计算机研究与发展　　ISSN 1000-1239/CN 11-1777/TP
Journal of Computer Research and Development　　48(8):1488-1496,2011

MRRbot:基于冗余机制的多角色P2P僵尸网络模型

李书豪[1,2]　云晓春[1,3]　郝志宇[1]　翟立东[1]

[1](中国科学院计算技术研究所信息安全研究中心　北京　100190)

[2](中国科学院研究生院　北京　100049)

[3](国家计算机网络应急技术处理协调中心　北京　100029)

(lishuhao0325@gmail.com)

MRRbot: A Multi-Role and Redundancy-Based P2P Botnet Model

Li Shuhao[1,2], Yun Xiaochun[1,3], Hao Zhiyu[1], and Zhai Lidong[1]

[1](*Information Security Research Center, Institute of Computing Technology, Chinese Academy of Sciences, Beijing* 100190)

[2](*Graduate University of Chinese Academy of Sciences, Beijing* 100049)

[3](*National Computer network Emergency Response Technical Team/Coordination Center of China, Beijing* 100029)

Abstract　As common platforms of cyber attacks, botnets cause great damage and bring serious threats. Though the defenses against current botnets are effective, botnets' evolution gives defenders a big challenge, which is worse with the development of tri-network integration. Therefore, it is indispensable to predict future botnets for timely defense. In this paper, we summarize the weaknesses of existing botnets, and present the design of a mutli-role and redundancy-based P2P botnet model (MRRbot). In this model, fake bots are created to be an important role that can help enhance bots' credibility and pertinence, and a redundancy mechanism and a selection algorithm are designed to improve the invisibility and robustness of the command and control channel. Furthermore, MRRbot is analyzed and evaluated on its controllability, efficiency, invulnerability, and its robustness is compared with others previous work. Both theoretical analysis and experimental results demonstrate that MRRbot's botmasters can quickly publish commands to each bot with the probability close to 100%, even suffering effective defenses. MRRbot is more dangerous with high controllability, efficiency, robustness and invulnerability, which is likely to be adopted by attackers in the future. Finally, a defense system against this advanced botnet, which is based on the volunteer network, is suggested.

Key words　network security; botnet; command and control; redundancy; multi-role

摘　要　对现有僵尸网络的防御已取得很大成效,但僵尸网络不断演变进化,尤其在三网融合不断推进的背景下,这给防御者带来新的挑战.因此,预测未来僵尸网络以及时应对,非常必要.提出了一种基于冗余机制的多角色P2P僵尸网络模型(MRRbot),该模型引入虚壳僵尸终端,能够很大程度上验证僵尸终端的软硬件环境,增强其可信度和针对性;采用信息冗余机制和服务终端遴选算法,使僵尸终端能够均衡、高效地访问服务终端,提高命令控制信道的健壮性和抗毁性.对MRRbot的可控性、时效性和抗毁性进行了理论分析和实验评估,并就其健壮性与前人工作进行了比较.结果表明,MRRbot能够高效下发指令,有效对抗防御,更具威胁.探讨了可能的防御策略,提出基于志愿者网络的防御体系.

关键词　网络安全;僵尸网络;命令控制;冗余机制;多角色

中图法分类号　TP393

收稿日期:2011-04-07;**修回日期**:2011-06-16

基金项目:国家自然科学基金项目(61003261);国家"八六三"高技术研究发展计划基金项目(2007AA010501)

僵尸网络(botnet)作为能够发起多种攻击的平台，已成为互联网安全的最大威胁之一. Arbor Networks 公司 2010 年 1 月发布的安全报告显示[1]：僵尸网络是全球互联网服务提供商(Internet Service Provider, ISP)的第二大运营威胁，位列分布式拒绝服务(distributed denial of service, DDoS)攻击之后，而 DDoS 主要是由僵尸网络发起的.

多年来，僵尸网络一直是信息安全领域的研究热点. 现有防御手段一定程度上减缓了僵尸网络的传播和危害，然而僵尸网络不断演变进化，结构从集中式发展到分布式，再到混合式；命令控制信道载体从简单的 IRC 协议发展到隐蔽的 HTTP 协议，再到复杂的 P2P 协议，使传统防御逐渐失效，给未来防御带来新的挑战. 僵尸网络已趋于复杂化和多样化，混合结构 P2P 模式是其主要演变趋势. 在三网融合不断推进和手机网民快速增长的背景下，攻击者很可能利用新型系统和应用漏洞，构建新型僵尸网络. 因此，防御者应把握僵尸网络的发展趋势，预先提出防御策略，防范于未然.

未来僵尸网络预测，作为僵尸网络重要研究方向，是从攻击者角度研究如何构建高对抗性僵尸网络，并提出防御对策. Vogt 等人[2]基于"network of botnets"思想，提出层次化的超级僵尸网络(super-botnet)模型，该模型具备了 P2P 的雏形，但传播上缺乏可控性. Starnberger 等人[3]进而提出一种基于 Kademlia 协议的僵尸网络命令控制协议(Overbot)，构建出隐蔽的命令控制信道，但其通信能力有限. Wang[4]和 Ying 等人[5]的工作更进一步，他们分析了攻击者需求，分别提出混合结构的 P2P 僵尸网络模型，这些模型具有较好的隐蔽性，难以被劫持和关闭，但无法探测蜜罐和对抗 P2P 主动防御①. Wang 等人[6]着眼于僵尸网络生存性，提出一种基于 UserID 的可重构僵尸网络体系，但其重构过程具有时空相似性，易被检测. 随着智能手机和移动网络的发展，手机僵尸网络(mobile botnet)引起了防御者的关注. Traynor 等人[7]提出一种 Cellular Botnet 攻击模型，能够利用多种命令控制信道，有效避免网络瓶颈，达到攻击目的. Singh 等人[8]提出一种手机僵尸网络模型，利用蓝牙(bluetooth)构建更为隐蔽的命令控制信道，能够有效对抗过滤检测防御，但这些模型的可控性和时效性均比较差.

