

加权局域网络上的病毒传播行为研究

李慧嘉, 马英红

LI Hui-jia, MA Ying-hong

山东师范大学 管理与经济学院, 济南 250014

School of Management and Economics, Shandong Normal University, Jinan 250014, China

E-mail: lihuu2000@126.com

LI Hui-jia, MA Ying-hong. Study of epidemic spreading in weighted local-world complex networks. *Computer Engineering and Applications*, 2009, 45(35): 80-83.

Abstract: The problem of epidemic spreading has always been concerned by many scientists all over the world. However, there are no studies in the field of weighted local-world network. Since a lot of the real-world networks are weighted, the virus spreading in weighted local-world complex networks is studied in this paper. With the SI model of virus spreading being adopted and the virus spreading speed between any two nodes being positive correlation with the corresponding weight between them, the effect of different weight distributions on the spreading behavior in three weighted evolving networks is investigated. Study shows that scale-free and Weighted local-world properties of weighted local-world network both have an great effect on the Epidemic spreading. As the weighted local-world model can well reflex the real world, these research has board practical background.

Key words: scale-free network; weighted local-world networks; epidemic spreading; SI model; immunization

摘 要: 病毒传播问题的研究一直是国际上科学家所关注的焦点,但是在加权局域网络中的病毒传播研究却是空白。由于实际存在的网络很大一部分是加权局域网络,因此研究了一种特定加权局域网络中的传播行为。采用病毒传播的 SI 模型,令病毒的传播速度和网络的连接权重正相关。对加权局域网络中病毒传播行为的研究表明:加权局域网络的无标度性质和加权局域世界性质对病毒的传播有深刻的影响。由于加权局域网络能够很好地反应实际世界,因此该研究具有很广的应用背景。

关键词: 无标度网络; 加权局域网络; 病毒传播; SI 模型; 免疫

DOI: 10.3778/j.issn.1002-8331.2009.35.025 文章编号: 1002-8331(2009)35-0080-04 文献标识码: A 中图分类号: N94; TP393

1 引言

Barabasi 和 Albert 提出的 BA 无标度模型^[1]考虑了实际网络的两个重要演化特性:一是增长特性;二是优先连接特性,或称“富者更富”现象。在 BA 无标度网络模型采用全局的优先连接,由此得到幂率形式的网络度分布。然而通过世界贸易网^[2](World Trade Web)的研究中发现,全局的优先连接机制并不适用于那些只与少数(小于 20 个)国家有贸易往来关系的国家。类似地,在科学家合作网络^[3-4]中,合作经常也仅限于各自的局域世界里。这些都说明在诸多实际的复杂网络中存在着局域世界。因此,在许多现实的网络中,每一个节点都有各自的局域世界,全局的优先连接不一定能实现,而是更经常采用局部优先连接。

李翔等人提出了局域世界演化模型(Local-World Evolving Network)以刻画这种现象^[5],然而,目前对于局域世界的研究工作中,大多只是考虑非加权局域世界的生长和演化^[5-8],而加权

局域世界的研究刚刚开始^[9]。权值演化机制^[10-12]在网络中拥有非常重要的地位,就使含权网络引入到局域网络成了必然。刘慧^[13]等人研究了局域权值演化,并为其建立了网络模型,该模型不仅考虑了网络拓扑对网络的影响,也考虑了权值在网络演化过程中的影响,发现网络节点的度分布和权值分布都服从幂律分布。此加权局域网络具有无标度特征和局域特征两种特性。

因为局域加权网络更加符合实际情况,因此对其上的病毒传播情况的研究成为必要。对文献^[13]中的加权局域模型上的病毒传播进行了研究,采用 SI 病毒传播模型,发现该网络具有的无标度特征和局域网络特征都对病毒传播有很大影响。无标度性质反映为网络的非均匀性上,而局域网络性质反映为局域世界的规模。对加权局域世界演化模型的病毒模型进行了仿真研究,从感染节点的平均强度、网络非均匀性、局域世界规模和免疫机制四方面入手考察了加权局域世界模型的病毒传播特征。实验表明,网络的不均匀性随局域世界的规模的增加而增

