复杂网络上的流行病传播

阮中远^{*}

浙江工业大学计算机学院, 杭州 310023

*联系人, E-mail: zyruan@zjut.edu.cn

收稿日期: 2019-04-15; 接受日期: 2019-05-10; 网络出版日期: 2019-09-02

国家自然科学基金青年项目(编号: 11605154)资助

摘要 流行病的传播伴随着整个人类社会的发展, 但历史上对于它的研究却进展缓慢. 直到最近几十年复杂网络这一学科的崛起, 流行病传播的研究才得到了迅猛发展. 本文简要介绍了近20年来复杂网络上流行病传播的一些重要研究成果, 主要关注四方面的内容: 基于接触网络模型的流行病传播; 基于集合种群模型的流行病传播; 流行病传播的预测; 流行病传播的控制. 这些研究成果从复杂网络的角度加深了对现实世界中流行病传播的理解, 使得对真实流行病传播的预测和控制成为可能.

关键词 复杂网络, 流行病传播, 接触网络模型, 集合种群模型

PACS: 64.60.aq, 87.19.X-, 87.23.Ge

1 引言

自古以来, 人类社会便遭受着各种各样的流行病的侵扰, 其中有些疾病对人类的发展历程产生了深远影响^[1,2]. 例如, 14世纪中期的黑死病, 曾席卷整个欧洲, 夺走当时欧洲约1/3人口的性命, 对整个西方文明造成了巨大冲击; 20世纪初, 一种称之为“西班牙流行性感冒”的传染病在全球范围内蔓延, 感染了近6亿人, 导致几千万人死亡, 成为第一次世界大战提前结束的原因之一; 我国东汉末年的“伤寒”大流行, 造成数千万人死亡, 引发了剧烈的社会动荡, 导致了著名的黄巾农民起义. 这样的例子举不胜数. 据相关统计, 自公元前7世纪至公元20世纪, 仅在中国, 较大规模的瘟疫就达七百多次^[2]. 正如美国历史学家麦克尼尔^[3]在《瘟

疫与人》一书中所论述: 传染病在历史上出现的年代早于人类, 未来也将会和人类长期地共存, 而且它也一定会和从前一样, 是人类历史中的一项基本参数以及决定因子.

尽管各种流行病一直与人类相伴, 但人们关于它们的认识却进展得相当缓慢. 最早关于流行病传播的研究开始于18世纪60年代, 瑞士著名数学家Bernoulli通过数学方法研究了天花接种的有效性^[1]. 19世纪中期, 英国流行病学家Farr采用经验分析的方法用正态曲线拟合了1837–1839年间英格兰和威尔士两个地区感染天花的死亡人数数据, 该描述性的方法后来被苏格兰医学统计学家Brownlee所进一步发展^[1]. 到20世纪初期, Hamer和Ross采用定量的方法研究了麻疹疾病的传播. 这一工作成为历史上采用精确数学语言来

引用格式: 阮中远. 复杂网络上的流行病传播. 中国科学: 物理学 力学 天文学, 2020, 50: 010507

Ruan Z Y. Epidemic spreading in complex networks (in Chinese). Sci Sin-Phys Mech Astron, 2020, 50: 010507, doi: [10.1360/SSPMA-2019-0128](https://doi.org/10.1360/SSPMA-2019-0128)

研究流行病传播的先例. 在研究中他们提出了著名的“质量作用原理”^[1], 即假定感染的净传播率(单位时间内新感染了多少人)正比于易感人群密度与感染人群密度的乘积. 这一原理后来成为数学流行病学的主要概念之一. 1927年, 受到之前工作的启发, Kermack和Mckendrick首次建立了著名的“仓室模型”, 得到了流行病传播的阈值理论^[4]: 在一个充满易感者的群体中引入少量的感染者并不会引发流行病的暴发, 除非该群体中初始易感者的数量高于某一临界值. 之后几十年几乎所有的相关研究都引用了仓室模型的核心假设, 即假定个体在空间中是均匀混合的.