从上述情况可以看出，现有僵尸网络模型还存在以下弱点：1)缺乏对僵尸终端的特征检验机制；2)命令控制消息的隐蔽性、完整性不能够得到保证；3)无法有效对抗 P2P 主动防御手段.

本文提出一种基于冗余机制的多角色 P2P 僵尸网络模型(multi-role and redundancy-based P2P botnet, MRRbot). 类似 MRRbot 的新型僵尸网络很可能在未来的网络中出现，带来更大危害. 本文旨在增加防御者对此类僵尸网络的了解，以及时应对. 本文的主要贡献如下：

1) 提出虚壳僵尸终端思想. FBot 作为 MRRbot 的重要角色，能够极大提高僵尸终端的可信度，有效防止僵尸网络被分析监测[9]，并使其具备有针对性的攻击能力.

2) 提出命令控制信道的冗余机制. 该冗余机制能够增强命令控制信道的可控性和时效性，并能够减少僵尸终端暴露或移除对僵尸网络的影响，保证高抗毁性.

3) 提出一种服务终端遴选算法. 该算法能够使僵尸网络中的普通终端访问优质服务终端，及时剔除不良服务终端，有效对抗 P2P 主动防御，避免服务终端失效.

1 MRRbot 设计

MRRbot 基于 P2P 混合结构，具有多种角色和命令控制冗余机制，并且采用服务终端遴选算法，能够有效对抗 P2P 主动防御手段.

1.1 多角色分级设计

随着三网融合的不断推进和移动网民的快速增长，本文提出的 MRRbot 能够应用于新的网络环境和多样化的智能终端，在未来异构网络环境中，其扩散性和生存性会更强. 在僵尸程序设计上，我们不仅需要考虑不同网络环境的安全防护能力和完善程度不同，还需要考虑不同类型终端的能耗、计算能力以及用户行为等因素. 僵尸终端可能是个人电脑、服务器、PDA、智能手机等，甚至未来可能出现的智能电视机顶盒，如图 1 所示.

定义 1. 终端信息. 本文简称"信息"，通过这些信息，其他终端能够与本机通信. 这些信息可以是〈IP,Port〉,BotID,UserID,或手机号等.

定义 2. 命令控制消息. 本文简称"消息"，是 MRRbot 内部通信数据的总称. 包括控制者(botmaster)

① P2P 主动防御是指防御者采用索引污染(index poisoning)、女巫攻击(sybil attack)等 P2P 攻击手段，使 P2P 僵尸网络的关键终端失效.

发布的命令、僵尸程序新版本以及 botmaster 回收的各种数据等.

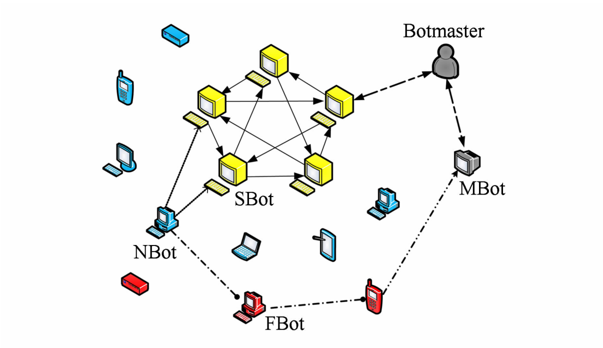


Fig. 1 MRRbot's framework.

图 1 MRRbot 架构

定义 3. 入侵感染链. 在入侵阶段,易感终端被感染之后,可能会继续入侵其他终端. 根据感染关系和顺序,若干感染终端形成了一条链.

MRRbot 包含 5 种角色:1) botmaster;2) 服务终端(ServiceBot, SBot);3) 普通节点(NormalBot, NBot);4) 虚壳终端;5) 监控终端(MonitorBot, MBot). botmaster 是僵尸网络的所有者,拥有发布攻击命令,更新僵尸程序等最高权限,并且能够获取所有 SBot 和 NBot 的信息. SBot 组成一个 P2P 网络,是 MRRbot 的核心层,负责转发 botmaster 发布的命令控制消息,回收和临时存储 NBot 返回的消息. Nbot 与相应 SBot 通信,获取 botmaster 的指令,进而执行相应攻击. FBot 是被成功入侵但未经验证的计算终端,这样的终端被植入了虚壳僵尸程序,它们不具备完整的通信能力,只具备感染传播能力. MBot 是 botmaster 的"耳目",参与对 FBot 的检测验证,以及统计 MRRbot 的各方面属性,如 SBot 和 NBot 的活跃数量,MRRbot 现有攻击能力等等.

在 MRRbot 中,我们把僵尸程序分为 2 类:一类是虚壳僵尸程序,运行在 FBot 上;另一类是真正的僵尸程序,运行在 SBot 和 NBot 上. 如图 2 所示,如果 botmaster 成功入侵正常主机,将注入虚壳僵尸程序,使其成为 FBot. 它不同于 NBot,不带有 SBot 的信息,因此无法与 SBot 通信. FBot 带有某一 NBot 和若干个 MBot 的信息,而这一 NBot 拥有所有由它直接感染和间接感染的终端信息. 一旦被感染,FBot 将向感染自己的 NBot 发送自己的信息,然后继续传播. 当 FBot 满足某些条件(见下文),botmaster 将发布升级命令,把真正僵尸程序发送给该 FBot,使其真正加入 MRRbot. 可见,FBot 只负责 MRRbot 的传播,类似于蠕虫感染别的计算机,它没有建立真正的命令控制信道,即使被防御者完全逆向,也不会暴露 MRRbot 的信息和攻击意图.