基金项目:国家自然科学基金(the National Natural Science Foundation of China under Grant No.60673047);山东省自然科学基金(the Natural Science Foundation of Shandong Province of China under Grant);山东省教育厅科技项目(the Science and Technology Project of Shandong Province under Grant No.J07YJ02)。

作者简介:李慧嘉(1985-),研究生,主要研究方向:复杂网络与复杂系统;马英红(1971-),副教授,主要研究方向:复杂网络与复杂系统,图论,粗糙集应用等。

收稿日期:2008-12-17 **修回日期:**2009-02-25

加, 并对病毒传播有明显的促进。最后, 免疫机制研究表明: 基于距离的局部免疫拥有非常好的病毒遏制效果。

2 加权局域网络演化模型

加权局域网络演化模型^[3]的基本思想是: 保留无局域模型中的局域增长机制, 加入权重机制和重连机制, 利用局部优先连接实现增长的局部优先连接机制。

模型的算法包括如下几步:

步骤 1 给一个具有 m_0 个节点, e_0 条边的初始网络。初始的 e_0 条边没有重连, 并给每条边赋予初始权重 1。

步骤 2 每个时间间隔, 随机选取 $M (M < m_0)$ 个点, 作为一个局域世界, 并且循环执行以下 3 步:

(1) 以概率 p_1 在局域世界与局域外之间添加 m 条边, 实现局域世界与外界的连接: 按强度分布优先从局域世界中中和局域世界外选取两个节点进行连接, 并允许重连。

(2) 以概率 p_2 向局域世界加入 m 条边, 实现局域内部的演化, 节点的选择按强度分布优先。允许重连。

(3) 以概率 p_3 向局域世界加入一个新节点和 m 条边, 实现局域世界的生长。边连接规则也是按强度分布优先进行的。这 m 条边不允许重连和自连。

其中, p_1, p_2, p_3 满足 $p_1 + p_2 + p_3 = 1$ 。重复进行第 2 步, 经过计算, 该加权局域网络模型的强度分布为:

$$p(s) \approx \frac{(m+B/A)^{1/A}}{A(s+B/A)^{1+1/A}} \quad (1)$$

其中 $A = \frac{2ap_1 + p_2 + p_3}{2(ap_1 + p_2 + p_3)}$, $B = \frac{p_3 m M}{p_3 (M-1)}$ 。

可以得出, 当 t 足够大时, 该模型的强度分布为幂律分布, 幂律分布的指数为 $1+1/A$ 。加权局域网络具有无标度和局域性质双重特性。

3 病毒传播模型介绍

采用 SI 病毒传播模型^[4]研究病毒爆发在加权局域网络中的动态演化情况。在此病毒传播模型中, 节点只有两种状态, 易感染状态和感染状态, 且已感染的节点是不能复原的。在病毒爆发初期, 由于病毒爆发的突然性和对病毒特性不了解, 一般会缺少相应控制措施, 而导致病毒在短期内得不到控制, 使其传播速度很快, 这也是采用模型的原因。虽然分析的结果是在 SI 模型得到的, 但是只许稍加改动就可以扩展到 SIS 模型或 SIR 模型中。

在复杂网络中, 节点代表不同的个体, 边代表的是现实物理连接或各种关系, 病毒通过这些边来进行传播。

把节点间边的权重看成是节点间的连接的紧密程度。因此, 紧密程度的不同将影响到节点间病毒传播速率的不同。为描述不同权重的边对连接的强弱不同这一现象, 假设网络中边的权重越大, 连接越紧密, 病毒的传播速度也越大。为简单起见, 假设节点 i 和 j 之间的权重为 W_{ij} , 则定义它们之间的病毒感染率 $\lambda_{ij}^{[15]}$ 为:

$$\lambda_{ij} = \frac{W_{ij}}{A}, A > 0 \quad (2)$$

其中, A 为可调参数, 保证 $\lambda_{ij} < 1$ 。

在该模型中, 做以下假设: 网络的大小 N 是连续的, 且保持不变。函数 $S(t)$ 和 $I(t)$ 分别代表网络中易感染节点的数量和已感染节点的数量, 因此 $N = S(t) + I(t)$ 。

