

小世界网络上随机 SIS 模型分析

李光正¹, 史定华²

(1. 江苏财经职业技术学院基础部, 淮安 223003; 2. 上海大学数学系, 上海 200444)

摘要: 考察小世界网络上疾病传播的随机 SIS 模型, 使用拟平稳分布计算方法得到疾病传播稳态时患病节点数的分布。取分布的均值, 得到与平均场方法相同的传播阈值。通过模拟所得的传染曲线解释现实传染过程中存在的波动性, 传播稳态结果和平均场结果拟合较好, 证实了平均场方法的合理性。

关键词: 小世界网络; SIS 模型; Markov 链; 拟平稳分布

Stochastic SIS Model Analysis on Small-world Network

LI Guang-zheng¹, SHI Ding-hua²

(1. Foundation Department, Jiangsu Vocational and Technical College of Finance & Economics, Huaian 223003;

2. Department of Mathematics, Shanghai University, Shanghai 200444)

【Abstract】 Stochastic SIS model of epidemic spreading on small-world network is studied. Using the computational method of quasi-stationary distribution, the distribution of the number of infective nodes at the steady state is obtained. Taking expectation of the distribution, the epidemic threshold is proved the same with the result of the mean-field method. Through simulating the epidemic curve of the stochastic SIS model, existing fluctuations in the real epidemic process is interpreted. The stationary epidemic results of stochastic SIS model fit well with the mean-field method, thus verifying the mean-field method is reasonable.

【Key words】 small-world network; SIS model; Markov chain; quasi-stationary distribution

随着近年来复杂网络研究的进展, 网络上的疾病传播问题成为复杂网络研究的一个重要组成部分^[1]。但到目前为止, 使用的研究方法主要是确定性方法, 即用平均场理论对模型进行解析分析, 并通过计算机模拟对解析结果进行验证。尽管疾病传播所具有的概率本性决定了随机模型比确定性模型更符合实际, 而且随机方法应用于疾病传播研究已经有比较长的历史, 但在复杂网络上的疾病传播研究中随机性方法基本没有介入, 针对复杂网络研究中这一比较奇特的现象, 本文在小世界网络^[2]上考察疾病传播模型 SIS, 研究随机方法应用于复杂网络上的疾病传播行为。

1 小世界网络上SIS模型的平均场方法分析结果

1.1 传染过程的精确解

SIS 模型根据节点是否处于患病状态把网络的全部节点分成 2 类: 易感者类 S 和染病者类 I , T 时刻 2 类节点的密度分别为 $s(t)$, $i(t)$, $s(t) + i(t) = 1$, $s(\infty)$, $i(\infty)$ 分别表示当传播达到稳定状态时 2 类节点的相对密度。易感节点如果与染病节点有边连接, 则以概率 β 变成染病节点; 染病节点以概率 γ 自动康复, 但没有免疫力, 可以立即被再次感染, 其中, $\lambda = \beta/\gamma$ ($\beta, \gamma > 0$) 称为有效传染率。

根据平均场方法^[3], 在小世界网络上 SIS 类疾病传播演化方程为

$$\begin{cases} \frac{ds(t)}{dt} = -\beta\langle k \rangle s(t)i(t) + \gamma i(t) \\ \frac{di(t)}{dt} = -\gamma i(t) + \beta\langle k \rangle s(t)i(t) \end{cases} \quad (1)$$

方程的平衡点为 $(s, i) = (1, 0)$ 或 $(\lambda^{-1}\langle k \rangle^{-1}, 1 - \lambda^{-1}\langle k \rangle^{-1})$, 令 $\lambda_c = \langle k \rangle^{-1}$, 则当 $\lambda > \lambda_c$ 时, 疾病爆发且稳定在地方病状态。

因此, 小世界网络上 SIS 模型的传播阈值为 $\lambda_c = \langle k \rangle^{-1}$, 只与 $\langle k \rangle$ 有关, 与传播参数 λ, β 及初始患病节点密度无关。传染曲线为 $i(t) = [Ce^{-(\beta\langle k \rangle - \gamma)t} + (1 - \lambda^{-1}\langle k \rangle^{-1})^{-1}]^{-1}$, 其中, C 是任意常数。

