

无标度网络中基于最短路径免疫策略的病毒传播研究

林 兵¹ 郭文忠¹ 陈国龙^{1,2} 林嘉雯¹

(福州大学数学与计算机科学学院 福州 350108)¹

(福州大学离散数学及其应用教育部重点实验室 福州 350003)²

摘 要 传统病毒免疫策略大多基于网络的全局拓扑信息。然而现实生活中的大部分复杂网络仅仅只能了解其局部拓扑信息。鉴于许多实际复杂网络具有无标度特性,研究了在无标度复杂演化网络中基于网络局部拓扑信息最短路径免疫策略的病毒传播现象。利用平均场理论建立含个体抵抗力重要因素的无标度网络病毒传播模型,并引入基于最短路径的免疫策略。比较了随机免疫、目标免疫和最短路径免疫 3 种策略对无标度复杂网络病毒传播的影响,结果表明了基于最短路径免疫策略的有效性。

关键词 复杂网络,病毒传播,个体抵抗力,免疫策略,局部拓扑

中图分类号 TP393.08 **文献标识码** A

Research of Virus Spreading Based on Shortest Path Immunization Strategy in Scale-free Network

LIN Bing¹ GUO Wen-zhong¹ CHEN Guo-long^{1,2} LIN Jia-wen¹

(College of Mathematics and Computer Sciences, Fuzhou University, Fuzhou 350108, China)¹

(Key Laboratory of Discrete Mathematics with Applications of Ministry of Education, Fuzhou University, Fuzhou 350003, China)²

Abstract Most of the traditional virus immunization strategies are based on global network topology information, however, most real-life complex networks are known to us with only the local topology information. In view of that scale-free property exists in many real-life complex networks, the virus spreading with shortest path immunization strategy based on local topology information in scale-free complex evolving network was studied. This article used the mean-field theory to build a virus spreading model in scale-free network introducing a immunization strategy based on shortest path and a key factor: individual resistance. After comparing the efficiency of random immunization, target immunization and shortest path immunization for virus spreading in scale-free complex networks, the result indicates that the immunization strategy based on shortest path has important effectiveness.

Keywords Complex network, Virus spreading, Individual resistance, Immunization strategy, Local topology

1 引言

无论是 Web 互联网、人类社会网络还是牲畜群网络,它们都是由众多复杂的节点和连接方式构成的庞大网络,都属于复杂网络^[1]。在人类的发展过程中,遭遇过许多毁灭性的传染病和计算机病毒,如流行性感冒、禽流感、SARS 以及千变万化的计算机病毒。大量研究表明,网络上的病毒传播与网络的拓扑结构有着密切的联系^[1-4]。仅仅研究网络的拓扑结构并不是重点,理解和解释这些网络上的系统行为才是最终目的,特别是理解网络拓扑结构对病毒传播动力学的影响。根据不同的网络拓扑结构,研究相应的病毒传播规律,从而制定出对应的免疫策略,具有重大的现实意义。

美国 Notre Dame 大学物理系的 Barabasi 教授及其博士生 Albert 于 1999 年 10 月在 Science 杂志上发表了题为《随机

网络中标度的涌现》的文章^[5],提出了一个无标度网络模型(BA 模型)。它考虑了现实网络中的两个重要特征:一是增长特性;二是优先连接特性,也称“马太效应”或“富者更富”现象。无标度网络可以是实例化的人类社会的社团组织、世界贸易网中的区域经济合作组织、计算机网络中的自治系统等。

在以往的复杂网络病毒传播研究中,大多数免疫策略都是基于网络的全局拓扑信息而做出的,然而现实生活中的大部分复杂网络仅仅只能了解其局部拓扑信息。基于了解网络局部拓扑信息,本文提出了基于网络局部拓扑信息的最短路径免疫策略。首先,基于无标度网络,建立引入个体抵抗力因子的病毒传播模型,并观察初始感染节点的选择和个体抵抗力差异对病毒传播的影响;其次,引入病毒免疫策略,比较目标免疫和随机免疫分别在病毒感染平衡态和快速增长期的作用效果;最后,提出一种基于最短路径的病毒免疫新策略,分

