

病毒在无标度网络上的传播及控制仿真研究^{*}

李 涛, 关治洪, 吴正平

(华中科技大学 控制科学与工程系, 武汉 430074)

摘 要: 网络病毒的爆发给计算机用户带来巨大的损失, 同时互联网被认为是无标度网络, 因此研究病毒在无标度网络上的传播及控制很有意义。通过构建一个 BA 无标度网络模型, 对病毒的传播行为及影响因素进行了仿真分析。研究表明, 采取恰当的策略可以有效地控制、预防病毒传播。

关键词: 网络病毒; 无标度网络; 幂律

中图分类号: TP309 文献标志码: A 文章编号: 1001-3695(2007)12-0177-02

Research on simulation of virus propagation and control in scale-free networks

LI Tao, GUAN Zhi-hong, WU Zheng-ping

(Dept. of Control Science & Engineering, Huazhong University of Science & Technology, Wuhan 430074, China)

Abstract: Network virus outbreak brought huge losses to computer users. It was found that Internet had the features of scale-free. So the research on simulation of virus propagation and control in scale-free networks was significant. Through building a BA scale-free network, studied the simulation of the propagation of virus and influencing factors. The fruitful simulation results show that appropriate policies can effectively control and prevent the propagation of virus.

Key words: network virus; scale-free networks; power-law

0 引言

1999 年 Faloutsos 等人^[1]发现, 互联网表现出很强的幂律分布特点, 节点的度(某节点的度是指与该节点相连的节点数目, 或者说与该节点关联的边的数目) 有很大的波动性^[1, 2]。目前为止主要从两种不同的角度来描绘互联网的结构^[1, 3]。无论将互联网中的路由器定义为节点, 路由器之间的通信链路定义为连接节点的边, 还是将互联网中的子域定义为节点, 域间连接定义为连接节点的边, 互联网节点的度都在双对数坐标图中呈现明显的线性分布。尽管随着时间的推移, 系统中的节点和边在不断增加, 但网络拓扑结构特性却不会发生很大的变化。

幂律分布也称为无标度分布, 具有幂律分布的网络也称为无标度网络^[4]。本文研究病毒在无标度网络上的传播及其控制, 通过仿真得到了一些有意义的结论。

1 BA 无标度网络模型

Barabasi 和 Albert(BA)^[5]提出的 SF(scale-free, 无标度) 网络模型, 遵循生长和择优连接两大机制。取初始 m_0 个顶点任意连接或完全连接, 然后:

a) 增长。在每个时间间隔增添一个具有 $m(\leq m_0)$ 条边的新节点, 连接这个新节点到 m 个不同的已经存在于系统中的节点上。

b) 择优连接。在选择新节点的连接点时, 假设新节点连

接到节点 i 的概率 Π 取决于节点 i 的度数, 即

$$\Pi(k_i) = k_i / \sum_j k_j$$

其中: 分母为各节点连通度的总和。经过 t 个时间步后, 形成一具有 $N = m_0 + t$ 个节点, mt 条边的网络模型。该网络节点的连通度服从 power-law 分布。Scale-free 网络的提出, 为仿真研究 Internet 上的病毒传播行为提供了基础。

本文利用如上所述的无标度网络生成算法生成一个具有 1 000 个节点, 平均度为 4 的无标度网络。其度分布如图 1 所示。在图 1 中, 节点度大于平均度 4 的节点占节点总数的 19.0%, 不妨定义这类节点为 A 类节点。显然在 A 类节点中, 包含了在无标度网络中被称为集散节点的节点。节点度小于等于平均度 4 的节点占节点总数的 81.0%, 不妨定义这类节点为 B 类节点。由图 1 可直观地看见该仿真网络的度分布服从幂律分布。

2 病毒传播仿真

病毒在无尺度网络上的传播受到很多因素影响。其中初始感染节点位置、免疫策略、节点的抗病毒能力对病毒传播的影响尤为显著。因此在这里主要考虑与仿真初始感染节点位置、免疫策略、节点的抗病毒能力这三种因素对病毒传播的影响。初始时刻网络中的节点均处于易感染状态, 病毒从其产生源出发, 沿网络连接向四周传播。