1.2 传染过程的仿真结果

取网络规模为 $N=10^4$, 重连概率 $p=0.1$, 网络的平均度 $\langle k \rangle = 6$; 初始患病节点密度为 $\alpha = i(0) = 0.1$; β 值依次取为 0.02, 0.03, 0.04; $\gamma = 0.1$ 。在小世界网络上模拟疾病的传播曲线如图 1 所示。其中, 平滑曲线是理论曲线; 波动曲线是模拟曲线, 从上到下, β 值依次为 0.02, 0.03, 0.04。

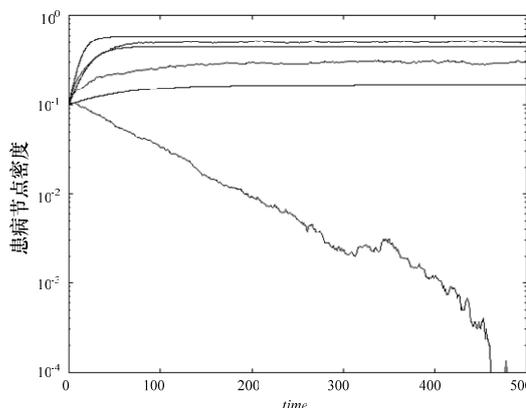


图 1 疾病的传播曲线

基金项目: 国家自然科学基金资助项目(10872119, 60874083)

作者简介: 李光正(1974—), 男, 讲师、硕士, 主研方向: 复杂网络传播动力学; 史定华, 教授、博士生导师

收稿日期: 2008-11-09 **E-mail:** guangzhengli@126.com

模拟曲线证实尽管模拟值和理论值有一定的差距,而且阈值 λ_c 的模拟值比理论值大,但能够充分说明小世界网络上疾病传播存在阈值现象。

1.3 重连概率 p 对传染过程的影响

图 2 的传染曲线显示重连概率 p 不影响小世界网络上的 SIS 类疾病的传播稳态,但会极大地影响到达稳态的时间,这说明在只考虑稳态时,可以不考虑 p 的影响。图 2 网络规模 $N=10^4$, 平均度 $\langle k \rangle=6$, 传染参数 $\beta=\gamma=0.1$ 。从左至右, p 依次为 0.1, 0.01, 0.001, 0.0001, 0 的传染曲线。

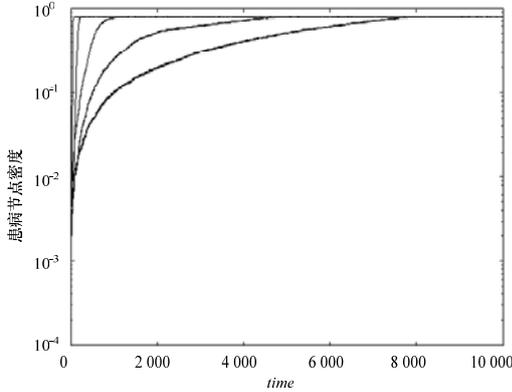


图 2 模拟的传染曲线重连概率

2 小世界网络上随机 SIS 模型分析

2.1 复杂网络上 SIS 模型研究现状和突出问题

SIS 模型的研究方法可以分为 2 大类:平均场方法和随机过程方法。通常先用平均场方法进行理论分析,并得到一些基本结论,如疾病传播阈值和稳态指标。但用平均场方法得到的稳态时患病节点密度是常数,而传染过程的概率本性决定了稳态时患病节点密度不可能是常数,应该具有一定的波动性,这种稳态时的波动性是平均场方法不能给出的。图 3 模拟了传染曲线的一个稳态现实,具有明显的波动性,是小世界网络上 SIS 模型的一次传播现实,其中, $N=10^4$; $\langle k \rangle=6$; $\beta=\gamma=0.1$; $p=0.1$ 。

