

基于拓扑结构的蠕虫防御策略仿真分析

王跃武 荆继武 向 继 刘 琦

(中国科学院研究生院信息安全国家重点实验室 北京 100049)

摘 要 提出了基于拓扑结构控制的蠕虫防御策略,并通过构建仿真模型对其进行了仿真验证分析.首先对蠕虫传播所依赖的拓扑结构的主要形式进行了分析,提出了相应的生成算法,并对算法的有效性进行了验证;随后提出了三种拓扑结构控制策略仿真模型;最后分别对这三种策略在不同拓扑结构下的蠕虫传播控制性能进行了仿真实验.实验结果证明:通过适当地控制拓扑结构,可以有效地遏制拓扑相关蠕虫传播.

关键词 仿真;拓扑相关蠕虫;拓扑结构;蠕虫防御策略

中图法分类号 TP393

A Simulation Analysis of Worm Defense Strategies Based on Topology Structure

WANG Yue-Wu JING Ji-Wu XIANG Ji LIU Qi

(State Key Laboratory of Information Security, Graduate University of Chinese Academy of Sciences, Beijing 100049)

Abstract Topology aware worms have been an important security threat on the Internet. They can spread across the Internet quickly, through topology structure information. If the topology structure were destroyed by defense strategies, the worm propagation can be held back effectively. Thus, in order to design effective topology aware worm defense strategies, it is necessary to analyze the relationship between worm defense strategies and topology structure. This paper provides a systemic analysis of worm defense strategies based on topology structure through packet level worm simulation. First the major topology structures used by topology aware worms and their generation algorithms are analyzed. Then, three defense strategy models are drawn from mainstream worm defense strategies. Finally, these defense strategies in different topology structure are analyzed with simulation experiments, and some interesting conclusions are drawn from these experiment results. These conclusions can provide valuable guidelines for real defense system implementation.

Keywords simulation; topology aware worm; topology structure; worm defense strategies

1 引 言

长期以来,拓扑相关蠕虫一直是网络安全的一个重要威胁.在 Internet 出现初期,网络节点稀疏,

主动扫描蠕虫难以在较短的时间内发现大量易感染主机,无法快速传播,因此蠕虫攻击多通过拓扑结构信息,发现目标主机,进行传播,如 Morris 蠕虫^[1]等.随着 e-mail 应用的普及,e-mail 地址形成了一个巨大的应用层网络,为蠕虫提供了一个非常适宜

收稿日期:2007-04-06;修改稿收到日期:2007-08-03.本课题得到国家自然科学基金(60573015)资助.王跃武,男,1975年生,博士研究生,主要研究方向为网络安全技术、大规模网络蠕虫仿真. E-mail: ywwang@lois.cn.荆继武,男,1964年生,教授,博士生导师,主要研究领域为网络与系统安全技术、PKI技术、入侵容忍技术和蠕虫仿真技术.向 继,男,1976年生,博士研究生,主要研究方向为网络安全技术、网络蠕虫仿真.刘 琦,女,1978年生,博士研究生,主要研究方向为网络安全技术、网络蠕虫仿真.

的传播环境,大量 e-mail 蠕虫相继出现,如 Melissa^①, Sobig^②等.目前,随着蠕虫自动检测防御技术的发展,蠕虫传播更加注重隐蔽性,而拓扑相关蠕虫传播由于不需要发送大量探测数据包,具有较高的隐蔽性.此外,随着 P2P,即时消息(Instant Message,IM)等网络应用的发展,新的拓扑相关蠕虫传播环境正在形成.所以今后一段时间,拓扑相关蠕虫的威胁将会更加严重.现在已经出现多种 IM 蠕虫和 P2P 蠕虫.因此研究拓扑相关蠕虫传播防御策略具有重要意义.

拓扑相关蠕虫传播与拓扑结构特性关系密切.其传播过程可简单描述如下:网络中一个节点被感染后,根据其所保存的相邻节点信息发送蠕虫数据包,感染其相邻节点,其相邻节点被感染后,同样根据其保存的相邻节点信息,继续发送蠕虫数据包,不断感染新的节点.可以看出拓扑相关蠕虫传播依赖于拓扑结构的连通性及鲁棒性.其中鲁棒性是指在存在节点感染失败可能的情况下,蠕虫传播能够持续的能力.如果能够通过一定的控制措施,破坏拓扑结构的连通性和鲁棒性,则可以有效地遏制蠕虫传播.然而由于拓扑结构的复杂性和蠕虫行为的随机性,现有的蠕虫传播分析方法如 Two-Factor 模型^[2]等,难以对其进行系统的分析研究. Briesemeister 等人对不同拓扑结构中的蠕虫防御效果进行了简单对比^[3]. Zou 等人通过 Monte Carlo 仿真方法分析了 e-mail 蠕虫传播的情况^[4].但是对于如何通过控制拓扑结构的连通性和鲁棒性阻断蠕虫传播,目前还没有一个全面、系统的研究.

本文基于我们此前提出的数据包级拓扑相关蠕虫仿真方法,对如何进行拓扑结构控制,阻断蠕虫传播进行了全面系统的研究.本文的主要工作包括:(1)分析了目前拓扑相关蠕虫传播所依赖的三种主要拓扑结构模型:Power-Law 模型、Small-World 模型以及混合结构模型,并在数据包级仿真系统中实现了三种拓扑结构模型的仿真生成算法;(2)分析了目前主要的蠕虫防御策略,提出了基于拓扑结构控制的主动防御、被动保护以及局部隔离等三种策略模型;(3)通过数据包级拓扑相关蠕虫仿真系统,分别对不同拓扑结构下的三种蠕虫防御策略进行了实验分析,并根据实验结果提出了这些防御措施实施过程中需要注意的一些重要原则.

本文第 2 节介绍蠕虫传播所依赖的主要拓扑结构模型及其仿真生成算法和基于拓扑结构控制的三种蠕虫防御策略仿真模型;第 3 节对基于拓扑结构

的主动防御策略进行仿真分析;第 4 节对基于拓扑结构的被动保护策略进行仿真分析;第 5 节对基于拓扑结构的局部隔离策略进行仿真分析;最后总结全文.

