具有变化感染率的僵尸网络传播模型

冯丽萍[1,2] 韩 琦[2] 王鸿斌[1]
(忻州师范学院计算机系 忻州 034000)[1] (重庆大学计算机学院 重庆 400030)[2]

摘 要 僵尸网络对网络安全的威胁已经引起了众多安全研究专家的高度重视。数学建模是研究僵尸网络传播特性的一种有效方法。现有的僵尸网络传播模型假设感染率是常量。事实上,大量僵尸病毒爆发时会引起网络拥塞,从而导致感染变慢。针对这一特点,提出一个具有变化感染率的僵尸网络传播模型。根据 Intetnet 的实际情况,在模型中加入了预先免疫特征项。最后通过 matlab 模拟对比了提出的模型与已有模型在反映僵尸网络感染时的差异。实验结果表明,具有变化感染率的僵尸网络传播模型能更准确地反映实际网络中僵尸程序的传播规律。

关键词 僵尸网络,传播模型,网络安全,计算机病毒
中图法分类号 TP309.5 **文献标识码** A

Botnet Propagation Model with Variable Infection Rate

FENG Li-ping[1,2] HAN Qi[2] WANG Hong-bin[1]
(Department of Computer, Xinzhou Normal College, Xinzhou 034000, China)[1]
(College of Computer Science, Chongqing University, Chongqing 400030, China)[2]

Abstract The threat of botnet on network security has been attached great importance by many security search experts. Mathematical modeling is a kind of effective method to study botnet propagation. Existing botnet propagation models assume that infection rate is constant. However in fact, network congestion will be caused when a large number of botnet viruses rampant. Against this backdrop, a botnet propagation model with variable infection rate was proposed, which considers pre-immunity according to the real situation of Internet. Finally, the difference of botnet infection between the model of this paper and existing models was depicted in matlab. Experimental results show that the botnet propagation model with variable infection rate is more accurate to reflect the practical propagation law of bots.

Keywords Botnet, Propagation model, Network security, Computer virus

1 引言

僵尸网络(botnet)是由大量被僵尸程序感染的主机(bot or zombie)构成的以恶意活动为目的的网络[1]。被僵尸主人控制的僵尸机通过分布式拒绝服务(DDoS)、发送垃圾邮件(Spam)、点击欺诈(click fraud)、窃取敏感信息(information theft)以及键盘记录(keylogging)等方式从事恶意网络活动[2,3]。相比传统的网络安全威胁,僵尸网络具有有组织性、传播速度快、传播途径多、隐蔽性强、技术含量高和影响破坏大等特点[4-6],已经成为 Internet 最大的威胁之一。

为了有效对抗僵尸网络,减轻其对网络安全的威胁,许多国内外安全研究专家已经对僵尸网络的工作机理与检测技术进行了深入的分析。文献[2]深入剖析了僵尸网络的工作机制与功能结构,并归纳总结了目前跟踪、检测和防御僵尸网络的最新进展。方滨兴等人在文献[7]中给出了僵尸网络的定义与分类,同时详细阐述了僵尸网络的检测方法,预测了未来僵尸网络的发展。Barford 等人在文献[8]中通过分析僵尸网络代码库来揭示它的攻击原理,进一步提出预防僵尸网络的策略。这些研究对僵尸网络的检测和消除奠定了良好的理论基础。通过监测僵尸网络的传播过程,Dagon 等人发现时区对僵尸网络蠕虫的传播具有较大的影响。根据这一发现,他们提出了考虑时区因素的僵尸网络的传播模型[9],通过与实际观察数据的对比发现该模型比传统的网络蠕虫传播模型更准确地反映了僵尸网络传播过程。但是 Dagon 在文献[9]中提出的模型假设整个僵尸网络传播过程中感染率都是常数,而且模型中没有考虑免疫措施对僵尸网络传播的影响。事实上,僵尸网络通过大量蠕虫扫描、发送垃圾邮件等方式进行恶意活动时会产生网络拥塞、资源紧张等现象,这种现象会导致感染率下降,因此在整个僵尸网络感染过程中感染率不应该是常量。另一方面,据调查,目前有 92% 的上网用户安装有反病毒软件(也就是采取预先免疫措施)[10],因此在建模时考虑预先免疫措施的影响更加符合现实情况。考虑到以上两个因素,结合文献[9],本文提出一个具有变化感染率和预先免疫措施的僵尸网络传播模型。

到稿日期:2012-02-20 返修日期:2012-05-11 本文受国家自然科学基金项目(61003247),山西省自然科学基金项目(2009011018-4),山西省忻州师范学院自然科学基金项目(201127)资助。
冯丽萍(1976—),女,博士生,副教授,主要研究方向为网络安全、复杂网络、动力系统,E-mail:fenglp@yeah.net;韩 琦(1981—),男,博士,主要研究方向为网络安全、细胞神经网络;王鸿斌(1972—),男,博士,副教授,主要研究方向为人工智能、无线传感器网络。

