第42卷 第6期 西 安 交 通 大 学 学 报 Vol.42 №6
2008年6月 JOURNAL OF XI′AN JIAOTONG UNIVERSITY Jun. 2008

一种基于子网保护的蠕虫传播模型及其应用

史榅，齐勇，杨斌侠
（西安交通大学电子与信息工程学院，710049，西安）

摘要：针对网络蠕虫传播特征，对主动蠕虫传播模型进行了修正和改进，进而提出了一种基于子网保护的蠕虫传播模型.该模型在保证实际使用地址数和利用保护比例的修正模型的情况下，考虑了不同的扫描策略、每个周期的有效扫描数以及本地优先扫描等因素，并用它们评价了基于本地网络的蠕虫检测定位方法对蠕虫在整个网络中的传播速率的抑制效果.计算结果表明，如果网络中的多数子网都能部署蠕虫防御系统，则能大幅降低蠕虫在全网中的传播速率.

关键词：网络蠕虫；扫描策略；传播模型
中图分类号：TP393.08 **文献标志码**：A **文章编号**：0253-987X(2008)06-0788-03

Local Network Based Worm Propagation Model and Its Application

SHI Yi, QI Yong, YANG Binxia
(School of Electronics and Information Engineering, Xi′an Jiaotong University, Xi′an 710049, China)

对网络蠕虫传播精确建模[1]，一方面可以更好地描述现代网络蠕虫传播的过程，另一方面可以验证蠕虫检测防御系统的有效性.

蠕虫传播是一种离散事件，如果能用离散模型为蠕虫传播建模，更能体现蠕虫传播的本质特性.为此，陈泽生等人提出了刻画随机扫描主动蠕虫传播的离散模型(AAWP)[2]，它比传统的免疫学模型更接近网络中蠕虫的真实传播情况，但AAWP模型没有区分不同的扫描策略，也没有考虑有效扫描数以及子网扫描等特殊情况.顾国飞等人[3]指出，在真实网络中并不是所有2^{32}个IPv4地址都被分配，只有大约10^9个地址是有效的，并认为如果网络中有D台主机位于部署了检测系统的子网内，只有D台之外被保护的主机才会引发新的感染，因此蠕虫的传播速率有所降低.

据此，本文在AAWP模型基础上进行了修正和改进，提出一种基于子网保护的蠕虫传播模型.

1 基于子网保护的蠕虫传播模型

基于AAWP模型，在考虑了网络中实际使用的地址数和利用了保护比例的修正模型的情况下，本文认为还有下述因素值得考虑.

(1)扫描策略.常见的随机扫描策略有：①统一随机扫描，如Code Red和Sapphire/Slammer，蠕虫可能生成随机地址的总数$\Omega=2^{32}$，文献[3]模型考虑的就是这种情形；②选择性随机扫描，$\Omega=160\times 2^{24}\approx 2.68\times 10^9$；③可路由扫描(routable scan)，$\Omega\approx T=10^9$.

AAWP模型只考虑了统一随机扫描的情况($\Omega=2^{32}$)，而且认为全部2^{32}个IPv4地址都是有效地址($T=2^{32}$)，忽略了网络中实际分配地址数小于IPv4地址总数的事实.

若忽略死亡率和补丁率，扫描策略①～③下的蠕虫传播模型为

$$n_{i+1}=n_i+(N-n_i)\left[1-\left(1-\frac{1}{T}\right)^{sn_i\frac{T}{\Omega}}\right] \quad (1)$$

式中：n_i表示时刻i受害主机数；N表示易受感染的主机总数；s表示单位时间内一台受感染主机扫描的平均次数；Ω表示蠕虫可能生成随机地址的总数，其取值依扫描方式而定.

(2)每个周期的有效扫描数.文献[3]认为，当发现蠕虫扫描后，所有被保护的网络都能够阻止本地网络中受害主机对其他主机的扫描.然而，检测系统都需要一定时间去判断数据包的特征或网络流量

收稿日期：2007-10-18. 作者简介：史榅(1972—)，女，在职博士生，讲师；齐勇(联系人)，男，教授，博士生导师. 基金项目：国家自然科学基金资助项目(60403028)；陕西省自然科学基金资助项目(2004F43).

特征，在系统检测到扫描行为之前，感染蠕虫的主机已经发出了一些扫描包，这不符合文献[3]中所有扫描包都被阻断的假设，已经发出的扫描包不应忽略，特别是在扫描速率较高的情况下. 为此，本文假设第 i 个时间周期新感染蠕虫的主机数为 I_i，蠕虫扫描了 τ(为了简化表述，取 τ 为整数)个时间周期后被检测系统发现，则第 i 个周期有效的扫描数为

$$q_i = n_i \frac{T}{\Omega}(1-\alpha) + \sum_{j=i-\tau+1}^{i} I_j \frac{T}{\Omega}\alpha \tag{2}$$

式中：$\alpha = D/T$. 式(2)的含义为，有效的扫描数等于未受保护网络中发出的扫描数与受保护网络中发现受害者之前新增的受害主机发出的扫描数之和. 因此，蠕虫传播模型可为

$$\left.\begin{aligned} q_i &= n_i \frac{T}{\Omega}(1-\alpha) + \sum_{j=i-\tau+1}^{i} I_j \frac{T}{\Omega}\alpha \\ I_{i+1} &= (N-n_i)\left[1-\left(1-\frac{1}{T}\right)^{sq_i}\right] \\ n_{i+1} &= n_i + I_{i+1} \end{aligned}\right\} \tag{3}$$

式中：$I_1 = n_1 = h$. 显然，文献[3]中的传播模型是 $\Omega = 2^{32}$，$\tau = 0$ 时的特殊情况.

