恶意代码传播效果的控制技术研究

曹莹莹[1,2],王绍棣[1],王汝传[1],张 伟[1]
(1.南京邮电大学 计算机学院,江苏 南京 210003;
2.盐城师范学院 信息科学与技术学院,江苏 盐城 224002)

摘 要:传统恶意代码在网络中通常尽最大能力进行传播,这种不受控制传播使得许多不具有攻击价值的主机被感染,从而影响了网络性能,另一方面,也增大了恶意代码被截获的概率。针对该问题,文中提出了一种恶意代码传播效果的控制方法。该方法从限定恶意代码的生存期、调节传播途径、限制传播范围三方面对恶意代码的传播效果进行综合调控。仿真实验详细分析比较了上述三因素对恶意代码的传播速度、感染主机数量、完成整个传播耗时等传播性能的影响。结果表明,与已有的方法相比,该方法明显提高了恶意代码的传播综合性能。

关键词:网络安全;恶意代码;受控传播;多途径传播;生存期控制
中图分类号:TP309 **文献标识码**:A **文章编号**:1673-629X(2010)08-0128-05

Research on Malware Code Propagation Effect Control Technology

CAO Ying-ying[1,2],WANG Shao-di[1],WANG Ru-chuan[1],ZHANG Wei[1]
(1. Institute of Computer Science and Technology, Nanjing University of Posts and Telecommunications,Nanjing 210003,China;
2. Institute of Information Science and Technology, Yancheng Normal College,Yancheng 224002,China)

Abstract:Traditional malware codes always try their best to propagate in the network. The uncontrolled propagation makes many worthless computers attacked, which affects the network performance. On the other hand, it also increases the probability of exposure. In this paper,a control method of malware code is proposed. The method systematically control the malware code propagation effect from setting survival time, regulating propagation approachs and restricting propagation range . As a case study, a comparative simulation is given, which shows the influence of above factors on malware code's propagation performance, such as, propagation velocity, quantity of affected computers and infectious period. The results prove that the method is more effective. Compared with the traditional ones,it improves the general propagation performance.

Key words:network security; malware code; controllable propagation; mutliple approach propagation; survival time control

0 引 言

目前,恶意代码综合病毒、网络蠕虫、特洛伊木马、网络攻击等多种手段于一体。主流的恶意代码多采用网络蠕虫的传播技术完成自身在网络中的大规模扩散。传统网络蠕虫传播技术的研究偏重于对传播途径、扫描方法、扩散算法的研究,以达到自身在网络中最快速、最广泛地传播的目标。新近出现的主动传播式网络蠕虫传播速度非常快,可以在几分钟内感染数百万台计算机[1]。这种不受控制传播使得许多不具有攻击价值的主机被感染,网络中大量恶意代码样本的扫描活动严重增加了网络负载,影响了网络性能。在信息战与网络安全压力测试中,攻击发起者迫切希望能够有效控制恶意代码的传播效果,但这方面的研究相对滞后。Ma J 等人提出了 Blind-k、Stop-k 等几种典型的自终止蠕虫启发式算法[1]。实验结果表明,这几种算法感染目标的速度与传统蠕虫相当,但传播控制能力远高于传统蠕虫。当所有目标被感染后,这几种算法均几乎能瞬间停止传播,而传统蠕虫需要经历较长时间才能最终停止传播。良性蠕虫的框架中也设计了对生存期控制的模块[2~5]。上述方法均单一采取缩短恶意代码生存期的方法干预恶意代码的传播过程,调整其传播能力。而在信息战与网络安全压力测试中,攻击发起者对恶意代码传播效果的可控性需求

收稿日期:2009-11-12;**修回日期**:2010-02-27
基金项目:国家自然科学基金(60973139,60773041);江苏省自然科学基金(BK2008451);博士后基金(20090451241)
作者简介:曹莹莹(1979-),女,江苏盐城人,硕士研究生,讲师,研究方向为网络安全、计算机在通信中的应用;王绍棣,教授,博士生导师,研究方向为计算机网络、信息安全和移动代理技术等。

