

无标度网络上的传播动力学^{*}

王 延 郑志刚[†]

(北京师范大学物理系, 北京 100875)

(2008 年 9 月 19 日收到, 2008 年 11 月 22 日收到修改稿)

介绍了无标度网络上的传播动力学, 在 susceptible-infected-susceptible (SIS) 模型的基础上考察了一般情况下无标度网络中疾病爆发的临界点问题, 得出了关于临界点一般性的表达式. 得到的结果在特殊情况下分别退化为已有的一些经典结论. 同时分别讨论了这些情况的建模意义和可靠性.

关键词: 无标度网络, 传播动力学, susceptible-infected-susceptible 模型, 临界点

PACC: 0545

通过将研究对象的个体抽象为网络的一个节点, 而把这些对象之间的相互作用抽象作边, 可以把复杂系统抽象为复杂网络. 这为研究复杂系统提供了新的思路, 一般来说, 不同的复杂系统千差万别, 每个系统都需要进行针对性地研究, 然而近年来大量的实证数据表明许多不同领域的复杂系统所对应的复杂网络具有许多统计性质上的相似之处. 这些统计性质包括度分布, 聚集系数, 以及最短路径等等. 其中最基本刻画网络特征的是网络的度分布 $P(k)$, 它表征任意一个节点与其余 k 个节点连接的概率. 基于度分布, 人们发现了一类广泛存在于科技, 生物, 社会各领域中的网络结构, 即无标度网络, 它的度分布满足 $P(k) \propto k^{-\gamma}$ ($2 < \gamma \leq 3$)^[1]. 这种网络具有很强的异质性, 少量的节点占据了大量的边, 因此这些少数节点对于网络的性质会有很大的影响, 实际上人们意识到网络的拓扑结构是极大影响其上进行的动力学行为的.

经常讨论的网络上的动力学包括同步行为^[2], 博弈过程^[3], 级联动力学^[4]和传播动力学^[5-7], 等等. 本文主要探讨在无标度网络上的传播动力学. 人们对于无标度网络上传播动力学的关注源自文献^[6]的提出, 该文作者得到无标度网络上病毒传播阈值为零的结论, 这个结果颠覆了以往基于随机网的传播阈值分析(传播阈值为一非零有限值), 因此该文引起广泛兴趣. 许多后续的研究可以大致分为如

下三个方向: 1) 研究无标度网络上不同的传播模型^[5-8]; 2) 从另一个侧面提出针对网络的免疫策略^[9-12]; 3) 网络性质本身对阈值的影响, 比如网络存在度度关联的影响^[13], 网络加权的影响^[14], 尺寸效应的影响^[15, 16], 以及团簇结构的影响^[17]等. 进一步了解网络上的传播动力学的基本情况可以参考文献^[18]. 本文研究了一般情况下随机无标度网络上的传播动力学, 之前的一些经典结论可以看成本文结果的特例.

网络上的传播动力学基于不同语境下的解释, 可以理解为计算机网络中的病毒传播, 人类社会病毒或者消息的传播等. 统一起见, 以下我们采取病毒传播的解释. 最基本的传播模型包括 SI, SIS, SIR 三种, 其中 S 代表未被感染, 而 I 相应表示已经感染. 在 SI 模型中, 所有的节点就只有这两种状态, 而且节点状态一旦为 I 就不会再改变; 而在后两种模型中人们考虑了更为实际的情况. SIS 模型的物理解释是经历一段时间后感染者痊愈, 并且在以后可以再次被感染, 这可以对应于流感等疾病, 在 SIR 模型中 R 依据不同的解释代表免疫或者死亡, 比如可以分别对应于麻疹和黑死病, 总之, 状态为 R 的节点将等效的从网络中移除.

给定一个由 N 个节点组成的网络, 每个节点有一定的免疫能力, 研究在什么样的情况下疾病会爆发, 即网络中被感染的点的个数为 $O(N)$ 是人们关

^{*} 国家自然科学基金(批准号: 70431002, 10575010, 10875011), 国家重点基础研究发展计划(973)项目(批准号: 2007CB814805)和教育部博士点基金(批准号: 20060027009)资助的课题.