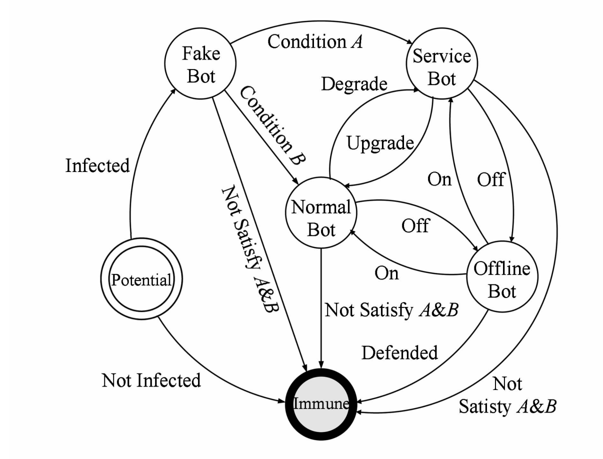


Fig. 2 A Bot's state transition.

图 2 僵尸终端的状态转移

FBot 满足一定条件后,会转化为真正的僵尸终端. 由于篇幅关系,本文只列出需要考虑的标准,不展开阐述.

1) 探测所感染的终端是否为防御者的控制环境[10],如蜜罐(honeypot). 比如,FBot 可以攻击一定数量的主机,其中包括 MBot,这样可以利用蜜罐的约束来进行探测.

2) 被感染主机的软硬件环境和资源是否满足 botmaster 的要求,如操作系统版本、计算能力、带宽等.

3) 终端用户的行为分析. 提取僵尸终端的活动行为,由 FBot 程序自身判断,该行为是否满足僵尸网络的攻击需求,如是否长时间在线,是否经常使用某款软件等.

4) 其他特殊的需求检测. 如要进行垃圾邮件攻击,则对带宽有一定的要求;收集敏感信息,则需要该终端用户有相关的信息;等等.

如果 FBot 满足上述条件,botmaster 则根据不同的转化条件[5](图 2 中的条件 A,条件 B. 条件 A, B 是满足上述 4 点标准的综合条件,而由于 SBot 需要提供更多的服务,所以条件 A 比条件 B 更苛刻,如对带宽、在线时长的要求),注入真正的僵尸程序,使其成为 SBot 或者 NBot. FBot 的引入使僵尸终端成为可信任的、满足攻击需求的终端,很大程度上保证了僵尸终端的安全性,同时也增加了 MRRbot 的危害程度.

1.2 命令控制信道设计

僵尸网络属于弱通信网络,与大多数 P2P 应用不同,其节点之间数据传输量不大、不频繁. 因此, botmaster 更加关心命令控制的完整性、可达性和时

效性.由于离线或被防御的原因，某些 SBot 会处于失效状态.为了保证在部分 SBot 失效的情况下 MRRbot 的通信能力，本文提出了命令控制信道的冗余机制.首先考虑传递方向不同，MRRbot 消息可以分为上行消息(bots to botmaster, B2M)和下行消息(botmaster to bots, M2B).上行消息用于 Bots 信息的汇总，下行消息用于命令下发、僵尸程序版本更新等.本文定义 botmaster 向 NBot 发布的消息为 M2NB(botmaster to NBots).在该模型中，作为攻击的执行者，NBot 占 MRRbot 的绝大部分，而 M2NB 包含所有攻击指令.因此，保证 M2NB 的时效性和完整性是命令控制信道设计的核心.

1.2.1 冗余机制基本思想

为应对 SBot 被劫持后消息被篡改，保持 M2NB 的下发能力，本文提出了冗余机制，如图 3 所示：

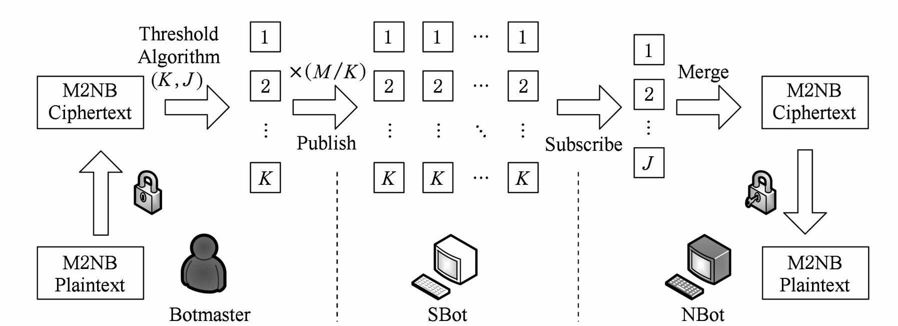


Fig. 3 The redundant command & control mechanism.

图 3 冗余命令控制机制

定义 4. Publish 方式.以推送方式发布消息，消息提供方发起通信，把消息推送给使用方.

定义 5. Subscribe 方式.以轮询方式“订阅”消息，消息使用方发起通信，访问提供方获取消息.