是多方位的,概括来说有三点:

①攻击者可以依据需求控制恶意代码生存时间;

②攻击者可以依据需求并结合实际情况控制恶意代码的传播途径;

③攻击者可以依据需求控制恶意代码扩散范围。

为了解决上述问题,文中提出了一种恶意代码传播效果的控制方法,该方法系统化地从恶意代码的传播途径、传播范围与生存期三方面对恶意代码的传播能力进行调控。试验结果表明,该方法明显提高了控制恶意代码传播的综合效果。

1　恶意代码传播性能指标与传播模型

1.1　恶意代码传播性能指标

(1)感染率 β。

感染率是指某时刻恶意代码感染的主机数量占主机总数的比率。感染率是衡量恶意代码传播速度的重要指标之一。感染率越高意味着单位时间内恶意代码感染的主机数量越多,恶意代码传播速度越快。

(2) 已感染主机被免疫的比率 γ。

已感染主机被免疫的比率是指某时刻恶意代码被清除的受感染主机数量占被感染过的主机总数的比率。已感染主机被免疫的比率是衡量恶意代码抗查杀能力的指标之一。恶意代码的隐蔽性与抗分析能力越好,已感染主机被免疫的比率越低。

(3) 易感染主机被免疫的比率 μ。

易感染主机被免疫的比率是指某时刻容易被感染某恶意代码的主机数量占主机总数的比率。易感染主机被免疫的比率是衡量恶意代码传播效果的重要指标之一。恶意代码传播利用的漏洞越多、利用的漏洞越新,其易感染主机被免疫的比率越低。

(4) 感染过的主机总数 J。

感染过的主机总数是指该恶意代码传播完成后被成功感染过的主机总数。感染过的主机总数是衡量恶意代码传播效果的重要指标之一。感染过的主机总数越多表示该恶意代码传播效果越好。

(5) 存活的感染主机数量 $I(t)$。

存活的感染主机数量是指某时刻在继续感染其他易感染对象的主机数量。存活的感染主机数量是衡量恶意代码传播效率的另一个重要指标。存活的感染主机数量越多意味着该恶意代码当前的传播能力越强;存活的感染主机数量递增越快说明传播速度越快。

(6) 易感主机数量 $S(t)$。

易感主机数量是指某时刻容易被感染该恶意代码的主机总数。初始状态时易感主机数量是衡量恶意代码传播能力的重要指标之一,其数量越多表示该恶意代码潜在的传播能力越强。易感主机数量的递减速度一定程度上反映了该恶意代码的传播效率,递减速度越快说明该恶意代码传播速度越快或者感染前被免疫的速度越快。

1.2　网络蠕虫传播模型

建立网络蠕虫传播模型的目的是借助数学工具分析在现代网络环境下网络蠕虫传播的规律,包括网络蠕虫传播速度随时间的变化规律、网络蠕虫传播速度与网络拓扑的相互影响关系。由于网络蠕虫与生物病毒有许多相似之处,因此人们借助生物病毒的传播模型去建立网络蠕虫的传播模型。已提出的网络蠕虫传播模型包括 SI 模型[6]、SIR 模型[7]、双因素模型[8]等。其中,双因素传播模型考虑了更多的外界影响因素和各 ISP 节点或用户的对抗措施。与经典的流行病模型相比,该模型中做了两点修正[9]:其一,考虑了由于易感主机因主动免疫而进入移去类这一事实,假定单位时间内易感主机主动进入移去类的数量与易感主机数成正比,与到目前为止所有感染主机总数成正比(包括已恢复与未恢复主机)。其二,由于网络蠕虫的快速传播导致一些路由器发生阻塞,从而降低恶意代码的传播速度,感染率由常数变为 t 的函数。很多学者选取双因素模型作为网络蠕虫传播的数学模型,研究网络蠕虫传播的规律。

