

考虑交通流量的 SIR-CA 病毒传播模型

任环¹, 蒋国平¹, 胡海宏²

(1. 南京邮电大学 自动化学院, 江苏 南京 210003;

2. 南京邮电大学 电子科学与工程学院, 江苏 南京 210003)

摘要:基于一维元胞自动机,考虑网络交通流量不均衡的特点,提出新的 susceptible-infected-removed (SIR) 病毒传播模型,研究病毒在复杂网络中的传播行为。研究表明,对于某些被治愈个体难以获得免疫能力的传染病,随着网络交通流量增大,病毒在网络中传播速度明显加快,并在更短的时间内达到稳定的更高的感染规模。研究还发现,对于某些被感染的个体一旦被治愈就立即获得了永久的免疫能力的传染病,增大网络通信流量可以加速病毒的消亡,提高网络中免疫节点的比例。

关键词:病毒传播;交通流量;元胞自动机;状态转换函数

中图分类号:TP391.9

文献标识码:A

文章编号:1673-629X(2011)08-0250-04

A SIR-CA Model for Epidemic Propagation Considering Traffic Flow

REN Huan¹, JIANG Guo-ping¹, HU Hai-hong²

(1. College of Automation, Nanjing University of Posts and Telecommunications, Nanjing 210003, China;

2. College of Electronic Science and Engineering, Nanjing University of
Posts and Telecommunications, Nanjing 210003, China)

Abstract:Based on the one-dimensional cellular automata and considering the features of the global interaction of information network's nodes and the unbalances of traffic flow, a new susceptible-infected-removed (SIR) model is proposed to study epidemic spreading in complex networks with traffic flow. Simulation results show that the propagation velocity increases obviously and the infection rate will reach a stable and higher scale in a shorter time with the traffic flow becoming large for some infectious diseases harder-to-get immunity ability. Moreover, the probability of virus disappearance is accelerated and the proportion of immune is improved with the traffic flow becoming large for some infectious diseases easier-to-get immunity ability.

Key words:epidemic spreading; communication flow; cellular automata; transition function

0 引言

人类社会的日益网络化促进了现代公共卫生体系的不断完善,以努力减小病毒的威胁;但另一方面,这种网络化进程也使得人们的流动日益频繁,从而极大地加快了病毒的扩散速度^[1-3]。与生物病毒相比,计算机病毒借助于庞大的 Internet,更加容易地跨国界而时时刻刻地入侵到世界上每个角落^[4,5]。到目前为止,

甲型 H1N1 流感、SARS、禽流感等生物和计算机病毒的肆虐给人们带来了以万亿美元计数的巨大损失。迅速发展的复杂网络理论正有效地增进人们对爆发大规模生物和计算机病毒流行的传染机制的认识^[6-8]。同时,大量实证研究表明,诸如 Internet、P2P^[9] 和移动网络等都具有小世界效应,并且其节点度服从幂率分布。复杂网络关于小世界效应^[10]、无标度特性^[11] 等统计特性的发现使人们认识到基于规则网络的经典传播模型所得出的传播规律与复杂网络上的传播规律存在根本的不同^[12]。借助于复杂网络理论研究病毒传播机理,进而有效控制病毒传播受到人们的极大关注,并取得了很重要的研究成果^[4,13]。

文中针对实际网络病毒传播过程中存在网络交通流量的问题,基于元胞自动机,构建病毒传播模型,应用复杂网络理论研究网络交通流量对病毒传播行为的影响。研究表明,对于某些被治愈个体难以获得免疫

收稿日期:2010-12-08;修回日期:2011-03-19

基金项目:国家自然科学基金项目(60874091);江苏省高校自然科学基金基础研究计划(08KJD510022);江苏省“六大人才高峰”高层次人才计划(SJ209006);南京邮电大学引进人才计划(NY209021);江苏省普通高校研究生科研创新计划(CX10B_193Z)

作者简介:任环(1984-),女,山东烟台人,硕士研究生,研究方向为复杂系统与网络控制;蒋国平,教授,博士生导师,研究方向为混沌系统与复杂动态网络、网络信息安全、混沌通信。

能力的传染病,随着网络交通流量增大,病毒在网络中传播速度明显加快,并在更短的时间内达到稳定的更高的感染规模。研究还发现,对于某些被感染的个体一旦被治愈就获得了永久的免疫能力的传染病,增大网络通信流量可以加速病毒的消亡,提高网络中免疫节点的比例。

