尸网络是攻击者通过一对多的控制结构,恶意组织大量受控网络结点形成的受控网络.记为 $Botnet=(CNS,CS,CA)$.

$CNS(ControlledNodeSet)$指僵尸网络受控结点集合,记为 $CNS=\{CN_1,CN_2,\cdots,CN_K\}$. $CN(ControlledNode)$指僵尸网络的受控结点.受控结点可以是计算机、手机,也可以是虚拟机等;不止包括非授权的网络结点(如通过入侵手段获取的僵尸主机等),还包括授权使用的网络结点(如用户志愿加入或自主部署到僵尸网络中的各类结点).记为 $CN=(ControlledProgram,AuthorityType)$.

$ControlledProgram$ 指运行于受控结点之上的受控程序. $ControlledProgram\in\{Infrastructure,BotProgram\}$,受控结点也相应分为两种类型. $Infrastructure$ 表示受控结点上运行的是僵尸网络操纵设施程序,此时该结点的作用在于帮助控制者向僵尸结点发布指令;此类结点称为 $CNInfrastructure$(操纵设施),通常体现为命令与控制服务器. $BotProgram$ 表示受控结点上运行的是僵尸程序,此时该结点的作用在于接收攻击者的指令并执行指定任务;此类结点称为 $CNBot$(僵尸结点),通常体现为僵尸主机.

$AuthorityType$ 指受控结点上 $ControlledProgram$ 运行权限的获取方式. $AuthorityType\in\{Authorized,Unauthorized\}$, $Authorized$ 表示 $ControlledProgram$ 的运行获得了用户授权, $Unauthorized$ 表示通过入侵手段获取了运行权限.

$CS(ControlStructure)$指僵尸网络的控制结构.攻击者为了实现对僵尸网络的控制,需要在自身和僵尸结点之间构建一对多的命令传递通道,本文将该通道的组织方式称为控制结构,记为 $CS=(botmaster\times CNInfrastructure^i\times CNBot^j)$. CS 刻画了僵尸网络中各角色在命令传递过程中的协作关系,$botmaster\times CNInfrastructure$ 表示由攻击者向操纵设施发布指令的路径,$CNInfrastructure\times CNBot$ 表示由操纵设施向僵尸结点传递指令的路径.i 表示操纵设施的层级数量,j 表示僵尸结点的层级数量.指令传递的路径可以基于互联网构建,也可以基于其他方式(如SMS、Bluetooth等)构建.

$CA(Command\text{-}Activity)$指僵尸网络的指令-响应模式,记为 $CA=(\langle Command_1,Activity_1\rangle,\langle Command_2,Activity_2\rangle,\cdots,\langle Command_n,Activity_n\rangle)$. $Command_i$表示第 i 种指令,$Activity_i$表示 $Command_i$ 触发的相关活动.

① Know your Enemy: Tracking Botnets. http://www.honeynet.org/papers/bots

与已有定义相比，本文定义在概念内涵上更关注僵尸网络的整体控制机制，在概念外延上适应了新型僵尸网络的发展需求.例如，通过分别定义僵尸网络的受控结点和控制结构，突出了网络的拓扑特性；通过区分受控程序的类型，指出了僵尸网络中的受控结点存在角色差异这一事实，明确地刻画了操纵设施（如控制服务器等）在僵尸网络命令分发中的作用；通过区分受控结点的权限获取方式，以非授权类型的僵尸网络表述了与已有定义相一致的概念，以授权类型的僵尸网络涵盖了志愿僵尸网络和自部署僵尸网络等新形态.

2.2　僵尸网络关键性能

从攻击者的视角，僵尸网络以下性能备受关注.

（1）透明性.透明性是指攻击者在操纵僵尸网络发起攻击或维护升级时可以将整个僵尸网络当作一个整体来操纵，不需要关心网络内部细节.透明性主要通过控制结构来实现，具体体现为僵尸网络控制者只需将命令与控制信息注入到控制结构中，由控制结构传递到各个僵尸结点，实现对僵尸网络的整体操纵.网络控制者不需要直接操纵僵尸结点，即僵尸结点对攻击者命令发布过程透明.

（2）攻击容量.攻击容量是指能够被僵尸网络控制者操纵的所有受控资源的总和，决定了控制者能够发起的最大攻击强度.攻击容量取决于网络规模、带宽资源等因素.网络规模越大，攻击者能够利用的网络地址就越多，攻击来源就越分散，受到的制约越小；带宽资源越丰富，攻击者能够发起的攻击流量就越强大.

（3）隐蔽性.隐蔽性主要是指传统僵尸网络生命周期中的 4 个主要阶段[30]（包括僵尸网络初始感染、二次注入、执行任务和维护升级等）的活动需要较为隐蔽，降低僵尸结点、操纵设施和僵尸网络整体产生的流量等被检测的可能性.隐蔽性要求僵尸结点通常不能显著占用 CPU、内存和带宽等资源或对宿主机可用性产生明显破坏，甚至需要使用 Rootkit 技术进行隐藏或主动禁止非授权行为检测工具的运行等.对志愿僵尸网络和自部署僵尸网络而言，隐蔽性更多是要防止被网络监管系统发现，而不注重防范终端用户查杀.

（4）抗毁性.抗毁性是指僵尸网络在面临部分僵尸结点被毁坏或清除的情况下，仍然能保持一定攻击能力（部分满足僵尸网络控制者预期）的特性，也可称为韧性.较高抗毁性可以提高僵尸网络的生存能力，可以为攻击者提供更充裕的时间调整僵尸结点行为特征，避免僵尸网络完全失效.僵尸网络抗毁性的提高主要通过构造更稳健的网络结构来实现.目前的研究[31-32]认为，采用结构化 P2P 构建的僵尸网络具有较好的抗毁性.对于志愿僵尸网络和自部署僵尸网络而言，僵尸结点被清除的威胁不大，主要的挑战来自于对其控制结构的破坏，因此相对于隐蔽性，控制结构的抗毁性对其更为重要.

3　类型划分

本节从网络结构、网络独立性和信息传递方式等不同维度划分僵尸网络的类型，如图 1 所示.其中从网络结构角度，根据僵尸网络是否具有中心控制结点，将僵尸网络划分为中心式、非中心式和复合式 3 种类型.其中，中心式僵尸网络指具有中心控制服务器的僵尸网络；非中心式僵尸网络指不具有中心控制服务器的僵尸网络；复合式僵尸网络指中心式和非中心式两种网络结构复合而成的僵尸网络.从网络独立性的角度，根据僵尸网络是否附着在其他网络中，将僵尸网络划分为寄生式和自生式两种类型.其中，寄生式僵尸网络指附着在其他网络中的僵尸网络；自生式僵尸网络指独立于其他网络的僵尸网络.从信息传递方式的角度，根据僵尸结点是否主动获取命令与控制信息，将僵尸网络划分为推送式和拉取式两种类型.其中，推送式僵尸网络指僵尸结点被动接收命令与控制信息的僵尸网络；拉取式僵尸网络指僵尸结点主动请求命令与控制信息的僵尸网络.

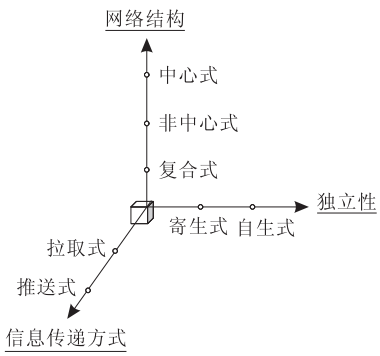


图 1　从网络结构、独立性和信息传递方式等维度对僵尸网络进行的类型划分

3.1　基于网络结构的划分

文献[33]将僵尸网络分为单服务器星形结构、多服务器结构、层次结构和随机结构等 4 种类型，随机结构主要是指采用 P2P 协议的网络，但是很多采用层次结构和多服务器结构的僵尸网络在某些层面

上也会采用随机结构，如 Waledac 等. Leder 等人[34]将僵尸网络分为中心式、非中心式和机动式，但很多机动式僵尸网络本质上仍然是中心式的. 为了提高分类的准确性，根据网络结构是否具有中心服务器，本文将当前的僵尸网络分为中心式、非中心式以及在这两种基本结构基础上形成的复合式结构.

中心式僵尸网络每次发布命令与控制信息所用的操纵设施位置相对固定，形成明显的中心服务器，多呈现出星型结构或多服务器结构[33]. 中心式僵尸网络采用最多的通信协议是 IRC 协议与 HTTP 协议. 其中，IRC 协议自从僵尸网络产生起一直是其主流的协议[35]，典型的基于 IRC 协议的僵尸网络有 SDBot、RBot、AgoBot 等. 典型的基于 HTTP 协议的僵尸网络有 BoBax、ClickBot、MegaD[36]、iKee.B[37]等. 另外，很多其他协议也都可以用于构建中心式的僵尸网络，如各种 IM 协议、FTP 协议、邮件协议[38]等. 中心式僵尸网络的主要优点在于便于控制，主要弱点在于单点失效.