2 拓扑结构及蠕虫防御策略仿真模型分析

拓扑相关蠕虫传播所依赖的逻辑拓扑结构可以用有向图 $G=\langle V, E \rangle$ 表示,任意一个顶点 $\forall v \in V$ 表示一个主机节点,任意一条有向边 $\forall e=\langle u, v \rangle \in E$, $u, v \in V$ 表示主机 u, v 之间存在联系.这种联系可以是节点 u 中保存了节点 v 的邮件地址或者 IM 联系人信息以及其它类似信息.根据拓扑结构生成规则的不同,可以有多种不同的结构形式.主要存在三种拓扑结构形式:Power-Law 结构、Small-World 结构以及混合结构.以下分别对三种拓扑结构形式及其仿真生成算法进行详细描述.本节最后对三种基于拓扑结构的蠕虫防御策略仿真模型进行描述.

2.1 Power-Law 拓扑结构及其仿真生成算法

在 P2P 网络系统生成中,为了通信方便,一个新增节点总是选择连接度较高的节点进行连接.这样形成的拓扑结构中,一部分节点拥有大多数的连接,而其余大部分节点则只占有一少部分连接.这种拓扑结构中节点的连接数分布可以表示为 $P(d)=cd^{-\alpha}$, $d=m, m+1, \dots, M$. 其中 $P(d)$ 表示节点拥有 d 个连接的概率, c 为一常数, m 表示拓扑结构中节点拥有的最小连接数, M 表示最大连接数, α 决定分布曲线的形状.具有这种特性的拓扑结构被称为 Power-Law 结构.多种网络拓扑结构测量结果显示其结构符合 Power-Law 结构模型,如 Internet AS 拓扑结构, $\alpha \approx 2.25^{[5]}$; Guntella 拓扑结构, $\alpha \approx 2.3^{[6]}$; Web 网页的链入拓扑结构, $\alpha \approx 2.09$; 链出拓扑结构, $\alpha \approx 2.72^{[7]}$.

在 Tian Bu 等人提出的 Power-Law 拓扑生成算法^[8]的基础上,本文提出了一个有向 Power-Law 拓扑生成算法.首先假定拓扑结构有 m_0 个节点,用 $2 \times (m_0 - 1)$ 条有向边连接.随后重复以下两个步骤,直至网络规模达到仿真要求:(1)以概率 P_e 添加 m 条边到拓扑结构中,对于每条边,分别以概率

① CERT. CERT Advisory CA-1999-04 Melissa Macro Virus. <http://www.cert.org/advisories/ca-1999-04.html>

② The W32/Sobig. E@MM Virus Spreads Steadily; Forges Its "From;" Field. <http://virusbusters.itsc.umich.edu/sobig-e.html>

$(d_a - \beta) / \sum_j (d_j - \beta)$ 和 $(d_b - \beta) / \sum_j (d_j - \beta)$ 选择不同的节点 v_a 和 v_b 作为该边的起始和终止节点, 并以概率 q 生成该边的逆向边; (2) 以概率 $1 - P_e$ 添加一个节点到拓扑中, 以该节点为起始节点生成 m 条有向边, 以概率 $(d_i - \beta) / \sum_j (d_j - \beta)$ 选择节点 v_i 作为每条边的终止节点, 并以概率 q 生成每条边的逆向边. 其中, d_a, d_b, d_j, d_i 分别代表每个步骤操作时, 节点 v_a, v_b, v_j, v_i 拥有的出度连接数. 由此生成的拓扑结构中节点出度连接数分布为 Power-Law 分布, 且 α 可以用下式计算:

$$\alpha = \frac{(1+q)m - (1-p_e)\beta}{(p_e+q)m} + 1 \quad (1)$$

一般取 $-\infty < \beta < 1$. 图 1 为该算法的仿真实验结果. 可以看出随着参数选择的不同, 该算法可以生成具有各种不同 α 值的拓扑结构, 并且对数坐标下, 节点出度连接分布基本上呈现一条直线, 符合 Power-Law 结构特性.

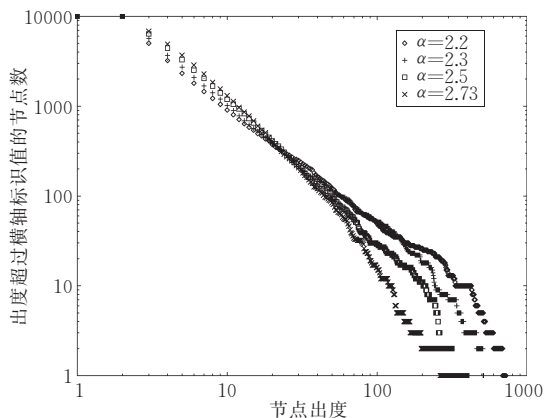


图 1 不同 α 值的 Power-Law 拓扑生成算法仿真结果

2.2 Small-World 拓扑结构及其仿真生成算法

许多网络应用系统的拓扑结构为人类社会联系在网络环境中的映射, 如 e-mail 地址网络和 IM 联

系人网络等. 因此可以用社会联系模型模拟这类网络拓扑结构. Watts 等人提出用 Small-World 模型模拟人类社会联系模型^[9], 所以本文用 Small-World 模型模拟这类拓扑结构.

Small-World 模型可以用图 2 进行解释. 开始为左边所示的规则图形. 对于规则图形中的每一条边, 以概率 p_r 重新生成该边. 保持一个端点不变, 在拓扑中随机地选择一个节点作为另一个端点, 形成新的替代边. 对规则图形中所有节点这样处理后, 形成 Small-World 模型, 如中间图所示. 特别地当 $p_r = 1$ 时, 拓扑结构变为完全随机模型, 如右边所示. Small-World 模型的两个重要属性为: 较高的聚合系数 C 和较小的特征路径长度 L . 我们根据该模型提出了一个有向 Small-World 模型生成算法. 每个节点拥有的平均邻接点数表示为 k , 图 2 中 $k=4$.

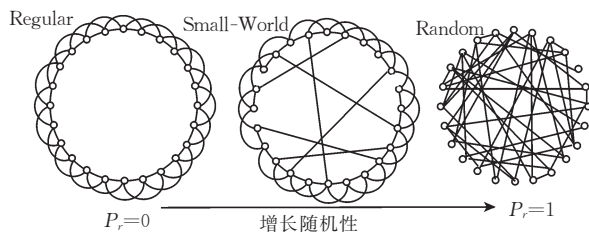


图 2 Watts 的 Small-World 模型

图 3 为该算法的仿真结果. 可以看出随着 p_r 的逐渐变大, L 迅速下降, 而 C 却一直直到 p_r 很大时才开始下降, 所以该算法可以通过调整 p_r 来实现较高的 C 和较短的 L 的 Small-World 结构特性组合. 由于少量的随机边可以极大地降低 L , 所以我们称这些边为“shortcut”.