2　恶意代码传播效果控制技术

为了满足在信息战与网络安全压力测试中,攻击发起者对恶意代码传播效果的可控性的综合性要求,文中提出的恶意代码传播效果的控制方法系统化地从恶意代码的传播途径、传播范围与生存期三方面对恶意代码的传播能力进行调控。

(1) 网络蠕虫生存期控制。

文献[1]介绍了 Blind－k、Stop－k 等几种典型的自终止蠕虫启发式算法并对这几种算法进行了性能分析与测试。实验结果表明,这几种算法感染目标的速度与传统蠕虫相当,但传播控制能力远高于传统蠕虫。但这几种算法均有一定的缺陷,不能直接运用于实际网络环境的蠕虫传播中。Blind－k 算法通过在每个时间片末尾以 $1/k$ 的概率停止活跃的蠕虫传播来实现传播控制,原理简单,但难于应用。Tree 算法要求每个蠕虫在传染了 k 个以前未被传染的易感主机后停止传播,由于传播网络规模存在随机性,确定合适的 k 值较为困难。k 值设定过大,容易造成蠕虫因长时间无法完成感染任务而无法停止传播,影响了传播控制性;k 值设定过小,容易影响感染速度。

文中进一步修改了 Blind－k 算法以适应于实际应

用。文中将在每个时间片末尾以 $1/k$ 的概率停止活跃的网络蠕虫传播的方案修改为设定每一个网络蠕虫在网络中的生存期，即允许攻击者按照攻击目标网络的规模与攻击目的，设定网络蠕虫的生存期，当生存期结束，网络蠕虫自行终止执行。根据概率统计知识可知，每一时刻生存期结束的网络蠕虫占所有存活网络蠕虫的比率是相同的，文中设该比值为 $1/m$。设定网络蠕虫在受害主机的生存期是一种间接提高已感染主机的免疫率的方法，即将此免疫率提高为 $\gamma + 1/m$。由双因素传播模型可知，已感染主机的免疫率值越大，网络中存活的网络蠕虫样本越少，网络蠕虫对网络的影响也越小。在网络蠕虫传播的后期，随着易感主机 $S(t)$ 急剧减少，网络蠕虫感染效果严重下降，即 $I(t)$ 急剧减少，但网络蠕虫对于网络负载不必要的影响却没有因为感染效果的降低而降低，从而增加了网络蠕虫被网络安全监控设备发现的概率。因此，主动提高已感染主机的免疫率，一定程度上减低了网络蠕虫被网络安全监控设备发现的概率。

另一方面，此生存期是一个攻击者可以修改的变量。攻击者需要综合考虑攻击目的、攻击目标的规模等多种因素后为每一次攻击任务设定该变量的大小。生存期越短，$1/m$ 越大，$\gamma + 1/m$ 越大，从而有效地调控了已感染主机的免疫率与网络中存活的网络蠕虫数量。

(2)恶意代码传播途径控制。

目前，恶意代码传播的途径主要包括电子邮件、移动存储设备、局域网络共享、即时通信、网页挂马、僵尸网络、主动漏洞攻击[10]。主流的恶意代码一般固定使用上述的一种或多种方式进行传播。而文中提出的恶意代码传播效果的控制方法提供了上述所有的传播方式，并允许攻击者依据攻击目的以及攻击目标网络特征选择其中的一种或多种进行恶意代码的传播。

由双因素传播模型可知，提高网络蠕虫的感染率 $\beta(t)$，尤其是初始感染率 β_0 是加快网络蠕虫扩散速度的关键，而通过多途径进行恶意代码的传播是提高恶意代码感染率的较为容易实现的方式之一。依据文献[6,13,14,15,16]的分析以及反病毒软件公司的统计，不同传播途径的蠕虫感染率大致相同，若设单一初始感染率 β_0 为 β_0'，则采用此传播途径控制技术的网络蠕虫的初始传播率 β_0'' 的取值范围是 $\beta_0' \leq \beta_0'' \leq 7\beta_0'$，从而加强了恶意代码传播率的可控性。