在 MRRbot 中，每个 NBot 拥有一个长度为 L 的 SBot 信息列表.botmaster 首先对 M2NB 用私钥加密，生成 M2NB 密文.然后采用(K,J)门限算法[11]，把 M2NB 密文分为 K 份消息片段，其中任意 J 份能够组成一个完整的 M2NB 密文.其约束条件为：

$$K=2J-1. \tag{1}$$

botmaster 把 K 份消息片段复制 $S_{\text{ativce}}(t)/K$ 份，采用匿名技术，用 Publish 方式发布给所有 SBot，每个 SBot 得到一份消息片段.NBot 用 Subscribe 方式访问自身 SBot 信息列表中的 SBot，获取其存储的 M2NB 密文消息片段.当 Nbot 得到 J 份不同的消息片段则可组成完整的 M2NB 密文，最后利用公钥解密，得到 M2NB 明文.

1.2.2 服务终端遴选算法

在实际网络环境中，P2P 僵尸网络若不采取相应措施，则无法避免 P2P 主动防御手段[12-13].对于 MRRbot，由于 P2P 主动防御的制约，终端用户行为的影响，以及网络状况不稳定等情况，SBot 可能处于离线(offline)状态，对 M2NB 传递造成影响.针对该问题，本文提出了一种服务终端的遴选机制，优化了 NBot 对于 SBot 的访问选择，以较大概率选择稳定的 SBot 通信，同时又兼顾 SBot 的服务能力.

算法 1. SBot 遴选算法子模块.

输入：长度为 L 的 *sbot_list*(SBot 信息列表)及相关参数(包括 SBot 的访问标记，初始为 True，访问成功次数，访问失败次数，访问时点和随机概率值)；

输出：符合条件的某个 SBot 信息.

1) 生成长度为 L 的随机概率列表 *rand_p_list*，与 *sbot_list* 对应，若 *sbot_list*[*i*]的访问标记为 False，则 *rand_p_list*[*i*]赋值 0；

2) 对 *rand_p_list* 进行归一化处理；

3) 若 *sbot_list*[*i*]的访问成功数、失败数均为 0，则 *sbot_list*[*i*].*rank*=*rand_p_list*[*i*]；若其成功数非 0，失败数为 0，则 *sbot_list*[*i*].*rank*=成功数×*rand_p_list*[*i*]；若为其他情况，则 *sbot_list*[*i*].*rank*=成功数/失败数×*rand_p_list*[*i*]；

4) 选择 *sbot_list* 中 *rank* 最大且为正值的 SBot，返回，不存在则返回空.

算法 2. SBot 遴选算法.

1) 构建链表 *M2NB_chain*，链表元素为存储 M2NB 消息片段的数组，长度为 J，表示一个 M2NB 消息；

2) 调用算法子模块(算法 1)，尝试获取某 SBot 信息；

3) 若成功，则用 Subscribe 方式访问该 SBot；若失败，则更新 SBot 列表，返回 2)；

4) 若成功，则查询该 SBot 是否有新的 M2NB

消息;若失败,则更新相关参数,返回2);

5) 若有新M2NB消息,则从SBot上获取一个或多个M2NB消息片段;否则更新相关参数,休眠时间t后,返回2);

6) 对于某个消息片段,若NBot没有同一消息的其他片段,则新建$M2NB_chain$元素;若存在,则进行匹配,重复则删除,否则保存;

7) 根据6)处理5)获取的所有M2NB消息片段,检测每个$M2NB_chain$元素,并统计已有M2NB消息片段是否达到J个;

8) 对于达到J个$M2NB_chain$元素,组成完整的M2NB消息,触发NBot执行模块处理,并删除该链表元素;

9) 若$M2NB_chain$为不为空,则返回2).

从算法2可以看出,对于一个M2NB消息,NBot至少需要访问J个SBot,才能够获取完整的M2NB消息.当访问过列表中所有SBot,仍未得到足够的消息片段时,NBot需要新SBot列表.这时,NBot依次访问L个SBot,统计失效SBot的数量,若存在有效SBot,则从其上获取其他新的SBot信息;若所有的SBot失效,则开启重构模式[7].SBot动态变化得越快,NBot的SBot信息列表的失效数就越多.NBot需要及时更新SBot列表信息,来保证有足够多有效的SBot可以被访问,使其以接近于100%的概率来获得完整的M2NB消息.

2 MRRbot评估与分析

2.1 相关参数

本文理论分析与实验评估所涉及的参数如表1所示:

Table 1 Relevant Parameters of MRRbot

表1 MRRBot相关参数

Symbol:Definition	Symbol:Definition
$S_{total}(t)$:SBots' number at t $S_{active}(t)$:active SBots' number at t $N_{total}(t)$:NBots' number at t $N_{active}(t)$: active NBots' number at t $F_{total}(t)$:FBots' number at t $F_{active}(t)$:active FBots' number at t (K,J):Threshold algorithm L:SBot list's length in each NBot	M: Stable active SBots' number N: stable active NBots' number p:M/N P:The probability that NBots get M2NB $T_{Publish}$:Publish time $T_{Subscribe}$:Subscribe time $T_{P2PRouting}$:P2P routing time $T_{QueryInterval}$: NBots' query interval

2.2 理论分析

假设MRRbot经过入侵传播,具备了一定的规模[14],此时可控性和时效性是其重要评价指标.设某一时刻t,SBot,NBot和FBot的总数和活跃数,如表1所示(MBot数量由botmaster设定),

$$BotNet_{totalsize}(t)=S_{total}(t)+N_{total}(t), \quad (2)$$

$$BotNet_{activesize}(t)=S_{active}(t)+N_{active}(t). \quad (3)$$

MRRbot在时刻t的攻击危害为$BotNet_{attack}(t)$,有:

$$BotNet_{attack}(t)\propto N_{active}(t). \quad (4)$$

在该模型中,SBot不参与传播和攻击,根据式(4)可知,NBot的活跃数量决定了MRRbot的攻击能力.因此,为了保证攻击力,p取值较小(详见2.3.2节).