直到20世纪末小世界网和无标度网的发现^[5,6], 为流行病传播的研究迎来了新的时代^[7]. 人们发现, 很多真实的系统都可以用复杂网络的框架加以描述^[8-12]. 这些系统是由大量的组元构成, 其中组元可表示为节点, 组元之间的相互作用可看作连边, 如社交网络、交通网络、生物网络、万维网等^[12]. 研究表明, 对于真实的社会系统, 人群之间的接触模式并不是均匀混合的, 而展现出了一些复杂的网络特性^[13-15], 如小世界性、无标度性等. 从此, 基于复杂网络的流行病传播的研究开始成为流行病传播这一领域的主流方向^[7]. 复杂网络的理论框架不仅把握了真实社会系统的本质特点, 而且具有良好的数学性能, 即能够进行严密的数学计算. 利用统计物理等学科的思想方法, 研究者提出了许多不同的理论方法, 如平均场理论、点对近似、主方程、动态消息传递等方法来研究复杂网络上的流行病传播行为^[16-23], 并且取得了许多深刻的结果, 它们甚至能为公共卫生机构政策的制定提供一些重要的理论依据.

总体来说, 现阶段关于流行病传播的研究主要集中在考虑个体之间的相互作用结构、接触模式以及群体的迁移运动模式这3个方面^[7]. 这些都是影响流行病传播的决定性因素, 并且需要在复杂网络的框架下加以考虑. 从这个角度来看, 复杂网络上流行病传播的研究大致可分为以下两类: (1) 基于接触网络模型^[17,24-28]. 在该模型中, 节点代表个体, 连边代表个体之间的相互联系(物理接触), 染病个体可以通过连边感染健康个体. 这一模型主要强调个体之间的相互作用结构对流行病传播的影响, 而忽略了个体的运动性. 其中一个标志性的成果是发现了无标度网络上的流行病传播阈值为零^[17]. (2) 基于集合种群模型^[29-37]. 在这

个模型中, 节点代表区域(或种群), 连边表示区域之间的人口迁移, 理论上一个区域可占据无穷多的个体. 这方面的一些典型工作结合了交通网络(如全球航空网络等)来研究流行病在种群层面上的传播情况, 特别在全球性的流行病传播的研究中取得了令人惊叹的成果^[38-41].

由上述论述可知, 真实社会系统中流行病的传播可以从不同的尺度上进行观察, 在相应的尺度上需要考虑对应的网络结构. 例如, 基于个体的流行病传播研究需要考虑社会关系网络, 而基于集合种群的流行病传播研究需要考虑种群的迁移网络等. 研究表明, 网络结构对于流行病传播的动力学有着巨大的影响^[17,26-28]. 另一方面, 流行病的传播反过来也会影响网络结构. 例如, 流行病暴发期间, 具有防范意识的健康个体可能会主动断开与染病个体之间的联系^[42-44], 同时也可能会改变运动迁移路线^[30,45]等. 事实上, 网络结构与流行病传播的动力学之间存在着相互作用, 这使得复杂网络上的流行病传播问题变得异常有趣和富有挑战性, 这也是目前这一领域的主要研究热点之一^[7]. 本文对基于复杂网络的流行病传播这方面的一些较重要的研究成果进行了介绍.

2 流行病的经典模型

不同的疾病具有不同的特点. 例如流行性感冒、淋病这类传染病, 患者在恢复之后仍存在被感染的风险; 而像麻疹、百日咳这类疾病, 患者一经治愈便对它们终身免疫. 这些不同类型的传染病需要用不同的数学模型加以描述. 模型主要是根据个体的染病状态对人群进行分类. 本节主要介绍了两个经典的流行病模型^[7]: SIS(易感态-感染态-易感态)模型和SIR(易感态-感染态-移除态)模型, 以便为本文后面部分的内容提供相应的基础.

在SIS模型中, 个体处于两种可能的状态之一: 易感态(Susceptible)和感染态(Infected). 用 i, j 表示不同的个体, S, I 分别表示个体处于易感态和感染态, 则其反应动力学过程如下:

$$\begin{aligned} S(i) + I(j) &\xrightarrow{\beta} I(i) + I(j), \\ I(i) &\xrightarrow{\mu} S(i). \end{aligned}$$

上式包含了两个转换过程: (1) 感染过程: 一个感

染态个体碰到一个易感态个体时会以概率 β 去感染它;
(2) 恢复过程: 一个感染态个体会以概率 μ 恢复为易感态. 这是一个自发恢复过程, 一般与网络结构无关, 并且恢复之后的个体可能会重新参与到感染过程. 因此, 在一定条件下, 系统中始终会存在感染态的个体.