非中心式僵尸网络每次发布命令与控制信息所用的网络结点不固定，不具有中心服务器，单点失效问题不明显. 这种僵尸网络一般采用 P2P 协议并呈现出 P2P 结构. 在这种僵尸网络中，僵尸结点也可以传播命令与控制信息，这为攻击者提供了灵活多变的操纵接口，而无需固定的操纵设施. 典型的采用结构化 P2P 协议的僵尸网络有 Storm 等，而非结构化 P2P 僵尸网络有 Nugache 等. Grizzard 等人[39]在总结 P2P 僵尸网络的基础上，对基于 Overnet 协议的 Storm 僵尸网络进行了分析；Dietrich 等人[40]对 Nugache 僵尸网络进行了详细分析，发现改进后的 Nugache 提高了加密强度，并通过博客和垃圾邮件等进行传播. 考虑到实际 P2P 网络中各个结点的异构性，Wang 等人[41]提出了一种新型非中心式僵尸网络——Hybrid P2P botnet. 该网络采用 P2P 协议，上层是 Servent 结点（一般具有独立 IP），下层是 Client 结点（一般是内网结点）；Client 之间没有连接，且需要通过 Servent 结点来连入整个 P2P 网络. 攻击者可以通过任一结点将命令与控制信息注入网络；网络中的任何一个结点在接收到新的命令与控制信息后，都会向其邻居列表中的 Servent 结点主动转发；而 Client 结点会定期地主动连接邻居列表中的 Servent 结点以获取命令与控制信息. 这样，命令与控制信息在 Servent 结点间以泛洪的方式传播，然后 Client 结点以拉取的方式从 Server 结点获取命令与控制信息. 非中心式僵尸网络的主要优点在于抗毁性高，主要弱点在于难以应对 Sybil[42]攻击.

复合式僵尸网络是指同时使用中心式和非中心式两种基本结构构造出的僵尸网络. Vogt 等人[43]基于僵尸网络小型化的趋势，提出了一种下层为中心式、上层为非中心式的复合式僵尸网络 Super Botnet. Super Botnet 僵尸程序在传播过程中，自动地构建出指定数量的小规模中心式僵尸网络，作为整个僵尸网络的底层；同时，利用传播过程中的感染关系以及重复感染关系，在各个子僵尸网络的中心服务器之间建立加密连接，形成上层的非中心式服务器网络. 命令与控制信息被注入上层服务器网络后，通过非中心式网络的连接关系传播到各个子网的中心服务器；各中心服务器接收到信息之后，完成信息在本子网内的传播，这样就实现了命令与控制信息的“一对多”传递和对复合式僵尸网络的透明操纵. Nunnery 等人[44]对曾在互联网中广泛分布的 Waledac 进行了研究，发现 Waledac 是一种具有四层结构的复合式僵尸网络，而 Conficker[45]也已被证明同时具备中心服务器和 P2P 连接. 复合式僵尸网络旨在同时借助两种控制方式提高控制便利性和生存能力，但往往也同时面临着单点失效和 Sybil 攻击的威胁.

3.2 基于网络独立性的划分

根据是否附着在其他网络上，僵尸网络还可以分为自生式和寄生式两种. 自生式僵尸网络，每个结点都是僵尸网络的结点，所用的通信协议也是独立的；寄生式僵尸网络是其他网络的子网，其命令与控制信息的发布需要利用被附着网络现有的通信协议，甚至命令与控制信息的传递过程也常常会用到被附着网络中的常规结点. 例如，2007 年以前的 Storm 僵尸网络[10]即属于寄生式，它附着于正常的 ED2K 网络中，借助 Overnet 协议通信，并且在搜索命令与控制信息的过程中，常常需要正常结点提供路由支持. 2007 年以后改用 StormNet 私有协议的 Storm 僵尸网络则属于自生式.

自生式僵尸网络的主要优点在于可以方便地采用各种控制机制，但其容易被防护方沿结点间连接关系、“顺藤摸瓜”地找到其他僵尸结点. 寄生式僵尸网络的主要优点在于结点活动的隐藏性更好，不易被网络监管系统发现. 然而，由于寄生式僵尸结点难以辨识当前的交互对象是否也为僵尸结点，因此往往只能通过预先约定“接头地点”的方式传递指令；而这种约定规律一旦被发现，很容易被反制.

3.3 基于信息传递方式的划分

操纵设施为操纵者向僵尸网络注入命令与控制信息提供了接口，但信息如何在操纵设施与僵尸结点之间传递还是另一个问题. Holz[11] 按照命令与控制信息传递的方式，将僵尸网络分为推送式(Push Mechanism)和拉取式(Pull Mechanism)两种类型. 推送式僵尸网络的命令与控制信息会被发送到各个结点，僵尸结点只需被动接收. IRC 僵尸网络一般都采用推送式机制，而 P2P 僵尸网络如果采用泛洪的通信方式，则也属于推送式. 拉取式僵尸网络中，僵尸结点需要主动请求命令与控制信息. 基于 HTTP 协议的僵尸网络一般采用拉取式机制，而如果基于 P2P 协议的僵尸网络采用基于发布/订阅机制(Publish/Subscribe)的通信方式，则也属于拉取式.

推送式僵尸网络的主要优点在于指令传递迅速，主要弱点在于如果采用 P2P 结构，泛洪信息会降低僵尸结点的隐蔽性. 拉取式僵尸网络的主要优点在于当采用 P2P 结构时，僵尸结点较为隐蔽，主要弱点在于指令传递的时效性受限.

Holz 结合基于网络结构的分类方法和基于信息传递方式的分类方法，形成了如下的交叉分类列表(表 1).

表 1　僵尸网络的交叉分类

传递方式	网络结构	
	中心式	非中心式
推送式	基于 IRC 机制	基于泛洪机制
拉取式	基于 HTTP 机制	基于 Publish/Subscribe 机制

3.4 典型僵尸网络的类型分析

一些典型僵尸网络在网络结构、网络独立性和信息传递方式等维度的特征比较如表 2 所示. 其中，Clickbot 主要用于进行点击欺诈，MegaD 和 Waledac 等主要用于发送垃圾邮件. 中心式僵尸网络与非中心式的区别主要体现在抗毁性上，中心式僵尸网络因存在单点失效问题而劣于非中心式僵尸网络. 推送式与拉取式僵尸网络的主要区别体现在隐蔽性上，推送式的僵尸网络在推送命令与控制信息的时候，非常容易暴露，因此劣于拉取式，但其在命令传递的时效性上优于拉取式. 寄生式僵尸网络因有被附着网络的掩护而增强了隐蔽性.

表 2　僵尸网络的分类

典型僵尸网络	僵尸网络类型						
	网络结构			网络独立性		信息传递方式	
	中心式	非中心式	复合式	自生式	寄生式	拉取式	推送式
SDBot	√irc			√			√
AgoBot	√irc			√			√
Rbot	√irc			√			√
Bobax	√http			√		√	
Clickbot	√http			√		√	
Torpig	√http			√		√	
MegaD	√http			√		√	
Storm(Overnet)		√p2p			√	√	
Storm(Stormnet)		√p2p		√		√	
Nugache		√p2p		√		√	
Hybrid P2P botnet		√p2p		√		Servent:泛洪;Client:拉取	
Super-Botnet			√irc	√			√
Waledac			√混合协议	√		√	

4 关键技术

由于我们以降低僵尸网络的危害为研究目标，因此本文关键技术主要涉及僵尸网络的防护型技术，包括僵尸网络检测技术、测量技术和反制技术，如图 2 所示. 检测技术所要解决的问题是如何找出互联网中活动的僵尸结点和僵尸网络，主要的思路有 3 种：流量分析、蜜网捕获和反向检测. 测量技术所要解决的问题是对僵尸网络的结点数量进行统

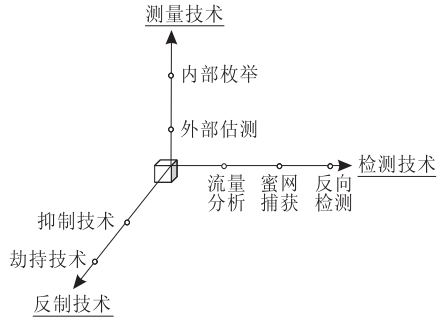


图 2　关键技术分类

计,主要思路可以分为渗透至僵尸网络内部进行枚举和从僵尸网络外部进行估测两种.反制技术所要解决的问题是在检测发现僵尸网络后,如何清除僵尸网络或减轻其危害,主要的思路可以分为对僵尸网络活动进行抑制和劫持僵尸网络控制权两种.

这3类技术之间的关系如图3所示,检测是测量和反制的基础,反制是检测和测量的目的,测量是反制的评估手段之一.

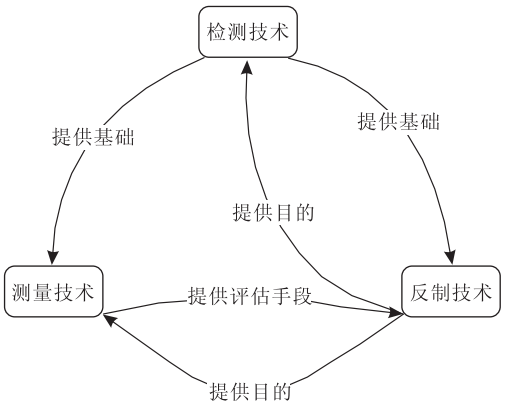


图3 关键技术之间的关系

由于目前鲜有对志愿僵尸网络和自部署僵尸网络的测量、检测和反制的研究成果,因此,如无特殊声明,本节内容所分析的技术只针对传统僵尸网络.

4.1 检测技术

僵尸网络的检测一直是僵尸网络研究的难点问题,目前主要的检测技术有基于流量分析的检测技术、基于蜜网捕获的检测技术和反向检测技术等.流量分析技术主要是对互联网中某个子网的主机集合实施主动持续监测,蜜网捕获技术对整个互联网进行被动抽样监测,而反向检测技术则根据危害后果反溯僵尸网络.