2.3 混合拓扑结构及其仿真生成算法

在社会联系模型从实际环境向网络环境转换时, 由于网络环境通信的便捷性, 会发生一些改变. 如人们的联系人数量可能会相对实际环境中

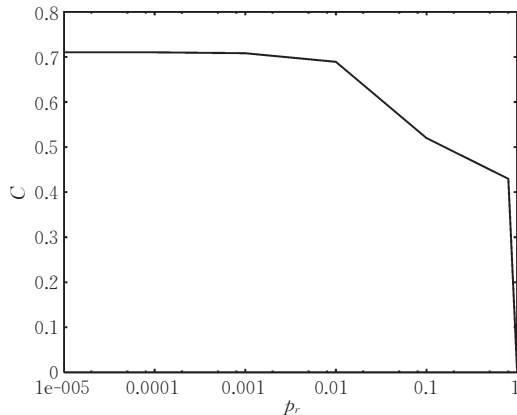
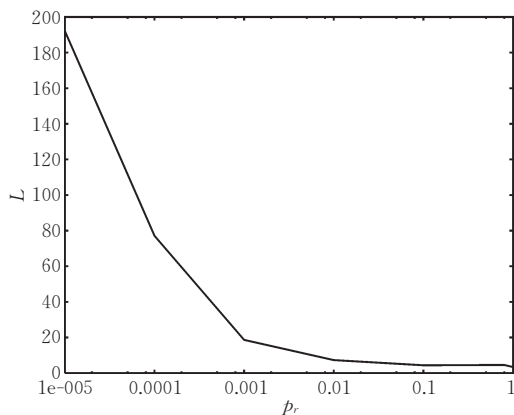


图 3 Small-World 拓扑中不同 p_r 下的 L 和 C 值

所增加. 其中对蠕虫传播影响较大的为用户群的出现, 如电子邮件群和 QQ 群等. 为了通信的方便, 人们在加入用户群时往往选择规模较大的群, 这样用户群的规模分布将呈现 Power-Law 分布, Zou 等人对 yahoo 邮件列表进行分析发现其规模分布近似 Power-Law 分布^[4]. 所以我们用混合拓扑结构模型模拟这类网络结构特性. 从本质上讲, Small-World 模型也可以看作规则模型和随机模型组成的一种混合结构特例.

在仿真系统中, 我们按照如下方法构造混合拓扑结构. 首先生成 Small-World 拓扑结构, 并生成初始规模呈 Power-Law 分布的用户群. 对于 Small-World 拓扑结构中的每个节点, 以概率 p_g , 从用户群中以概率 $S_i / \sum_j S_j$ 选择用户群 i 加入. 其中 S_i, S_j 分别表示节点加入时, 用户群 i, j 的规模. 可以证明, 随着节点的不断加入, 用户群规模将保持 Power-Law 分布, 并且 α 值将保持不变. 图 4 为该算法的仿真实验结果, 其中用户群个数为 500, $\alpha = 2.2$. 可以看出随着节点规模的不断增加, 用户群规模分布始终保持 Power-Law 分布, 并且分布曲线基本保持平行, 证明了该算法可以有效地生成具有混合结构的仿真拓扑.

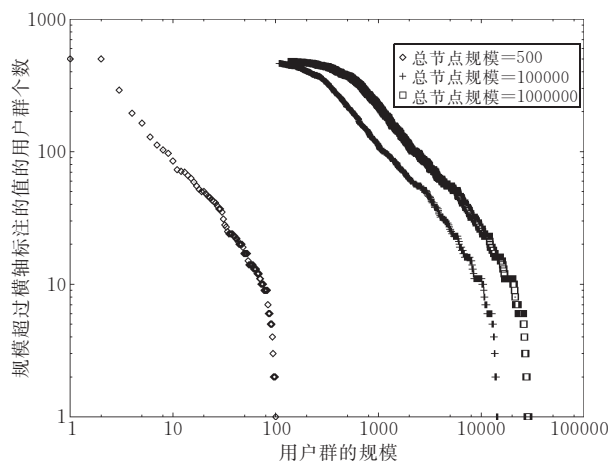


图 4 不同总节点规模下的用户群规模分布

2.4 基于拓扑结构的蠕虫防御策略模型分析

目前存在多种蠕虫防御策略, David Brumley 等人对其进行了系统的分析总结^[10]. 忽略各种防御策略的具体实现技术, 蠕虫防御策略可以分为如下三种: 主动防御策略、被动保护策略和局部隔离策略. 我们将分别对拓扑结构环境下的三种蠕虫防御策略的仿真模型进行描述.

主动防御策略, 即在蠕虫传播之前采取一定措施, 加固节点, 使节点可以免受蠕虫感染, 如打补丁

等措施. 在拓扑结构中, 主动防御措施相当于将一部分节点从拓扑结构中去除. 当这样的节点不断增多, 能够将拓扑结构的连通性破坏时, 即可成功阻断蠕虫传播. 在主动扫描蠕虫传播过程中, 蠕虫传播途径可以抽象为全连通结构, 所以无法通过节点控制, 有效地破坏蠕虫传播途径. 仿真中, 我们将主动防御策略抽象为: 从拓扑结构中去除部分节点. 拓扑结构下的主动防御措施需要解决: 如何进行节点控制才能最有效地破坏蠕虫传播网络的连通性.

被动保护策略, 即在蠕虫爆发后, 根据特定的蠕虫特性, 采取相应的措施, 降低节点被感染的概率. 目前存在多种被动保护策略的实现方法, 如 Signature 过滤、地址过滤等. 本文忽略过滤措施的具体实现细节, 将其抽象为节点的被感染概率. 若过滤措施有效, 则节点以较小的概率被感染; 反之节点以较大的概率被感染. 根据中心极限定理, 我们用正态分布表示每个节点的感染概率: $p_i \sim N(\mu_i, \delta_i^2)$, 其中 μ_i 表示节点被感染概率的平均值, δ_i 表示节点被感染率的方差. 本文主要研究拓扑蠕虫传播的鲁棒性和被动保护策略的关系.

局部隔离策略, 为“Good Neighbor”技术思想的一个实例. 它通过减少已感染节点转发蠕虫数据包, 阻断蠕虫传播. Williamson 等人提出了一个局部隔离策略的具体实现形式: virus throttle^[11], 它通过限制节点的对外连接率, 阻断蠕虫数据包的转发. 在拓扑相关蠕虫传播中实施局部隔离措施, 存在一定的难度, 因为拓扑相关蠕虫传播过程中, 被感染节点总是向固定的节点发送蠕虫数据包, 蠕虫流量特性和正常流量特性差别不大. 本文忽略局部隔离措施的具体实现细节, 假定将限制节点转发蠕虫数据包的方式, 作为拓扑结构下局部隔离措施的一种抽象. 本文将每个节点成功转发蠕虫数据包的概率表示为 $p_s \sim N(\mu_s, \delta_s^2)$. 该策略与被动保护策略相似, 本文主要分析拓扑结构下两种策略在蠕虫传播控制上的差异.