本文受国家自然科学基金项目(10871221),福建省科技创新平台计划项目(2009J1007),福建省自然科学基金重点项目(A0820002),福建省教育厅科技项目(JK2010001)资助。

林 兵(1986-),男,硕士生,主要研究方向为复杂网络,E-mail: CatherineWheel520@gmail.com;郭文忠 男,博士,副教授,主要研究方向为计算智能及其应用;陈国龙 男,博士,教授,博士生导师,主要研究方向为计算智能、计算机网络等;林嘉雯 女,硕士,助教,主要研究方向为计算机网络。

别和目标免疫、随机免疫在病毒感染平衡态和快速增长期做比较,观察其是否有效。

2 无标度网络上的病毒传播

在本文中,我们研究在无标度复杂网络上的 SIR 病毒传播模型^[6]。与传统传播模型相比,本文模型中引入了抵抗力差异因子,并利用 n_1, n_2, s, r 4 个参数分别改变初始个体的差异程度。实验证明,该传播模型更贴近病毒在无标度网络下的传播特征。此模型中,个体被划分为 3 类:第一类是易感种群(S),他们不会感染其他的个体,但是有可能被感染;第二类是感染种群(I),他们已经患病,具有传染性;第三类是免疫种群(R),他们是经过免疫策略作用后已获得免疫能力的个体,不具传染性,也不会被再次感染。在不加入任何免疫措施的情况下,定义 SIR 模型自然演化成 SIS 模型^[7]。在本研究中,定义病毒的有效传播率为 $\lambda = \alpha(1-s)^{n_2} / \beta(1+r)^{n_1}$, 其中 $\alpha(1-s)^{n_2}$ 为被感染概率, $\beta(1+r)^{n_1}$ 为被治愈概率(自我恢复概率)。由于引入了个体抵抗力差异因子 n_1, n_2 , 因此 $\alpha(1-s)^{n_2} / \beta(1+r)^{n_1}$ 并非传统定义的常量,而是根据个体自身抵抗力差异而定义的变量,感染概率为 $\alpha(1-s)^{n_2}$, 其中 α, s 为常量, n_2 为个体经历从感染状态到易感状态变化的次数;被治愈概率为 $\beta(1+r)^{n_1}$, 其中 β, r 为常量, n_1 为个体经历从感染状态到易感状态变化的次数。随着病毒传播的演化, n_1, n_2 不断增加,被感染概率 $\alpha(1-s)^{n_2}$ 不断减小,被治愈概率 $\beta(1+r)^{n_1}$ 不断增加,这符合现实生活中个体经过与病毒的长期斗争后,抵抗力不断增强而产生的结果。由于 n_1, n_2 根据个体的抵抗力差异而不一,因此 λ 也会随着个体的抵抗力差异而变化。

在无标度网络中,没有任何免疫策略的情况下,根据平均场理论,对于任意一个度为 k 的节点,定义 $\rho_k(t)$ 是 t 时刻网络中度数为 k 且被感染节点占度数为 k 节点的比例, $f(\rho_k(t))$ 表示该节点的邻居节点是感染节点的概率。由此可得到:

$$\frac{\partial \rho_k(t)}{\partial t} = -\beta(1+r)^{n_1} \rho_k(t) + \alpha(1-s)^{n_2} \cdot k \cdot (1-\rho_k(t)) \cdot f(\rho_k(t)) \quad (1)$$