2 具有变化感染率的僵尸网络传播模型

僵尸程序是网络蠕虫、病毒和特洛伊木马的进一步发展，根据一些经典的蠕虫传播模型[11,12]可以构建出僵尸网络的传播模型。但是与蠕虫传播不同的是，僵尸程序在传播过程中是有组织的，受僵尸主人控制的。因此，我们只能借鉴已有的蠕虫传播模型，结合僵尸网络感染的特点，构建新的模型。为了方便，列出经典的简单 K-M 蠕虫传播模型[13]。

$$\begin{aligned}\frac{dS(t)}{dt}&=-\beta S(t)I(t)\\ \frac{dI(t)}{dt}&=\beta S(t)I(t)-\gamma I(t)\\ \frac{dR(t)}{dt}&=\gamma I(t)\end{aligned}\tag{1}$$

式中，符号 $S(t)$，$I(t)$和 $R(t)$分别表示 t 时刻网络中易感染主机数、已感染主机数和已恢复(免疫)主机数。β 表示 S 被 I 感染的概率，γ 表示已感染主机的恢复率。

文献[9]的观察数据表明，网络中主机的上下线状态对僵尸网络病毒传播有明显影响，所以为了更准确地反映僵尸网络病毒传播机理，需对式(1)作改进。

2.1 模型改进的两个关键因素

第一，主机的在线和下线状态。用 $\alpha(t)$表示 t 时刻主机在线的比例。相应地，t 时刻在线易感染主机数和已感染主机数可以表示为：

$$S'(t)=\alpha(t)S(t)\tag{2}$$

$$I'(t)=\alpha(t)I(t)\tag{3}$$

由 $\alpha(t)$的定义知，$\alpha(t)$是一个周期函数，它的变化周期为 24 小时。通常 $\alpha(t)$的值在白天达到最高，而在晚上当许多用户关闭计算机时它的值最低。

第二，感染率 β。大量僵尸病毒在网络中爆发时，会导致 Internet 路由器工作负荷严重加大、网络资源短缺等情况，从而引起网络拥塞。最终，病毒传播速度也会减慢。所以在整个僵尸网络感染过程中感染率是随时间逐渐减小的。因此，引入 $\beta(t)$表示在 t 时刻易感染主机被已感染主机感染的有效感染率。

2.2 具有变化感染率的僵尸网络感染模型

根据以上分析，可以构建下列新的僵尸网络感染模型：

$$\begin{aligned}\frac{dS(t)}{dt}&=-\alpha^2(t)\beta(t)S(t)I(t)-\delta\alpha(t)S(t)\\ \frac{dI(t)}{dt}&=\alpha^2(t)\beta(t)S(t)I(t)-\gamma\alpha(t)I(t)\\ \frac{dR(t)}{dt}&=\alpha(t)(\delta S(t)+\gamma I(t))\end{aligned}\tag{4}$$

式中，δ 表示由于防火墙、打补丁、路由器等反病毒措施产生的主机从易感染状态转化为免疫状态的比率。下面主要讨论 $\beta(t)$的确定。

令 β_0 表示病毒的初始感染率，用下列方程表示随时间逐渐降低的感染率 $\beta(t)$。

$$\beta(t)=\beta_0 f_1(I(t))\tag{5}$$

式中，f_1 是 I 的非线性递减函数。令 $f(I(t))=f_1(I(t))I(t)$，那么式(4)可重写为：

$$\begin{aligned}\frac{dS(t)}{dt}&=-\alpha^2(t)\beta_0 f(I(t))S(t)-\delta\alpha(t)S(t)\\ \frac{dI(t)}{dt}&=\alpha^2(t)\beta_0 f(I(t))S(t)-\gamma\alpha(t)I(t)\\ \frac{dR(t)}{dt}&=\alpha(t)(\delta S(t)+\gamma I(t))\end{aligned}\tag{6}$$

3 模型仿真与分析

本节主要通过 Matlab 仿真来观察微分式(6)反映的僵尸网络病毒传播过程，重点是观察 $\alpha(t)$、$\beta(t)$和预先免疫措施对模型的影响。实验中，S，I 和 R 的初始值取值如下：$I(0)=2$，$S(0)=N-I(0)$，$R(0)=0$，$N=10000$。非线性函数取：$f(I(t))=\beta_0\left[1-\frac{I(t)}{N}\right]^{\eta}$。当 $\eta=0$ 时，$f(I(t))=\beta_0$。这时感染率 $\beta(t)$成为一常量。