在SIR模型中, 个体处于三种可能的状态之一: 易感态(Susceptible)、感染态(Infected)和移除态(Removed). 用 i, j 表示不同的个体, S, I, R 分别表示个体处于易感态、感染态和移除态, 则其反应动力学过程如下:

$$S(i) + I(j) \xrightarrow{\beta} I(i) + I(j),$$

$$I(i) \xrightarrow{\mu} R(i).$$

这里与SIS模型不同的地方在于恢复过程——恢复为 R 态的个体不再参与任何反应过程. 显然, 在SIR模型中, 稳态后处于感染态的个体数量总是趋于0.

根据上述两个模型的描述过程可知, 个体之间的相互作用模式在病毒传播过程中扮演着至关重要的角色. 一种简单且基本的想法是假设人群是均匀混合的, 即每个个体随机地与其他个体发生相互作用. 这意味着一个易感个体被感染的可能性正比于人群中处于感染态的个体数量. 引入感染率 α 来衡量个体被感染的可能性(或风险性)^[7]. 其数学定义为: $\alpha = \beta' N_I / N$, 其中 N_I 表示感染人数, N 表示总人数, β' 取决于特定疾病的传染性以及个体间的接触模式. 在某些情况下 β' 可表示为如下两个参数的乘积: $\beta' = \beta k$, 这里 β 为单次接触的感染概率, k 表示接触频次^[7]. 当接触频次取不同的值, α 具有不同的形式. 例如, 对于单次接触的情形, 即 $k=1$, 感染率 $\alpha = \beta N_I / N$, 与系统的尺寸成反比; 对于全接触的情形, 即 $k=N$, 感染率 $\alpha = \beta N_I$, 与系统尺寸无关.

根据平均场假设, 处于同一类状态的个体可近似认为是等同的. 此时, 描述系统动力学行为的关键变量为处于每类状态的个体数量 N_u (或密度 $\rho_u = N_u / N$), 其中 u 表示个体所有可能的状态. 利用感染力的概念(这里考虑单次接触的情形), 可以得到SIS或SIR模型的动力学演化方程^[7]:

$$\begin{aligned} d\rho_I(t) / dt &= \beta \rho_I(t) \rho_S(t) - \mu \rho_I(t), \\ d\rho_S(t) / dt &= -\beta \rho_I(t) \rho_S(t) + \chi \rho_I(t). \end{aligned} \quad (1)$$

当 $\chi = \mu$, 对应于SIS模型; $\chi = 0$ 对应于SIR模型. 假定系统中总人数是守恒的, 则有 $\rho_I + \rho_S = 1$ (SIS模型)

或 $\rho_I + \rho_S + \rho_R = 1$ (SIR模型). 这样, 就得到一套描述系统动力学行为的完备微分方程组. 假设初始时刻 $\rho_I(0) \approx 0$, 则在初始阶段(1)式可线性化为如下微分方程:

$$d\rho_I(t) / dt \approx (\beta - \mu) \rho_I(t), \quad (2)$$

其解为

$$\rho_I(t) \approx \rho_I(0) e^{(\beta - \mu)t}. \quad (3)$$

由此可知, 当 $\beta - \mu > 0$ 或 $\beta / \mu > 1$, 系统中感染态人数随时间的增加呈指数增长. 定义有效传染率(或基本再生数) $\lambda = \beta / \mu$, 它表示一个感染者在平均患病周期 $1/\mu$ 内, 在一个充满易感者的人群中能够传染的平均个体数. 这是经典流行病传播理论中最重要的概念之一. 根据以上分析, 可以得到均匀混合条件下流行病的传播阈值. 以SIS模型为例, 当 $\lambda > 1$, 稳态时系统中处于感染态的人数密度维持在一个大于0的有限值; 反之, 稳态时感染态的人数密度为0 (图1).