(3)恶意代码传播范围控制。

传统的恶意代码通常希望在整个 Internet 环境中尽可能地传播，但信息战、网络安全性压力测试通常是对特定对象发起的，其 IP 地址范围是有限的。若采用传统的全网扩散方法容易引起对“无辜”对象的破坏以及增加不必要的网络负载，既影响了计算机系统的正常使用与网络环境的稳定运行，又增大了恶意代码被捕获的概率。文中提出的恶意代码传播效果的控制方法允许攻击者依据攻击目的设置传播范围以禁止其在目标范围外传播，从而有效地控制了恶意代码对公共网络环境的影响。

3 性能分析

文中提出的恶意代码传播效果的控制方法所使用的传播技术继承于网络蠕虫，因此网络蠕虫的传播模型同样适用于该控制方法。文中同样选取双因素模型对该控制方法的传播性能进行分析。该控制方法兼具生存期受限与多途径传播的特性，为了有效地分析每一种特性对网络蠕虫性能的影响，文中分别对这两种特性进行了实验。文中称使用该传播控制方法的恶意代码为 DWorm；称仅具有多途径传播能力而生存期不受限的网络蠕虫为多途径传播网络蠕虫；相应称生存期受限而仅具有单一途径传播能力的网络蠕虫为生存期受控网络蠕虫。本章分别模拟了“冲击波”网络蠕虫、多途径传播网络蠕虫、生存期受控网络蠕虫、DWorm 的性能，并进行了性能对比。

3.1 参数设置

(1)“冲击波”网络蠕虫参数。

“冲击波”网络蠕虫 Worm. msBlast 是第一个利用 DCOM RPC 漏洞进行攻击和传染的网络蠕虫，该蠕虫会下载并运行文件 Msblast. exe，使计算机出现系统重启、无法正常上网等现象。“冲击波”网络蠕虫每 1.8 秒尝试连接 20 个随机 IP，传播能力强。依据文献[15]，“冲击波”网络蠕虫是 2003 年 8 月 8 日最先在互联网络中被发现的。在 8 月 19 日的时候，金山毒霸推出了“冲击波”专杀工具，据此推断此源“冲击波”的病程时间大约为 12 天。根据《2004 年电子商务及其发展状况》中的数据估计，在 2003 年 8 月全球大约有 7 亿互联网用户，而根据搜狐 IT 在 2003 年 8 月 13 日的报道，计算机防毒软件厂商趋势科技(Trend Micro) 说：“冲击波”可能感染了全球一、两亿台计算机。根据文献[16]，“冲击波”网络蠕虫每 1.8 秒随机扫描 20 个 IP 新连接，则其平均扫描率 $w = 667/s$。根据文献[6,17]，取 $I_0 = 1$，$\eta = 3$，$\gamma = 0.03$，$\mu = 0.06/N$。根据文献[2]，$\beta_0 = \eta/\Omega$ 表示初始感染率，$\Omega = 2^{32}$ 表示初始传染率，则 $\beta_0 = 0.00000016$。由上文的分析可见，文中主要关注网络蠕虫在局部范围内的传播特性，故取 $N = 4 \times 10^5$。

(2) 多途径传播网络蠕虫参数。

多途径传播网络蠕虫综合了电子邮件、移动存储设备、局域网络共享、即时通信、网页挂马、僵尸网络、主动漏洞攻击中的一种或多种传播方式,则其初始感染率 $0.00000016 < \beta_0 < 0.0000011$,文中取 $\beta_0 = 0.0000006$。由于多途径传播网络蠕虫可以利用的漏洞较一般网络蠕虫多,因而易感染主机被免疫率相应下降,文中设 $\mu = 0.03/N$,同样取 $N = 4 \times 10^5$,$I_0 = 1$,$\eta = 3$,$\gamma = 0.03$。