首先观察 $\alpha(t)$对模型的影响，取下列参数值：$\beta_0=0.0006$；$\delta=0.0059$；$\gamma=0.0059$；$\eta=0$。$\alpha(t)$取从 0.1 到 1 的 90 个数，每个模拟运行 500 次，观察从 0～1000(以小时为单位)这段时间内的感染情况，然后取平均值，得到的模拟结果如图 1 所示。

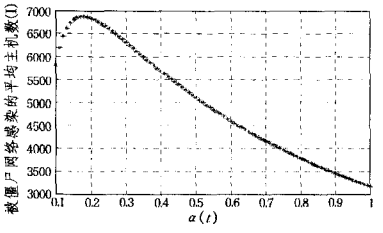

图 1 被僵尸网络感染的平均主机数随 $\alpha(t)$的变化图

图 1 反映了主机上下线状态对僵尸网络病毒传播的影响。我们发现，并不是在线主机数越多，被僵尸网络感染的平均主机数就越多，相反，在线主机的比例大约在 18%左右时，被感染的平均主机数达到最多。这是由于有反病毒措施之故，当大量主机被感染时，人们的预防意识就会增强，相应地会采取隔离、断网、打补丁等预防措施。图 1 的情况可以用图 2 的曲线走向来反映。

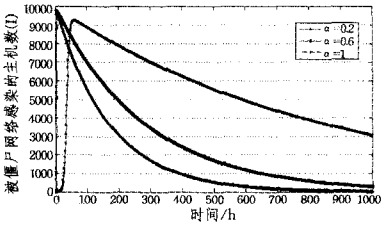

图 2 被僵尸网络感染的主机数随时间的变化

图 2 表明了 $\alpha(t)$分别取 0.2、0.6 和 1 时，僵尸网络感染主机的情况。当 $\alpha(t)$等于 0.2 时，感染暴发相对较慢，恢复的也较慢。随着 $\alpha(t)$的逐渐增大，被感染的主机数会在短时间内上升，但是由于反病毒技术的因素，感染达到最高峰时恢复的速度也明显加快。产生这种现象的另外一个因素是，模型(6)中被感染主机数是随时间单调递减的，当易感染主机数减小到临界值($\frac{\gamma}{\beta}$)时，就会导致 $I(t)$的值逐渐减小。

图 1 和图 2 的结果可以帮助我们评估未来僵尸网络的感染情况，从而提出相应的对抗方案。

下面观察变化感染率 $\beta(t)$对模型的影响。

取参数值 $\beta_0=0.0009$，$\delta=0.0059$；$\gamma=0.0059$；$\alpha(t)=0.6$。图 3 显示了当 $\eta=0$，3 和 6 时，被僵尸网络感染的主机数的变化情况。

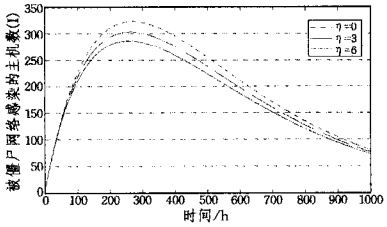

图 3　被僵尸网络感染的主机数随时间的变化图

图 3 中，$\eta=0$ 对应的曲线即为经典的感染率为常数的情况。显然当 $\eta=3$ 和 6 时，被感染主机的最大值要小于 $\eta=0$ 时的感染数。因此具有常量感染率的模型在预测僵尸网络的传播行为时会高估感染规模。

最后，为了观察预先免疫措施对模型的影响，取参数 $\delta=0$，其它参数值与图 3 的取值相同。得到的模拟结果如图 4 所示。

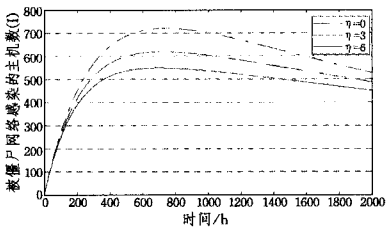

图 4　被僵尸网络感染的主机数随时间的变化图

对比图 3 和图 4 发现，在考虑预先免疫措施的模型中，感染的最大主机数远远小于不考虑免疫措施的模型，而且被感染主机的减少速度也远远大于不考虑免疫措施的模型。说明不考虑预先免疫措施的模型会高估被感染主机数，而且会低估主机的恢复速度。