(3)本地优先扫描. 该优先扫描又称子网扫描，是指受害机器以 P_0 概率扫描所有的地址，以 P_1 概率扫描与本地地址前 8 位相同的地址，以 P_2 概率扫描与本地地址前 16 位相同的地址，以 P_3 概率扫描与本地地址前 24 位相同的地址. 根据文献[3]，一个时间周期内扫描到这 4 类网络的次数为

$$\left.\begin{aligned} k_1 &= P_3 sV_1 + P_2 sg_1/2^8 + P_1 sg_2/2^{16} + P_0 sg_3/b \\ k_2 &= P_3 sV_2 + P_2 sg_1/2^8 + P_1 sg_2/2^{16} + P_0 sg_3/b \\ k_3 &= P_3 sV_3 + P_2 sV_3 + P_1 sg_2/2^{16} + P_0 sg_3/b \\ k_4 &= P_3 sV_4 + P_2 sV_4 b/2^{24} + P_1 sV_4 b/2^{24} + P_0 sg_3/b \end{aligned}\right\} \tag{4}$$

式中

$$\left.\begin{aligned} g_1 &= V_1 + (2^8-1)V_2 \\ g_2 &= V_1 + (2^8-1)V_2 + (2^{16}-2^8)V_3 \\ g_3 &= V_1 + (2^8-1)V_2 + (2^{16}-2^8)V_3 + (b-2^{16})V_4 \end{aligned}\right.$$

其中 V_j 表示第 j 类网络中的平均受害主机数. 若 $V_{j,i}$ 表示在第 i 个时间周期第 j 类网络中的受害主机数，则在没有任何保护的情况下

$$V_{j,i+1} = V_{j,i} + \left(2^8 \frac{N}{\Omega} - V_{j,i}\right) \cdot \left[1-\left(1-\frac{1}{2^8}\right)^{k_j}\right] \tag{5}$$

时刻 i 平均每个地址为/24 的子网受害主机数为

$$V_{1,i} = V_{1,i}/b + (2^8-1)V_{2,i}/b + (2^{16}-2^8)V_{3,i}/b + (b-2^{16})V_{4,i}/b \tag{6}$$

所以，网络中受害主机的总数为

$$n_i = V_{1,i} + (2^8-1)V_{2,i} + (2^{16}-2^8)V_{3,i} + (b-2^{16})V_{4,i} \tag{7}$$

在这种情况下，假设网络中检测系统均匀分布，部署率为 α($\alpha = D/T$)，第 j 类网络新增的受害主机数为 I_j，检测系统用 1 个时间周期($\tau = 1$)发现并停止自己子网内的扫描，则各类子网收到的扫描包数为

$$\left.\begin{aligned} k_1 &= (1-\alpha)(P_3 sV_1 + P_2 sg_1/2^8 + P_1 sg_2/2^{16} + P_0 sg_3/b) + \alpha(P_3 sI_1 + P_2 sg'_1/2^8 + P_1 sg'_2/2^{16} + P_0 sg'_3/b) \\ k_2 &= (1-\alpha)(P_3 sV_2 + P_2 sg_1/2^8 + P_1 sg_2/2^{16} + P_0 sg_3/b) + \alpha(P_3 sI_2 + P_2 sg'_1/2^8 + P_1 sg'_2/2^{16} + P_0 sg'_3/b) \\ k_3 &= (1-\alpha)(P_3 sV_3 + P_2 sV_3 + P_1 sg_2/2^{16} + P_0 sg_3/b) + \alpha(P_3 sI_3 + P_2 sI_3 + P_1 sg'_2/2^{16} + P_0 sg'_3/b) \\ k_4 &= (1-\alpha)(P_3 sV_4 + P_2 sV_4 b/2^{24} + P_1 sV_4 b/2^{24} + P_0 sg_3/b) + \alpha(P_3 sI_4 + P_2 sI_4 b/2^{24} + P_1 sI_4 b/2^{24} + P_0 sg'_3/b) \end{aligned}\right\} \tag{8}$$

式中

$$\left.\begin{aligned} g_1 &= V_1 + (2^8-1)V_2 \\ g_2 &= V_1 + (2^8-1)V_2 + (2^{16}-2^8)V_3 \\ g_3 &= V_1 + (2^8-1)V_2 + (2^{16}-2^8)V_3 + (b-2^{16})V_4 \end{aligned}\right\} \tag{9}$$

$$\left.\begin{aligned} g'_1 &= I_1 + (2^8-1)I_2 \\ g'_2 &= I_1 + (2^8-1)I_2 + (2^{16}-2^8)I_3 \\ g'_3 &= I_1 + (2^8-1)I_2 + (2^{16}-2^8)I_3 + (b-2^{16})I_4 \end{aligned}\right\} \tag{10}$$