4.1.1 基于流量分析的检测技术

基于流量分析的检测技术通过对网络数据报文进行持续监测和关联分析来发现僵尸网络的活动.关联分析可以从3个层面进行[46]:纵向关联、横向关联和因果关联.

纵向关联.Gu等人[12]为僵尸程序的多阶段感染过程构建了"基于状态的感染序列模型",并基于Snort设计实现了僵尸结点检测系统BotHunter.该系统通过在网关上追踪内部主机与外部网络间的双向通信流量来识别感染过程的不同阶段(包括指向内部网络的扫描、攻击与恶意代码下载,指向外部网络的协同会话、对外攻击与传播等),并将识别结果汇总到关联分析引擎.关联分析引擎对汇总结果进行纵向关联形成可疑会话链,然后与感染序列模型进行匹配以判断是否存在僵尸结点.这种方法的局限性在于:感染序列模型是预定义的,且对于感染过程各个阶段的检测手段是基于特征的.

横向关联.横向关联检测方法基于如下认识:僵尸结点活动规律是由其共同代码逻辑定义的,因而同一僵尸网络中的结点行为应具有时空相似性.横向关联方法正是通过分析不同主机流量间的时空相似性来检测僵尸结点.基于这种思想实现的系统有BotSniffer[47]和BotMiner[48].BotSniffer主要针对中心式僵尸网络,从时空特性和消息/响应特性等角度对多个主机的网络流量进行关联分析.BotMiner是一种不限于中心式僵尸网络的通用检测框架,它先分别根据通信流量相似性和恶意行为相似性对主机组进行两维度聚类,然后对聚类结果进行跨维度关联分析,以期发现既具有相似通信模式又具有相似恶意行为模式的主机集合.与BotHunter相比,这种方法的好处是不需要僵尸网络的先验知识(如僵尸程序行为特征等).

因果关联.BotHunter、BotSniffer和BotMiner都是通过对网络流量的收集分析来发现僵尸结点,文献[49]还进一步提出了一种主动探测IRC僵尸网络的系统BotProbe.BotProbe系统基于典型僵尸程序具有确定的命令-响应模式且不能容忍会话中发生拼写错误的这一理性认识,利用假设检验的思想,从网络流量中分辨出哪些是IRC僵尸程序与外部的命令与控制会话,哪些是人与人的会话.在检测过程中,BotProbe会向网络中注入一些探针报文,因而是一种主动探测技术.这种方法较之前两种,更为积极.

4.1.2 基于蜜网捕获的检测技术

基于蜜网捕获的检测技术是指通过设置蜜网系统,对所捕获程序的行为进行分析,判断是否存在僵尸结点.德国的蜜网项目组①最早利用蜜网技术对僵尸网络进行检测和跟踪研究.该项目组利用蜜罐蜜网技术捕获了大量僵尸程序,并对这些程序进行了详细研究,为僵尸网络的检测研究提供了第一手资料.Holz等人[11]详细介绍了利用蜜罐蜜网技术对中心式僵尸网络和非中心式僵尸网络的检测、分析和跟踪方法.北京大学的诸葛建伟等构建了一种基于高交互式蜜罐技术的恶意代码捕获器——

① The honeynet project. http://www.honeynet.org/

HoneyBow[50]，获得了优于 Nepenthes① 的检测效果. ShadowServer 组织②也利用蜜罐蜜网技术来对互联网中僵尸网络的活动进行检测. 蜜网技术本身“守株待兔、愿者上钩”的特点保证了检测结果的准确性，但无法感知不访问蜜网系统的大量恶意代码.

4.1.3　反向检测技术

反向检测技术是指通过对互联网中大规模恶意行为的危害后果进行数据挖掘，寻找隐藏其后的僵尸网络. Husna 等人[51]基于时域特征对邮件进行聚类分析，获得了对发送垃圾邮件的僵尸网络检测方法. Sroufe 等人[52]提出了一种基于邮件外形轮廓特征聚类以检测僵尸网络的方法. Xie 等人[53]提出了一种基于内容特征的僵尸网络识别系统 AUTORE，该系统提取海量垃圾邮件中包含的 URL 信息并据此对邮件进行聚类分析，追踪可能构建于部分邮件源地址之上的僵尸网络. 这种方法基于危害结果进行反溯，针对性较强，结果一般较为可信.

4.1.4　检测技术比较分析

表 3 对各种检测技术进行了比较. 三类技术都可以实现对中心式和非中心式僵尸网络的检测，但是能力各有不同. BotHunter 等系统可以通过对本地网络的实时检测快速发现僵尸结点，但其误报率问题尚待解决. 蜜网技术可以通过捕获僵尸程序并进行逆向分析来获得僵尸网络的详细信息，检测准确率较高，但由于方式过于被动，导致检测范围受限. 而且具备蜜罐感知能力的僵尸网络[54]也对蜜网的检测能力构成严重挑战. 反向检测技术基于充足的危害数据，结果较为可信，但必须在大规模危害发生之后才能够进行检测，时效性差，且误报率较高. 综合来看，在僵尸网络检测研究方面，Gu 等人的方法虽然在误报率问题上(尤其是对非中心式僵尸网络的检测上)不及蜜网技术，但既克服了蜜网技术的被动性带来的高受限度，又能够在僵尸网络的传播和活动阶段实施检测，时效性较高，实用性较强.

表 3　僵尸网络检测技术比较

方法类型	检测系统	中心式适用	非中心式适用	低误报率	高时效性	低受限性
流量分析	BotHunter	√			√	√
	BotSniffer	√				
	BotMiner	√	√			
	BotProbe	√				
蜜网捕获	HoneyBow	√	√	√		
反向检测	AUTORE	√	√			√

4.2　测量技术

美国约翰霍普金斯大学的 Rajab 等在文献[13]中介绍了一种对僵尸网络的规模进行测量的思路：

(1) 收集恶意代码；

(2) 分析通信规律；

(3) 利用已知的通信规律对僵尸网络进行测量.

在该思路下，主要的测量方法有 DNS cache 探测、C&C 信道监测、DNS 重定向和 P2P 网络爬虫. 另外，研究人员还提出了 DNSBL、垃圾邮件数据挖掘等其他的规模测量思路. 这些测量技术可以归纳为两类，一类是内部枚举技术，即通过渗透进入僵尸网络，从其信道内部测算结点数量；一类是外部估测技术，即对僵尸网络的外部活动数据进行分析，估算网络规模.

4.2.1　内部枚举技术

内部枚举的主要途径包括 C&C 信道监测和 P2P 网络爬虫.

C&C 信道监测. 对于中心式僵尸网络，进入其 C&C 服务器信道内部可以高效枚举网络结点. Freiling 等人[7]使用了一种基于 IRC 信道监测的僵尸网络规模测量方法，该方法通过潜入 IRC 信道中观察僵尸结点的加入信息来统计网络规模. 但这种方法适用范围非常有限[30]，因为对于配置良好的 IRC 服务器，普通客户端几乎都没有权限获取僵尸结点的加入信息. 实际上在大部分情况下，C&C 信道监测方法都需要在一定程度上监控服务器，因此其实施难度较大.

P2P 网络爬虫. 对于 P2P 僵尸网络，利用网络爬虫进行结点枚举是一种有效的测量方法. Holz 等人[10]针对 Storm 僵尸网络，提出了利用 P2P 网络爬虫来获得结点信息的方法. 该爬虫利用 P2P 协议查询机制，首先通过向 P2P 僵尸网络中的结点发送路由请求，获知其他僵尸结点信息；然后反复迭代地向新获知的结点发送路由请求，就可以估测出整个网络的结点数量. 他们在 2007 年 12 月到 2008 年 1 月对 Storm 僵尸网络进行了测量，每次测量时间 30 min，发现 Storm 僵尸结点分布范围超过 200 个国家和地区，同时在线的僵尸结点约为 5000 至 40 000. 这种方法非常适用于自生式的 P2P 僵尸网络，对于寄生式的僵尸网络，还需要对正常结点进行

① http://nepenthes.carnivore.it/
② http://www.shadowserver.org/

区分.

4.2.2 外部估测技术

外部估测的主要途径包括 DNS cache 探测、DNS 重定向、DNSBL 统计和垃圾邮件数据挖掘等.

DNS cache 探测. Rajab 在对僵尸网络进行研究的过程中,提出了利用 DNS cache 探测结果对僵尸网络规模进行下限估测的方法[13]. 现实中大多数的僵尸结点在与 IRC 服务器进行连接时,都要进行域名解析. 如果使用 IRC 服务器的域名对 DNS 服务器的 cache 进行探测并命中,则说明至少有一个僵尸结点曾向该 DNS 服务器发起过请求. 通过对海量 DNS 服务器的 cache 进行探测,就可以得到 cache 命中的服务器数量,并以该数量作为对僵尸网络规模下限的估测值. 这种方法的问题在于,只适用于通过域名解析定位命令与控制服务器的中心式僵尸网络.

DNS 重定向. Dagon 等人[55]提出了通过 DNS 重定向测量僵尸网络规模的方法. 该方法首先分析捕获的僵尸程序,发现中心服务器的域名;然后通过域名服务商将对该中心服务器的域名解析重定向到特定网络位置,进而统计僵尸网络中的结点数量. 其缺陷在于,如果僵尸结点不再需要进行 DNS 解析或者相关的域名服务商不愿合作,则难以进行统计.