2.1 初始感染节点位置对病毒传播的影响

在这里, 讨论病毒源为 A 类节点和 B 类节点两种情况下

收稿日期: 2006-08-08; 修返日期: 2006-10-27 基金项目: 国家自然科学基金资助项目(60573005)

作者简介: 李涛(1974-), 男, 博士研究生, 主要研究方向为复杂网络、网络通信、网络安全(tao20012004@163.com); 关治洪(1955-), 教授, 博导, 主要研究方向为非线性复杂网络与系统、脉冲混合系统与混沌控制、网络系统的控制与安全; 吴正平(1968-), 副教授, 博士研究生, 主要研究方向为复杂网络、复杂环境下的多机器人协调控制。

病毒传播仿真,如图 2 所示。

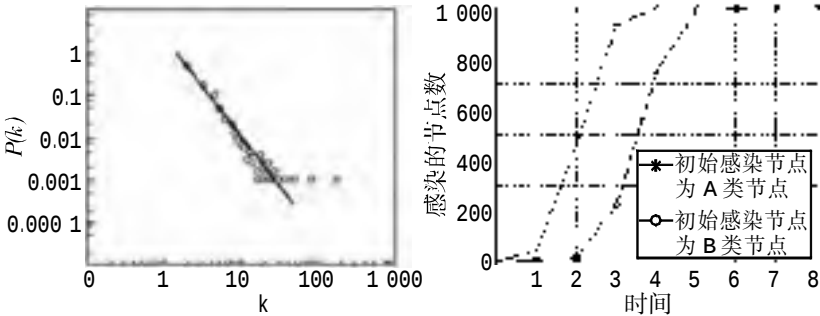


图 1 无标度仿真网络度分布
($m_0=6, m=2, n=1\ 000$)

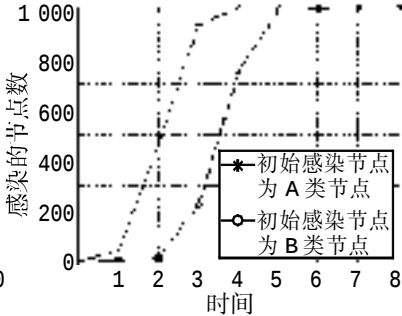


图 2 初始感染节点位置对病毒传播的影响

由图 2 可以看到,网络病毒在无标度网络上的传播与初始感染节点的位置关系密切。当初始感染节点为节点度较小的 B 类节点时,传播时间是七个单位时间;当初始感染节点为度较大的类 A 节点时,平均传播速度增加约一倍,传播时间减少了近一半,约为四个单位时间。因此,可以说,防止度较大的节点成为病毒源是防止病毒快速传播的第一步,尤其是防止集散节点成为病毒源,可以有效地延缓病毒的扩散,为进一步采取措施控制、消灭病毒大面积的传播赢得时间。

2.2 不同的免疫策略对病毒传播的影响

免疫是抑制计算机病毒的一种重要方法。免疫节点是指当病毒传播到该节点时,该节点既不会被病毒感染,也不会将病毒向其他节点传播。对网络节点进行免疫,既可以防止被免疫的节点感染病毒,又可以隔断病毒通过该节点感染其他节点。所以,可以看做是从网络中删除该节点。较好的免疫策略应该是在对较少的节点进行免疫的情况下,能够最大程度地延缓病毒的传播,为消灭病毒赢得时间。目前,主要有随机免疫和选择免疫两种免疫策略^[6]。随机免疫不考虑节点间的差别,对所有的节点平等对待,进行免疫时随机地选择节点,没有优先顺序。选择免疫多为根据某一标准有选择地取一类节点进行免疫,以达到最好的免疫效果。下面笔者对病毒传播过程中,对不采取任何免疫策略、随机免疫策略和选择免疫策略三种情况进行仿真分析,仿真结果如图 3 所示。这里选择免疫原则是度大的节点优先进行免疫。在两种免疫策略下分别选择 3% 的节点进行免疫。