[†] 通讯联系人, E-mail: zgzheng@bnu.edu.cn

心的问题.为了考察这个问题,我们先具体地介绍一下所采用的 SIS 模型上病毒传播的规则(在研究无标度网络临界点问题的情况下,SIR 模型也有类似的结果^[8]):1)在每一时刻,每个节点的状态不是 S 就是 I;2)所有 t 时刻状态为 I 的节点在 $t+1$ 时刻状态变为 S;3)在 t 时刻状态为 S 的节点 i 如果与一个状态为 I 的节点相连,则在 $t+1$ 时刻它的状态以概率 $\lambda(k_i)$ 变为 I, k_i 表示节点 i 的度, $\lambda(k_i)$ 实际表征其免疫能力或者说染病概率的大小.我们接着定义 ρ_k 为系统达到稳态后状态为 I 的节点中度为 k 的节点占有网络节点的平均比例,而定义 $\rho = \sum \rho_k$ 即为网络中所有感染节点的平均比例.基于平均场的思想,我们可以得到描述这个过程的主方程^[6,9]

$$\partial_t \rho_k(t) = -\rho_k(t) + \lambda(k)k \times [1 - \rho_k(t)]\Theta(\rho(t)), \quad (1)$$

其中湮没项为 $-\rho_k$, 后面的为产生项.产生项有如下几个因素决定:1) $\lambda(k)$, 表示一个节点与一个感染的邻居接触染病的概率;2) k , 节点的度,即邻居的个数;3) $[1 - \rho_k(t)]$, 度为 k 的节点中未被感染的比例;以及4) $\Theta(\rho(t))$, 它表示某节点给定一边指向状态为 I 的邻居的概率.考虑如果网络是均匀混合的,即没有明显的度关联,在这种情况下我们可以写出:

$$\Theta = \sum_k \frac{k' P(k') \rho_{k'}(t)}{k}, \quad (2)$$

其中 k 代表网络的平均度,形式上显然 $\Theta \in [0, 1]$, 极值分别对应 $\rho_k = 0$, 网络没有被感染,和 $\rho_k = 1$, 网络所有节点都被感染的情况.前者在这个问题中是平庸的情况,而后者根据规则是达不到的.我们接着考虑系统达到稳态后的情况,即令 $\partial_t \rho_k(t) = 0$, 则此时 ρ_k 和 Θ 都是时间无关的状态量.我们得到

$$\rho_k = \frac{\lambda(k)k\Theta}{1 + \lambda(k)k\Theta}. \quad (3)$$

将 Θ 的表达式代入上式,我们可以得到一个关于 Θ 的自洽方程,即

$$\Theta = \sum_k \frac{P(k)}{k} \frac{\lambda(k)k^2\Theta}{1 + \lambda(k)k\Theta}. \quad (4)$$

如果存在非平庸的稳态解,即 $\rho_k \neq 0$, 则要求上式在 $\Theta \in [0, 1]$ 的区间内有解.令

$$\Omega = \sum_k \frac{P(k)}{k} \frac{\lambda(k)k^2\Theta}{1 + \lambda(k)k\Theta},$$

则不难发现 Ω 是关于 Θ 的单调增函数,但是增幅越

来越小,即

$$\frac{d\Omega}{d\Theta} > 0, \frac{d^2\Omega}{d\Theta^2} < 0,$$

所以有非平庸解的必要条件相当于是

$$\left. \frac{d\Omega}{d\Theta} \right|_{\Theta=0} \geq \left. \frac{d\Theta}{d\Omega} \right|_{\Theta=0}. \quad (5)$$

上式中等号意味着临界点,标志着疾病爆发的开始,将具体表达式代入,我们有

$$\frac{\lambda(k)k^2}{k} = 1. \quad (6)$$

这是一个普遍成立的式子,对于不同的网络结构和不同的节点染病概率 $\lambda(k)$ 的形式,我们可以得到很多结果.