DNSBL 统计. Ramachandran 等人[56]提出了通过分析 DNSBL① 的查询记录来测量僵尸网络规模的方法. DNSBL 是一种垃圾邮件黑名单,登记了具有不良记录的垃圾邮件源地址,可以为用户屏蔽垃圾邮件提供帮助. 由于黑名单中的源地址都以域名方式记录,因而称为 DNSBL. 该测量方法基于这样一种假设:使用僵尸网络发送垃圾邮件的攻击者,可能会对 DNSBL 进行查询以了解僵尸结点是否在 DNSBL 中,进而调整僵尸结点的发送行为. 因此,通过对 DNSBL 的查询情况进行统计分析,就可以对僵尸网络的结点规模形成粗略估测. 这种方法的缺陷在于,无法揭示所发现的僵尸结点是否属于同一个僵尸网络,并且只有在僵尸网络用于发送垃圾邮件而攻击者进行 DNSBL 查询时,才能进行测量.

垃圾邮件数据挖掘. Zhuang 等人[57]提出了一种通过对垃圾邮件数据进行分析,发现僵尸网络并对网络规模进行计算的方法. 方法首先对海量邮件进行指纹计算及比较,实现邮件聚类(每个类可以看作一次垃圾邮件大规模发送行动);然后对聚类后的邮件源 IP 地址信息进行分析,映射出对应的僵尸网络,从而估算出僵尸网络的规模. 这种方法也可用于僵尸网络的检测.

4.2.3 测量技术比较分析

表 4 给出了各种规模测量技术的比较. 在僵尸网络规模测量方面,内部枚举技术的准确度明显高于外部估测技术.

表 4 僵尸网络规模测量技术比较

方法类型	测量技术	中心式适用	非中心式适用	高准确度	局限性
内部测量	P2P 网络爬虫		√	√	对于寄生式网络,需要识别正常结点
	C&C 信道监测	√		√	一般需要监控 C&C 服务器,实施难度较大
外部测量	DNS cache 探测	√			适用于成员结点通过 DNS 查询中心服务器的僵尸网络
	DNS 重定向	√			适用于成员结点通过 DNS 查询中心服务器的僵尸网络,且一般需要服务商支持
	DNSBL 方法	√	√		适用于发送垃圾邮件的僵尸网络,且需要攻击者查询 DNSBL 记录
	垃圾邮件数据挖掘	√	√		只适用于发送垃圾邮件的僵尸网络

对于中心式僵尸网络而言,由于 C&C 信道监测方法往往实施难度较大,因而外部估测方法更可行. 但外部估测方法中各种技术也都有其局限性. DNS cache 探测技术和 DNS 重定向技术都利用了僵尸节点的 DNS 查询行为,且常常需要网络服务商的支持,因此其适用性同时受到僵尸网络行为特征和服务商的制约. 垃圾邮件数据挖掘技术虽然同时适用于中心式与非中心式僵尸网络,但只对发送垃圾邮件的僵尸网络有效. DNSBL 方法的适用对象具有较强的特殊性,且无法区分所观察到的僵尸结点是否属于同一僵尸网络,因而实用性较差.

对于非中心式僵尸网络而言,P2P 爬虫技术受限度低、准确度高,因而综合性能最好. 但对于寄生式的僵尸网络,仍然需要解决好识别正常 P2P 结点的问题.

4.3 反制技术

检测到僵尸网络之后,如何有效反制以缓解或消除其危害,也是僵尸网络研究的一个重要方面. 当前的僵尸网络反制技术可以分为网络抑制技术和网络劫持技术两类.

① http://www.dnsbl.info

4.3.1　网络抑制技术

网络抑制技术的相关研究主要有 Sybil 攻击技术、内容污染技术和结点清除技术 3 类.

Sybil 攻击技术[42]. 对于非中心式僵尸网络，针对传统 P2P 网络的攻击技术可以起到反制作用. Sybil 攻击有两种方式，一种是内容屏蔽，一种是查询抑制. 以 Storm 网络为例，内容屏蔽技术首先向 Storm 网络中加入一定数量的 Sybil 结点，使其 DHT ID 值接近某目标关键字 Key；然后 Sybil 结点在网络中声明自己的存在，试图对多数 Storm 结点的路由表造成影响；当 Storm 结点查询该关键字 Key 时，就有可能被误导到 Sybil 结点，从而屏蔽真实搜索目标. 查询抑制是指利用 DHT 技术中的查询停止机制，令 Sybil 在收到任何查询请求时，都向请求结点返回特定信息以使其停止进一步搜索，导致请求结点无法获得目标信息. 文献[10]针对 Storm 僵尸网络，分析了内容屏蔽技术. 不过 Holz 认为，由于 Storm 网络信息发布机制不同于 KAD 网络，内容屏蔽方法对 Storm 网络的抑制效果不佳. Davis 等人[14-15]定量研究了查询抑制方法对 Storm 僵尸网络的抑制效果. 他们以模拟的方式在 Storm 网络中加入大量的 Sybil 结点，这些节点接收到来自 Storm 结点的任何搜索请求时，都返回错误的应答消息，使搜索失败. 分析结果表明，查询抑制方法可以在较大程度上抑制 Storm 僵尸网络命令与控制信息的传递. 由于非中心式系统很难解决 Sybil 问题，因此 Sybil 攻击技术始终是抑制 P2P 僵尸网络的有效手段.

内容污染技术. 内容污染技术主要用于非中心式的僵尸网络，通过覆盖目标结点中所存储的命令与控制信息，阻止命令与控制信息在僵尸网络中传播. 以 Storm 为例，为阻止以关键字 Key 为索引的命令与控制信息在网络中传播，内容污染技术首先搜索所有可能存储了该信息的结点，然后重新构造以 Key 为索引的假信息并以一定的频率反复地向这些结点发布，覆盖原有信息. Wang 等人[58]进一步研究了内容污染方法，并指出内容污染能够高效抑制 Storm 是因为 Storm 网络在搜索中体现出了一定程度的中心化. Starnberger 等人[59]提出了一种新型的结构化 P2P 僵尸网络 Overbot，解决了 Storm 网络所残留的中心化问题. Overbot 控制者针对每个僵尸结点采用不同的关键字来发布命令与控制信息，因此要对 Overbot 类型的僵尸网络进行内容污染，几乎需要向所有的僵尸结点发布污染内容. 实际上，Overbot 僵尸网络因其管理复杂和维护困难，目前几乎无法有效部署. 因此，在能够获得命令与控制信息发布规律的情况下，内容污染方法仍然是一种高效的抑制手段.

结点清除技术. 清除僵尸网络结点可以达到直接抑制僵尸网络的效果. 僵尸网络结点清除有 3 种策略：定向清除、树形清除和随机清除[31]. 定向清除按照连接度由高到低的顺序清除僵尸网络结点；这种方式对中心式僵尸网络的抑制效果显著，通过直接停止中心服务器的运转或者联系域名服务商拒绝对僵尸网络提供服务，可以达到瘫痪僵尸网络的目的. 树形清除，是防护方利用已经捕获的僵尸结点，分析出与之有连接关系的其他僵尸结点，“顺藤摸瓜”地进行结点清除. 随机清除指随机地对僵尸网络中的结点进行清除；僵尸网络的结点被宿主机用户或杀毒软件发现并查杀，是随机清除的典型情况. 各种清除策略的效果与僵尸网络采用的网络结构有关. Dagon 等人[32]从复杂网络的研究角度出发，提出了描述僵尸网络结构的效能、效率、健壮性这 3 个属性的指标-最大分支规模及带宽、网络直径、聚类系数；然后以这 3 个指标对各种网络结构进行了衡量. 研究结果表明，结构化 P2P 网络和 ER(Erdös-Rényi model)随机网络[60]对于定向清除的韧性要高于非结构化 P2P 和 BA(Barabási-Albert model)无标度网络[61]，而对于随机清除策略则相反. Davis 等人[31]进一步提出了两个新的性能指标-网络对单个结点的可达性和最短路径集分布，并通过模拟实验发现：与 ER 网络相比，结构化 P2P 网络具有更高的抗毁性.

4.3.2　网络劫持技术

网络劫持技术主要包括域名抢注、Sybil 欺骗和内容替换等技术.

域名抢注. 很多中心式僵尸网络为了提高隐蔽性，会频繁更换其中心服务器的域名，这就为劫持僵尸网络提供了机会. Stone-Gross 等人[16]通过分析 Torpig 僵尸程序，获得了其中心服务器域名变换规律；通过预测并抢注 Torpig 中心服务器将要采用的域名，使僵尸结点误将防护方的服务器作为其中心服务器，并接受其指令. 由于不少僵尸网络都设计有自卸载功能(如 AgoBot、SDbot 等)，因此在劫持这类僵尸网络后，可以通过发布自卸载命令来高效清除宿主机器上的僵尸程序[62].

Sybil 欺骗. 在 P2P 僵尸网络中，Sybil 结点如果对任何僵尸结点的搜索请求返回特意构造的信息，就有可能实现对僵尸网络的劫持.

内容替换. 在对 P2P 僵尸网络的抑制技术中，内容污染通过将被目标关键字索引的内容都覆盖掉，导致僵尸网络无法有效传递命令与控制信息. 如果使用特意构造的信息来覆盖目标内容，僵尸结点在获取并执行相应命令后，就有可能被劫持.

目前有关网络劫持技术的研究尚未涉及 Sybil 欺骗和内容替换. 但 Sybil 欺骗和内容替换将是未来网络劫持的重要手段.