(3) 生存期可控网络蠕虫参数。

设定网络蠕虫在受害主机的生存期是一种间接提高已感染主机的免疫率 γ 的方法,即经过时间 ζ 后,第一批感染主机上的多途径传播可控蠕虫执行自删除程序而自行清除,则模型中的已感染主机的免疫率增加,依据上文分析,免疫率增加 $1/m$,文中设 $1/m = 0.02$,则 $\gamma' = 0.05$,同样取 $\beta_0 = 0.00000016$,$\mu = 0.06/N$,$N = 4 \times 10^5$,$I_0 = 1$,$\eta = 3$。

(4)DWorm 参数。

DWorm 兼具生存期受限与多途径传播的特性,因此初始感染率 β_0、易感染主机被免疫率 μ 与多途径那个传播网络蠕虫相同,取 $\beta_0 = 0.0000006$、$\mu = 0.03/N$。为了进一步分析生存期长短对传播性能的影响,文中设定 DWorm1 的生存期与传播可控蠕虫相同,则 DWorm1 的已感染主机被免疫的比率与生存期受控蠕虫的已感染主机被免疫的比率相同,即 $\gamma_1' = 0.05$;设定 DWorm2 的生存期短与 DWorm1 的生存期,则 DWorm2 的已感染主机被免疫的比率 γ_2' 大于 DWorm1 的已感染主机被免疫的比率,取 $\gamma_2' = 0.08$。同样,$N = 4 \times 10^5$,$I_0 = 1$,$\eta = 3$。

3.2　性能比较

为了更为清楚地比较上述几种恶意代码的传播性能,文中分别对它们的传播效率与传播效果进行了实验与对比。几种恶意代码的 $I(t)$ 性能比较如图 1 所示。DWorm1、DWorm2 传播速度明显高于"冲击波"网络蠕虫与生存期可控蠕虫,在约 80 秒时网络中存活的 DWorm1、DWorm2 样本均达到峰值,此性能与多途径传播网络蠕虫相当,而"冲击波"网络蠕虫、生存期可控蠕虫分别在 350 秒、650 秒时网络中存活的样本达到峰值。DWorm1、DWorm2 在 250 秒内完成大部分的传播任务,此项性能优于其他三种蠕虫;多途径传播网络蠕虫、"冲击波" 网络蠕虫、生存期可控蠕虫分别需要 350 秒、700 秒、1300 秒完成主要传播任务。DWorm1、DWorm2 的初始感染率与多途径传播网络蠕虫相同,但它们的 $I(t)$ 曲线变化比多途径传播网络蠕虫的 $I(t)$ 曲线变化缓和,可见通过设定 DWorm1、DWorm2 的生存期有效地调控了恶意代码的感染率,存活主机数量也因此得到调控。另一方面,由于 DWorm2 的生存期短于 DWorm1 的生存期,因此其感染率与存活主机数量少于 DWorm1。