1 考虑通信流量的 SIR-CA 模型

使用元胞自动机建立的病毒传播随机模型如下:

$$CA = (C, Q, V, f) \quad (1)$$

其中, C 表示元胞空间, Q 表示有限状态集, V 表示节点的邻域, f 代表状态转换规则函数。网络 $G = (N; E)$, N 表示网络中节点的个数, E 表示网络中节点与节点的边, 令 A 表示网络的邻接矩阵, 它反映网络的拓扑信息。根据元胞自动机四要素, 建立模型。

元胞空间 C : 建立包含 N 个元胞的一维元胞空间, 一维元胞空间中的一个元胞即代表网络中的一个节点;

邻域 V : 本模型中以网络的邻接矩阵 A 定义各元胞邻居关系, 所以节点 i 的邻域 V_i 就被定义为 A 中的第 i 行的向量, 即 $V_i = \{a_{ij} \mid a_{ij} \in A, j = 1, 2, \dots, N\}$; 若 $a_{ij} = 1$, 表示节点 i 和 j 之间存在连接。

考虑免疫机制后, 节点状态处于健康状态 (susceptible)、感染状态 (infected) 和免疫状态 (removed) 之一, 节点状态变换关系: susceptible \rightarrow infected \rightarrow removed。特别应该注意的是, 文中所提的考虑通信流量的 SIR-CA 模型中的免疫状态的含义与经典的 SIR-CA 模型^[14]中免疫状态 (removed) 的含义有所差别: 在经典的 SIR-CA 模型中, 免疫节点意味着不会再遭受感染也不会再去感染其它健康节点, 所以网络中的免疫节点意味着对病毒的传播是没有贡献的, 可以把此种节点从网络中移除; 而在文中所提考虑通信流量的 SIR-CA 模型中显然不能把免疫节点从网络中移除, 因为所提模型中健康节点能否被其邻居感染是由其邻居是否发送了带病毒的数据包决定, 而不是由其邻居自身的状态决定。虽然免疫节点其本身不会产生带病毒的数据包也不会被带病毒的数据包感染, 但是它也可以转发带病毒的数据包给它的邻居, 从而感染它的邻居。为此定义状态集 $Q = \{0, 1, 2\}$, 其中 0 表示健康状态, 1 表示感染状态, 2 表示免疫状态, 节点 i 在 t 时刻的状态变量用 $s_i(t)$ ($s_i(t) \in Q$) 表示, 则有:

$$\begin{cases} s_i(t) = 0, \text{node } i \text{ is susceptible at time } t \\ s_i(t) = 1, \text{node } i \text{ is infected at time } t \\ s_i(t) = 2, \text{node } i \text{ is removed at time } t \end{cases} \quad (2)$$

另外, 在经典的 SIR-CA 模型中, 感染节点在单位

时步以相同的概率感染它的所有邻居, 但实际并不如此, 例如在 Internet 网络中, 两个主机之间虽然有通信连接但是没有数据包的传输, 那么也不会传播病毒。考虑到网络中交通流量的不均衡特性以及数据包传输路径的选取与具体的路由转发协议有关的特点, 对上述 SIR-CA 模型进行改进, 并提出改进的 SIR-CA 模型。为了便于问题的分析, 假设网络中每个节点生成数据包的平均速率为 α , 则在每一个时步, 网络中生成新数据包的总量即网络交通流量^[15,16]为 $N\alpha$ 。在每个时步, 网络中任何一个节点生成的数据包都随机选择目的节点进行发送, 数据包传输路径的选取则采用最短路径优先算法^[17]实现, 基于此定义网络中节点算法介数^[18] b_{alg} 为穿过一个节点的所有最短数据包传输路径的总和。那么, 网络中任一健康节点 i 在单位时步被感染的概率取决于单位时步穿过这个节点的带病毒的数据包的个数 M_i 。显然 M_i 的值与网络交通流量 $N\alpha$ 和节点算法介数 b_{alg} 有关: 网络交通量越大, 带病毒的数据包数目发送量越多, M_i 的值就会越大; 节点算法介数 b_{alg} 越大, 穿过这个节点的数据包越多, M_i 的值就会越大。假设一个健康节点接收或者转发一个带病毒的数据包, 这个健康节点被感染的概率和一个健康数据包经过一个感染节点被感染的概率都为 β 。那么一个健康节点 i 在单位时步被感染的概率为 $1 - (1 - \beta)^{M_i}$ 。定义 μ 为: 在 t 时刻为健康状态的节点 i 在 $t+1$ 时刻得以康复的概率。定义 θ 为: 在 t 时刻为感染状态的节点 i 在 $t+1$ 时刻以康复率 μ 获得治愈的条件下, 节点 i 在 $t+1$ 时刻获得免疫的概率。