除了上面介绍的这两种流行病模型, 在文献当中还会经常遇到一些其他常见的模型, 例如SI模型、SEIR模型、SIRS模型等. SI模型相比于SIS模型少了 $I \rightarrow S$ 的恢复过程, 意味着个体一旦被感染就会始终处在感染态^[46,47]. 该双态模型还被广泛地应用于社会传播领域. SEIR模型是在SIR模型中加入了潜伏态 E . 处在 E 态的个体已被感染, 但由于处在潜伏期, 还不具备感染其他个体的能力. 这个模型比较适合于流感类的疾病^[48]. SIRS模型考虑的是感染个体在治愈

图1 (网络版彩图)SIS模型的相图, 其中 $\lambda = \beta / \mu$, $\lambda_c = 1$, ρ 为感染态人数密度. 引自文献^[7]

Figure 1 (Color online) Phase diagram for SIS model. Here, $\lambda = \beta / \mu$, $\lambda_c = 1$, and ρ is the infected density. Cited from ref. ^[7].

后暂时获得免疫能力的情形, 此时处在移除态的个体在一段时间后会重新转变为易感态(失去免疫效果). 这一模型可以用来解释季节性的病毒传播^[49,50].

3 基于接触网络模型的流行病传播

从前面的论述可知, 传统流行病传播的研究是基于个体均匀混合的假设. 然而真实情况下群体的接触模式展现出了复杂的网络结构. 归功于计算机科学领域的飞速发展, 科学家获取了海量的数据来研究个体之间的交互模式. 他们发现, 真实的网络结构呈现出丰富的特性, 如小世界性、无标度性、社区性、度度关联性、时序性和层次性等. 为了准确理解和描述复杂网络上的疾病传播动力学过程, 科学家提出了许多不同的理论分析方法. 通过这些方法, 可以精确求解疾病的暴发阈值、暴发规模以及感染人数随时间的演化行为特征等. 值得注意的是, 由于SIS模型和SIR模型(或其他模型)的感染机制不同, 所涉及的理论方法也有所不同. 以下主要以SIS模型为例来介绍一些比较重要的研究成果.

首先, 考虑最简单的均匀随机网络的情形. 此时, 网络中每个节点的度近似等于平均值 $\langle k \rangle$. 为了方便起见, 令恢复概率 $\mu=1$. 根据平均场理论(所有处于感染态的节点都是等价的), 可得到感染节点密度 $\rho(t)$ 随时间的演化方程^[17]:

$$\frac{d\rho(t)}{dt} = -\rho(t) + \beta \langle k \rangle \rho(t)(1 - \rho(t)), \quad (4)$$

其中等式右边第一项表示感染节点恢复的速率, 第二项表示易感节点被感染邻居传染的速率. 达到稳态时有 $d\rho(t)/dt = 0$, 可得到

$$\rho(\infty)[-1 + \beta \langle k \rangle (1 - \rho(\infty))] = 0. \quad (5)$$

显然, 方程(5)有两个实数根, 第一个是平凡解 $\rho(\infty) = 0$, 第二个是非平凡解 $\rho(\infty) > 0$, 其存在的必要条件为有效传染率 β 大于疾病暴发的阈值 $\beta_c = 1 / \langle k \rangle$. 疾病暴发阈值 β_c 将系统的稳态解划分为两个区域(类似于图1, 不同的地方在于此时 $\lambda_c = 1 / \langle k \rangle$): 当 $\beta > \beta_c$, 系统处于活跃态, 最终疾病将大规模暴发; 当 $\beta < \beta_c$, 系统处于吸收态, 疾病最终将消亡.

对于异质度分布的随机网络, 上述方法过于粗糙.