4.3.3 反制技术比较分析

如表 5 所示是对僵尸网络反制技术的比较. Sybil 攻击方法阻止僵尸结点寻找存储有命令与控制信息的结点，使得僵尸网络出现“有命令，取不到”的现象，抑制了僵尸结点的命令获取能力. 内容污染通过直接破坏命令与控制信息，抑制了僵尸网络的命令发布与传递能力. 结点清除方法通过直接清除僵尸结点或操纵设施来抑制僵尸网络，破坏了僵尸网络的结点生存能力. 网络劫持技术都是通过误导僵尸结点，劫持攻击者对僵尸网络的控制权. 对于中心式僵尸网络，由于其存在单点失效问题，因此应当主要考虑用定向结点清除或域名抢注的方法实施反制. 对于非中心式的僵尸网络，虽然结点清除能够降低其网络规模，但由于其抗毁性一般较高，因此应当更多考虑用 Sybil 攻击、内容污染、Sybil 欺骗或内容替换的方法实施反制. 内容污染和内容替换需要首先分析僵尸网络的命令与控制信息发布规律，而分析发布规律具有一定的难度，因此 Sybil 攻击和 Sybil 欺骗的可行度更高. 目前对基于 Sybil 欺骗的劫持技术尚无研究，而定量研究[61-62]已证明 Sybil 攻击技术的有效性，因此对非中心式僵尸网络而言，Sybil 攻击是最可靠的反制技术. 尽管如此，目前利用 Sybil 攻击技术抑制僵尸网络的研究仍然不够充分，既缺乏对抑制效果的定量预测模型，又缺乏更加灵活的抑制策略，因此应当受到研究人员的更多关注.

表 5 僵尸网络反制技术比较

方法类型	反制技术	中心式适用	非中心式适用
网络抑制	Sybil 攻击		√
	内容污染		√
	结点清除	√	√
网络劫持	域名抢注	√	
	Sybil 欺骗		√
	内容替换		√

5 僵尸网络演化及研究趋势

基于对僵尸网络分类和关键技术研究现状的系统介绍与分析，本节提炼出了僵尸网络自身的演化趋势和僵尸网络未来研究方向. 其中，僵尸网络演化趋势是对僵尸网络自身结构和功能发展方向的预测，僵尸网络未来研究方向是对僵尸网络研究未来亟需解决的技术问题的判断.

5.1 演化趋势

僵尸网络在与安全组织和厂商的博弈过程中，为了解决隐蔽性、生存性等问题，自身也在不断地演化，从而呈现出如下的发展趋势.

非中心化. 非中心式结构因其在抗毁性方面的优势，正逐渐被越来越多的僵尸网络采用. Dittrich 等人[40]通过分析指出，在僵尸网络多年来的演化过程中，由于现有的方法对 P2P 僵尸网络的检测和抑制效果相对有限，僵尸网络逐渐趋于采用 P2P 类型的命令与控制信息传递机制. Wang 等人[58]指出，尽管当前大多数僵尸网络仍然是中心式的，但是没有中心服务器的 P2P 僵尸网络比传统的僵尸网络表现出更好的抗毁性和生存性. 更多的研究[41,63]也都将 P2P 看作下一代高级僵尸网络将会采用的机制. 另外，志愿僵尸网络和自部署僵尸网络基本上不受反病毒软件的影响，其最大的威胁在于其可能会造成单点失效问题的中心控制服务器，因此，这类僵尸网络的进一步发展，必然会走向非中心化.

小型化. 僵尸网络的规模在一定程度上呈现出缩小的趋势. 密歇根大学的 Cooke 等[64]认为，这种趋势的出现，可能是因为防御措施获得改进，使新的大规模僵尸网络难以构建. 但更可能的原因是随着带宽的不断增加，更小规模的僵尸网络完全可以胜任原先低带宽环境中需要大规模僵尸网络完成的任务，因此攻击者为了提高隐蔽性，主动地减小了单个僵尸网络的规模. 另外，对于采用 P2P 信道的僵尸网络，在网络规模扩大之后，命令与控制信息的传递效率会受到一定影响，这可能也会迫使网络规模小型化. Damballa 公司①通过对 600 多个僵尸网络长达 3 个月的监测发现，这些僵尸网络中结点规模超过 10000 的只占 5%，规模在 500 到 10000 之间的占 17%，规模在 100 到 500 之间的占 21%，而规模在 100 以下的则占到了 57%.

智能化. 未来僵尸网络在演化过程中将更具智能性. Zou 等人[54]提出了一种具有反蜜罐能力的僵尸网络. 这种僵尸网络在传播过程中，通过要求目标

① Gunter Ollmann. Botnet Size within the Enterprise. http://blog.damballa.com/?p=361

主机执行某些恶意的网络攻击行为来判断该目标是否是蜜罐；而蜜罐在设计中考虑到法律因素，一般不会执行这些攻击行为. Hund 等人[63]提出了一种具有较高智能和较强防反制能力的高级僵尸网络. 这种僵尸网络采用信誉积分体系建立僵尸结点的信誉档案，用于识别伪造结点；要求新入网结点执行复杂耗时的计算任务，阻止 P2P 爬虫的高效运行，并在一定程度上遏制 Sybil 攻击；提供对脱网僵尸结点的回收机制，使僵尸结点可以借助现有的 P2P 网络重新入网.

强化加密传输. 早期的僵尸网络大多不对传输信道进行加密，缺乏认证机制，这使僵尸网络很容易被分析、渗透和劫持. 然而在演化过程中，更多僵尸网络(如 Peacomm[65])正在采用越来越复杂的加密技术来提高其反渗透能力. Dittrich 等人[66]通过分析僵尸程序发现，Nugache 僵尸网络已经采用了非常强的加密机制，而 Storm 僵尸网络的加密机制也在不断演变. Hund 等人[63]提出的下一代僵尸网络更利用了 DH 密钥协商方法来实现僵尸结点之间的密钥交换，采用了较强的对称加密措施保护传输信道和隐藏流量特征，并通过公钥加密实现了僵尸网络的分片出租.

隔离化. 更多的僵尸网络将会采用 Fast-Flux 技术来隔离僵尸结点与命令服务器的直接通信，为服务器提供一层安全保护. Fast-Flux 技术主要可以分为两种类型[33]，一类是 IP-Flux，一类是 Domain-Flux. IP-Flux 是一种由一定数量受控主机所组成的服务网络，利用自身的 IP 地址资源池，以一定的速率轮转地对指定域名设置具有短生命周期的解析映射，从而达到为一个域名分配不断变换的多个(几百甚至上千)IP 地址的目的. Fast-Flux 网络中的大量 IP 地址一般不是内容请求的最终目的地址，而是仅仅被部署成为一层重定向代理服务器，转发僵尸结点与真正服务器之间的请求与响应. 蜜网组织[①]进一步将 IP-Flux 服务网络分为两种不同的类型：Single-Flux 和 Double-Flux. 与 Single-Flux 网络相比，Double-Flux 网络多了一个保护层，这一层通过不断地改变授权名字服务器的 IP 地址来增强保护能力. 与 IP-Flux 技术相反，Domain-Flux 技术的目的是为一个 IP 提供不断变换的多个域名. 僵尸网络在 Fast-Flux 技术上的发展趋势体现在以下两个方面. 一方面是僵尸网络通过改造自身功能，对外提供 Fast-Flux 服务. Holz 等人指出，在 2007 年 6 月 Storm 僵尸网络组织被发现时，该组织正在对 Storm 进行改进以使之具备 Fast-Flux 服务能力. Hu 等人[67]利用在边界路由器上的 Netflow 数据，对提供 Fast-Flux 服务的僵尸网络进行了研究，而 Perdisci 等人[8]的工作进一步指出，Fast-Flux 服务网络的结点通常以僵尸网络功能模块的形态出现. 另一方面，僵尸网络也会利用 Fast-Flux 技术来完成自身组网. Stone-Gross 等人[16]在对 Torpig 僵尸网络的研究过程中发现，该僵尸网络采用了Domain-Flux 技术来提高自身的生存能力，而 Waledac 更在对外提供 IP-Flux 服务的同时，利用自身的 Fast-Flux 机制对僵尸结点进行检查，以避免虚假结点的加入[44].

移动化. 移动计算技术的蓬勃发展为僵尸网络提供了新的生存空间，基于移动平台构建的僵尸网络从种类和数量上都呈现出快速增长态势. SymbOS. Yxes[68]通过感染塞班平台构建了一个通过 HTTP 服务器进行控制的僵尸网络；Geinimi[②] 通过感染安卓平台构建了一个中心式的僵尸网络；甚至在较为封闭的苹果平台上，iKee. B[37]也利用 iPhone 越狱造成的安全漏洞，通过扫描传播构建了大规模的僵尸网络. Mulliner 等人[69]和 Traynor 等人[18]已经实证性地研究了移动僵尸网络对移动网络的安全所能造成的严重危害. 很多安全研究人员从不同角度对移动僵尸网络的发展做出了预测. Dimitrios 等人[70]提出了基于 iPhone 的移动僵尸网络高级构建技术 iSAM. 崔翔等人[71]设计了基于 HTTP 和 URL-Fulx 技术的中心化移动僵尸网络 Andbot. Weidman 等人[20]基于 SMS 设计了一套推送式移动僵尸网络的组网技术；Geng 等人[72]基于 SMS 设计了一种层次式移动僵尸网络的组网技术；Zeng 等人[19]基于 SMS 设计了构建 P2P 结构移动僵尸网络的方法. Singh 等人[21]研究了基于蓝牙技术构建移动僵尸网络控制信道的可行性. 2012 年初，赛门铁克公司指出中国境内有超过 10 万部安卓手机感染了 Andrio. Bmaster 僵尸程序[③]，进一步说明移动僵尸网络的发展趋势已经十分明显.