因此, 建立考虑通信流量的 SIR-CA 模型的状态转换规则如下:

```
(1) if  $S(t-1, i) = 0$  //节点  $i$  在  $t-1$  时刻是健康的
     $p1 = \text{rand}(1, 1)$  //产生一个随机数  $p1$ 
    if  $p1 < p$  //若该随机数  $p1$  小于感染概率  $p$ ,
         $S(t, i) = 1$ ; //则节点  $i$  在  $t$  时刻被感染, 即  $S(t, i) = 1$ 
    end
end
(2) if  $S(t-1, i) = 1$  //节点  $i$  在  $t-1$  时刻是感染的
     $p2 = \text{rand}(1, 1)$  //产生一个随机数  $p2$ ,
    if  $p2 < \mu$  //若  $p2$  小于康复率  $\mu$ , 则节点  $i$  在  $t$  时刻得以康复
         $p3 = \text{rand}(1, 1)$  //产生一个随机数  $p3$ 
        if  $p3 < \theta$  //若  $p3$  小于获得免疫的概率  $\theta$ ,
             $S(t, i) = 2$ ; //则节点  $i$  在  $t$  时刻变为免疫的, 即  $S(t, i) = 2$ ;
```

```

else          //否则,节点 i 仅仅是健康的
S(t,i)=0;      //即 S(t,i)=0;
end
end
end

```

(3) if $S(t-1,i)=2$ //节点 i 在 $t-1$ 时刻是免疫的,表示节点 i 不会产生带病毒的数据包

```

S(t,i)=2;    //也不会被带病毒的数据包感染
end          //节点 i 的状态不随时间变化

```

2 仿真结果与讨论

本节以 BA 网络为例,通过数值仿真研究文中提出新的 SIR-CA 模型在小世界网络^[19]上的传播特性。仿真过程中选取参数如下:网络的节点数为 $N=150$,病毒在网络中传播的初始时刻为 0,随机选择一个节点作为感染节点;在 WS 网络中随机化重连概率 p 的取值为 0.2,网络的平均度 $\langle k \rangle = 6$;假设初始时刻网络中感染病毒的节点数为 1, $I(t)$ 表示 t 时刻网络中受感染节点数目,即 $I(0)=1$,其它参数取值如下: $\beta=0.3$, $\mu=0.5$, α 分别取 1、2、3、4。仿真过程中转发数据包的路由选择协议采用开放最短路径优先协议。

在 WS 网络中,针对不同的 α 值分别执行 100 次仿真试验, $i(t)$ 表示 t 时刻网络中受感染节点的比率, $r(t)$ 表示 t 时刻网络中免疫节点的比率。 $s(t)$ 表示 t 时刻网络中健康节点的比率。对 100 次仿真实验所得数据进行统计平均,得到病毒的传播演化曲线如图 1、2 所示。

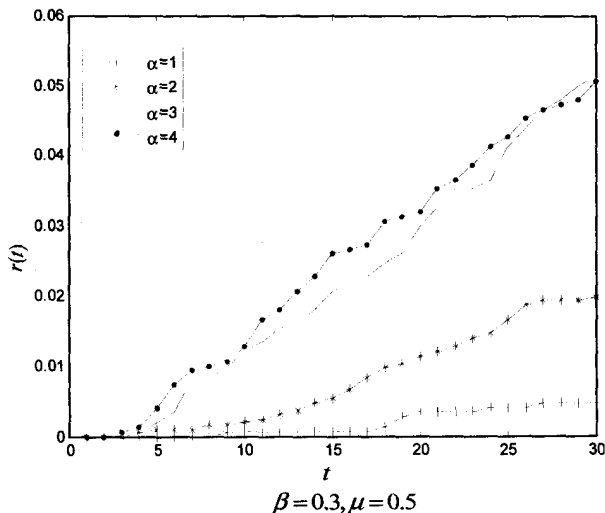


图 1 感染节点密度 $i(t)$ 与通信流量 α 的关系
($\theta=0.01$)

由图 1、2 可见,网络交通流量对病毒传播行为的影响是显著的:随着网络交通流量增大,病毒在小世界网络中传播速度明显加快,病毒在更短的时间内达到稳定的更高的感染规模。这告诉我们,减少网络的交