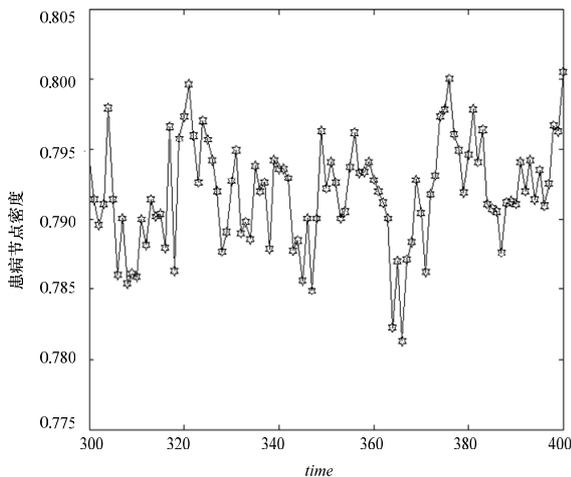


图 3 小世界网络上 SIS 模型

为克服平均场方法的不足,必须考虑传播过程中的随机因素,下面建立随机 SIS 模型。

2.2 复杂网络上的随机 SIS 模型

对规模为 N 的小世界网络,由于传染概率 β 和康复概率 γ 的影响,在 $t(\geq 0)$ 时刻患病节点个数 $X(t)$ 是不确定的整数,

又由于未来的患病节点个数只与当前患病节点个数有关,与过去患病节点个数无关,这是典型的 Markov 过程特征,因此可以把 t 时刻患病节点个数 $X(t)$ 描述为一个连续时间有限状态的 Markov 链。

用 $p_i(t) = Pr(X(t)=i)$ 表示 t 时刻有 i 个患病节点的概率,则经过充分小时间 Δt 后,由于传染作用, $t+\Delta t$ 时刻网络中有一个易感节点患病的概率为 $\lambda_i \Delta t = i(N-i)N^{-1}\beta\langle k \rangle \Delta t > 0$, 同时,有一个患病节点康复的概率为 $\gamma_i \Delta t$ 。这样,上述 Markov 链的状态转移概率为

$$Pr(X(t+\Delta t)=i+1 | X(t)=i) = \beta\langle k \rangle i(N-i)N^{-1}\Delta t = \lambda_i \Delta t$$

$$Pr(X(t+\Delta t)=i-1 | X(t)=i) = i\gamma\Delta t = \mu_i \Delta t$$

$$Pr(X(t+\Delta t)=i | X(t)=i) = 1 - (\lambda_i + \mu_i)\Delta t$$

这里, $0 < i < N$ 。特别地,对 $i=0$ 和 $i=N$, 有

$$Pr(X(t+\Delta t)=0 | X(t)=0) = 1$$

$$Pr(X(t+\Delta t)=N-1 | X(t)=N) = \mu_N \Delta t$$

$$Pr(X(t+\Delta t)=N | X(t)=N) = 1 - \mu_N \Delta t$$

为使上述等式具有概率意义,必需 $(\lambda_i + \mu_i)\Delta t \leq 1$, 对充分小的 Δt , 这些要求总能满足。由此得到 $X(t)$ 的状态转移概率矩阵为

$$T = \begin{pmatrix} 1 & 0 & 0 & 0 & \dots & 0 & 0 \\ \mu_1 \Delta t & 1 - (\lambda_1 + \mu_1)\Delta t & \lambda_2 \Delta t & 0 & \dots & 0 & 0 \\ 0 & \mu_2 \Delta t & 1 - (\mu_2 + \lambda_2)\Delta t & \lambda_3 \Delta t & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & 0 & \dots & 1 - (\mu_{N-1} + \lambda_{N-1})\Delta t & \lambda_N \Delta t \\ 0 & 0 & 0 & 0 & \dots & \mu_N \Delta t & 1 - \mu_N \Delta t \end{pmatrix}$$