由图 3 可以看出,不采取任何免疫策略,病毒用六个单位时间感染了 99.2% 的节点;用七个单位时间感染了所有的节点;采取随机免疫策略,病毒传播用六个单位时间感染了 95.9% 的节点,也是用七个单位时间感染了所有的节点;采取选择免疫策略,病毒传播花费 13 个单位时间感染了 90.2% 的节点,最终也是感染了 90.2% 的节点。可见,采取随机免疫策略,传播时间未发生明显地变化,传播速度略微减小;采取选择免疫策略,传播时间延长近一倍,传播速度显著降低。

2.3 节点的抗病毒能力对病毒传播的影响

在这里不妨以电子邮件病毒传播为例说明节点的抗病毒能力对病毒传播的影响。在电子邮件病毒传播过程中,用户总在定期地检查他们的电子邮件。当一位用户检查信箱并且遇到带病毒附件的邮件时,他有可能丢弃这封信(如果对这封信表示怀疑或者反病毒软件检测出该信带有病毒),也有可能打开这封信的附件。而打开概率则由用户关于电子邮件病毒的知识意识和意识所决定。打开概率可以看成节点的抗病毒能力,打开概率越大,相当于该节点的抗病毒能力越强;打开概率越小,

相当于该节点的抗病毒能力越强。当然每个节点的抗病毒能力是不同的。这里可看成是平均抗病毒能力。节点具备不同抗病毒能力对病毒传播的影响如图 4 所示。

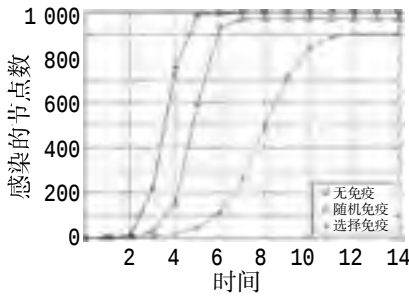


图 3 不同的免疫策略对病毒传播的影响

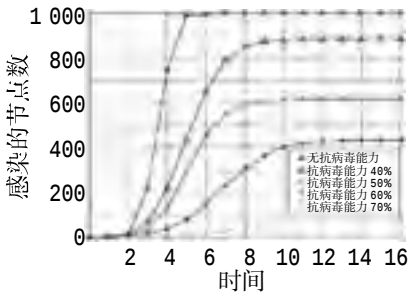


图 4 节点的不同抗病毒能力对病毒传播的影响

由图 4 可以看出,当每个节点具备 60% 的抗病毒能力时,病毒传播时间延长一倍,传播范围降低至 43.3%;当每个节点具备 70% 的抗病毒能力时,病毒传播时间延长一倍,传播范围降低至 8.34%。对于电子邮件病毒传播来说,也就是提高用户的警惕和防病毒能力。当用户打开带有病毒的附件的平均概率降低到 30% 时,可以本质性地抑制病毒的广泛传播。

2.4 病毒传播预防控制仿真

病毒在网络中传播时,笔者对初始感染节点位置、免疫策略、节点的抗病毒能力(用户的警惕和防病毒能力)对病毒传播速度和范围的影响分别进行了仿真分析。从仿真结果可以看出初始感染节点位置、免疫策略、节点的抗病毒能力分别对病毒传播的影响。以下综合分析病毒传播受这三种因素共同影响的仿真。考虑到免疫的节点不可能占有总节点数太大的比例,否则会增加太大的人为工作量;节点的抗病毒能力也不可能太高,要求所有节点具有很高的平均抗病毒能力也是不现实的。因此笔者选取以下各种因素值:病毒源避免是节点度较大的 A 类节点;根据度大的节点进行优先免疫的原则分别选择 2%、4% 的节点进行免疫;节点分别具有 40%、50%、60% 的抗病毒能力。病毒传播的仿真图如图 5、6 所示。