第一种情况 假定 $\lambda(k) = \lambda$ 为常数,即所有节点的染病概率都是一样的,这时候我们发现^[13]

$$\lambda_c = \frac{k}{k^2}. \quad (7)$$

正是这个结果说明了病毒在无标度网络上传播和在其他类型的网络传播的不同.对于无标度网络,我们将它的最小度记为 m ,并且令

$$A = \int_m^\infty P(k) \lambda dk = \int_m^\infty k^{-\gamma} dk = \frac{m^{1-\gamma}}{\gamma-1},$$

则 k^α 的平均值可以写成

$$k^\alpha = \frac{1}{A} \int_m^\infty P(k) k^\alpha dk = \frac{\gamma-1}{\gamma-\alpha-1} m^\alpha,$$

其中 $\alpha < \gamma-1$. 而当 $\alpha \geq \gamma-1$ 时 k^α 的值发散,这也是无标度网络名称的由来.有了这些结果,我们可以看出对于人们感兴趣的那一类无标度网络, $\lambda_c = 0$, 这由 k^2 的发散直接得出.对于别的类型的网络,比如规则网,随机网或者小世界网络, k^2 是个有限值,所以 $\lambda_c > 0$. 可见,在被感染这个意义上,无标度网络比其他类型的网络要脆弱得多.从结构上看,一旦那一小部分具有很高连接度的点被感染,整个网络就很容易被影响.另一方面,这也说明不同的网络结构确实会带来迥然不同的动力学后果.当然,这样的分析建立在所有节点的染病概率都是一致的前提下,实际上这是一种很粗略的近似,很明显真实生活中不同节点的抵抗能力是不一样的.而我们可以预期以上的这个结果在染病概率呈高斯分布且具有较小的半高宽时是可以相信的.假如 $\lambda(k)$ 的分布满足

$$P(\lambda) = \frac{1}{\sqrt{2\pi\sigma}} \exp\left[-\frac{(\lambda - \lambda_0)^2}{2\sigma^2}\right],$$

我们可以等效地将 $\lambda(k)$ 写成 $\lambda(k) = \lambda_0 + \sigma\xi(k)$, 其

中 $\xi(k) \propto N(0,1)$ 是个标准的高斯随机数,则由(6)式得

$$\lambda_0 = \frac{k - \sigma \xi(k)k^2}{k^2}.$$

由于涨落项 $\sigma \xi(k)k^2$ 的存在,如果 σ 不是一个很小的量,则 λ_0 的值对于不同的网络实现将出现很大的涨落,涨落项成为主导项,我们因此不能肯定必有 $\lambda_0 = 0$. 所以(7)式的成立要求在染病概率这一点上所有节点都是差不多的,然而这点并不总能满足,因此需要对模型进行推广,我们将在下面讨论.

第二种情况 $\lambda(k) \propto k^\alpha (\alpha < 0)$ ^[14] 在 [14] 文中作者考虑了 BA 网络 ($\gamma = 3$) 上的情况,而我们将考虑更为一般的无标度网络 ($2 < \gamma \leq 3$). 此时每个节点的染病概率不再是一样的,而是加权的,这种加权方式是与度相关的. 我们认为度越大的点其染病概率越小,这是考虑到实际上的一些情况. 比如一般来说年轻人的社交圈子比幼儿或者老年人要更广,每天接触到的人会更多,而他们的身体素质是最好的,抵抗力也是最强的;又比如政府或者大型网站所使用的计算机服务器的连接数要远远大于家用电脑的连接数,相应的,这些电脑的病毒防护措施也远比个人所能采取的要好得多. 所以这样的加权考虑是有一定道理的. 在这种情况下,每个节点的染病概率都不一样,所以 λ 不再是一个有效的考察病毒爆发的量,我们需要寻找一个新的量来刻画. 首先我们来写出 $\lambda(k)$ 的确切表达式:

$$\lambda(k) = \frac{\lambda k^\alpha}{k^\alpha}.$$

当然,为了保证染病概率在 (0,1) 之间以及度越大染病概率越小的单调变化关系,我们需要保证两个条件: 1) $\lambda < \frac{\gamma-1}{\gamma-\alpha-1}$ 和 2) $m > \frac{\gamma-\alpha-3}{\gamma-2}$, 实际上这两个条件是不难满足的,所以在数学上这样的一种形式也是可取的. 现在我们可以看到 λ 是一个合适的参量,它表示网络节点的平均染病概率. 随着它的变化,网络将出现从未被感染到被感染的相变. 实际上我们不难看出

$$\lambda_c = \frac{k k^\alpha}{k^{\alpha+2}}. \quad (8)$$

注意到 $\alpha < 0$, 则 $k^{\alpha+2}$ 和 k^α 都是有限值,所以 $\lambda_c > 0$. 我们得出的这个式子适用于一般的无标度网络,在 BA 无标度网络化简就得到以前的结果^[11]. 在图 1 中我们也采用了 BA 无标度网络,网络由 10000 个节点构成,其最小度 $m = 4$ 结果由 100 次网