2.2.1 可控性分析

僵尸网络的可控性表示botmaster对僵尸网络控制能力的强弱.MRRbot的可控性等价于M2NB消息被NBot接收并执行的概率.假设僵尸程序设计上无误,可控性的问题归约为:botmaster发布M2NB消息后,NBot能够以多大概率(即P)接收到.

不失一般性,假设在某段时间内,botmaster仅发布一条M2NB消息,$S_{active}(t)$和$N_{active}(t)$保持不变,且活跃NBot的SBot信息列表中的SBot均活跃.那么,设$M=S_{active}(t)$,$N=N_{active}(t)$,且$M\gg L$.P_k表示访问k个SBot后获取完整M2NB的概率:

$$P=\sum_{k=J}^{K}P_k,\ k\in[J,K], \quad (5)$$

m_i表示NBot收到第i个M2NB消息片段的次数,

$$L=\sum_{i=1}^{k}m_i,\ m_i\geqslant 1. \quad (6)$$

对于式(6),向量$(m_1,m_2,\cdots,m_k)$有C_{L-1}^{k-1}种取值,对于每种取值,

$$P_k(m_1,m_2,\cdots,m_k)=\frac{\prod_{i=1}^{k}C_R^{m_i}}{C_M^L}, \quad (7)$$

进而可得:

$$P=\sum_{k=J}^{K}P_k=\sum_{k=J}^{K}\left(C_K^k\times\sum_{j=1}^{C_{L-1}^{k-1}}P_{jk}\right), \quad (8)$$

整理之后:

$$P=\frac{1}{C_M^L}\times\sum_{k=J}^{K}\left(C_K^k\times\sum_{j=1}^{C_{L-1}^{k-1}}\left(\prod_{i=1}^{k}C_R^{m_{ijk}}\right)\right). \quad (9)$$

由式(9)可知,在M不变的情况下,P值与(J,L)的大小相关.(J,L)的合理取值对MRRbot的可控性至关重要,可由botmaster预先设定或动态调

整. 根据式(10)，在 $M=1\,000$ 的条件下，P 值随(J,L)的变化如图 4 所示：

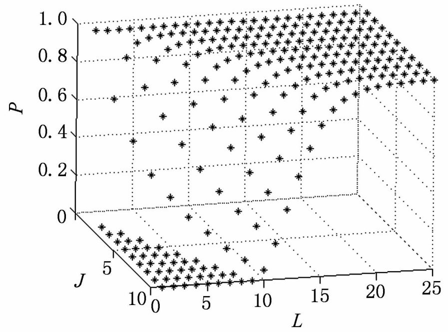


Fig. 4 The discrete changes of P value.

图 4 P 值离散变化曲面

从图 4 可以看出，当 J 取某一固定值，P 随 L 的增大而增大，并快速趋近于 1；当 L 取值固定，P 随着 J 的增大而减小；当 $J>L$ 时，$P=0$.

在满足 MRRbot 可控性的情况下，选取合理的 J 值，使 L 值适当大，保证 M2NB 消息的冗余性；同时选取合理的 L 值，尽量减少 NBot 拥有的 SBot 信息数，保证 MRRbot 的不可测量性. 因此，(J,L)的合理取值是 MRRbot 的可控性、冗余度和不可测量性三者之间的平衡. 如果设定 P 值不小于 95%，那么根据图 4，当$(J,L,M)=(5,10,1\,000)$时，$P(J,L,M)=97.11\%$. 此时，$(K,J)=(9,5)$，具有较高的冗余度；每个 NBot 拥有 10 个 SBot 信息，使 MRRbot 具有较高的不可测量性.

2.2.2 时效性分析

僵尸网络的时效性是指僵尸网络对命令控制的反应速度. MRRbot 的时效性等价于从 botmaster 发布 M2NB 开始，到 NBot 接收到完整 M2NB 消息所用的时间. 即：

$$BotNet_{\text{efficiency}} \propto T_{\text{Publish}} + T_{\text{Subscribe}}, \tag{10}$$

T_{Publish} 与 $T_{\text{P2PRouting}}$ 相关，表示为：

$$T_{\text{Publish}} = \alpha T_{\text{P2PRouting}}, \tag{11}$$

其中，α 由 botmaster 采用的策略决定. $T_{\text{Subscribe}}$ 与 $T_{\text{QueryInterval}}$ 相关，表示为：

$$T_{\text{Subscribe}} = \beta T_{\text{QueryInterval}},\ \beta \in (0,2]. \tag{12}$$

虽然不同 P2P 协议寻找节点所用时间不尽相同，但 $T_{\text{P2PRouting}} \ll T_{\text{QueryInterval}}$. 因此，MRRbot 的时效性主要与 NBot 采用的轮询机制有关.

2.3 实验评估

首先，本文利用 MATLAB 开发了一个离散事件模拟器，模拟了 MRRbot 和其他有代表性的僵尸网络模型[4-5]，对健壮性进行了比较分析. 然后，本文利用 OverSim 平台[15]，对 MRRbot 进行了模拟实现，在考虑多种实际因素的条件下，对 MRRbot 的可控性、时效性、抗毁性进行了评估.

2.3.1 健壮性分析

由于无法获知其他模型的具体实现，本文仅根据其系统构架，提取通信特征，进行模拟. 在离散事件模拟平台上，本文设置了 10 000 个易感节点，其中包括 1% 的蜜罐，10% 的行为受限节点和 20% 的不满足攻击需求节点，三者独立分布. 本文模拟漏洞传播和垃圾邮件传播 2 种常用入侵传播方式. 为接近实际情况，本文假设防御者可以通过蜜罐实施有效防御. 由于 MRRbot 和其他模型均为混合 P2P 僵尸网络模型，所以相关参数一致，保证客观性.