3 基于拓扑结构的主动防御策略分析

在本节, 本文将通过数据包级拓扑相关蠕虫仿真系统, 对三种拓扑结构下的主动防御措施进行分析. 假定蠕虫传播过程中, 节点收到蠕虫数据包后, 以 1.0 的概率被感染, 延时一段时间后, 按照其保存的相邻接点地址信息, 转发蠕虫数据包, 延时时间服从指数分布, $\lambda = 1/30.0$. 蠕虫数据包发送可以采

用两种方式:Non-reinfection 方式和 Reinfection 方式. Non-reinfection 方式指节点仅转发一次蠕虫数据包,其后在任何情况下将不再转发蠕虫数据包. Reinfection 是指节点无论在何情况下,收到蠕虫数据包后,都将按照邻接点地址信息转发蠕虫数据包. 因为在不考虑节点被感染概率的情况下,两种蠕虫数据包转发方式对拓扑相关蠕虫传播的影响基本一致,所以本节分析仅考虑 Non-reinfection 传播方式.

3.1 Power-Law 结构下的主动防御策略

Power-Law 拓扑结构下,主动防御策略的有效性与拓扑结构的连通性关系密切. Cohen 等人分析了 Power-Law 结构中节点随机故障与拓扑连通性之间的关系^[12]. 在该研究基础上,我们对 Power-Law 结构中的主动防御策略进行分析. 首先我们假定完全随机地对拓扑结构中的节点实施主动防御策略. 导致蠕虫传播被成功阻断的控制节点比率的阈值 p_c 可以按如下公式计算:

$$1 - p_c = \frac{1}{\kappa_0 - 1} \quad (2)$$

其中 κ_0 表示为

$$\kappa_0 \rightarrow \left| \frac{2-\alpha}{3-\alpha} \right| \times \begin{cases} m, & \alpha > 3 \\ m^{\alpha-2} M^{3-\alpha}, & 2 < \alpha < 3 \\ M, & 1 < \alpha < 2 \end{cases} \quad (3)$$

考虑蠕虫传播环境中的拓扑结构通常有 $2 < \alpha < 3$, 并且 M 可以估计为 $mN^{1/\alpha-1}$, N 表示拓扑结构中的节点总数^[15]. 所以, p_c 又可以表示为

$$p_c = 1 + \left(1 - m^{(\alpha-2) \cdot (3-\alpha)} N^{\frac{3-\alpha}{\alpha-1}} \frac{\alpha-2}{3-\alpha} \right)^{-1} \quad (4)$$

由式(4)可以看出,Power-Law 结构的拓扑连通性与连接边的数量和分布关系密切,并且在 $2 < \alpha < 3$ 时拓扑结构有较强的连通性. 在 $N=10000$, $m=2$ 时, p_c 均在 90% 以上. 即通过随机控制,必须对 90% 以上的节点进行主动免疫控制才能有效遏制蠕虫传播. 而如果能够选择连接度较多的节点进行控制,则可以通过控制较少的节点,减少尽可能多的连接边,从而最大限度地破坏拓扑结构的连通性. 所以选择性控制为有效阻断 Power-Law 结构中蠕虫传播的一个重要途径. 图 5 为 Power-Law 结构中,不同控制比率下的蠕虫传播仿真结果, $N=10000$, $\alpha=2.5$. 可以看出通过完全随机防御控制,在控制比率达到 50% 时,蠕虫仍可以迅速感染全部未受保护的宿主,只有在控制比率达到 90% 时才可以有效阻断蠕虫传播,仿真结果与理论分析基本一

致. 当通过选择连接度较多的节点进行控制时,阻断效果明显增加,控制比率到达 50% 时,即可有效阻断蠕虫传播.

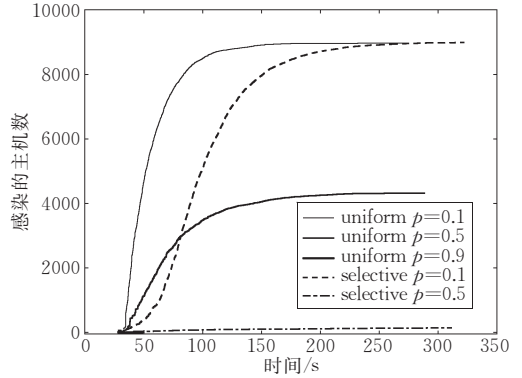


图 5 Power-Law 结构中不同控制比率下的蠕虫传播

3.2 Small-World 结构下的主动防御策略

Newman 等人对 Small-World 拓扑的连通性进行了分析^[13]. 我们在此研究基础上对 Small-World 拓扑结构下的主动防御策略进行分析. 如果不考虑“shortcut”, Small-World 拓扑结构为一个规则拓扑结构,如图 2 左图所示. 假定规则拓扑结构中存在 k 个相互连接的实施主动防御的节点群,则拓扑结构被分割成两个互不连接的区域. 假定拓扑结构中未实施主动防御措施的节点的比率为 p_{uc} ,则在两个相互连接的未受控节点群之间存在 k 个相互连接的受控节点群的概率为 $p_{uc}(1-p_{uc})^k$. 这样整个拓扑结构被分割成 n 个互不连通的区域, n 表示为

$$\begin{aligned} n &= p_{uc}(1-p_{uc})^k(N-kn) \Rightarrow n \\ &= N \frac{p_{uc}(1-p_{uc})^k}{1+kp_{uc}(1-p_{uc})^k} \end{aligned} \quad (5)$$

当受控节点增多, $n > 1$ 时,拓扑结构即可被破坏.