络平均得来. 我们给出了不同 α 取值的情况下 ρ 随着 λ 的变化情况,我们同时给出临界点理论和数值模拟结果的比较,图中箭头上所标数字即为根据主方程得到的理论值 ($\alpha = 0$ 时 $\lambda_c = \lambda_c = 0$, 理论值在图上没有标出),可以看出两者相符得很好. 由此可见在 k^α 这种加权模式下,疾病的流行需要有个阈值,这与不加权的理想情况相比存在定性的区别. 但同样这个结果也是和网络的异质性极大相关的,如果那些小部分具有较高连接度的点被保护,则网络整体相对会安全很多. 这从一个侧面提出了一种免疫策略,即优先保护度大的节点. 但是这种加权方式存在两个问题: 一是出于免疫控制的角度,它的耗费太大,越是度大的节点越要投入更多; 二则更为本质一些,作为一种假设,度大的节点染病概率更低时可以接受的,但是否一定在度趋近于无穷的时候染病概率趋近于零是一个很关键的问题,这点我们将在下面进行讨论.

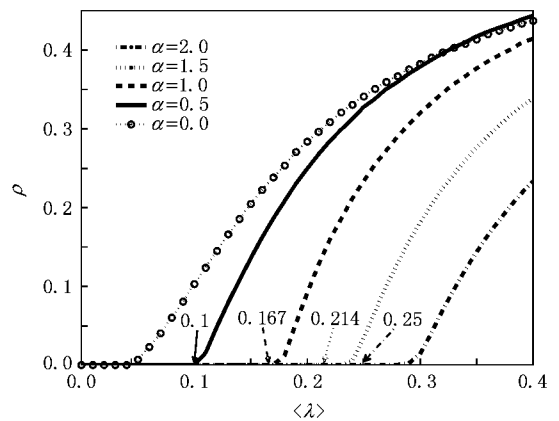


图 1 在不同的加权形式下节点平均染病概率与定态平均染病节点比例的关系图(箭头所示为(8)式给出的临界平均染病概率,理论与数值结果符合得很好)

第三种情况 也是最一般的情况,我们认为 $\lambda(k)$ 就是一般的非线性函数,出于如上第二种情况的考虑可以要求它关于 k 单调减. 则我们将它展开为 $\lambda(k) = \sum_{\beta} \lambda_{\beta} k^{\beta} (\beta \leq 0)$, 在这个情况下,我们发现临界点有

$$\lambda_0 \frac{k^2}{k} + \sum_{\beta} \lambda_{\beta} \frac{k^{\beta+2}}{k} = 1, \quad (9)$$

其中 $\sum_{\beta} \lambda_{\beta}$ 表示求和不对 0 进行,显然等式的成立要求 $\lambda_0 = 0$, 即疾病传播的阈值依然为 0. 举个例子,假如 $\lambda(k) \propto \lambda_0 + \lambda_{-1} k^{-1}$, 我们依然发现 $\lambda(k_1) > \lambda(k_2)$ 如果 $k_2 > k_1$, 但在此时,决定疾病流行与

否的关键在于常数项 λ_0 , 而不在于是否度大的节点是否比度小的节点更难以被感染. 即便 λ_0 是一个并不大的数它就足以定性影响传播动力学的行为, 就像在第二种情况中所讨论的, 即便 α 是个不大的负数, 它依然可以引起与第一种情况下截然不同的动力学后果.