5.2　未来研究方向

结合僵尸网络当前的研究现状及其演化趋势，

① The Honeynet Project. Know your enemy：Fast-flux servicenetworks. http：//www. honeynet. org/papers/ff

② Tim Wyatt. Security Alert：Geinimi，Sophisticated New Android Trojan Found in Wild. http：//blog. mylookout. com/blog/2010/12/29/geinimi_trojan/

③ Cathal Mullaney，Android. Bmaster：A Million-Dollar Mobile Botnet. http：//www. symantec. com/connect/blogs/androidbmaster-million-dollar-mobile-botnet

本文认为僵尸网络的未来研究趋势可从如下三类传统研究方向中提炼概括,具体包括僵尸网络认识型研究新技术、僵尸网络防护型研究新技术和僵尸网络攻击型研究新技术.

5.2.1 认识型研究新技术

认识型研究新技术主要包括传播模型技术、协议模型自动构建技术和僵尸网络性能评价技术等.

传播模型构建技术.僵尸网络传播模型刻画僵尸网络生长发展的行为规律,为僵尸网络防护型研究提供指导,有助于把握僵尸网络的本质特点.目前对僵尸网络传播模型的研究中最重要的工作是Dagon等人[55]基于时区的传播模型,但其仅适用于扫描传播的僵尸网络.僵尸网络的传播途径是灵活多样的,除扫描传播外,垃圾邮件传播、网站挂载传播、植入盗版软件传播和操纵软件下载站点的下载评分进行传播等也都是重要途径.针对这些途径的传播模型技术尚有待研究.

协议模型自动逆向构建技术.协议模型以形式化方法描绘僵尸结点活动和命令控制信息之间的关系,是认识和分析僵尸网络活动规律的重要手段,能够为检测、测量和反制提供重要的指导.僵尸网络泛滥肆虐,变种众多,需要协议模型自动逆向构建技术来为快速认清僵尸网络的控制规律提供手段.Cho等人[73]基于在线测试和自动协议分析技术,针对MegaD僵尸网络设计实现了一个协议分析引擎,逆向构建了MegaD的Mealy[74]自动机模型.但该系统仅仅针对特定僵尸网络,具有更高通用性和更高自动化程度的僵尸网络协议模型自动逆向构建平台,是未来研究的一个重要方向.

僵尸网络性能评价技术.僵尸网络性能评价主要指从可用性、隐藏性、鲁棒性等各个方面对僵尸网络危害能力进行综合度量,为僵尸网络威胁评估、反制效果评估等提供支撑.Dagon等人[32]指出,仅通过规模来衡量僵尸网络的性能是远远不够的,因此构建综合指标评价僵尸网络是一个重要的研究方向.但是这方面的现有研究较少,而且已有成果大多从复杂网络抗毁性角度进行分析[31-32],鲜有从僵尸网络特有属性出发的性能评价技术研究.

5.2.2 防护型研究新技术

防护型研究新技术包括全球范围的监测技术、针对寄生式僵尸网络的规模测量技术和借助Sybil欺骗反制僵尸网络的新技术等.

全球范围的监测技术.全球范围的检测系统要通过对整个互联网中的网络活动进行监测和汇聚,从中找出僵尸网络.非中心化是僵尸网络结构演化的重要趋势,会降低单个结点因流量异常而被检测的概率.如果非中心僵尸网络进一步采用反蜜罐技术,并寄生在正常P2P网络中,则基于主机的检测难度会进一步增大.因此,僵尸网络检测的总体思路应当放在充分利用僵尸结点活动的时空相似性上,最大范围地汇聚其恶意性,克服其隐蔽性,而全球范围的监测系统是一个重要选择.

针对寄生式僵尸网络的规模测量技术.在对P2P僵尸网络的规模测量技术中,DNSBL方法实用性较差,垃圾邮件数据挖掘方法准确度较低,P2P网络爬虫技术综合效果最好.虽然P2P网络爬虫技术可以对独立式P2P僵尸网络进行较为准确的规模测量,但对于寄生式的P2P僵尸网络,还需要解决如何区分正常结点与僵尸结点的问题.

借助Sybil结点反制僵尸网络的技术.对于中心化的僵尸网络,往往可以在ISP的支持下,采用关闭中心服务器的方法进行反制.但对非中心化的僵尸网络,目前主要的反制手段仍然是通过各种恶意代码查杀工具进行清除,时效性差,不确定性高.进一步,随着志愿僵尸网络和自部署僵尸网络的出现,防护方必须要在无法清除僵尸结点的前提下对僵尸网络实施反制.因此,如何对非中心化的僵尸网络进行高效反制,成为一个非常重要的课题.考虑到P2P网络协议较强的扩展性所带来的开放特点,利用Sybil结点对P2P僵尸网络进行渗透,并通过发布虚假消息抑制僵尸结点间的通信、甚至劫持僵尸网络的控制权,应为反制技术研究的一个重要方向.

5.2.3 攻击型研究新技术

本文以研究僵尸网络防御技术为主,但僵尸网络攻击技术的研究仍然是一个重要方向.深度把握僵尸网络的攻击技术研究趋势可以为研究僵尸网络防御技术提供支持.根据本文对僵尸网络内涵的把握和演化趋势的分析,我们认为如下两类攻击型技术将会对僵尸网络的未来演化产生重要影响:

僵尸网络协同技术.僵尸网络的协同技术可以分为两个层面,一是受控结点间协同技术,即同一僵尸网络内部结点间分工协作完成任务的技术;二是僵尸网络间协同技术,即不同僵尸网络间分工协作完成任务的技术.目前僵尸网络结点间协同主要体现为操纵设施与僵尸结点之间的信息传递,但能够胜任更复杂攻击任务的僵尸网络要求在设计中进一步细分角色,这需要更先进的结点间协同技术的支持.另外,文献[75]指出,在僵尸网络小型化趋势下,

大规模僵尸网络倾向于采用协同管理模式：整个僵尸网络划分为若干小网络，各自构建独立的命令与控制机制，同时接受攻击者统一控制. 这使得网络间协同技术变得非常重要.

僵尸网络先进组网技术. 僵尸网络的非中心化会大幅提高其抗毁性，但 Wang 等人[58]的研究指出，目前 P2P 僵尸网络在命令与控制机制中仍然存在中心化现象，这严重威胁了僵尸网络的生存性和可用性. Overbot[59]试图解决这个问题，然而其实用性较差. 因此，完全非中心化且实用性强的先进组网技术，很可能会成为攻击技术研究者下一步的重点关注方向.

除上述技术因素之外，未来还应当从道德、法律方面，为降低僵尸网络的危害提供重要手段. 著名科幻作家阿西莫夫曾提出机器人三定律，为机器人的设计指出了伦理边界. “盗亦有道”，僵尸网络作为一种人造的受控网络，也应当对其在伦理上进行限定，从而为法律约束提供依据，以此从一定程度上通过威慑来保护重要的网络基础设施不被僵尸网络破坏.

6　结束语

僵尸网络作为一种复杂、灵活和高效的网络攻击平台，对个人、企业和国家等在各个层面构成了巨大的、潜在威胁. 目前，僵尸网络已引起了网络安全研究领域的高度关注，并已经成为研究热点. 本文首先深入剖析了僵尸网络概念的内涵和外延，提出了僵尸网络的新定义，并从网络结构、网络独立性和信息传递方式等维度对僵尸网络进行了划分；然后系统分析了僵尸网络检测技术、测量技术和反制技术等三类关键技术的最新研究进展；最后探讨了僵尸网络自身结构和功能方面的演化趋势，并结合僵尸网络的本质内涵和关键技术的最新研究进展等，提炼出了僵尸网络的未来研究方向，包括传统技术的新方向和小型化、智能化等僵尸网络演化趋势下的新技术.

参　考　文　献

[1] Chiang Ken, Lloyd Levi. A case study of the rustock rootkit and spam bot//Proceedings of the 1st Workshop on Hot Topics in Understanding Botnets. Cambridge, USA, 2007: No. 11

[2] Thonnard Olivier, Dacier Marc. A strategic analysis of spam botnets operations//Proceedings of the 8th Annual Collaboration, Electronic messaging, Anti-Abuse and Spam Conference. Redmond, USA, 2011: 162-171

[3] Kokkodis Marios, Faloutsos Michalis. Spamming botnets: Are we losing the war?//Proceedings of the 6th Conference on Email and AntiSpam. Moutain View, USA, 2009: No. 16

[4] Bailey Michael, Cooke Evan, Jahanian Farnam, Xu Yunjing, Karir Manish. A survey of botnet technology and defenses//Proceedings of the Cybersecurity Applications & Technology Conference for Homeland Security. Washington, USA, 2009: 299-304

[5] Gilou Tenebro. W32. waledac threat analysis. Cupertino, CA, USA: Symantec Corporation, Technical Report: W32_Waledac, 2009

[6] Kshetri N. The economics of click fraud//Proceedings of the 31st IEEE Symposium on Security & Privacy. Oakland, USA, 2010: 45-53

[7] Freiling Felix C, Holz Thorsten, Wicherski Georg. Botnet tracking: Exploring a root-cause methodology to prevent distributed denial-of-service attacks//Proceedings of the 10th European Symposium on Research in Computer Security. Milan, Italy, 2005: 319-335

[8] Perdisci Roberto, Corona Igino, Dagon David, Lee Wenke. Detecting malicious flux service networks through passive analysis of recursive DNS traces//Proceedings of the 25th Annual Computer Conference Security Applications. Honolulu, USA, 2009: 311-320