一种改进的方法是基于度的平均场理论(或异质平均场理论). 这一理论假定具有相同度的节点在结构特征和动力学特征上是完全等价的. 因此, 用 $\rho_k(t)$ 来替换 $\rho(t)$, 其中 $\rho_k(t)$ 表示 t 时刻度为 k 的所有节点处于感染态的比例(也表示一个度为 k 的节点处于感染态的概率). 设网络是度无关联的, 则 $\rho_k(t)$ 随时间的演化方程可写为^[16,17]

$$\frac{d\rho_k(t)}{dt} = -\rho_k(t) + \beta k [1 - \rho_k(t)] \Theta(t), \quad (6)$$

其中,

$$\Theta(t) = \frac{1}{\langle k \rangle} \sum_k P(k) k \rho_k(t) \quad (7)$$

表示任选一条连边, 该连边指向感染节点的概率. 这里 $P(k)$ 指网络度分布, $\langle k \rangle$ 为平均度. 方程(6)的等式右边第一项表示度为 k 的感染节点的恢复速率, 第二项表示度为 k 的易感节点被感染态邻居所感染的速率. 设初始时刻 $\rho_k(t) \approx 0$, 则在初始阶段方程(6)可线性化为^[16]

$$\frac{d\rho}{dt} = A\rho, \quad (8)$$

其中 $\rho(t) = [\rho_1(t), \rho_2(t), \dots, \rho_{k_{\max}}(t)]^T$, $k_{\max}$ 为最大度值, 矩阵 A (又称雅可比矩阵)的元素为

$$A_{kk'} = \beta \frac{kk' P(k')}{\langle k \rangle} - \delta_{kk'}, \quad (9)$$

其中 $\delta_{kk'}$ 为狄拉克函数, 当 $k=k'$, $\delta_{kk'} = 1$; 否则 $\delta_{kk'} = 0$. 当矩阵 A 的最大特征值 $\beta \langle k^2 \rangle / \langle k \rangle - 1$ 大于0时($\langle k^2 \rangle$ 表示度分布的二阶矩), $\rho(t) = \sum_k P(k) \rho_k(t)$ 将随时间呈指数增长, 即流行病最终会大规模暴发. 由此可得到疾病的暴发阈值 $\beta_c = \langle k \rangle / \langle k^2 \rangle$. 对于度分布满足幂律分布 $P(k) \sim k^{-\gamma}$ 且 $\gamma \leq 3$ 的异质网络, 在热力学极限下(即网络大小趋于无穷时), $\langle k^2 \rangle$ 是发散的, 意味着疾病的暴发阈值趋于0; 而当 $\gamma > 3$, 疾病传播具有一个有限的暴发阈值. 这与均匀随机网络的情形有所不同.

值得注意的是, 对于一些更为真实的异质性网络, 如度度关联网络、权重无标度网络和(异质)多层网络等, 同样可以用异质平均场理论进行近似求解^[51-53]. 特

别地, Liu等人^[54]研究了活动驱动网络上的SIS传播过程, 这是一类典型的时变异质网络. 在这个模型中, 每个节点*i*对应一个活跃率 a_i (a_i 满足幂律分布), 表示单位时间内该节点与其他节点建立联系的概率. 在每个时间步*t*, 网络都是从*N*个孤立的节点开始构建; 每个节点*i*以概率 $a_i \Delta t$ 变为活跃态, 并伸出*m*条连边连向任意*m*个其他节点; 在下个时间步*t* + Δt , 所有连边被删除, 然后重复前面的步骤. 定义 I_a^t 和 S_a^t 分别表示*t*时刻活跃率为*a*的感染态节点数和易感态节点数, 根据平均场近似可得到如下动力学方程^[54]:

$$\begin{aligned} I_a^{t+1} = I_a^t - \mu I_a^t + \beta m (N_a - I_a^t) a \int da' I_a^t / N \\ + \beta m (N_a - I_a^t) \int da' I_a^t a' / N, \end{aligned} \quad (10)$$

其中 N_a 表示活跃率为*a*的节点数. 式(10)右边第三项表示在活跃率为*a*的易感节点中, 处于活跃状态并连向感染节点从而导致被感染的节点数量; 最后一项表示活跃率为*a*的易感节点被处于活跃态的感染节点连接从而被感染的数量. 通过求解上述方程, 可得到流行病的暴发阈值为

$$(\beta / \mu)_c = \frac{2\langle a \rangle}{\langle a \rangle + \sqrt{\langle a^2 \rangle}}.$$

上式表明, 活动驱动网络上流行病是否暴发只与节点的平均活跃程度及其涨落有关.