令 $p(t) = (p_0(t), p_1(t), \dots, p_N(t))$, 有 $p(t+\Delta t) = p(t)T$, 即状态转移概率 $p_i(t)$ 满足下面的差分方程:

$$p_i(t+\Delta t) = p_{i-1}(t)\lambda_{i-1}\Delta t + p_i(t)(1 - \lambda_i\Delta t - \mu_i\Delta t) + p_{i+1}(t)\mu_{i+1}\Delta t$$

其中,对于 $i \in \{0, 1, \dots, N\}$, $0 < i < N$, 并且 $p_i(t) = 0$ 。对于 $i=0$ 和 $i=N$, 有

$$p_0(t+\Delta t) = p_0(t) + p_1(t)\mu_1\Delta t$$

$$p_N(t+\Delta t) = p_{N-1}(t)\lambda_{N-1}\Delta t + p_N(t)(1 - \mu_N\Delta t)$$

注意到状态 0 是转移概率矩阵 T 的唯一吸收状态。根据 Markov 过程理论,有 $\lim_{t \rightarrow \infty} p_0(t) = 1$, 所以,过程的平稳分布为 $p = (1, 0, 0, \dots, 0)$ 。这说明经过长时间演化之后,不论传染过程的参数如何,以概率 1 网络中不再有患病节点,这与平均场理论给出的传染阈值有本质的区别。

2.3 随机 SIS 模型的拟平稳分布

吸收状态状态 0 导致疾病传播以概率 1 必然会自动灭绝,为避免其影响,考察过程 $X(t)$ 的拟平稳分布。对非灭绝的情况,令 $q_i(t) = p_i(t)(1 - p_0(t))^{-1}$, 则 $q_i(t)$ 满足差分方程:

$$q_i(t+\Delta t)(1 - q_i(t)\gamma\Delta t) = q_{i-1}(t)\lambda_{i-1}\Delta t + q_i(t)[1 - \lambda_i\Delta t - \mu_i\Delta t] + q_{i+1}(t)\mu_{i+1}\Delta t$$

其中, $i=1, 2, \dots, N$, $q_i(t) + q_2(t) + \dots + q_N(t) = 1$ 。记 $q = (q_1, q_2, \dots, q_N)$ 是上述差分方程的平稳解,由上式得 q 满足等式:

$$-q_i q_i \mu_i = q_{i-1} \lambda_{i-1} - q_i (\mu_i + \lambda_i) + q_{i+1} \mu_{i+1}$$

这是关于 q 的非线性二次方程组,任意的 $q_i (i \geq 2)$ 都可以由 q_1 唯一表示,再由归一性条件 $q_1(t) + q_2(t) + \dots + q_N(t) = 1$, 可得 q 的唯一解。但由于计算机的 CPU 运行时间或内存的制约,对于较大的 $N (\geq 10^4)$, 很难得到上述方程的显式解或数值解。

为克服上述计算困难, Näsell^[4] 在考虑传统的随机 SIS 模型时,假定当 $X(t)=1$ 时患病节点个数不会减少,即

$P(X(t+\Delta t)=0|X(t)=1)=0$, 并去掉矩阵 T 的第 1 行和第 1 列, 重新设定 $T_{2,2}=1-\lambda_1\Delta t$, 得到拟平稳分布的一个近似良好的算法。由上述近似得到 Markov 链的概率转移矩阵:

$$\tilde{T} = \begin{pmatrix} 1-\lambda_1\Delta t & \lambda_1\Delta t & 0 & \dots & 0 & 0 \\ \mu_2 & 1-(\mu_2+\lambda_2)\Delta t & \lambda_2\Delta t & \dots & 0 & 0 \\ 0 & \mu_3\Delta t & 1-(\mu_3+\lambda_3)\Delta t & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 1-(\mu_{N-1}+\lambda_{N-1})\Delta t & \lambda_{N-1}\Delta t \\ 0 & 0 & 0 & \dots & \mu_N\Delta t & 1-\mu_N\Delta t \end{pmatrix} \quad (2)$$