考虑“shortcut”,如果“shortcut”的两个端点均在互不连接的两个未受控节点群,则这两个未受控节点群可以通过“shortcut”连通. 当有 n 个这样的“shortcut”时,被切断的拓扑结构又可以形成一个连通的网络,所以这种“shortcut”的存在是 small-World 拓扑结构连通性增强的一个重要因素. 这种“shortcut”存在的数量近似为 $p_{uc}^2 p_r k N$. 所以可以得出:满足如下方程的 p_{uc} 为 Small-World 拓扑结构连通性破坏的阈值.

$$p_{uc}^2 p_r k = 0.5 \frac{(1-p_{uc})^k}{1+kp_{uc}(1-p_{uc})^k} \quad (6)$$

当未受控节点比率多于 p_{uc} 时,拓扑连通性未

受破坏,反之拓扑连通性将被破坏,蠕虫传播被阻断.由式(6)可以看出:虽然 Small-World 的连通性小于 Power-Law 结构,但是仍然具有较强的连通性,只有当 p_r 降至较低的水平时, $1-p_{uc}$ 才可能有较小的值.图 6 是 Small-World 结构下,不同节点控制比率的蠕虫传播仿真结果, $N=10000$, $k=20$.从图 6(a)可以看出,当 $p_r=1\times 10^{-2}$ 时,Small-World

拓扑结构有着较强的连通性,只有当 80% 的节点受到主动防御控制后,蠕虫传播才能被有效遏制,而在 $p_r=1\times 10^{-10}$ 时拓扑的连通性变得非常脆弱,如图 6(b)所示,在控制比率为 50% 的情况下,蠕虫传播即可被有效阻断.所以如何有效控制拓扑结构的“shortcut”是 Small-World 拓扑结构中主动防御实施的关键.

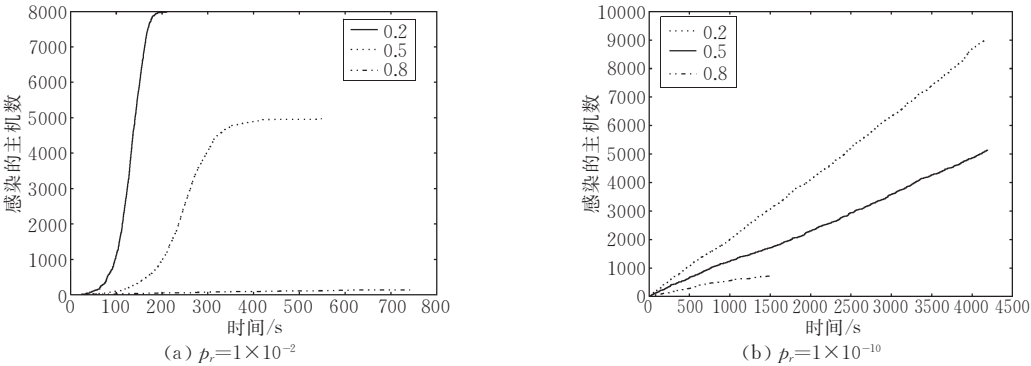


图 6 Small-World 拓扑结构下不同控制率的蠕虫传播

3.3 混合拓扑结构下的主动防御策略

混合拓扑结构由 Small-World 拓扑结构和规模呈 Power-Law 分布的用户群组成. 所以其主动防御策略可以分为两个部分:节点控制和用户群控制. 用户群的出现相当于引入大量的“shortcut”. 假定用户群的规模为 S ,则由该用户群引入的“shortcut”数量为 $S(S-1)/2$. 所以用户群的出现将极大地增加拓扑结构的连通性. 如果能够有效地控制用户群,则混合拓扑结构可以退化为 Small-World 拓扑结构,并可以进一步通过控制 p_r ,进行有效的蠕虫传播控制. 图 7 是混合拓扑结构下的主动防御策略仿真结果,其中 $k=10$, $N=10000$,用户群规模 Power-Law 分布 $\alpha=2.2$.

从图 7(a)可以看出用户群的引入增强了拓扑结构的连通性. 在 $p_r=1\times 10^{-7}$,“shortcut”很少,50% 的节点实施主动防御策略的情况下,保持 $p_g=$

0.1,蠕虫仍能够像 $p_r=0.1$ 的情况下那样,在未控制节点之间快速传播. 只有当 p_r 和 p_g 很小时,节点控制才能使蠕虫传播得到有效遏制. 所以混合拓扑结构中蠕虫传播控制必须考虑用户群的控制. 图 7(b)为不同用户群控制设置下的蠕虫传播仿真结果. 可以看出:完全随机用户群控制,即使 95% 的用户群被控制,蠕虫传播也几乎没有任何影响. 如果选择规模较大的用户群进行控制,当 30% 的用户群被控制后,蠕虫传播过程几乎与 Small-World 下的蠕虫传播过程一致.

4 基于拓扑结构的被动保护策略分析

拓扑相关蠕虫通过特定的传播途径在拓扑结构中扩散. 当引入被动保护策略后,每个节点的感染率降低. 如果拓扑结构中任意两个节点之间经过的中

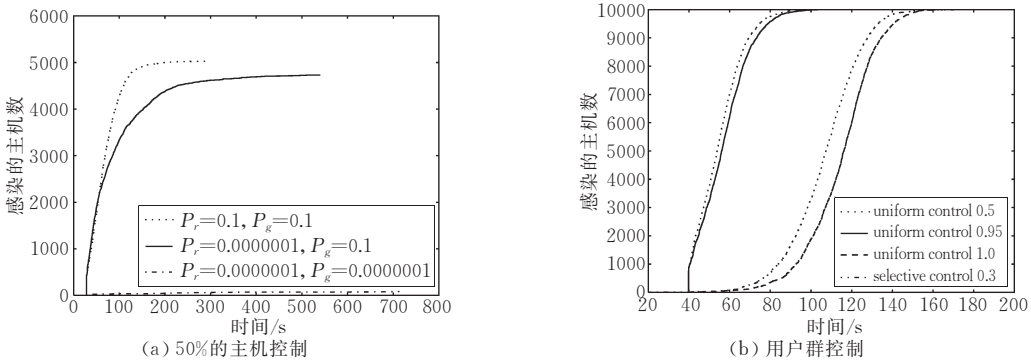


图 7 混合拓扑结构中不同主动控制设置下的蠕虫传播

间节点数较多,节点感染率经过累加,可以降至较低的水平.如节点的感染率为 0.5,任意两点之间的路径平均经过 10 个节点,蠕虫传播平均感染率可降至 $0.5^5=0.000977$. 拓扑结构特征路径长度 L 即为拓扑结构中任意两点之间的路径经过的节点数的平均值. 所以特征路径越长,被动保护策略效果越明显,反之被动保护策略效果越差. 以下分别对三种拓扑结构特性下的被动保护策略进行仿真分析. 蠕虫传播过程除节点感染率外,其余设置与第 3 节一致.