对于任意度为 k 的节点,在没有任何免疫策略情况下,它不是易感节点,就是感染节点。如果它是感染节点的概率为 $\rho_k(t)$, 并且其自我恢复概率为 $\beta(1+r)^{n_1}$, 则感染率减少 $\beta(1+r)^{n_1} \rho_k(t)$; 否则是易感节点的概率为 $1-\rho_k(t)$ 。因为其度为 k , 且其邻居节点为感染节点的概率为 $f(\rho_k(t))$, 所以邻居感染节点的个数是 $k \cdot f(\rho_k(t))$ 。所有邻居节点都会影响该节点,所以感染率增加 $\alpha(1-s)^{n_2} \cdot k \cdot (1-\rho_k(t)) \cdot f(\rho_k(t))$ 。

无标度网络中,度为 j 的节点出现的次数与 j 和度 j 成比例。因此,邻居节点是感染状态且度为 j 的概率为 $j \cdot n(j) \cdot \rho_j(t) / \sum_i i \cdot n(i)$ 。度为 k 的节点, $f(\rho_k(t)) = \eta_k + \delta_k \cdot \rho_k(t)$, 其中 $\eta_k = \sum_{j \neq k} j \cdot n(j) \cdot \rho_j(t) / \sum_i i \cdot n(i)$, $\delta_k = k \cdot n(k) / \sum_i i \cdot n(i)$ 。代入式(1)可得:

$$\frac{\partial \rho_k(t)}{\partial t} = -\beta(1+r)^{n_1} \rho_k(t) + \alpha(1-s)^{n_2} \cdot k \cdot (1-\rho_k(t)) \cdot (\eta_k + \delta_k \cdot \rho_k(t)) \quad (2)$$

2.1 加入免疫策略的病毒传播

抑制病毒传播的一种重要方法是引入免疫策略。目前,两种主要的免疫策略是随机免疫(均匀免疫)和选择免疫(目

标免疫)^[8]。

随机免疫完全随机地选取网络中的一部分节点进行免疫。它不考虑网络中节点之间的差异,平等对待度大的节点(被感染风险高)和度小的节点(相对安全)。随机免疫虽然能局部抑制感染,但要彻底消灭感染则需要对网络中大部分节点进行免疫,这显然不现实。目标免疫选取少量度大的节点进行免疫,实验结果表明,它是有效的方法^[9]。

我们采用 SIR 模型来模拟引入免疫策略的病毒传播过程,将已免疫的节点转化为免疫状态(R)。本文提出了新的免疫策略:基于最短路径的免疫策略。基于最短路径的免疫策略的关键步骤如下:

Step1 对于任意被感染节点 i , 利用最短路径算法求出与 i 距离不大于 len (与 i 相邻的节点 len 为 1) 且度大于 deg 的所有临近节点集 Set;

Step2 对 Set 集中的 $n(\text{Set})$ 个节点按与节点 i 的距离大小 len_j 进行从小到大排序,选择前 m 个节点进行免疫。如果 $n(\text{Set}) < m$, 则对 Set 中所有节点进行免疫,再随机对临近的 $m - n(\text{Set})$ 个节点进行免疫;否则对 Set 中的前 m 个节点进行免疫。如果 $n(\text{Set})$ 为非 0, 转 Step1, 否则,结束计算。