结束语　僵尸网络作为一种有组织的网络攻击平台，已经对网络安全造成了很大的威胁。数学建模对研究已有僵尸网络传播的特性以及预测未来僵尸网络的传播动力学是一种非常有效的手段。本文结合僵尸网络传播的特点以及 Internet 的实际情况，提出了新的具有变化感染率的僵尸网络传播模型。模型同时考虑了预先免疫措施对僵尸网络传播的影响。模拟实验对比了本文提出的考虑时区因素和变化感染率的模型与已有模型的差异。结果表明，具有变化感染率的僵尸网络传播模型能更准确地反映实际 Internet 上僵尸病毒的传播特性。因而本文提出的模型对预测未来僵尸网络的传播行为有很好的指导意义。网络安全人员可以根据模型中相关参数反映的情况，采取相应的对策来维护网络安全。今后将结合网络拓扑结构的特点对模型进行进一步改进。

参 考 文 献

[1] Song L P, Jin Z, Sun G Q. Modeling and analyzing of botnet interactions [J]. Physica A, 2011, 390(2): 347-358

[2] 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

[3] 邓君，张兆心，张淼，等. 基于僵尸网络的攻击验证模型研究[J]. 微计算机信息，2011，27(1)：187-189

[4] 李润恒. 大规模网络中僵尸网络分析技术研究[D]. 长沙：国防科学技术大学，2010

[5] 杜勤，吕光丽，苏忠，等. P2P 僵尸网络研究与进展[J]. 计算机科学，2011，38(10A)：66-71

[6] 诸葛建伟，韩心慧，周永林，等. 僵尸网络研究[J]. 软件学报，2011，19(3)：702-715

[7] 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

[8] Barford P, Yegneswaran V. A Inside Look at Botnet [J]. Malware Detection, 2007, 27: 171-191

[9] Dagon D, Zou C C, Lee W. Modeling Botnet Propagation Using Time Zones [C]// Proceedings of 13th Annual Network and Distributed System Security (NDSS'06). 2006

[10] Bridwell L M, Tippett P. ICSA Labs 9th Annual Computer Virus Prevalence Survey [M]. ICSA Labs, 2004

[11] 冯丽萍，王鸿斌，冯素琴. 改进的 SIR 计算机病毒传播模型[J]. 计算机应用，2011，31(7)：1891-1893

[12] 冯丽萍，王鸿斌，冯素琴. 基于生物学原理的计算机网络病毒传播模型[J]. 计算机工程，2011，37(11)：155-157

[13] 马知恩，周义仓. 传染病动力学的数学建模与研究[M]. 北京：科学出版社，2004

(上接第 44 页)

[7] Venkataraman V, Yoshida K, Francis P. Chunkyspread: Heterogenous Unstructured Tree-Based Peer-to-Peer Multicast[C]// 14th IEEE International Conference on Network Protocols. IEEE, 2006: 2-11

[8] Hales D. From Selfish Nodes to Cooperative Networks-Emergent Link-based Incentives in Peer-to-Peer Networks [C]// Fourth International Conference on Peer-to-Peer Computing. IEEE, 2004: 151-158

[9] Feldman M, Papadimitriou C, Chuang J. Free-riding and Whitewashing in Peer-to-Peer Systems[J]. IEEE Journal on Selected Areas in Communications, 2006, 24(5): 1010-1019

[10] Yang M, Feng Q, Dai W F. A Multi-dimensional Reputation System Combined with Trust and Incentive Mechanisms in P2P File Sharing Systems[C]// 27th International Conference on Distributed Computing Systems Workshops. IEEE, 2007: 29-29

[11] Lian Q, Zhang Z, Yang M. An Empirical Study of Collusion Behavior in The Maze P2P File-sharing System[C]// 27th International Conference on Distributed Computing Systems. Institute of Electrical and Electronics Engineers, 2007: 56-56

[12] Fedotova N, Veltri L. Byzantine Generals Problem in The Light of P2P Computing[C]// 2006 Third Annual International Conference on Mobile and Ubiquitous Systems: Networking & Services. IEEE, 2006: 1-5

[13] 孙玉星，赵燕飞，李娅，等. 基于概率博弈的无线自组网信任推荐激励策略的研究[J]. 计算机科学，2011，38(4)：65-71

[14] Xie F, Liu F M, Yang R R, et al. Game-Based Incentive Mechanism for Cooperation in P2P Networks[C]// Fourth International Conference on Natural Computation. Computer Society, 2008: 498-501

[15] 林玉清，朱琦，朱洪波. 认知 Ad hoc 网络中基于市场的三级频谱分配方案[J]. 电子与信息学报，2010，32(3)：643-648

[16] Forges F, Orzach R. Core-stable Rings in Second Price Auctions with Common Values[J]. Journal of Mathematical Economics, 2011, 47(6): 760-767

[17] Amor D R, Fort J. Effects of Punishment in a Mobile Population Playing the Prisioner's Dilemma Game[J]. Physical Review, 2011, 84(6): 1-10