4 加权局域网络上的病毒传播研究

下面给出一些在加权局域网络上的重要的病毒传播性质。网络参数固定在 $p_1 = 0.01$ 、 $p_2 = 0.29$ 、 $p_3 = 0.7$ 、 $m = 3$ 和 $a = 10$ 的情况下。

4.1 加权局域网络感染节点的平均强度变化

为了分析不同节点在病毒传播中的作用, 给出了 $N = 5000$ 、 M 分别等于 10 和 80 时的加权局域网络中感染节点的平均强度变化(图 2)。定义新感染节点在 t 时刻的平均强度^[15]为:

$$\bar{S}_{inf}(t) = \frac{\sum_s s [I_s(t) - I_s(t-1)]}{I(t) - I(t-1)} \quad (3)$$

其中 $I_s(t)$ 为强度为 s 的感染节点的数量。图 1 展示了 $\bar{S}_{inf}(t)$ 随时间 t 的变化态势。

从图 1 可以看出, $\bar{S}_{inf}(t)$ 呈现幂率分布, 因此节点强度越大, 被感染后的危险性就越大。因此从免疫的角度分析, 如果要对整个网络实行有效的防护, 应该优先选择强度大的节点。

4.2 初始感染节点选择的比较

网络中的节点开始均处于易染状态(S), 选择其中 1% 的节点使之被感染。初始感染节点选择的不同, 在不同拓扑结构的网络中会不会影响病毒传播的性质? 图 2 给出了 $N = 5000$ 的初始感染网络中强度最大的节点与强度最小的节点的比较, 分别是 $M = 10$ 和 $M = 80$ 的情况。

由图 2 可以看出, 后者两条曲线的差异比前者显著, 而且 $M = 10$ 时病毒传播到达稳态的时间明显长于 $M = 80$ 时的情况。这说明初始被感染的节点的地位(以度的大小衡量)在 $M = 80$ 的网络中比 $M = 10$ 的网络中更重要。这种现象本质上是由网络的拓扑结构造成的, 随着局域世界模型演化参数 M 的变化, 生成不同拓扑特性的网络, 初始被感染节点的选择影响到病毒的传

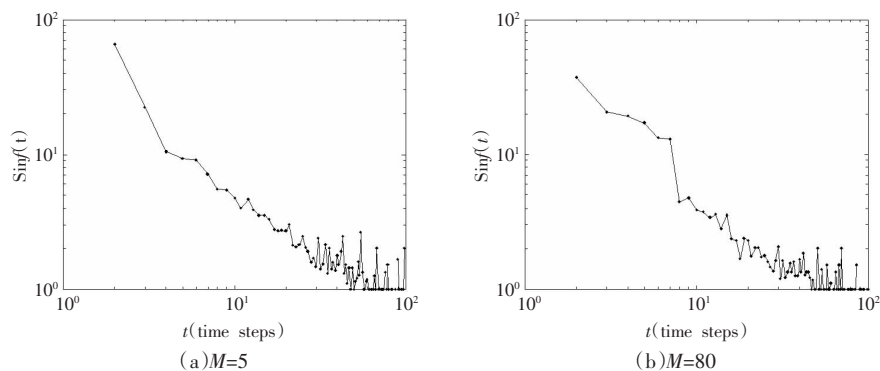
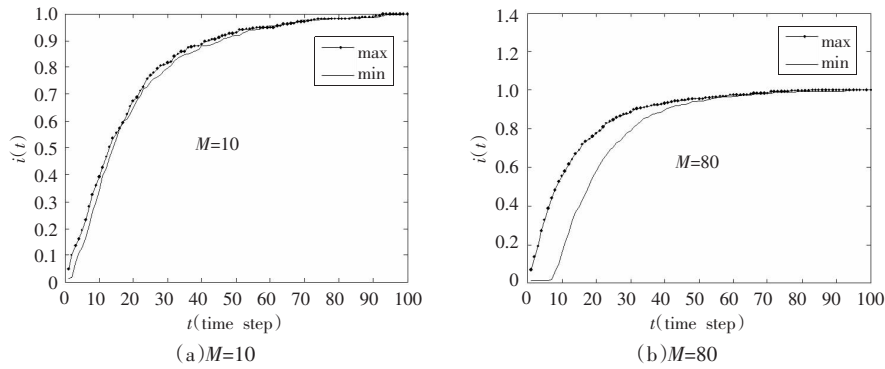
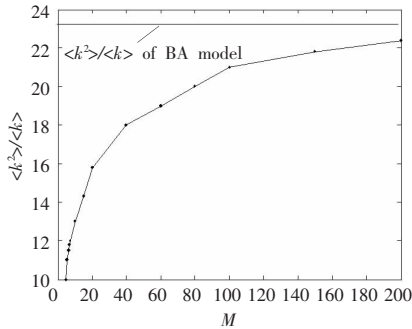
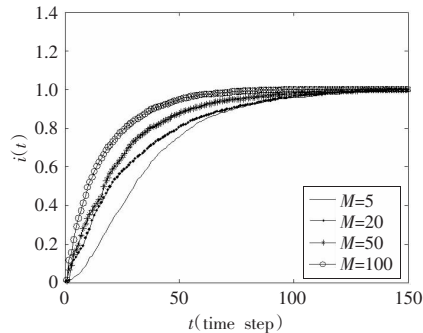
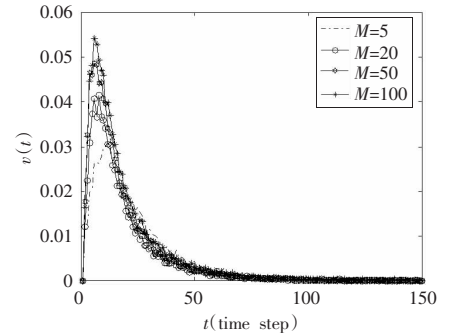