2.2 病毒传播模型

无标度网络中基于最短路径免疫策略的病毒传播模型如图 1 所示。

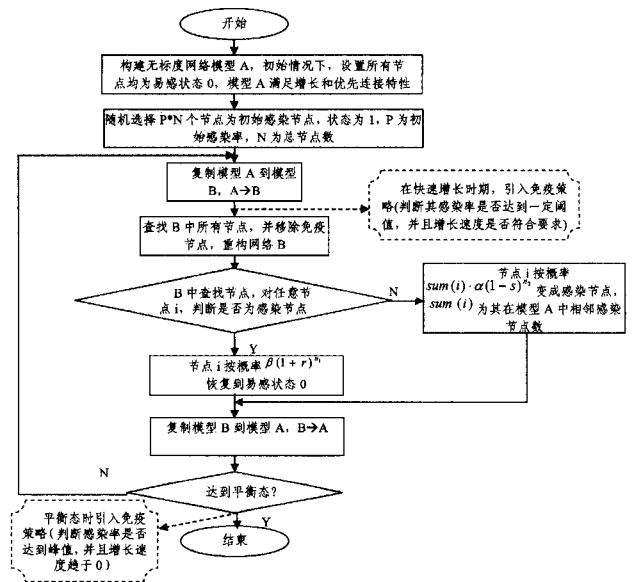


图 1 病毒传播模型

3 仿真

在以下仿真过程中,无标度网络规模为 $N=8000, k=3$ (这足以体现无标度复杂网络模型的主要病毒传播特征)。

3.1 无免疫策略的病毒传播仿真

3.1.1 个体抵抗力差异的比较

我们知道,在现实世界中,个体对同一种病毒的抵抗力是不同的,当周围的邻居处于感染状态时,对于自身素质好的个体,其对病毒的抵抗力就强,不容易被感染上;而对于抵抗力弱的个体,则其容易染上病毒。本文的个体抵抗力差异主要通过修改式(1)的参数 n_1, n_2 , 从而改变某个节点 i 的自恢复概率 $\beta(1+r)^{n_1}$ 和被感染概率 $\text{sum}(i) \cdot \alpha(1-s)^{n_2}$, 其中 sum

(i)为节点 i 相邻的感染节点总数。当 r, s 均取 0 时,任意节点 i 的自恢复能力为 β ,而被感染概率为 $\text{sum}(i) \cdot \alpha$,二者均为常量,不存在个体抵抗力差异。仿真中, $\alpha=0.1, \beta=0.02, r=0.02, s=0.02$,第一组中取 $n_1=1, n_2=1$,第二组中取 $n_1=15, n_2=15$,初始感染率 P 为 0.1。由式(1)可知,第二组节点的整体初始抵抗力强于第一组,仿真结果如图 2 所示。

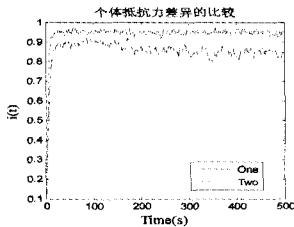


图 2 个体抵抗力差异

从图 2 可以看出,节点的个体抵抗力差异影响到病毒的传播行为,它不但影响病毒的扩散速率,而且影响最终的感染比例。节点整体抵抗能力强的不易受到感染,即使受到感染,整体的感染速度也相对缓慢。由此可知,提高个体本身的抵抗力将在一定程度上有利于遏制病毒的扩散速度和最终的感染比例。

3.1.2 初始感染节点选择的比较

仿真病毒传播过程中,所有节点刚开始均处于易感状态(S),即初始感染率 P 为 0。式(1)中,取 $s=0, r=0$,即忽略个体抵抗力差异, n_1, n_2 可取任意值,不影响自恢复概率和被感染概率。假设 $\alpha=0.1, \beta=0.02$ 。分别选择度最大节点(如果不唯一则选择序号最小者)和度最小节点(如果不唯一则选择序号最大者)作为初始感染节点,观察病毒传播情况。图 3 给出了二者之间病毒传播的差异。

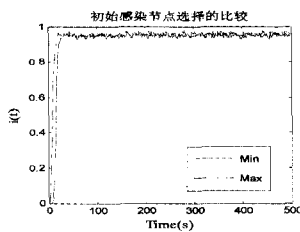


图 3 初始感染节点的选择

从图 3 可知,无标度网络的度幂律分布特性,使初始被感染节点的选择影响到病毒的传播行为,并且在病毒传播的前期影响较大,而在病毒传播中后期影响较小。初始感染度最大的节点能加快病毒的传播速率,而初始感染度最小的节点病毒扩散传播明显缓于前者。因此在病毒传播防范过程中,度大的节点应当是重点考虑对象。