通流量可以有效地降低病毒爆发的可能性。随着网络中交通流量的增加,即在单位时步里网络中产生并发送的数据包数量的增加,病毒爆发的可能性明显增大,但是当网络交通流量达到一定的值后,再增加网络交通流量,对病毒感染规模的影响却不再明显,可见病毒最终感染规模并不是随网络交通流量的增加而线性增加的。

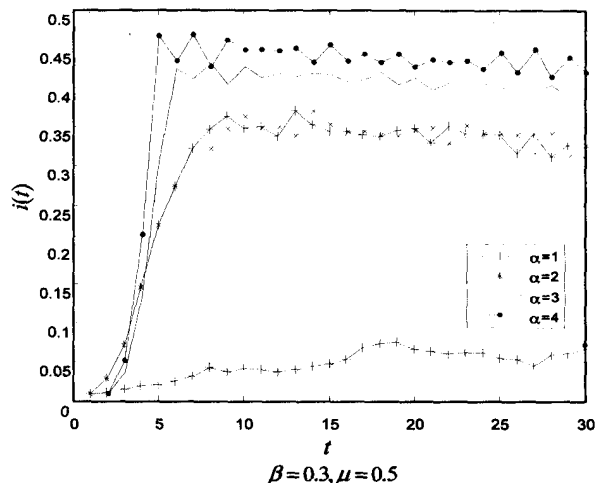


图 2 免疫节点密度 $r(t)$ 与通信流量 α 的关系
($\theta=0.01$)

下面改变康复节点获得免疫能力的概率 θ 的值: $\theta=1$,进一步研究网络通信流量对病毒传播行为的影响。对 100 次仿真实验所得数据进行统计平均,得到病毒的传播演化曲线如图 3、4 所示。由图 3、4 可见,随着网络通信流量的增大,病毒的传播速度明显加快,同时病毒的消亡速度也明显加快,病毒在更短的时间内消亡。由图 3、4 还可以观察到,当到达第 18 时步时,不论通信流量 α 为多少,网络中感染病毒的节点数都为 0,而获得免疫的节点数目随 α 的不同而稳定在不同的规模: α 越大,获得免疫能力的节点的密度就越

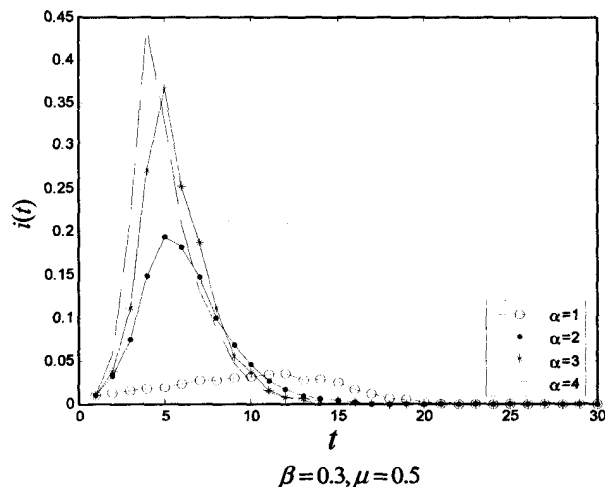


图 3 感染节点密度 $i(t)$ 与通信流量 α 的关系
($\theta=1$)

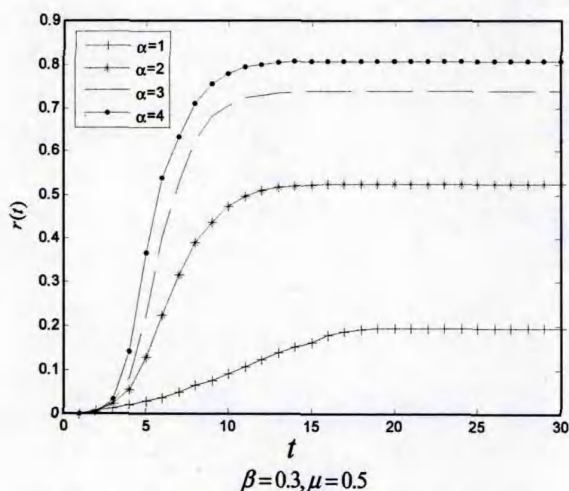


图4 免疫节点密度 $r(t)$ 与通信流量 α 的关系
($\theta=1$)