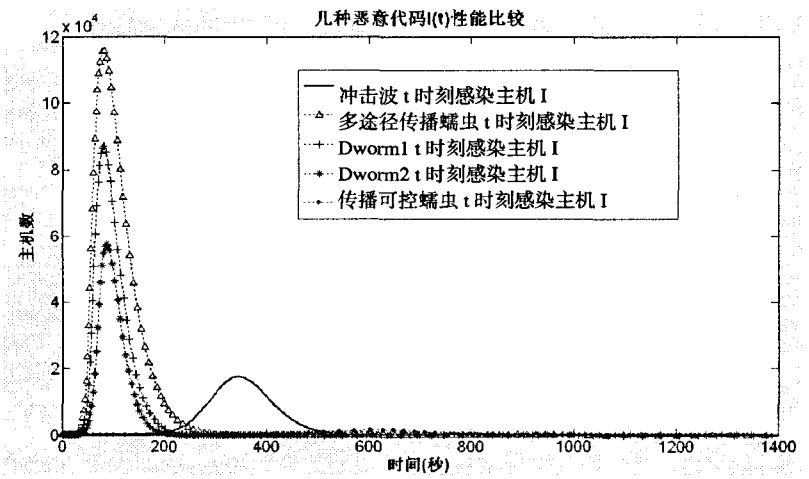

图 1　几种恶意代码的 $I(t)$ 性能比较

几种蠕虫的 $R(t)$ 性能比较如图 2 所示,DWorm1 在指定的范围内传播效果明显好于"冲击波"网络蠕虫与生存期可控蠕虫,略高于多途径传播网络蠕虫。传播完成后,DWorm1 共感染约 1.171×10^5 台主机,多途径传播网络蠕虫、"冲击波" 网络蠕虫、生存期可控蠕虫感染的主机数分别为 1.163×10^5、0.8×10^5、20500 台。而 DWorm2 共感染约 1.379×10^5 台主机,明显多于 DWorm1,可见通过缩短恶意代码的生存期可进一步加强其感染效果。

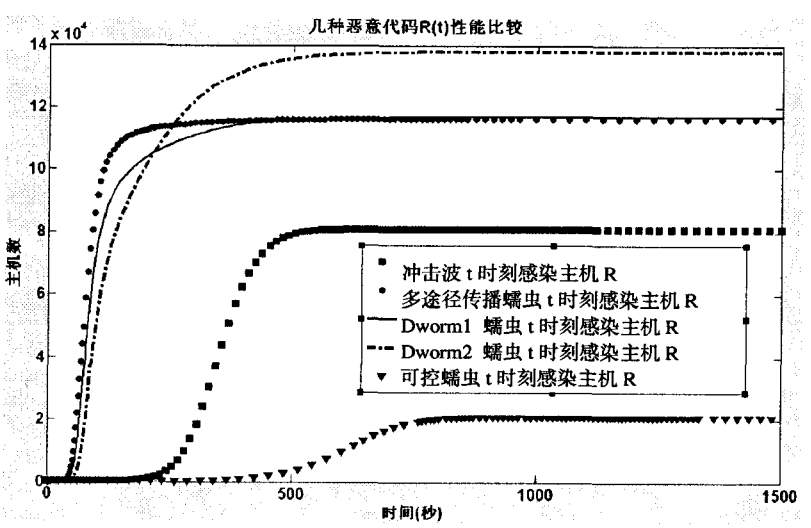

图 2　几种恶意代码的 $R(t)$ 性能比较

几种蠕虫的初始时刻感染率 β_0、易感染主机被免疫的比率常量 μ、已感染主机被免疫的比率常量 γ，存活的感染主机数量峰值、达到存活的感染主机数量峰值所需时间、感染的主机总数、传播完成耗时具体情况如表 1 所示。DWorm 在传播速度、感染主机数量、完成整个传播耗时等各方面的性能均明显优于“冲击波”网络蠕虫与生存期可控网络蠕虫。DWorm 的传播速度比多途径传播网络蠕虫略差，但感染的主机总数、传播完成耗时均好于多途径传播网络蠕虫。因此，可以得到结论，上述几种网络蠕虫中，DWorm 的综合性能最好。考虑到恶意代码从出现到被相关机构捕获有一定的时间间隔，DWorm 传播周期短，而且传播后期对网络流量影响较小，因此其被捕获的概率小于传统恶意代码被捕获的概率。

4 结束语

恶意代码在 Internet 范围内不受控制的传播增加了网络负载，影响了网络性能并感染了许多不具有攻击价值的主机。文中提出的恶意代码传播效果控制技术明显缩短了恶意代码传播周期，提高了感染主机总数，改善了传播率。同时，该技术允许攻击者通过调整恶意代码的传播途径与生存期进一步调控恶意代码传播过程的综合性能。