图 1 M 取不同值时的加权局域网络感染节点的平均强度比较

图2 初始感染 $M=10$ 和 $M=80$ 的网络中强度最大的节点和强度最小的节点引起的感染密度的变化图3 $\langle k^2 \rangle / \langle k \rangle$ 随 M 的变化态势

(a) 网络中稳态感染密度



(b) 病毒传播速率

图4 不同 M 值的网络中稳态感染密度和病毒传播速率的变化趋势

播行为。当 M 较大时,网络非均匀性增强,初始感染度最大的节点能加快病毒的传播,初始感染度最小的节点病毒传播扩散明显缓于前者。

为了给出网络非均匀性^[6]的准确性质,用数值仿真模拟出网络的非均匀性(由统计量 $\langle k^2 \rangle / \langle k \rangle$ 来表示)随 M 的变化情况(图3)。

结果显示 M 值越大, $\langle k^2 \rangle / \langle k \rangle$ 随之增加,网络的非均匀性越高,初始被感染的节点的地位也就随之提高。这与上面的分析相同,体现了加权局域世界存在非均匀网络在面对病毒攻击时存在着脆弱性。因此可见, M 的大小必然对于病毒传播有着重要的影响。下面分别通过理论分析和数值仿真对其进行详细研究。

4.3 局域世界规模 M 对病毒传播的影响

对于不同 M 下的加权网络模型上的病毒传播,首先给出不同 M 下的稳态感染密度变化趋势和传播速率变化态势。网络规模 $N=2000$ 。定义病毒传播的速率为网络中被感染个体的变化速率为:

$$v_{\text{inf}}(t) = \frac{di(t)}{d(t)} \approx \frac{I(t) - I(t-1)}{N} \quad (4)$$