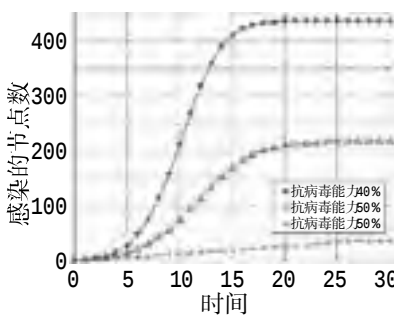


图 5 选择免疫 2% 的节点

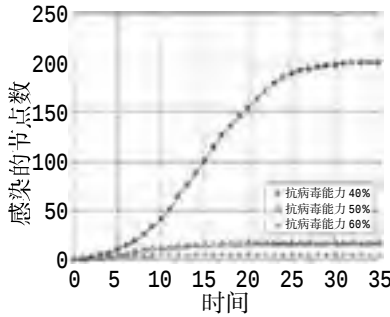


图 6 选择免疫 4% 的节点

图 5 中当节点的平均抗病毒能力为 40% 时,病毒的传播范围被控制在 43.59%;当节点的平均抗病毒能力为 50% 时,病毒的传播范围被控制在 21.75%;当节点的平均抗病毒能力为 60% 时,病毒的传播范围被控制在 3.63%。图 6 中当节点的平均抗病毒能力为 40% 时,病毒的传播范围被控制在 20.10%;当节点的平均抗病毒能力为 50% 时,病毒的传播范围被控制在 1.70%;当节点的平均抗病毒能力为 60% 时,病毒的传播范围被控制在 0.47%。可见,在合理的范围内提高免疫节点的百分比,增强节点的抗病毒能力,可以有效地预防和控制病毒传播的范围。

3 结束语

中心服务节点计算全局推荐信任度。恶意节点分别 10、20、30、40、50 个,即分别占总节点数的 10%、20%、30%、40%、50%。恶意节点提供正确文件下载的概率是 50%。最初其他节点提供正确文件下载的概率是 95%。实验中对于全局推荐信任度的计算没有采用分布式计算方式,而是采用集中计算全局推荐信任度,直接信任度由各个节点自己计算。

实验过程:100 个节点,每个节点下载 100 次数据,每个节点每次下载数据后把每次交易评价记录下来。下载 10 次后,将 10 次下载的评价序列发送给计算全局推荐信任度的节点。计算全局推荐信任度的节点计算所有节点的全局推荐信任度;然后再将全局推荐信任度发送给各个节点。每个节点接收到全局推荐信任度后,计算节点的综合信任度。根据综合信任度,按照不同的服务选择策略,选择服务节点下载数据。下载 10 次后,再向信誉度节点发送评价序列。这样,在实验过程中信誉度计算 10 次。不必在某个节点下载完数据后就更新信誉度,可以减少计算量。实验中没有采用请求竞争策略。因为请求竞争策略只是在资源有限的情况下限定请求者使用资源,不能直接反映对下载成功率的影响。

实验运行的硬件环境为:PC 机,CPU 是 Intel Pentium 4 处理器,其主频是 1.8 GHz;内存 512 MB。软件环境为:操作系统 Windows XP SP2;实现语言为 Java,JDK 为 1.4.2。

图 3 为实验结果比较。图中方块点组成的线表示采用随机选择下载策略(即不参考任何信任度的策略);圆点组成的线表示采用设置信任度阈值的策略;三角点组成的线表示采用相近信任度策略。从三条曲线可以看出,随着恶意节点的增加,三种服务选择策略的下载成功率均有所下降。其中:设置信任度阈值策略的曲线下降最慢,相对比较平滑;随机选择下载策略和相近信任度策略下降幅度相差不大。但是相近信任度策略下载成功率大于随机选择下载情况。在恶意节点占 10% 情况下,设定信任度阈值策略和相近信任度策略均能识别出隐藏的恶意节点。