[9] Schuchard Max, Mohaisen Abedelaziz, Vasserman Eugene Y. Losing control of the internet: Using the data plane to attack the control plane//Proceedings of the 17th ACM Conference on Computer and Communications Security. Chicago, USA, 2010: 726-728

[10] Holz Thorsten, Steiner Moritz, Dahl Frederic, Biersack Ernst, Freiling Felix. Measurements and mitigation of peer-to-peer-based botnets: A case study on storm worm//Proceedings of the 1st USENIX Workshop on Large-Scale Exploits and Emergent Threats. San Francisco, USA, 2008: No. 9

[11] Holz Thorsten. Tracking and mitigation of malicious remote control networks[Ph. D. dissertation]. University Mannheim, Mannheim, Germany, 2009

[12] Gu Guofei, Porras Phillip, Yegneswaran Vinod, Fong Martin, Lee Wenke. BotHunter: Detecting malware infection through IDS-driven dialog correlation//Proceedings of the 16th USENIX Security Symposium. Boston, USA, 2007: 167-182

[13] Rajab Moheeb Abu, Zarfoss Jay, Monrose Fabian, Terzis Andreas. A multifaceted approach to understanding the botnet phenomenon//Proceedings of the 6th ACM SIGCOMM Conference on Internet Measurement. Rio de Janeriro, Brazil, 2006: 41-52

[14] Davis Carlton R, Fernandez Jose M, Neville Stephen, McHugh John. Sybil attacks as a mitigation strategy against the storm botnet//Proceedings of the 3rd Internal Conference on Malicious and Unwanted Software. Alexandria, USA, 2008: 32-40

[15] Davis Carlton R, Fernandez Jose M, Neville Stephen. Optimising sybil attacks against P2P-based botnets//Proceedings

of the 4th International Conference on Malicious and Unwanted Software. Montreal, Canada, 2009: 78-87

[16] Stone-Gross Brett, Cova Marco, Cavallaro Lorenzo, Gilbert Bob, Szydlowski Martin, Kemmerer Richard, Kruegel Christopher, Vigna Giovanni. Your botnet is my botnet: Analysis of a botnet takeover//Proceedings of the 16th ACM Conference on Computer and Communications Security. Chicago, USA, 2009: 635-647

[17] Reed Theodore, Geis Joseph, Dietrich Sven. SkyNET: A 3G-enabled mobile attack drone and stealth botmaster//Proceedings of the 5th USENIX Workshop on Offensive Technologies. San Francisco, USA, 2011: 4-13

[18] Traynor Patrick, Lin Michael, Ongtang Machigar, Rao Vikhyath, Jaeger Trent, McDaniel Patrick, La Thomas Porta. On cellular botnets: Measuring the impact of malicious devices on a cellular network core//Proceedings of the 16th ACM Conference on Computer and Communications Security. Chicago, USA, 2009: 223-234

[19] Zeng Yuanyuan, Hu Xin, Kang G S. Design of SMS commanded-and-controlled and P2P-structured mobile botnets. Ann Arbor, Michigan, USA: University of Michigan, Technical Report: CSE-TR-562-10, 2010

[20] Weidman Georgia. Transparent botnet command and control for smartphones over SMS//Proceedings of the Conference Shmoocon. Washington, USA, 2011: No. 41

[21] Singh Kapil, Sangal Samrit, Jain Nehil, Traynor Patrick, Lee Wenke. Evaluating bluetooth as a medium for botnet command and control//Proceedings of the International Conference on Detection of Intrusions and Malware, and Vulnerability Assessment. Bonn, Germany, 2010: 61- 80

[22] Steve Mansfield-Devine. Anonymous: Serious threat or mere annoyance? Network Security, 2011, (1): 4-10

[23] Pras Aiko, Sperotto Anna, Moura Giovane C M, Drago Idilio, Barbosa Rafael, Sadre Ramin, Schmidt Ricardo, Hofstede Rick. Attacks by "anonymous" wikileaks proponents not anonymous. University of Twente, Dutch: CTIT Technical Report: 10.41, 2010

[24] Clark Kassidy, Warnier Martijn, Brazier Frances M T. Botclouds the future of cloud-based botnets? //Proceedings of the International Conference on Cloud Computing and Services Science. Noordwijkerhout, Netherlands, 2011: 597-603

[25] Thomas Kurt, Nicol David M. The Koobface botnet and the rise of social malware//Proceedings of the 5th International Conference on Malicious and Unwanted Software. Nancy, France, 2010: 63-70

[26] Boshmaf Yazan, Muslukhov Ildar, Beznosov Konstantin, Ripeanu Matei. The socialbot network: When bots socialize for fame and money//Proceedings of the 27th Annual Computer Security Applications Conference. Orlando, Florida, USA, 2011: 93-102

[27] Ratkiewicz Jacob, Conover Michael, Meiss Mark, Gonçalves Bruno, Patil Snehal, Flammini Alessandro, Menczer Filippo. Truthy: Mapping the spread of astroturf in microblog streams//Proceedings of the 20th International Conference Companion on World Wide Web. Hyderabad, India, 2011: 249-252

[28] Gu Guofei, Yegneswaran Vinod, Porras Phillip, Stoll Jennifer, Lee Wenke. Active botnet probing to identify obscure command and control channels//Proceedings of the 2009 Annual Computer Security Applications Conference. Honolulu, USA, 2009: 241-253

[29] Zhuge Jian-Wei, Han Xin-Hui, Zhou Yong-Lin, Ye Zhi-Yuan, Zou Wei. Research and development of botnets. Journal of Software, 2008, 19(3): 702-715(in Chinese)
(诸葛建伟, 韩心慧, 周勇林, 叶志远, 邹维. 僵尸网络研究与进展. 软件学报, 2008, 19(3): 702-715)

[30] Zhu Zhaosheng, Lu Guohan, Chen Yan, Fu Zhi Judy, Roberts Phil, Han Keesook. Botnet research survey//Proceedings of the 32nd Annual IEEE International Computer Software and Applications Conference. Turku, Finland, 2008: 967-972

[31] Davis Carlton R, Neville Stephen, Fernandez Jose M, Robert Jean-Marc, McHugh John. Structured peer-to-peer overlay networks: Ideal botnets command and control infrastructures? //Proceedings of the 13th European Symposium on Research in Computer Security. Malaga, Spain, 2008: 461-480

[32] Dagon David, Gu Guofei, Lee Christopher P, Lee Wenke. A taxonomy of botnet structures//Proceedings of the 23rd Annual Computer Security Applications Conference. Miami, USA, 2007: 325-339

[33] Ollmann Gunter. Botnet communication topologies. Atlanta, GA: Damballa Inc, Technical Report: 2009-06-04, 2009

[34] Leder Felix, Werner Tillmann, Martini Peter. Proactive botnet countermeasures — An offensive approach//Proceedings of the 1st CCDCoE Conference on Cyber Warfare. Tallinn, Estonia, 2009: 211-225

[35] Schiller Craig A et al. Botnets: The Killer Web App. Waltham United States: Syngrass Publishing, 2007

[36] Cho Chia Yuan, Caballero Juan, Grier Chris, Paxson Vern, Song Dawn. Insights from the inside: A view of botnet management from infiltration//Proceedings of the 3rd USENIX Workshop on Large-Scale Exploits and Emergent Threats. San Jose, USA, 2010: No. 2

[37] Porras Phillip, Saidi Hassen, Yegneswaran Vinod. An analysis of the iKee.B iPhone botnet. Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering, 2010, 47(5): 141-152

[38] Singh Kapil, Srivastava Abhinav, Giffin Jonathon, Lee Wenke. Evaluating email feasibility for botnet command and control//Proceedings of the 38th Annual IEEE/IFIP International Conference on Dependable Systems and Networks. Anchorage, USA, 2008: 376-385

[39] Grizzard Julian B, Sharma Vikram, Nunnery Chris, Dagon David. Peer-to-peer botnets: Overview and case study//Proceedings of the 1st Workshop on Hot Topics in Understanding Botnets. Cambridge, USA, 2007: No. 1

[40] Dittrich David, Dietrich Sven. P2P as botnet command and control: A deeper insight//Proceedings of the 3rd International Conference on Malicious and Unwanted Software. Fairfax, USA, 2008: 46-63

[41] Wang Ping, Sparks Sherri, Zou Cliff C. An advanced hybrid peer-to-peer botnet//Proceedings of the 1st Workshop on Hot Topics in Understanding Botnets. Cambridge, USA, 2007: No. 2

[42] Douceur John R. The sybil attack//Proceedings of the 1st International Workshop on Peer-to-Peer System. Cambridge, USA, 2002: 251-260

[43] Vogt Ryan, Aycock John, Michael J Jacobson Jr. Army of botnets//Proceedings of the 14th Annual Network and Distributed System Security Symposium. San Diego, USA, 2007: 111-123

[44] Nunnery Chris, Sinclair Greg, Kang Brent Byung Hoon. Tumbling down the rabbit hole: Exploring the idiosyncrasies of botmaster systems in a multi-tier botnet infrastructure//Proceedings of the USENIX Workshop on Large-Scale Exploits and Emergent Threats. San Jose, USA, 2010: No. 1

[45] Shin Seungwon, Gu Guofei. Conficker and beyond: A large-scale empirical study//Proceedings of the 2010 Annual Computer Security Applications Conference. Austin, USA, 2010: 151-160

[46] Gu Guofei. Correlation-based botnet detection in enterprise networks[Ph. D. dissertation]. Atlanta, Georgia, USA: Georgia Institute of Technology, 2008