令 $\tilde{q} = (\tilde{q}_1, \tilde{q}_2, \dots, \tilde{q}_N)$ 为式(2)的平稳分布, 重复上面的分析, 得到拟平稳分布的近似结果:

$$\tilde{q}_n = \tilde{q}_1 \frac{\lambda_{n-1}\lambda_{n-2}\dots\lambda_1}{\mu_n\mu_{n-1}\dots\mu_2}$$

其中, $n=2, 3, \dots, N$; $\sum_{i=1}^N \tilde{q}_i = 1$ 。

2.4 随机 SIS 模型传染曲线模拟

转移概率矩阵(式(2))对应的密度矩阵为

$$Q = \begin{pmatrix} -\lambda_1 & \lambda_1 & 0 & \dots & 0 & 0 \\ \mu_2 & -(\mu_2+\lambda_2) & \lambda_2 & \dots & 0 & 0 \\ 0 & \mu_3 & -(\mu_3+\lambda_3) & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & -(\mu_{N-1}+\lambda_{N-1}) & \lambda_{N-1} \\ 0 & 0 & 0 & \dots & \mu_N & -\mu_N \end{pmatrix}$$

假定在 t 时刻系统中有 i ($1 \leq i \leq N$) 个患病节点, 则系统中有 i 个患病节点的逗留时间 τ_i 是随机的, 且服从分布 $F_i(t) = P_i(\tau_i \leq t) = 1 - e^{-Q_i t}$, 其中, $t \geq 0, 1 \leq i \leq N$; Q_i 是矩阵 Q 对角线上第 i 元素。

由于 τ_i 服从指数分布 $F_i(t)$, 因此 $F_i(\tau_i)$ 服从区间(0,1)上的均匀分布, 从而可以通过生成(0,1)上的均匀分布随机数来模拟系统中有 i 个患病节点的逗留时间 τ_i 。当经过时间 τ_i 后, 系统中患病节点个数开始变化, 但每次改变只能增加或减少一个节点, 不能增加或减少多个。特别是当 $i=1$ 时, 系统中患病节点个数只增不减; 当 $i=N$ 时, 系统中患病节点个数只减少不增加。此时, 系统中增加一个患病节点的概率为

$$P_a = \begin{cases} 1 & i=1 \\ \lambda_i/(\mu_i+\lambda_i) & 2 \leq i \leq N-1 \\ 0 & i=N \end{cases}$$

减少一个患病节点的概率为

$$P_b = \begin{cases} 0 & i=1 \\ \mu_i/(\mu_i+\lambda_i) & 2 \leq i \leq N-1 \\ 1 & i=N \end{cases}$$

取参数值分别为: $N=10^4, \langle k \rangle=6, p=0.1$, 初始患病节点密度 $\alpha=0.1$, 传染参数 $\beta=\gamma=0.1$, 模拟的传染曲线见图 4, 从上到下分别为平均场方法、Markov 链方法和模拟曲线。

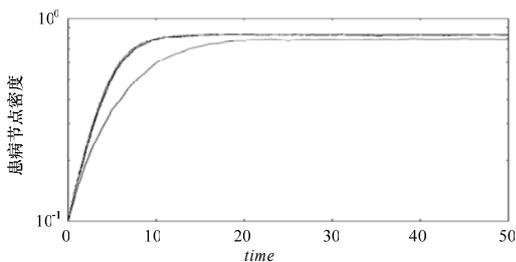


图 4 平均场方法、Markov 链方法和模拟曲线

由图 4, 平均场方法和 Markov 链方法的结果拟合非常好, 但两者与模拟结果有一定差距, 原因有 2 个方面: (1) 模拟曲

线对时间的离散化处理; (2) 平均场方法和 Markov 链方法分析过程中的近似处理使理论分析结果和实际结果有一定的差距。由图 5, Markov 链方法体现了传播过程患病节点个数变化的波动性, 这与现实的疾病传播是一致的。图中, 上方直线是平均场方法曲线, 上方波动曲线是随机方法曲线, 下方是模拟曲线。