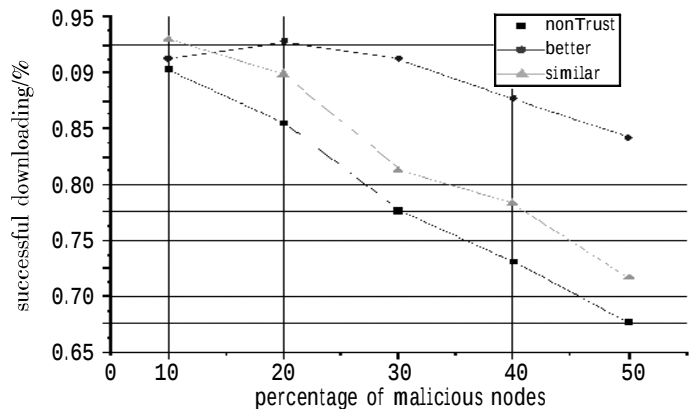


图 3 随着恶意节点数比率增大,两种策略成功下载变化图

实验结论:基于信任度的两种服务选择策略均能提高交易完成的成功率。相对来说,设置信任度阈值策略优于相近信任度策略。

4 结束语

本文描述了网格系统中信任关系的计算。通过对分布式环境下授权模型的研究,在 CAS 的基础上提出了基于信任度的访问控制策略。服务选择策略和请求竞争策略使得网格用户合理使用资源,排除恶意使用资源的用户,同时这种策略的应用为那些信誉好的资源提供者提供了更多使用资源的机会,

按照对网络的贡献多少来使用资源。为了合理地利用资源,防止集中访问信任度高的节点,冷落信任度低的节点,提出了 Ticket 机制。通过模拟实验得出结论:信任度能够有效提高网格作业完成的成功率和网格系统的安全性。

参考文献:

[1] EARLMAN L, WELCH V, FOSTER I, *et al.* A community authorization service for group collaboration[C] //Proc of the 3rd International Workshop on Policies for Distributed Systems and Networks. Washington DC: IEEE Computer Society, 2002: 50-59.

[2] KAMVAR S D, SCHLOSSER M T, MOLINA H G. The eigenTrust algorithm for reputation management in P2P Networks[C] //Proc of the 12th International World Wide Web Conference. New York: ACM Press, 2003: 640-651.

[3] FU Y, CHASE J, CHUN B, *et al.* SHARP: an architecture for secure resource peering[C] //Proc of the 19th ACM Symposium on Operating Systems Principles. Bolton Landing, NY: ACM Press, 2003: 133-148.

[4] DAMIANI E, VIMERCATI D C D, PARABOSCHI S, *et al.* A reputation-based approach for choosing reliable resource in peer-to-peer networks[C] //Proc of the 9th ACM Conference on Computer and Communications Security. Washington DC: ACM Press, 2002: 207-216.

[5] WANG Y, VASSILEVA J. Bayesian network-based trust model in peer-to-peer networks[C] //Proc of IEEE/WIC International Conference on Web Intelligence(WI 2003) . Halifax: [s. n.], 2003: 372-378.

[6] 刘宏月, 范九伦, 马建峰. 访问控制技术研究进展[J] . 小型微型计算机系统, 2002, 25(1): 56-59.

(上接第 178 页)通过构建一个无标度仿真网络,研究了病毒在该网络上的传播行为及其影响因素。仿真实验表明,避免度较大的节点,尤其是集散节点成为病毒源;合理地选择免疫节点,提高每个节点的抗病毒能力,可以最大程度地预防和控制病毒的传播。这对于防范计算机病毒在互联网上的传播,具有重要的现实意义。

参考文献:

[1] ALOUTSOS M, FALOUTSOS P, FALOUTSOS C. On power-law relationships of the Internet topology[J] . Computer Communication Review, 1999, 29(4): 251-262.

[2] V ZQUEZ A, PASTOR-SATORRAS R, VESPIGNANI A. Large-scale topological and dynamical properties of the Internet[EB/OL] . (2002) . <http://arxiv.org/abs/cond-mat/0112400>.

[3] CHEN Guan-rong, FAN Zheng-ping, LI Xiang. Modelling the complex Internet topology[C] //Proc of Complex Dynamics in Communication Networks. Berlin: Springer-Verlag, 2004.

[4] 汪小帆, 李翔, 陈关荣. 复杂网络理论及其应用[M] . 北京: 清华大学出版社, 2006: 12-13.

[5] BARABASI A L, ALBERT R. Emergence of scaling in random networks[J] . Science, 1999, 286(5439): 509-512.

[6] PASTOR-SATORRAS R, VESPIGNANI A. Epidemic spreading in scale-free networks[J] . Physical Review Letters, 2001, 86: 3200-3203.