[47] Gu Guofei, Zhang Junjie, Lee Wenke. BotSniffer: Detecting botnet command and control channels in network traffic//Proceedings of the 15th Annual Network and Distributed System Security Symposium. San Diego, USA, 2008: No. 17

[48] Gu Guofei, Perdisci Roberto, Zhang Junjie, Lee Wenke. BotMiner: Clustering analysis of network traffic for protocol- and structure-independent botnet detection//Proceedings of the 17th USENIX Security Symposium. San Jose, USA, 2008: 139-154

[49] Gu Guofei, Yegneswaran Vinod, Porras Phillip, Stoll Jennifer, Lee Wenke. Active botnet probing to identify obscure command and control channels//Proceedings of the 2009 Annual Computer Security Applications Conference. Honolulu, USA, 2009: 241-253

[50] Zhuge Jian-Wei, Han Xin-Hui, Zhou Yong-Lin, Song Cheng-Yu, Guo Jin-Peng, Zou Wei. HoneyBow: An automated malware collection tool based on the high-interaction honeypot principle. Journal on Communications, 2007, 28(12): 8-13(in Chinese)
(诸葛建伟，韩心慧，周勇林，宋程昱，郭晋鹏，邹维. HoneyBow：一个基于高交互式蜜罐技术的恶意代码自动捕获器. 通信学报，2007，28(12)：8-13)

[51] Husna Husain, Phithakkitnukoon Santi, Palla Srikanth, Dantu Ram. Behavior analysis of spambotnets//Proceedings of the 3rd International Conference on Communication Systems Software and Middleware and Workshops. Bangalore, India, 2008: 246-253

[52] Sroufe Paul, Phithakkitnukoon Santi, Dantu Ram, Cangussu Joao. Email shape analysis for spam botnet detection//Proceedings of the 6th IEEE Consumer Communications and Networking Conference. LasVegas, USA, 2009: 1-2

[53] Xie Yinglian, Yu Fang, Achan Kannan, Panigrahy Rina, Hulten Geoff, Osipkov Ivan. Spamming botnets: Signatures and characteristics//Proceedings of the ACM Annual Conference of the Special Interest Group on Data Communication. Seattle, USA, 2008: 171-182

[54] Zou Cliff C, Cunningham Ryan. Honeypot-aware advanced botnet construction and maintenance//Proceedings of the International Conference on Dependable Systems and Networks. Philadelphia, USA, 2006: 199-208

[55] Dagon David, Zou Cliff, Lee Wenke. Modeling botnet propagation using time zones//Proceedings of the 13th Annual Network and Distributed System Security Symposium. San Diego, USA, 2006: No. 15

[56] Ramachandran Anirudh, Feamster Nick, Dagon David. Revealing botnet membership using dnsbl counter-intelligence//Proceedings of the Workshop on Steps to Reducing Unwanted Traffic on the Internet. San Jose, USA, 2006: 49-54

[57] Zhuang Li, Dunagan John, Simon Daniel R, Wang Helen J, Tygar J D. Characterizing botnets from email spam records//Proceedings of the USENIX Workshop on Large-Scale Exploits and Emergent Threats. San Francisco, USA, 2008: 1-9

[58] Wang Ping, Wu Lei, Aslam Baber, Zou Cliff C. A systematic study on peer-to-peer botnets//Proceedings of the International Conference on Computer Communications and Networks. San Francisco, USA, 2009: 1-8

[59] Starnberger Guenther, Kruegel Christopher, Kirda Engin. Overbot—A botnet protocol based on kademlia//Proceedings of the 4th International Conference on Security and Privacy in Communication Networks. Istanbul, Turkey, 2008: 13-22

[60] Erdös P, Rényi A. On the evolution of random graphs. Publication of the mathematical institute of the Hungarian Academy of Sciences, 1960, (5): 17-61

[61] Barabási A L, Albert R. Emergence of scaling in random networks. Science, 1999, 286(5439): 509-512

[62] Vinoo Thomas, Nitin Jyoti. Bot countermeasures. Journal in Computer Virology, 2007, 3(2): 103-111

[63] Hund Ralf, Hamann Matthias, Holz Thorsten. Towards next-generation botnets//Proceedings of the 4th Annual European Conference on Computer Network Defense. Dublin, Ireland, 2008: 33-40

[64] Cooke Evan, Jahanian Farnam, McPherson Danny. The zombie roundup: Understanding, detecting, and disrupting botnets//Proceedings of the Workshop on Steps to Reducing Unwanted Traffic on the Internet. Cambridge, USA, 2005: 39-44

[65] Feng Yong-Liang. Research on the detection of structured P2P botnets[M. S. dissertation]. Huazhong University of Science and Technology, Wuhan, 2008(in Chinese)
(冯永亮. 结构化 P2P 僵尸网络检测技术的研究[硕士学位论文]. 华中科技大学，武汉，2008)

[66] Stover Sam, Dittrich Dave, Hernandez John, Dietrich Sven. Analysis of the storm and nugache trojans: P2P is here. USENIX login, 2007, 32(6): 18-27

[67] Hu Xin, Knysz Matthew, Shin Kang G. Rb-seeker: Auto-detection of redirection botnets//Proceedings of the 16th Annual Network & Distributed System Security Symposium. San Diego, USA, 2009: No. 10

[68] Apvrille Axelle. Symbian worm yxes: Towards mobile botnets?//Proceedings of the 19th EICAR Annual Conference. Paris, France, 2010: 31-54

[69] Mulliner Collin, Golde Nico, Seifert Jean-Pierre. SMS of death: From analyzing to attacking mobile phones on a large scale//Proceedings of the 20th USENIX Conference on Security. San Francisco, USA, 2011: 24-40

[70] Damopoulos Dimitrios, Kambourakis Georgios, Gritzalis Stefanos. iSAM: An iphone stealth airborne malware//Proceedings of the 26th IFIP International Information Security Conference. Lucerne, Switzerland, 2011: 17-28

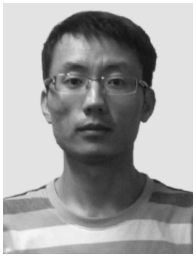
[71] Cui Xiang, Fang Binxing, Yin Lihua, Liu Xiaoyi, Zang Tianning. Andbot: Towards advanced mobile botnets//Proceedings of the 4th USENIX Workshop on Large-Scale Exploits and Emergent Threats. Boston, USA, 2011: 11-18

[72] Geng Guining, Xu Guoai, Zhang Miao, Guo Yanhui. An improved SMS based heterogeneous mobile botnet model. Journal of Computers, 2012, 7(1): 235-243

[73] Cho Chia Yuan, Babic Domagoj, Shin Eui Chul Richard, Song Dawn. Inference and analysis of formal models of botnet command and control protocols//Proceedings of the 17th ACM Conference on Computer and Communications Security. Chicago, USA, 2010: 426-439

[74] Mealy George H. A method for synthesizing sequential circuits. Bell System Technical Journal, 1955, 34(5): 1045-1079

[75] Li Run-Heng. Analysis technologies research on botnets in large scale network[Ph. D. dissertation]. National University of Defense Technology, Changsha, 2011(in Chinese)
(李润恒. 大规模网络中僵尸网络分析技术研究[博士学位论文]. 国防科学技术大学, 长沙, 2011)



WANG Tian-Zuo, born in 1982, Ph. D. candidate. His current research interests include distributed computing and information security.

WANG Huai-Min, born in 1962, Ph. D., professor. His research interests include distributed computing, information security and computer software.

LIU Bo, born in 1973, Ph. D., associate professor. His research interests include distributed computing and information security.

SHI Pei-Chang, born in 1981, Ph. D. candidate. His current research interests include distributed computing and information security.

Background

The botnet is actually a complex, flexible and effective platform for the controller to launch many kinds of large-scale malicious activities in the Internet. Today, more and more botnets are found in the Internet, and a sophisticated chain of interests has already been established around it. So botmasters can make profits from renting it, which can be used to offer Fast-Flux service, launch spamming campaigns, launch DDoS attacks and make fraud clicking attacks, et al. Botnets are posing serious threats to individuals, enterprises and countries, so that great attentions are aroused in both academic and industrial communities. However, there is still a long way to solve this problem well. Firstly, the development of botnets is challenging the existing concepts with the emergence of volunteer botnets and self owned botnets. Secondly, the taxonomy of botnets needs to be further analyzed to get better understanding of botnets. Thirdly, the detection, measurement and mitigation techniques are still not enough to counter attack botnets. Lastly, we must try to predict the evolving trend of botnets so that we can identify which technologies should be pay more attention. We call the four aspects of the research of botnets the four critical problems.

In this paper, we mainly reviewed the existing work related to this topic. This paper also contains some original insights, such as the definition of botnet, taxonomy of botnet and a set of research directions we believe to be the most important ones in the near future. We hope that they can help and inspire the researchers and practitioners who concern this issue. This work is supported by the National Natural Science Foundation of China under Grant No. 90818028: "Behavior Monitoring and Trustworthiness-Oriented Evolution of Large-Scale Distributed Software Systems". This project aims at the methodology and a set of necessary techniques for online evolution of software services. This work is also supported by National Basic Research Program (973 Program) of China under Grant No. 2011CB302600: "Basic Research on Effective and Trustworthy Internet-Based Virtual Computing Environment (iVCE)", whose purpose is to design basic models and mechanisms for effective and trustworthy virtual computing environment. This work is also supported by the National Science Fund for Distinguished Young Scholars under Grant No. 60625203: "Computer Software", whose purpose is to design basic method and mechanisms for trustworthy software.