表 1 几种恶意代码传播效果比较

恶意代码名称	β_0 ($\times 10^{-6}$)	$\mu(/N)$	γ	I_{max} ($\times 10^5$ 台)	T_{Imax} (秒)	T_{Rmax} (秒)	R_{max} ($\times 10^5$ 台)
“冲击波”网络蠕虫	0.16	0.06	0.03	0.018	350	700	0.808
多途径传播网络蠕虫	0.6	0.03	0.03	0.118	80	350	1.163
生存期可控网络蠕虫	0.16	0.06	0.05	0.016	650	1300	0.205
DWorm1	0.6	0.03	0.05	0.086	80	300	1.171
DWorm2	0.6	0.03	0.08	0.575	80	250	1.379

参考文献：

[1] Ma J, Voelker G, Savage S. Self - Stopping Worms[C]// Proc. of the 2005 ACM Workshop on Rapid Malcode. New York: ACM Press, 2005: 12 - 21.

[2] 王佰玲. 基于良性蠕虫的网络蠕虫主动遏制技术研究[D]. 哈尔滨：哈尔滨工业大学，2006.

[3] 罗军宏，张有为，谢余强，等. 受控蠕虫框架结构研究[J]. 计算机应用研究，2006，23 (8)：74 - 82.

[4] 张殿旭，张 怡. 基于可控蠕虫的 W - A - W 传播模型研究[J]. 计算机工程与科学，2008，30(5)：15 - 22.

[5] 崔惠鹏. 良性网络蠕虫研究与进展[J]. 计算机与数字工程，2007，35(10)：77 - 79.

[6] Zou C C, Gong W, Towsley D. Code red worm propagation modeling and analysis[C]//Proc. of the 9th ACM Conf. on Computer and Communications Security (CCS 2002). New York: ACM Press, 2002: 138 - 147.

[7] Kim J, Radhakrishnan S, Dhall S K. Measurement and analysis of worm propagation on internet network topology[C]//Proc. of the IEEE Int'l Conf. on Computer Communications and Networks (ICCCN 2004). [s. l.]: [s. n.], 2004: 495 - 500.

[8] Zou C C, Gong W, Towsley D. Worm propagation modeling and analysis under dynamic quarantine defense[C]//Proc. of the ACM CCS Workshop on Rapid Malcode (WORM 2003). New York: ACM Press, 2003: 51 - 60.

[9] 彭俊好. 信息安全风险评估及网络蠕虫传播模型[D]. 北京：北京邮电大学，2007.

[10] 诸葛建伟，韩心慧，周勇林，等. 僵尸网络研究[J]. 软件学报，2008，19(3)：702 - 715.

[11] Zou C C, Towsley D, Gong Weibo. Email Worm Modeling and Defense[C]//Proceedings of 13th International Conference on Computer Communications and Networks. [s. l.]: [s. n.], 2004: 409 - 414.

[12] Hu Huaping, Wei Jianli. Instant Messaging Worms Propagation Simulation and Counter Measures[J]. Wuhan University Journal of Natural Sciences, 2007, 12(1): 95 - 100.

[13] Zhou Ying, Wu Zhongfu, Ye Chunxiao, et al. Proactive Worm Prevention Based on P2P Networks[C]//IET International Conference on Wireless Mobile and Multimedia Networks Proceedings. Hangzhou, China: [s. n.], 2006: 405 - 406.

[14] 王跃武，荆继武，向 继，等. Contagion 蠕虫传播仿真分析[J]. 计算机研究与发展，2008，45(2)：207 - 216.

[15] 张丽萍，洪 龙，王惠南. 一种网络病毒传播的时滞微分方程模型[J]. 南京邮电大学学报，2008，27(5)：78 - 83.

[16] 张运凯. 网络蠕虫的传播与控制研究[D]. 西安：西安电子科技大学，2005.

[17] 马 铭. 蠕虫模拟研究[D]. 北京：中国科学院研究生院，2005.