4.1 Power-Law 结构下的被动保护策略

Power-Law 结构中,少数连接度较多的节点与拓扑结构中的很多节点存在直接连接,当一个节点与这样的节点存在连接时,它可以通过该节点,很快地到达拓扑结构中的多个节点,所以特征路径长度减小. 当极少量的这种连接度较多的节点被控制后,虽然不能造成拓扑连通性破坏,却能大大增加特征路径长度,这样被动保护策略可以有效地遏制蠕虫传播. Reinfection 传播方式可以增加节点收到蠕虫数据包的数量,从而可以增加节点被感染的总概率. 但是拓扑结构中,Reinfection 传播方式总是将蠕虫数据包发送到确定的几个节点,所以其对被动保护策略效果的削弱作用是有限的,特别是当聚合系数 C 较大时,蠕虫数据包将在几个已感染节点之间转发,而 Power-Law 结构的聚合系数较完全随机模型要大. 图 8 为 Power-Law 结构下被动防御策略的仿真结果. 仿真默认设置为 $N=10000$, $m=2$, 节点被感染率 $p_i \sim N(0.3, 0.01)$, 选择性控制率 $sc=0.0$, Non-reinfection 传播方式.

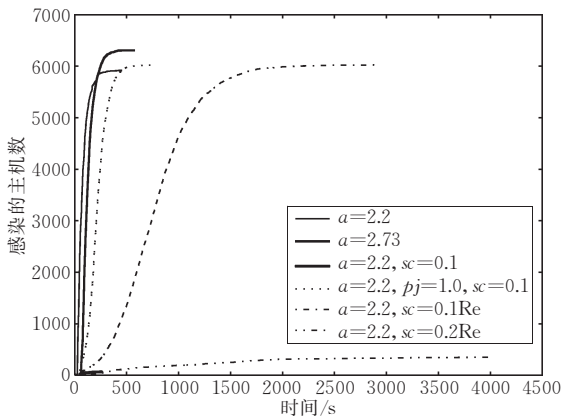


图 8 Power-Law 中被动保护策略下的蠕虫传播

可以看出,不对连接度较大的节点控制时,被动保护策略效果不明显,在 $\alpha=2.2$ 和 2.73 时蠕虫都能够在网络中迅速传播开来. 当控制 10% 的连接度最多的节点后, L 变大,被动保护策略可以有效地阻

断蠕虫传播. 为了对比,我们将感染率 p_i 设定为 1.0,进行仿真,可以看到蠕虫又可以快速传播,说明前面蠕虫传播得到遏制不是因为拓扑连通性被破坏. 相同设置下,Reinfection 传播方式虽然速度有所降低,但是仍可以在拓扑结构中顺利传播开来,但是提高选择控制率 sc 到 20%,蠕虫传播即可被有效地遏制,所以 Reinfection 传播方式可以削弱被动保护策略的效果,但是其作用是有限的. 仿真结果与我们前面的分析一致.

4.2 Small-World 结构下的被动保护策略

由 2.2 节分析可知,Small-World 拓扑结构的特征路径长度主要决定于 p_r . 随着 p_r 的增加,“short-cut”数量增加,特征路径长度不断减少,被动保护策略的作用随之削弱. k 值的增加也可以减小特征路径长度,但是其对特征路径长度的影响远不如 p_r 显著,并且随着 p_r 增加, k 变化对特征路径长度的影响逐渐减小. 参考文献[9]对特征路径长度和 k 之间的关系进行了详细描述. 所以 Small-World 拓扑结构中被动保护策略的实施需要对 p_r 进行有效控制. 同样 Reinfection 传播方式可以增强蠕虫传播的鲁棒性,但是由于 Small-World 拓扑结构具有较高的聚合系数,所以 Reinfection 传播方式对被动保护策略的影响将非常有限. 图 9 为 Small-World 拓扑结构中被动保护策略的仿真结果. 仿真实验默认设置: $N=10000$, $k=10$, 节点感染率 $p_i \sim N(0.3, 0.01)$, Non-Reinfection 传播方式.

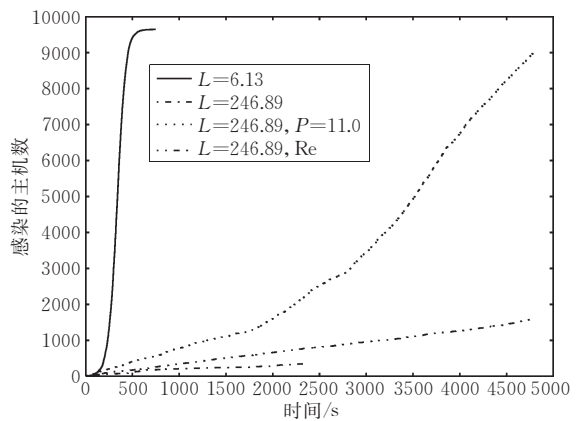


图 9 Small-World 中被动保护策略下的蠕虫传播

由仿真结果可以看出,Small-World 结构中,特征路径长度较小的情况下,被动保护策略对蠕虫传播的抑制作用不是十分有效. 并且同 Power-Law 结构相比,Small-World 结构下蠕虫感染范围更广. 这是因为 Small-World 结构中,节点连接度分布均

匀,每个节点都可以收到多个蠕虫数据包,在被动保护策略下,能够使大量节点被感染,而 Power-Law 结构中,少数节点“浪费”了大量的连接,多数节点之间连接数较少,在被动保护策略下蠕虫传播较为脆弱,所以感染范围较小.当 L 增加到 246.89 时,同样的被动保护策略可以使蠕虫传播得到有效遏制.为了对比,我们将节点被感染率 p_i 设为 1.0 进行仿真实验.可以看到虽然传播速度较慢,蠕虫仍可以感染拓扑结构中的几乎全部节点.在特征路径长度较大的情况下,Reinfection 传播方式对被动防御策略的削弱作用非常有限,仅仅感染大约 10% 的节点.仿真结果与分析结论一致.

4.3 混合拓扑结构下的被动保护策略

混合拓扑结构特征路径长度与用户群关系密切.通过用户群,蠕虫数据包可以直接传送到多个节点,大大降低了拓扑结构的特征路径长度,相应地削弱了被动保护策略的作用效果.当大多数规模较大的用户群被控制,不能传播蠕虫数据包后,混合拓扑结构特征路径长度增加到 Small-World 结构特征路径水平,再进一步控制“Shortcut”数量,增加特征路径长度,才能使被动保护策略有效地切断蠕虫传播. Small-World 结构下的被动保护策略在 4.2 节中已有详细分析,所以本节主要分析用户群结构对被动保护策略的影响.用户群的聚合系数接近 1.0,所以 Reinfection 传播方式对混合结构中的被动保护策略影响更小.图 10 为混合拓扑结构中被动保护策略的仿真结果.默认设置为 $N=10000$, $k=10$, $p_g=0.1$, 用户群分布 $\alpha=2.2$, 节点被感染率 $p_i \sim N(0.3, 0.01)$, 为了减少 Small-World 结构影响,设定 $p_r=0.0001$.