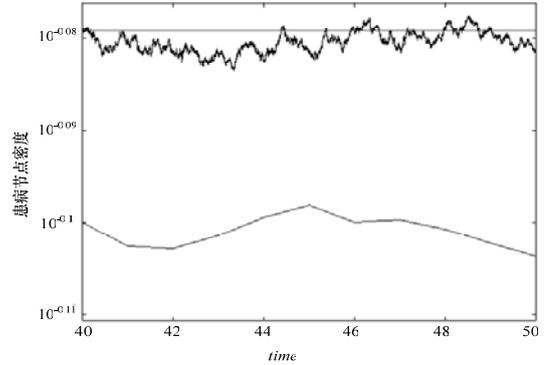


图 5 图 4 的局部曲线微视图

2.5 随机 SIS 模型的传播阈值

平均场方法在理论上给出了传播过程的阈值, 传播过程模拟结果也证实了阈值的存在性, 而模拟曲线是对多次传播现实的算术平均, 这启发人们可以通过考察概率分布的均值来研究随机 SIS 模型的阈值情况。

令概率分布 $p(t)$ 和 $q(t)$ 的均值分别为

$$m(t) = \sum_{i=0}^N i p_i(t), m^*(t) = \sum_{i=0}^N i q_i(t)$$

为考察均值的时间演化行为, 考虑两者的差分方程:

$$m(t+\Delta t) - m(t) = \sum_{i=1}^N (\frac{\lambda_i}{i} \Delta t - \gamma \Delta t) i p_i(t) = \gamma \Delta t [\lambda \langle k \rangle \xi(t) - 1] m(t)$$

$$m^*(t+\Delta t) [1 - \gamma q_1(t) \Delta t] - m^*(t) = \gamma \Delta t [\lambda \langle k \rangle \xi^*(t) - 1] m^*(t)$$

其中, $\xi(t) = \frac{\sum_{i=0}^N \lambda_i p_i(t)}{\beta \langle k \rangle \sum_{i=0}^N i p_i(t)}$; $\xi^*(t) = \frac{\sum_{i=1}^N \lambda_i q_i(t)}{\beta \langle k \rangle \sum_{i=1}^N i q_i(t)}$, 显然, $\xi(t)$

和 $\xi^*(t)$ 在任意时刻都严格小于 1。

当且仅当 $\lambda \langle k \rangle \xi(t) - 1 > 0$ 时, $m(t+\Delta t) - m(t) > 0$, 即在均值意义下, 疾病会大规模流行, 此时有 $(\lambda \langle k \rangle)^{-1} < \xi(t) < 1$, 这意味着传播阈值 $\lambda_c > \langle k \rangle^{-1}$ 。当 $\lambda_c < \langle k \rangle^{-1}$ 时, 在均值意义下疾病会自动灭绝, 不需要人为控制。这个阈值和平均场方法的结果是一致的。

取定规模 $N=10^4$ 、平均度 $\langle k \rangle=6$ 、重连概率 $p=0.1$ 的小世界网络, 初始患病节点密度 $\alpha=0.1$, 传染参数 $\beta=\gamma=0.1$, 计算平均场方法的 $i(\infty)$ 和随机方法的分布均值。表 1 给出的对比结果显示稳态时患病节点密度在理论上比较接近, 说明随机 SIS 模型不但和平均场方法一样能够较好地描述传播稳态, 而且能够刻画传播过程的波动性, 这是随机 SIS 模型相对于平均场方法的最大优势。

表 1 平均场方法与随机过程方法结果对比

λ	平均场方法	Markov 链方法
1.00	0.833 33	0.833 13
0.25	0.333 33	0.331 30
0.20	0.166 67	0.161 27

3 结束语

本文的 Markov 链方法是对实际传播过程的一种较强的近似, 在分析时只使用平均度这一个网络结构信息, 因此,

(下转第 124 页)