3.2 加入免疫策略的病毒传播仿真

基于最短路径的免疫策略,对无标度网络上的病毒传播进行免疫,并比较随机免疫、目标免疫和最短路径免疫 3 种策略对无标度复杂网络病毒传播的影响。仿真中,不失一般性,与传统的病毒免疫策略进行了比较,取 $\alpha=0.1, \beta=0.02, r=0.02, s=0.02, n_1=1, n_2=1$,初始感染率 P 为 0.1, Step1 中取 $len=3, deg=2 * \langle k_i \rangle$,其中 $\langle k_i \rangle$ 表示与节点 i 相邻的所有节点的平均度, Step2 中取 $m=5$ 。图 4 是基于最短路径免疫在平衡态与随机免疫和目标免疫的防治效果比较;图 5 是基于最短路径免疫在快速增长期与随机免疫和目标免疫的防治

效果比较。实验中,平衡态的判别方式是:感染率增长速度绝对值小于 1%;快速增长期的判别方式是:感染率增长速度绝对值大于 20%。

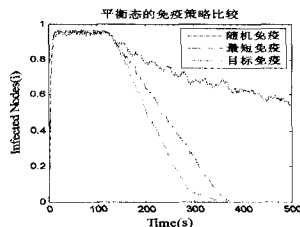


图 4 平衡期的免疫效果

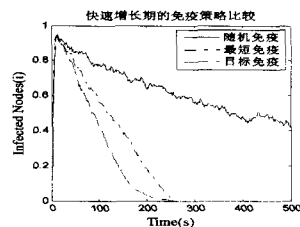


图 5 快速增长期的免疫效果

从图 4 和图 5 可以看出,目标免疫策略和最短路径免疫策略与随机免疫相比都是很有有效的方法,最短路径免疫策略的效果稍微弱于目标免疫策略。从现实考虑,了解全局的网络拓扑信息是比较困难的,特别是对于节点数目众多、结构复杂的网络。目标免疫策略必须在了解复杂网络全局信息的基础上才能做出正确的免疫措施,选择度最大或某种最优进行免疫,实施起来难度系数明显大于最短路径免疫策略。相反,最短路径免疫策略只要了解局部的网络拓扑信息,就能采取免疫措施,具有现实意义。

结束语 病毒传播日益严重,持续不断地威胁着网络安全。不同的免疫策略产生不同的有效性,在不同的现实网络中同一种免疫策略的效果也不一。在现实工作中,可结合基于最短路径的免疫策略和病毒遏制技术,使病毒难以传播,从而提高整个网络的系统安全性和可靠性。

本文的研究中,假定网络的拓扑结构是保持不变的,但实际上个体间的关系会随着时间而发生改变。因此,动态网络上的病毒还有待于进一步研究。

参考文献

- [1] 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
- [2] Wang C, Knight J C, Elder M C. On Computer Viral Infection and the Effect of Immunization [A]//IEEE 16th Annual Computer Security Applications Conf [C]. New Orleans, 2000; 246-256
- [3] Vazquez A, Satorras R P, Vespignani A. Large-Scale Topological and Dynamical Properties of the Internet [J]. Physical Review E, 2002, 65(6): 130-142
- [4] Li X, Chen G, Li C G. Stability and Bifurcation of Disease Spreading in Complex Networks [J]. Int J of Systems Science, 2004, 35(9): 527-536
- [5] Barabasi A L, Albert R. Emergence of scaling in random networks [J]. Science, 1999, 286: 509-512
- [6] Pastor-Satorras R, Vespignani A. Epidemic Spreading in Scale-free Networks [J]. Phys. Rev. Lett., 2001, 86(14): 3200-3203
- [7] Albert R, Jeong H, Barabasi A-L. Diameter of the World Wide Web [J]. Nature, 1999, 401(6749): 130-131
- [8] Cohen R, Havlin S. Scale-free networks are ultra-small [J]. Phys. Rev. Lett., 2003, 90: 058701(1-4)
- [9] Fronczak A, Fronczak P, Holyst J A. Mean-field theory for clustering coefficients in Barabasi-Albert networks [J]. Phys. Rev. E, 2003, 68: 046126(1-4)