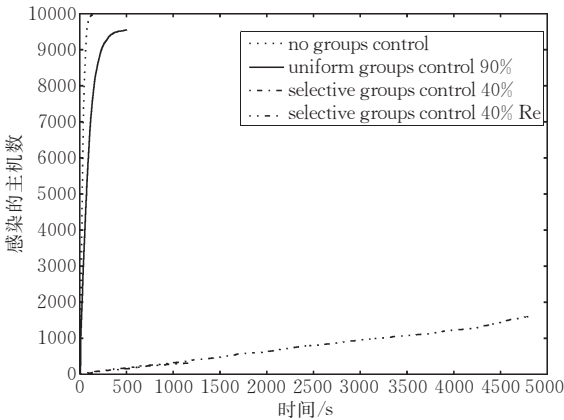


图 10 混合拓扑中被动保护策略下的蠕虫传播

由仿真设置可知,构成混合拓扑结构的 Small-World 拓扑特征路径长度为 246.89. 不考虑用户群

时,在被动保护策略下,蠕虫传播非常脆弱.从图 10 可以看出:当有 10% 的节点加入用户群后,蠕虫即可在拓扑中迅速传播,被动保护策略的作用近乎为 0. 采用完全随机方式控制用户群,即使控制 90% 用户群停止转发蠕虫数据包,被动保护策略也只能稍微降低蠕虫传播速度.通过选择性地控制 40% 规模最大的用户群后,被动保护策略可以有效地遏制蠕虫传播.此外图 10 显示混合拓扑结构下,Reinfection 传播方式对被动保护策略的影响较小,蠕虫传播曲线与 Small-World 下的蠕虫传播曲线基本一致.所以混合结构下主动防御策略的有效实施需要对用户群进行有效控制.

5 基于拓扑结构的局部隔离策略分析

在拓扑相关蠕虫传播过程中,实施局部隔离策略存在一定难度,所以本文忽略局部隔离策略的具体实施细节,如第 2 节描述,将其抽象为被感染节点成功发送蠕虫数据包的概率.这样局部隔离策略虽然实现方式与被动保护策略不同,但两者与拓扑结构的关系却存在一定的相似性,都与拓扑结构的特征路径长度关系密切.本节主要分析三种拓扑结构下,局部隔离策略与被动保护策略作用的不同.图 11 为三种拓扑结构下,两种策略对蠕虫传播的影响的仿真结果对比.拓扑结构与被动保护策略参数设置同第 4 节,局部隔离策略 $p_s \sim N(0.3, 0.01)$. R(Reactive)表示被动保护策略,LC(Local Containment)表示局部隔离策略.

由图 11 可以看出:三种拓扑结构下, p_i 和 p_s 均服从 $N(0.3, 0.01)$ 分布时,局部隔离策略对蠕虫传播的遏制效果明显好于被动保护策略.图 11(a) 显示在 Power-Law 结构下,不进行拓扑结构控制,被动保护策略感染了大约 6000 节点,而局部隔离策略感染大约 5500 节点.当控制 10% 连接度最多的节点,增加了特征路径长度后,两种策略都有效地遏制了蠕虫传播,在 Reinfection 传播方式下,局部隔离策略也明显好于被动保护策略.图 11(b) 显示在 Small-World 结构下,两种策略对蠕虫控制的作用差异更大,在特征路径为 6.13 时,局部隔离策略可以减少感染大约 2000 节点.混合拓扑结构下,这种差异进一步增加,局部隔离策略感染范围为 5000 节点,而被动保护策略感染近 10000 节点,如图 11(c) 所示.