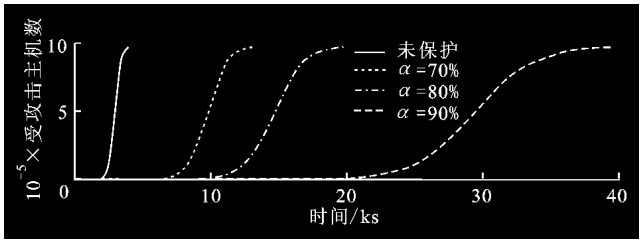
假设第 i 个时间周期第 j 类网络新增的受害主机数为 $I_{j,i}$，则

$$\left.\begin{aligned} I_{j,i+1} &= \left(2^8 \frac{N}{\Omega} - V_{j,i}\right)\left[1-\left(1-\frac{1}{2^8}\right)^{k_j}\right] \\ V_{j,i+1} &= V_{j,i} + I_{j,i+1} \end{aligned}\right\} \tag{11}$$

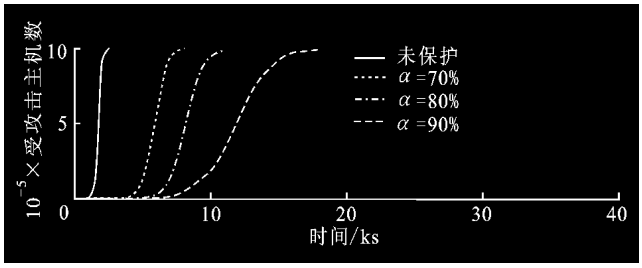
上述 3 种随机扫描又可以看成是子网优先扫描中 $P_0 = 1$，$P_1 = P_2 = P_3 = 0$，Ω 分别为 2^{32}、2.68×10^9 和 10^9 时的特殊情况.

本文将以上修正和改进后的模型称为基于子网保护的蠕虫传播模型，在此之前对蠕虫的研究工作中曾提出了一种基于本地网络的蠕虫检测定位方法

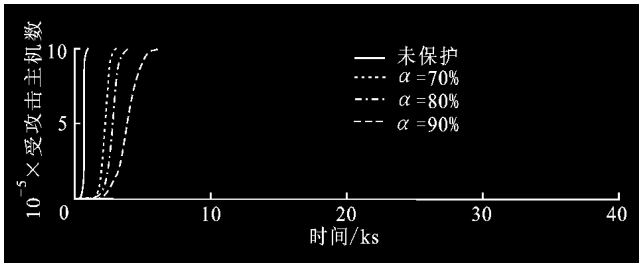
LTBWD(论文待发),下面将用改进后的模型评价配置有 LTBWD 的蠕虫防御系统在降低蠕虫传播速率方面的有效性.



(a)统一随机扫描



(b)选择性随机扫描



(c)可路由扫描

图1 网络中主机被感染的情况

2 效果评价

不失一般性,令 $s=20$,$T=10^9$,$\Omega=2^{32}$,$N=10^6$,$P_0=1$,$P_1=P_2=P_3=0$. 为了简化运算,认为 $\tau=1$,则在不同部署率下整个网络中受感染的主机数如图1所示. 从中可以看出,如果 LTBWD 系统能及时发现所管辖子网内的受害主机,并阻止它们扫描,就能有效地减缓蠕虫在整个网络中的传播速率. 其原因在于,只有在保护网络之内暂时未被检测到的受害主机和被保护网络之外的受害主机才会进行蠕虫传播,所以引起传播的受感染主机数比较少. 如果在靠近主机端的网络设备上部署了检测系统,仅通过检测自己管辖的子网内的主机感染情况,阻止子网内的受害主机扫描其他主机,就能有效减缓蠕虫的传播,从而减轻蠕虫对整个网络的危害.

3 结 论

本文改进了 AAWP 模型,提出了一种基于子网保护的蠕虫传播模型,并用它评价了基于本地网络的蠕虫检测定位方法对蠕虫在整个网络中的传播速率的抑制效果. 实验、计算表明,如果网络中的多数子网都能部署这种蠕虫防御系统,则能大幅降低蠕虫在全网中的传播速率.

参考文献:

[1] FRAUENTHAL J C. Mathematical modeling in epidemiology [M]. Berlin, Germany: Springer-Verlag, 1980.

[2] CHEN Zesheng, GAO Lixin, KEVIN K. Modeling the spread of active worms[C]// IEEE Conference on Computer Communications. Piscataway, NJ, USA: IEEE, 2003: 1890-1900.

[3] GU Guofei, SHARIF M, QIN Xinzhou, et al. Worm detection, early warning and response based on local victim information [C]// 20th Annual Computer Security Applications Conference. Los Alamitos, CA, USA: IEEE Computer Society, 2004:136-145.

(编辑 苗凌)