综合如上结果, 在理论上处理网络上的传播问题时已经有了比较成熟的理论, 通过建立适当的主方程可以很好地研究传播的动力学临界行为, 我们得到了描述临界点的一般公式(6式), 并且考察了几种特殊的情况, 还原了以前一些经典的结果. 通过这样一种综合的考察我们发现在描述传播问题时怎样的模型是真正反映实际的依然需要认真甄别. 这

样的问题很微妙, 对于热力学极限的情况, 不同的理论导致不同的结果, 极细微的参数差别可能导致定性上的差别. 然而实际上不论理论如何, 总是在实际生活中发现有限的流行病传播阈值, 因为真实情况有更多的因素左右, 没有热力学极限, 尺寸效应^[12]会体现出来, 而且有不可忽略的涨落, 那么复杂网络的结果有什么意义呢? 复杂网络的研究至少可以从理论上使人们确信异质性极高的无标度网络是在抵抗流行病方面是脆弱的, 而有效的控制方法至少在目前看来是代价高昂的. 在不改变网络结构的前提下如何低耗有效提高网络的抵抗能力是值得进一步探讨的问题.

-
- [1] Barabási A L , Albert R 2002 *Rev. Mod. Phys.* **74** 47
- [2] Barahona M , Pecora L M 2002 *Phys. Rev. Lett.* **89** 054101
Liu F , Guan Z H , Wang H 2008 *Chin. Phys. B* **17** 2405
- [3] Pacheco J M , Traulsen A , Nowak M A 2006 *Phys. Rev. Lett.* **97** 258103
- [4] Motter A E 2004 *Phys. Rev. Lett.* **93** 098701
- [5] Barthélemy M , Barrat A , Pastor-Satorras R , Vespignani A 2004 *Phys. Rev. Lett.* **92** 178701
- [6] Pastor-Satorras R , Vespignani A 2001 *Phys. Rev. Lett.* **86** 3200
- [7] Newman M E J 2002 *Phys. Rev. E* **66** 016128
- [8] Moreno Y , Pastor-Satorras R , Vespignani A 2002 *Eur. Phys. J. B* **26** 521
- [9] Pastor-Satorras R , Vespignani A 2002 *Phys. Rev. E* **65** 036104
- [10] Xu D , Li X , Wang X F 2007 *Acta Phys. Sin.* **56** 1315 (in Chinese)
[许 丹 , 李 翔 , 汪小帆 2007 物理学报 **56** 1315]
- [11] Pei W D , Chen Z Q , Yuan Z Z 2008 *Chin. Phys. B* **17** 373
- [12] Wang L , Yan J R , Zhang J G , Liu Z R 2007 *Chin. Phys.* **16** 2498
- [13] Boguñá M , Pastor-Satorras R , Vespignani A 2003 *Phys. Rev. Lett.* **90** 028701
- [14] Dezsö Z , Barabási A L 2002 *Phys. Rev. E* **65** 055103 (R)
- [15] May R M , Lloyd A L 2001 *Phys. Rev. E* **64** 066112
- [16] Li Y , Liu Y , Shan X M , Ren Y , Jiao J , Qiu B 2005 *Chin. Phys.* **14** 2153
- [17] Serrano M Á , Boguñá M 2006 *Phys. Rev. Lett.* **97** 088701
- [18] Dorogovtsev S N , A V Goltsev 2008 *Rev. Mod. Phys.* **80** 1275

Spreading dynamics on scale-free networks^{*}

Wang Yan Zheng Zhi-Gang[†]

(*Department of Physics ,Beijing Normal University ,Beijing 100875 ,China*)

(Received 19 September 2008 ; revised manuscript received 22 November 2008)

Abstract

The spreading dynamics on scale-free networks is considered. Based on the SIS (susceptible-infected-susceptible) model , we investigate the critical point at which the epidemic breaks out in a very general setting. Some previous results , each of which is analyzed , can be seen as the special cases of our result , and the modeling reliability in each case is also discussed.

Keywords : scale-free networks , spreading dynamics , susceptible-infected-susceptible model , critical point

PACC : 0545

^{*} Project supported by the National Natural Science Foundation of China (Grant Nos. 70431002 , 10575010 and 10875011) , the 973 Project (Grant No. 2007CB814805) , and the Foundation of Doctoral Training (Grant No. 20060027009) .

[†] Corresponding author. E-mail : zgzhen@bnu.edu.cn