模拟结果如图5所示,其中均为50次传播的平均。

从图4中可以清楚地看出,病毒传播速度随 M 的增大而增大。从上面的分析得知, M 的大小直接影响网络的不均匀性。局域世界规模 M 越大,新增点连接原有网络的边的范围越广,网络中的边分布的越不均匀。这反映在网络的聚类系数的大小上。聚类系数衡量网络的紧密程度,一般对于网络来说,聚类系数和平均路径长度有紧密联系,类似于社会关系网络中“物以类聚,人以群分”的特性。这里,聚类系数 C ^[7]按 Watts 和 Strogatz 提出的定义:

$$C = \frac{1}{N} \sum_i C_i \quad (5)$$

其中,局部值 C_i 定义为:

$$C_i = \frac{\text{包含顶点 } i \text{ 的三角形个数}}{\text{以顶点 } i \text{ 为中心的三点组的个数}}$$

计算了 M 分布为 5、10、20、50、100、200 下的 50 个网络的平均聚类系数,见表1。

表1 不同 M 下的聚类系数比较

M	5	10	20	50	100	200
聚类系数	0.191 6	0.142 3	0.113 2	0.106 4	0.096 9	0.090 8

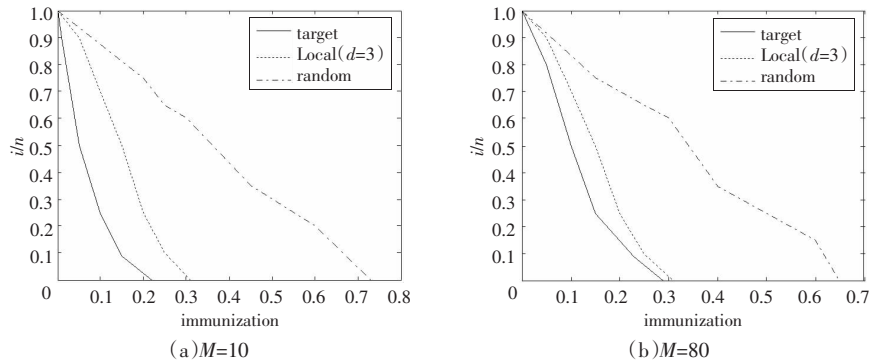
由表1可见,局域世界规模越大,聚类系数越小,因而网络中点间的路径长度也就越小,而病毒传播更容易,因此病毒传播速率更快。

4.4 加权局域网络模型上的病毒免疫传播过程比较

免疫是抑制病毒传播的一种重要方法。目前探讨最多的有三种免疫策略:随机免疫、目标免疫^[8]和基于距离的局域免疫^[9]。随机免疫对度大的节点和度小的节点平等看待,进行免疫时随机地选择节点,没有优先顺序。目标免疫法则利用网络的非均匀性特点,选取权大的节点进行免疫。一旦节点被免疫,则意味着它们所连的边可从网络中去除,这些边的去除重构了网络结构,切断了病毒的传播路径。但是目标免疫需要知道网络的全局信息才能实施。局域免疫是在网络全局信息未知的情况下,利用被感染节点周围的局部信息,控制被感染节点周围一定距离 d 内的节点,从而达到控制病毒在整个网络大规模传播的目的。

针对局域世界网络研究这三种免疫策略。如图5,对于 $M=10$ 和 80 时的加权局域网络,为完成整个网络的完全病毒控制,采用随机免疫、以强度为目标免疫和 $d=3$ 的局域免疫下的病毒传播情况,其中横坐标为免疫节点占总节点的比例,纵坐标为稳态感染节点数的比例。

从结果可以看出,对于加权局域网络,随机免疫所需要的

图5 $M=10$ 和 $M=80$ 时三种免疫策略条件下病毒传播情况

免疫节点数量最多,目标免疫最少,而 $d=3$ 的局部免疫和目标免疫在 M 比较大的时候相差不大。这是由加权局域网络特有的无标度性质和局域网络性质所引起的。通过文献[18]的分析,无标度网络中随机免疫的性能最差,目标免疫性能最好。因为加权局域网络中节点的不均匀性,当 M 比较大时,控制感染节点 3 步以内的节点,这些节点同是感染节点的概率很大,因而总的被控制节点数的比例相对较小,时间相对较短。因此,在不知道网络全局信息的情况下,对局部集聚的节点进行控制显然要比对分散稀疏的节点进行随机控制更为容易,代价也更低,更适合各种类型的实际的局域网络,如社会网络和因特网等。