对于文献[4-5]中的模型(Model 1，Model 2)，该实验模拟服务终端和普通终端，易感节点被感染后成为服务终端或普通终端(比例为 1∶9)，两者均具有感染能力和攻击能力，但无法检测蜜罐、行为受限和不满足攻击需求的节点. 对于 MRRbot，该实验模拟 SBot，NBot 和 FBot，易感节点被感染后成为 FBot，FBot 用于检测蜜罐，行为受限或不满足攻击的节点，设其检测能力为 c(表示 FBot 能够成功检测出比例为 c 的非目标终端)，然后 FBot 以一定比例($p=1/9$)转化为 SBot 或 NBot. 其中，FBot 不计入 MRRBot 活跃规模，SBot 不参与传播.

上述 3 种模型的健壮性比较如图 5 所示：

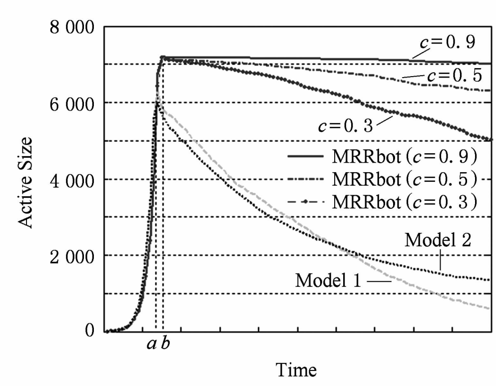


Fig. 5 The comparison of different P2P botnet models.

图 5 不同 P2P 僵尸网络模型的比较

从图 5 中可以看出，时刻 a 前，3 种模型活跃规模迅速增长，而 MRRbot 略低于其他模型；时刻 a 后，MRRbot 超越其他模型，活跃规模继续增长，在时刻 b 达到顶峰，之后缓慢下降且保持较高水平. 而其他模型在时刻 a 后活跃规模快速减少，一段时间后降至较低水平. 从图 5 中还可以看出，时刻 b 后，

对于 MRRbot,c 值越高,其活跃规模下降速度越缓慢,当 $c=0.9$ 时,其规模下降极不明显,处于稳态.

在其他模型中,随着活跃规模的增长,蜜罐、行为受限终端和不满足需求终端的数量也增加,防御者通过蜜罐获取更多僵尸终端信息进行防御,行为受限终端和不满足需求终端使僵尸网络的通信能力下降,这样大量僵尸终端失效或失控,当损失速度超过感染速度时,僵尸网络的规模就不断减少.在 MRRbot 中,FBot 用于检测蜜罐、行为受限终端和不满足需求终端,这一过程需要一定时间,使 MRRbot 的传播速度受到一定影响,所以初期(时刻 a 前)其活跃规模略低.随着时间的推移,FBot 检测机制排除了一定比例的蜜罐等非目标终端,大大减少了暴露的终端数,提高了命令控制信道的通信效率,使 MRRbot 具备高健壮性.FBot 检测能力越高,蜜罐等非目标终端在 MRRbot 中的数量越少,对 MRRbot 规模的影响就越低.单位时间内,如果防御者通过蜜罐获取的僵尸终端数量少于 MRRbot 的感染数量,MRRbot 规模就不会减少.

以上实验结果表明,在考虑多种实际因素的情况下,引入 FBot 的 MRRbot 能够在高对抗的情况下保证僵尸网络的健壮性,优于其他模型;提高 FBot 的检测能力,能够增强 MRRbot 的健壮性.

2.3.2 可控性、时效性与抗毁性评估

本文采用 OverSim,对 MRRbot 多种角色进行模拟实现,重点对 MRRbot 的演化阶段进行了模拟.假设 MRRbot 已达到一定规模,SBot 和 NBot 的活跃量维持在稳定水平.该组实验同样设置了 10000 个节点,$p=1/9$ $M=1000$,$N=9000$,并加入了终端用户行为的模拟.该组实验的目标是验证理论分析的结果,并评估 MRRbot 在不同防御程度下的抗毁性.

首先,我们利用多种 P2P 协议作为载体,构建命令控制信道,取 $T_{\text{QueryInterval}}=600$ s,$(J,L)=(5,10)$,进行模拟实验,结果如表 2 所示:

Table 2 Results of Different P2Ps

表 2 不同 P2P 协议的实验结果

C&C Protocol	Success:Total	Controllability/%	Efficiency/s
Chord	8730:9000	97.0	1251
Kademlia	8726:9000	96.9	1252
Pastry	8746:9000	97.2	1254
GIA	8741:9000	97.1	1258

实验结果表明,MRRbot 适用于多种 P2P 协议,其可控性和时效性都能够满足 botmaster 的需求.同时也验证了 2.2.2 节对时效性的理论分析结论.

然后,我们选取 Kademlia 协议,对 (J,L) 的不同取值,分别进行模拟实验,结果如图 6 所示:

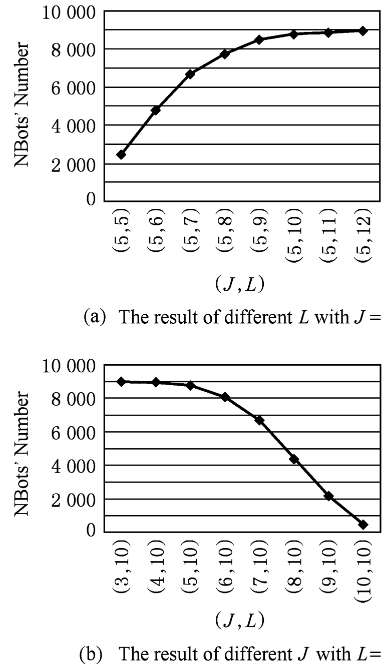


Fig. 6 The result of different (J,L).