大。由此可见,对于某些被感染的个体一旦被治愈就具有了永久的免疫能力的传染病,增大网络通信流量可以加速病毒的消亡,提高网络中免疫节点的比例。

3 结束语

在许多复杂网络病毒传播过程中,病毒的传播特性会受到网络交通流量的影响。为此文中提出了一种新的 SIR-CA 传播模型,研究了该模型在小世界网络上的传播行为。经研究发现:对于某些被治愈个体难以获得免疫能力的传染病,随着网络交通流量增大,病毒在网络中传播速度明显加快,并在更短的时间内达到稳定的更高的感染规模。研究还发现,对于某些被感染的个体一旦被治愈就获得了永久的免疫能力的传染病,增大网络通信流量可以加速病毒的消亡,提高网络中免疫节点的比例。

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] Faloutsos M, Faloutsos P, Faloutsos C. On power-law relationships of the internet topology[C]//Proc of the conference on applications, technologies, architectures, and protocols for computer communication. Cambridge, Massachusetts, United States: ACM, 1999.
- [3] 2005 年全国信息安全状况与计算机病毒疫情调查分析报告[EB/OL]. 2005. <http://blog.hexun.com/glaciers/1550753/rss/viewarticle.html>.

- [4] Vázquez A, Pastor-Satorras R, Vespignani A. Large-scale topological and dynamical properties of the internet[J]. Physical Review E, 2002, 65(6): 66130-66142.
- [5] 方锦清,汪小帆,刘曾荣. 略论复杂性和非线性复杂网络系统的研究[J]. 科技导报, 2004(2): 9-12.
- [6] 许丹,李翔,汪小帆. 复杂网络理论在互联网病毒传播研究中的应用[J]. 复杂系统与复杂性科学, 2004(3): 10-26.
- [7] 许丹. 复杂网络中的病毒传播研究[D]. 上海:上海交通大学, 2006.
- [8] 南湖浩,陈钟. 网络安全技术概论[M]. 北京:国防工业出版社, 2003.
- [9] Khambatti M, Ryu K D, Dasgupta P. Structuring peer-to-peer networks using interest-based communities[C]//Proc. of the 1st Int. DBISP2P Workshop. Berlin, Germany: [s. n.], 2003: 154-166.
- [10] Watts D J, Strong S H. Collective dynamics of 'small-world' networks[J]. Nature, 1998, 393(6684): 440-442.
- [11] Barabási A L, Albert R. Emergence of scaling in random networks[J]. Science, 1999, 286(5439): 509-512.
- [12] Faloutsos M, Faloutsos P, Faloutsos C. On power-law relationships of the internet topology[C]//Proceedings of the conference on applications, technologies, architectures, and protocols for computer communication. [s. l.]: [s. n.], 1999: 251-262.
- [13] Albert R, Jeong H, Barabási A L. Internet: Diameter of the world-wide web[J]. Nature, 1999, 401(6749): 130-131.
- [14] Jiang Guoping, Song Yurong. Research of malware propagation in complex networks based on 1-D cellular automata[J]. Chin. Phys. Soc., 2009, 58(9): 5911-5918.
- [15] Farley D, Slade R, Gattiker U E. 计算机病毒揭秘[M]. 朱代祥,贾建勋,史西斌,译. 北京:人民邮电出版社, 2002.
- [16] Meloni S, Arenas A, Moreno Y. Traffic-driven epidemic spreading in finite-size scale-free networks[J]. Proceedings of the National Academy of Sciences of the United States of America, 2009, 106: 16897-16902.
- [17] Boguñá M, Krioukov D, Claffy K C. Navigability of complex networks[J]. Nature Physics, 2009, 5(1): 74-80.
- [18] Guimerà R, Díaz-Guilera A, Vega-Redondo F, et al. Optimal network topologies for local search with congestion[J]. Physical Review Letters, 2002, 89(24): 248701-248705.
- [19] Barabási A L, Albert R, Jeong H. Scale-free characteristics of random networks: The topology of the world-wide web[J]. Physica A: Statistical Mechanics and its Applications, 2000, 281(1-4): 69-77.

计算机技术与发展友情提示: 本刊为中国科技核心期刊, 中国科技论文统计源期刊。《中国核心期刊数据库收录期刊》、《中国学术期刊综合评价数据库统计源期刊》、《中国期刊全文数据库收录期刊》、《万方数据资源系统数字化期刊群上网期刊》、《中国学术期刊(光盘版)》。不愿意通过上述媒体发行者, 请在来稿首页注明。