异质平均场方法一般适用于具有无穷大拓扑维度的网络, 但是它忽略了两个重要的因素^[16]: (1) 该方法并没有完全描述接触网络的结构, 而是用度分布来粗略进行描述, 忽视了节点之间的淬火连接信息. (2) 忽略了节点之间的动力学关联性. 这些简化使得这种方法可以较准确地描述退火网络上的传播动力学过程, 但并不适用于淬火网络. 对于第一种缺陷, 一种改进的方法是基于个体的平均场理论(或淬火平均场理论). 淬火平均场方法的核心思想是用邻接矩阵*C*来完整表示网络结构: 当节点*i*和*j*相连时, 邻接矩阵元素 $C_{ij} = 1$; 否则 $C_{ij} = 0$. 令 $\rho_i(t)$ 表示*t*时刻节点*i*处于感染态的概率, 则可以得到 $\rho_i(t)$ 随时间的演化方程^[16,19]:

$$\frac{d\rho_i(t)}{dt} = -\rho_i(t) + \beta [1 - \rho_i(t)] \sum_{j=1}^N C_{ij} \rho_j(t), \quad (11)$$

其中*N*表示网络的大小. 方程(11)等式右边第一项表示节点*i*的恢复速率, 第二项表示节点*i*被感染邻居所传染

的速率. 在*t*时刻, 网络中处于感染态的节点比例为 $\rho(t) = 1/N \sum_{i=1}^N \rho_i(t)$. 假设初始时刻只有极少数节点处于感染态, 即 $\rho_i(0) \approx 0$, 则在初始阶段方程(11)可做线性化近似:

$$\frac{d\rho}{dt} = -\rho + \beta C \rho, \quad (12)$$

其中 $\rho(t) = [\rho_1(t), \rho_2(t), \dots, \rho_N(t)]^T$. 类似于方程(8)的求解方法(求矩阵的最大特征值), 可以得到流行病的暴发阈值 $\beta_c = 1/\Lambda_c$, 其中 Λ_c 为邻接矩阵*C*的最大特征值. 由此可知, 对于无度关联的无标度网络, 当幂律指数 $\gamma \leq 2.5$ 时, 疾病的暴发阈值为 $\beta_c \propto \langle k \rangle / \langle k^2 \rangle$, 这个结果与异质平均场的结果一致; 当 $\gamma > 2.5$, 疾病的暴发阈值为 $\beta_c \propto 1/\sqrt{k_{\max}}$, 在热力学极限下该值同样趋于0, 这一结论显然与异质平均场的理论结果不同.

淬火平均场理论得到了比异质平均场理论更准确的结果, 但与模拟结果仍然存在偏差, 有时候偏差甚至很大, 这是因为它仍然忽略了网络中节点状态之间的动力学关联性, 使得易感节点被感染的概率被高估^[16]. 一种较为精确地刻画节点状态之间的动力学关联性的方法为点对近似方法^[55]. 点对近似方法不仅考虑节点状态的变化, 同时也考虑了点对状态随时间的变化. 引入参量 $\varphi_{x_i}(t)$ 表示*t*时刻节点*i*处于状态 x_i 的概率, $\varphi_{x_i x_j}(t)$ 表示*t*时刻节点*i*和*j*分别处于状态 x_i 和 x_j 的概率, 以此类推, 这里 $x_i, x_j \in \{S, I\}$. 令 $\rho_i(t) = \varphi_{I_i}(t)$ 表示*t*时刻节点*i*处于感染态的概率. 不难得出, 对于任意一对节点以下关系成立^[16,56]:

$$\begin{aligned} \varphi_{I I_j} + \varphi_{S I_j} &= \rho_j(t), \\ \varphi_{I I_j} + \varphi_{I S_j} &= \rho_i(t), \\ \varphi_{S S_j} + \varphi_{S I_j} &= 1 - \rho_i(t), \\ \varphi_{S S_j} + \varphi_{I S_j} &= 1 - \rho_j(t). \end{aligned}$$

类似于方程(11)的思路, 节点*i*态演化方程可写为^[16,56]

$$\frac{d\rho_i(t)}{dt} = -\rho_i(t) + \beta \sum_{j=1}^N C_{ij} \varphi_{S I_j}(t), \quad (13)$$

其中 C_{ij} 为邻接矩阵*C*的元素. 方程(13)右边第一项表示节点*i*的恢复速率, 第二项表示节点*i*被邻居节点感染的

概率. 如果忽略节点之间的动力学关联性, 则有 $\varphi_{S_i I_j}(t) = (1 - \rho_i)\rho_j$, 此时方程(13)退化为方程(11).