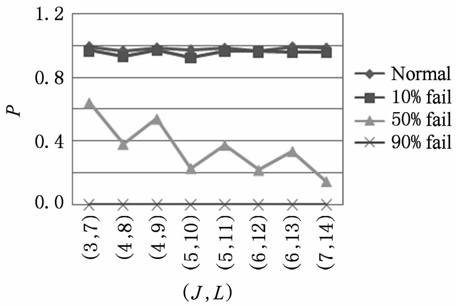
图 6 不同(J,L)的实验结果

从图 6(a)可以看出,在 J 不变的情况下,随着 L 的增大,MRRbot 的可控性增强;从图 6(b)可以看出,在 L 不变的情况下,随着 J 的增加,MRRbot 的可控性下降(验证 2.2.1 节的结论).另外,通过实验,本文发现(J,L)合理取值范围较大,多组(J,L)取值均能够保证 95%以上的 NBot 接收到控制命令,且满足 MRRbot 对冗余度和不可测量性的要求.

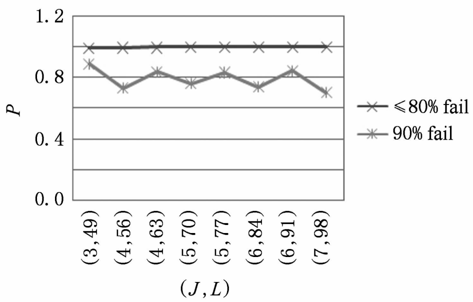
最后,本文假设 MRRbot 受到 P2P 主动防御的情况,使一定比例的活跃 SBot 长时间处于失效状态.本文分别对活跃 SBot 集合不同程度的失效情况进行了模拟实验,结果如图 7 所示.

图 7 中,P 值表示接收到攻击命令的 NBot 比例,等价于 MRRbot 的攻击能力;不同的线段表示不同比例的 SBot 失效.从图 7 中可以看出,在大量活跃 SBot 失效时,MRRbot 的攻击能力将明显下降(如图 7(a)所示);而适当减小 J 值,或增大 L 值,能够有效应对大量活跃 SBot 失效的情况,保证 MRRbot 持续攻击能力(如图 7(b)所示).在受到 P2P 主动防御的情况下,本文提出的 SBot 遴选算法,能够快速有效地剔除不稳定的 SBot,保证 NBot 能够及时获

取 M2NB. 另外，botmaster 可以通过动态调整(J，L)，最大程度保证 MRRbot 的冗余度和不可测量性，提高抗毁性.



(a) The result before adjustment



(b) The result after adjustment

Fig. 7 The effects of different SBot's failure ratios.

图 7 SBot 不同失效程度的影响

综合上述实验结果，MRRbot 能够利用多种 P2P 协议作为载体，保证命令控制信道的可控性和实效性，并有效对抗 P2P 主动防御策略，具有高抗毁性和高危害性.

3 基于志愿者网络的综合防御策略

防御者若能够获取大量的 SBot 信息，MRRbot 的通信能力和控制能力将大大减弱；如果能够检测出大量的 NBot，MRRbot 的攻击能力将被大大削弱. 防御策略需要检测并定位 MRRbot 的大量终端，由于 FBot 的引入，现有防御手段无法做到. 因此，文本提出了一种基于志愿者网络的综合防御策略.

志愿者网络是由自愿加入僵尸网络防御的人员组成，防御者在他们的计算终端上安装探测器，第一时间获取可疑恶意程序的信息. 志愿者终端为普通计算机，FBot 检测无法区分，使其为 SBot 或 NBot. 志愿者终端探测器能够获取到真实的僵尸程序信息，提取出命令控制信息和特征，然后通过志愿者网络汇报给防御服务器进行处理. 志愿者防御网络思想突破了传统蜜罐蜜网的限制和约束，能够及时有效地获取可疑程序样本，并且能够深度进行监测和挖掘，如图 8 所示：

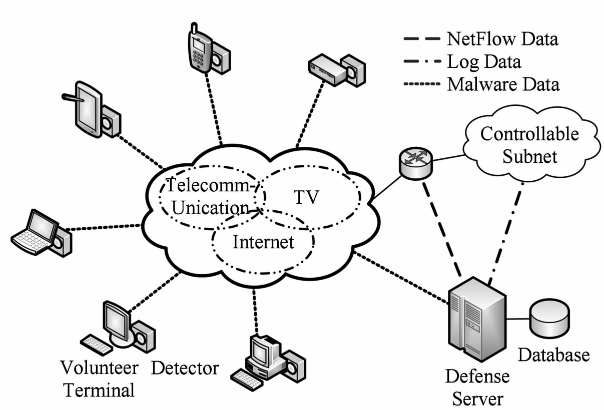


Fig. 8 The defense architecture of the volunteer network.

图 8 志愿者网络防御架构

防御服务器首先通过综合分析恶意程序信息，提取僵尸网络的网络流量特征和本地行为特征，根据这些特征分析处理网络流量数据，定位可控子网内的可疑终端，然后对这些终端的日志和流数据进行分析处理，进一步确认僵尸终端，最后向已经感染的可控子网终端和志愿者终端发布防御策略，清除志愿者终端上的可疑僵尸程序，有效削弱僵尸网络的规模和危害，为志愿者终端和可监控子网提供安全保障.