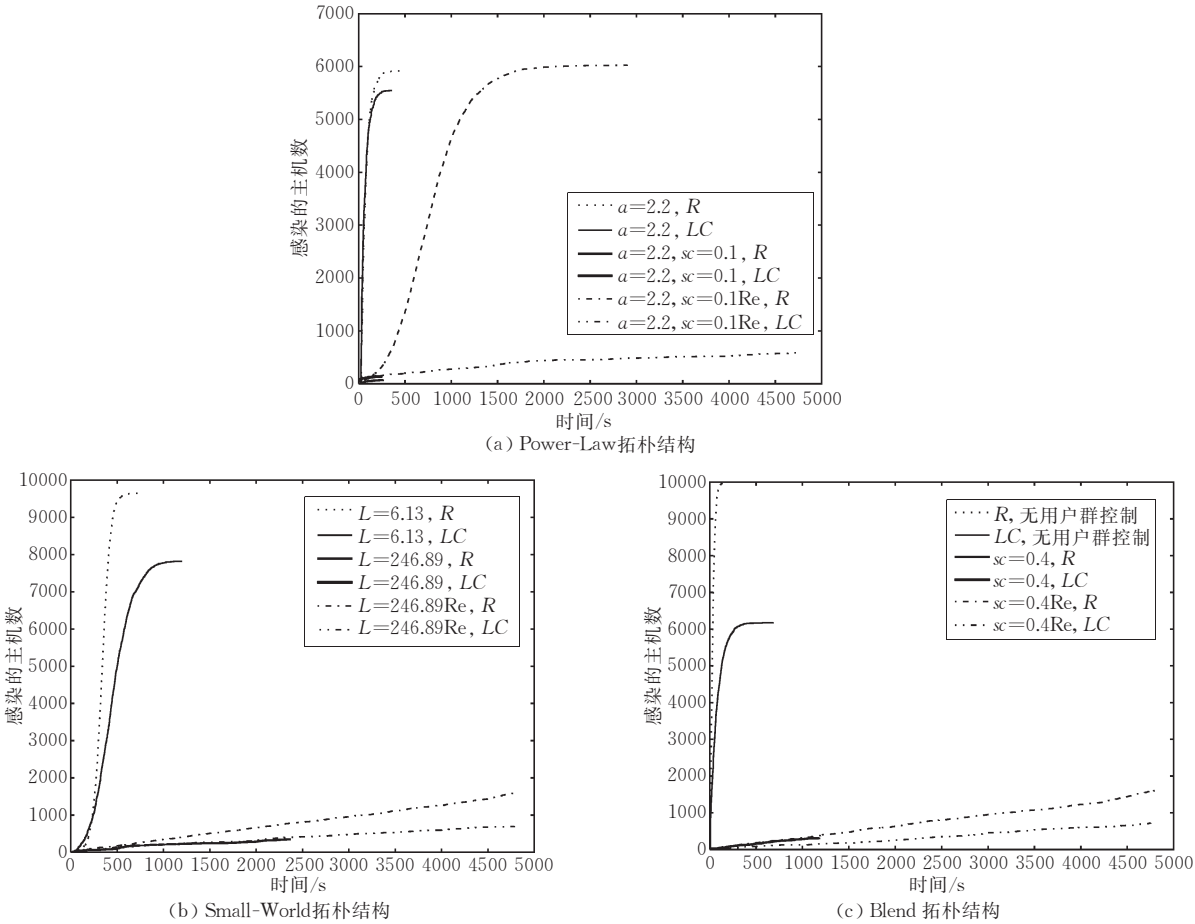


图 11 不同拓扑结构下实施被动保护策略和实施局部隔离策略的蠕虫传播对比

6 结束语

本文首先分析了拓扑相关蠕虫传播所依赖的三种主要拓扑结构形式,并提出了它们的数据包级仿真生成算法。随后,根据目前主要的蠕虫防御策略提出了三种基本的蠕虫传播防御策略仿真模型。最后,通过我们此前开发的数据包级拓扑相关蠕虫传播仿真系统,对三种拓扑结构下的这些蠕虫传播策略进行了仿真分析。仿真结果显示:(1)主动防御策略与拓扑结构连通性关系密切,完全随机的实施主动防御不能有效阻断蠕虫传播,对于 Power-Law 结构下的主动防御策略需要选择连接度最多的节点进行控制,对于 Small-World 结构需要减小“shortcut”数量才能提高主动防御策略效率,而混合结构下的主动防御策略实施的重点在于用户群的选择性控制;(2)被动保护策略有效实施的关键在于如何增加拓扑结构特征路径的长度;(3)局部隔离策略与被动保护策略一样,与拓扑结构特征路径长度关系密切,但是在相同控制力度下,局部隔离策略对蠕虫

传播的遏制作用优于被动保护策略。这些结论可以用于具体拓扑相关蠕虫防御策略设计。

参 考 文 献

[1] Eichin Mark, Rochlis Jon. With microscope and tweezers: An analysis of the Internet virus of November 1988//Proceedings of the 1989 IEEE Computer Society Symposium on Security and Privacy. Oakland, 1989: 326-343

[2] Zou C C, Gong W, Towsley D. Code red worm propagation modeling and analysis//Proceedings of the 9th ACM Symposium on Computer and Communication Security. Washington DC, 2002: 138-147

[3] Briesemeister Linda, Lincoln Patrick, Porras Phillip. Epidemic profiles and defense of scale-free networks//Proceedings of the 2003 ACM CCS Workshop on Rapid Malcode. Washington DC, 2003: 67-75

[4] Zou C C, Towsley D, Gong W. Email virus propagation modeling and analysis. University of Massachusetts CSE, Amherst, Mass: Technical Report TR-CSE-03-04, 2003

[5] Faloutsos M, Faloutsos P, Faloutsos C. On power-law relationships of the Internet topology//Proceedings of the ACM SIGCOMM: Applications, Technologies, Architectures, and

- Protocols for Computer Communication. Cambridge, Mass, 1999; 251-262
- [6] Saroiu Stefan, Gummadi P K, Gribble S D. A measurement study of Peer-to-Peer file sharing systems. University of Washington CSE, Seattle, Wash; Technical Report UW-CSE-01-06-02, 2002
- [7] Broder A, Kumar R, Maghou F, Raghavan P, Rajagopalan S, Stata R, Tomkins A, Wiener J. Graph structure in the Web//Proceedings of the 9th ACM-WWW International Conference. Amsterdam, 2000; 309-320
- [8] Tian Bu, Don Towsley. On distinguishing between Internet power law topology generators//Proceedings of the IEEE INFOCOM 2002. New York, 2002; 638-647
- [9] Watts D, Strogatz S. Collective dynamic of small-world networks. *Nature*, 1998, 393(4): 440-442
- [10] David Brumley, Liu Li-Hao, Pongsin Poosankam, Dawn Song. Taxonomy and effectiveness of worm defense strategies. CMU CS, Pittsburgh, PA; Technical Report CMU-CS-05-156, 2005
- [11] Twycross Jamie, Williamson M M. Implementing and testing a virus throttle//Proceedings of the USENIX Security Symposium. Washington DC, 2003; 285-294
- [12] Cohen R, Erez K, Ben-Avraham D, Havlin S. Resilience of the Internet to random breakdowns. *Physical Review Letters*, 2000, 85(21): 4626-4628
- [13] Newman M E J, Watts D J. Scaling and percolation in the small world network model. *Physical Review E*, 1999, 60(6): 7332-7342



WANG Yue-Wu, born in 1975, Ph.D. candidate. His research interests include network security, large-scale worm simulation, etc.

JING Ji-Wu, born in 1964, Ph.D., professor. His re-

search interests include network security, PKI, worm simulation and intrusion tolerance.

XIANG Ji, born in 1976, Ph.D. candidate. His research interests include network security, worm simulation, etc.

LIU Qi, born in 1978, Ph.D. candidate. Her research interests include network security, worm simulation, etc.

Background

This paper is supported by the National Natural Science Foundation of China project under grant No. 60573015. The main goal of this project is to develop a large-scale Internet worm propagation simulation platform used to the research of Internet worm propagation characters and defense strategies. Topology aware worms as one kind of the mainstream Internet worms must be given most attentions in this simulation platform. The work of this paper mainly studies the relationship between topology aware worm defense strategies and topology structure with simulation method. It not only is one of the most important parts of the project, but also provides a lot of common methods for the simulation of strategies used in other kind of worms.

Topology aware worm propagation depend on the topology structure. In order to hold back the worm propagation in topology structure effectively, it is of great importance to understand the impact of topology structure on the worm de-

fense strategies. However, because of the complexity of the topology structure and the randomness of worm behavior, it is hard to do this work well with existent worm analysis methods. Thus, this paper proposes a packet level simulation method to analyze systemically the worm defense strategies based on topology structure. Topology aware worm simulation system design and implementation have been described clearly in the authors' another paper. The main contents of this paper include: (1) analysis of the major topology structure model and their simulation generation algorithms; (2) the simulation model construction of defense strategies; (3) a systemic analysis of worm defense strategies in different topology structure with simulation experiments, and some conclusions are drawn from the analysis. These conclusions can provide valuable guidelines for real worm defense system design.