5 小结

无标度性质和局域网络性质是加权局域网络特有的两种性质,无标度性质体现为网络的非均匀性,而局域网络性质导致了局域世界规模的大小差别。对加权局域世界演化模型的病毒模型进行了仿真研究,从感染节点的平均强度、网络非均匀性、局域世界规模和免疫机制四方面入手考察了加权局域世界模型的病毒传播特征。由于加权局域网络的非均匀性,初始选择感染节点对病毒传播有着重要影响。强度大的节点对于传播有着巨大的促进作用。局域世界规模 M 对传播过程的作用主要体现在传播速度上。这也反过来验证了网络的非均匀性质影响病毒传播的正确性。

加权无标度网络的这两条重要特性对其上的免疫也产生了影响。加权无标度网络的无标度性质使得目标免疫在其上拥有最好的效果,相反,随机免疫效果最差。而局域网络特性使得基于距离的局部免疫也拥有非常好的效果,图 5 表明局部免疫拥有几乎和目标免疫同样好的效果。这对实际加权局域网络的病毒防治也有重要的指导作用。

该研究也从一个侧面揭示了实际网络中人际关系方式和个人的联系范围对病毒传播趋势的影响。通过研究表明,降低人群和计算机链接的密集程度,减少联系和链接范围,再加上行之有效的免疫方式和医疗措施,如一定范围内的免疫和隔离,就会在短时间内实现较好的防御。

参考文献:

[1] Barabási A L, Albert R. Emergence of scaling in random networks[J]. Science, 1999, 286: 509-512.

[2] Ma Ángeles Serra no, Marián Boguñá. Topology of the world trade web[J]. Phys Rev, 2003, E 68.

[3] Newman M E J. Proc Natl Acad Sci[C]. 2001, USA 98, 404.

[4] Barabási A L, Jeong H, Neda Z, et al. Statistical, nonlinear, and soft matter physics[J]. Physica A, 2002, 311: 590.

[5] Li X, Chen G R. A local-world evolving network model[J]. Physica A, 2003, 328(1-2): 274-286.

[6] Wang Binghong, Tang Huanwen, Zhang Zhongzhi, et al. Evolving scale-free network model with tunable clustering[J]. Int J Mod Phys B, 2005, 19(26): 3951-3959.

[7] Chen Guanrong, Fan Zhengping, Li Xiang. Modeling the complex Internet topology[M]//Complex Dynamics in Communication Networks. Berlin: Springer-Verlag, 2004.

[8] 宋莉雅, 李翔, 汪小帆. 互联网的局域世界演化模型仿真研究[J]. 计算机仿真, 2006(10): 103-108.

[9] Pan Zaofeng, Li Xiang, Wang Xiaofang. Generalized local-world models for weighted networks[J]. Phys Rev E, 2006, 73.

[10] Barrat A, Barthélemy M, Vespignani A. Weighted evolving networks: Coupling topology and weight dynamics[J]. Phys Rev Lett, 2004, 92(22).

[11] 汪秉宏, 周涛, 何大韧. 统计物理学与复杂系统研究最新发展趋势分析[J]. 中国基础科学, 2005(3): 37-43.

[12] Wang W X, Wang B H, Hu B, et al. General dynamics of topology and traffic on weighted technological networks[J]. Phys Rev Lett, 2005, 94(18).

[13] 刘慧, 李增扬, 陆君安. 局域演化的加权网络模型[J]. 复杂系统与复杂性科学, 2006(3): 36-43.

[14] Murray J D. Mathematical Biology[M]. Berlin: Springer Verlag, 1993.

[15] Yan Gang, Zhou Tao, Wang Jie, et al. Epidemic spread in weighted scale-free networks[J]. Chinese Physics letters, 2005, 22(2): 510.

[16] 夏承遗, 刘忠信, 陈增强, 等. 局域世界网络上流行病传播的动态行为研究[J]. 天津师范大学学报: 自然科学版, 2007(2): 66-68.

[17] Watts D J, Strogatz S H. Collective dynamics of small-world networks[J]. Nature, 1998, 393: 440-442.

[18] Satorras R P, Vespignani A. Immunization of complex networks[J]. Phys Rev E, 2002, 65.

[19] 许丹, 李翔, 汪小帆. 复杂网络病毒传播的局域控制研究[J]. 物理学报, 2007, 56(3): 1313-1317.