4 总 结

在三网融合不断推进和手机网民快速增长的背景下，僵尸网络越来越多样化，越来越复杂化，而混合结构的 P2P 僵尸网络更具威胁. 为了把握僵尸网络发展趋势，为防御者提供更多的帮助，本文提出了 MRRbot，MRRbot 引入 FBot，能够剔除具有约束行为的蜜罐，防御者很难捕获到真正的僵尸程序；利用冗余机制，增加了命令控制信道的安全性和不可测量性；采用服务终端遴选算法则很好地解决了 SBot 动态变化和失效的问题. 随后本文对 MRRbot 的健壮性、可控性、时效性和抗毁性进行了理论分析和实验评估. 结果表明，现有的防御手段无法有效地防御该僵尸网络. 最后本文提出了一个基于志愿者网络的综合防御策略.

下一步，我们将把工作重点放在新型僵尸网络攻防的研究上，包括：1)完善 MRRbot 和 FBot 检测机制；2)新型僵尸网络传播方式和传播模型的研究；3)僵尸网络评价体系和评估方法的研究；4)基于志愿者网络防御策略具体防御技术的研究.

参 考 文 献

[1] McPherson D, Dobbins R, Hollyman M, et al. Worldwide infrastructure security report, Vol 5 [R/OL]. Chelmsford, MA: Arbor Networks, 2010. [2011-01-14]. http://www.arbornetworks.com/report

[2] Vogt R, Aycock J, Jacobson M. Army of botnets [C] //Proc of the 2007 Network and Distributed System Security Symposium (NDSS'07). Reston, VA: ISOC, 2007

[3] Starnberger G, Kruegel C, Kirda E. Overbot—A botnet protocol based on Kademlia [C] //Proc of the 4th Int Conf on Security and Privacy in Communication Networks. New York: ACM, 2008: 1-9

[4] Wang P, Sparks S, Zou C C. An advanced hybrid peer-to-peer botnet [C] //Proc of the 1st USENIX Workshop on Hot Topics in Understanding Botnets (HotBots'07). Berkeley, CA: USENIX, 2007: No 2

[5] Ying Lingyun, Feng Dengguo, Su Purui. P2P-based super botnet: Threats and defenses [J]. Acta Electronica Sinica, 2009, 37(1): 31-37 (in Chinese)
(应凌云,冯登国,苏璞睿. 基于P2P的僵尸网络及其防御[J]. 电子学报,2009,37(1):31-37)

[6] Wang W, Fang B, Cui X, et al. A UserID-centralized recoverable botnet: Structure research and defense [J]. International Journal of Innovative Computing, Information and Control, 2010, 6(4): 4307-4317

[7] Traynor P, Lin M, Ongtang M, et al. On cellular botnets: Measuring the impact of malicious devices on a cellular network core [C] //Proc of the 16th ACM Conf on Computer and Communications Security (CCS'09). New York: ACM, 2009: 61-80

[8] Singh K, Sangal S, Jain N, et al. Evaluating bluetooth as a medium for botnet command and control [C] //Proc of the Int Conf on Detection of Intrusions and Malware, and Vulnerability Assessment (DIMVA'10). Washington, DC: IEEE Computer Society, 2010: 61-80

[9] Wang Hailong, Gong Zhenghu, Hou Jie. Overview of botnet detection [J]. Journal of Computer Research and Development, 2010, 47(12): 2037-2048 (in Chinese)
(王海龙,龚正虎,侯婕. 僵尸网络检测技术研究进展[J]. 计算机研究与发展,2010,47(12):2037-2048)

[10] Wang P, Wu L, Cunningham R, et al. Honeypot detection in advanced botnet attacks [J]. International Journal of Information and Computer Security, 2010, 4(1): 30-51

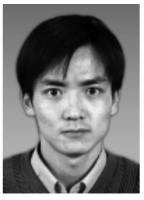
[11] Shamir A. How to share a secret [J]. Communications of the ACM, 1979, 22(11): 612-613

[12] naoumov N, Ross K. Exploiting P2P systems for DDoS attacks [C] //Proc of the 1st Int Conf on Scalable Information Systems. New York: ACM, 2006: No 47

[13] Leder F, Werner T, Martini P. Proactive botnet countermeasures—An offensive approach [R/OL]. Tallinn, Estonia: Cooperative Cyber Defence Centre of Excellence (CCDCOE). 2009. [2011-01-14]. http://www.ccdcoe.org/publications/virtualbattlefield/15_LEDER_Proactive_Coutner measures.pdf

[14] Rajab M, Zarfoss J, Monrose F, et al. My botnet is bigger than yours (maybe, better than yours): Why size estimates remain challenging [C] //Proc of the 1st USENIX Workshop on Hot Topics in Understanding Botnets (HotBots'07). Berkeley, CA: USENIX, 2007: No 5

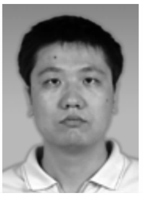
[15] Baumgart I, Heep B, Krause S. Oversim: A flexible overlay network simulation framework [C] //Proc of the 10th IEEE Global Internet Symposium (GI'07) in Conjunction with IEEE INFOCOM. Washington, DC: IEEE Computer Society, 2007: 79-84



Li Shuhao, born in 1983. PhD candidate. Student member of China Computer Federation. His current research interests include information security and network attack and defense.

Yun Xiaochun, born in 1971. PhD. Professor and PhD supervisor. His main research interests include information security and network attack and defense.

Hao Zhiyu, born in 1980. PhD. Assistant professor. His main research interests include information security and network attack and defense.



Zhai Lidong, born in 1982. PhD. Assistant professor. His main research interests include network security and information processing.