方程(13)中 $\varphi_{S_i I_j}(t)$ 随时间的演化方程为

$$\frac{d\varphi_{S_i I_j}(t)}{dt} = -\varphi_{S_i I_j}(t) - \beta \varphi_{S_i I_j}(t) - \beta \sum_{l \in N(i) \setminus j} \varphi_{I_l S_i I_j}(t) + \varphi_{I_l I_j}(t) + \beta \sum_{l \in N(j) \setminus i} \varphi_{S_i S_j I_l}(t), \quad (14)$$

等式右边的负数项表示 $\varphi_{S_i I_j}(t)$ 的减少, 它来自三方面的贡献: (1) 节点 j 从感染态恢复为易感态(第一项); (2) 节点 i 以一定概率被邻居节点 j 感染(第二项); (3) 节点 i 以一定概率被除 j 之外的其他邻居节点所感染(第三项). 这里 $\varphi_{I_l S_i I_j}(t)$ 表示节点 l , i 和 j 分别处于感染态、易感态和感染态的概率, $\pi(i)$ 表示节点 i 的所有邻居集合. 等式右边的正数项表示 $\varphi_{S_i I_j}(t)$ 的增加, 它们来自两方面的贡献: (1) 节点 i 从感染态恢复为易感态(第四项); (2) 节点 j 以一定概率被除 i 之外的其他邻居节点所感染(第五项).

利用点对近似, 即仅考虑点对之间的动力学相关性, 三元态 $\varphi_{x_i x_j x_l}(t)$ 可近似表示为^[56]

$$\varphi_{x_i x_j x_l}(t) \approx \frac{\varphi_{x_i x_j}(t) \varphi_{x_j x_l}(t)}{\varphi_{x_j}(t)}. \quad (15)$$

将(15)式代入方程(14)并结合方程(13), 可得到一套完整地描述SIS动力学过程的微分方程组. 设初始条件为 $\varphi_{I_i I_j}(0) \rightarrow 0$ 及 $\varphi_{S_i S_j}(0) \rightarrow 1$ (即假设只有少数个体处于感染态), 沿用前面求解方程的思路, 将方程在初始阶段作线性化近似, 最终流行病爆发的阈值归结为求雅可比矩阵 $\mathbf{L}$ 的最大本征值问题, 其中 $\mathbf{L}$ 矩阵的元素为

$$L_{ij} = -\left(1 + \frac{\beta^2 k_i}{2\beta + 2}\right) \delta_{ij} + \frac{\beta(2 + \beta)}{2\beta + 2} C_{ij}. \quad (16)$$

模拟结果表明, 点对近似方法比异质平均场方法或淬火平均场方法对于流行病爆发的刻画更为准确, 但是其代价是计算复杂度更高.

点对近似方法可以运用到许多不同的接触网络模型上. 其中一个典型的应用是自适应网络上的流行病传播^[42]. Gross等人^[42]在经典SIS模型的基础上考虑了这样一种情形: 对于随机网络中的SI连边(表示一个易

感态节点和一个感染态节点相连), 处于易感态的个体会以一定的概率 w 断开与 I 态节点的连边, 并重新连接到另一个随机选择的 S 态节点(不允许双重连接或自连). 利用点对近似的方法, 可以得到网络中感染态人数、 II (表示两个感染态节点相连)连边数、 SS (表示两个易感态节点相连)连边数随时间的演化方程. 通过求解这组微分方程, 可得到一些有趣的结果. 如图2所示, 当不存在重连, 流行病的暴发是一个连续相变过程, 其临界阈值为 p_c . 当考虑重连机制时, 会出现两个阈值. 这两个阈值对应于非连续的一级相变, 在它们中间是双稳态, 形成一个迟滞回线. 其主要原因在于自适应重连倾向于孤立感染节点, 使得流行病的传播得到抑制; 另一方面, 重连会导致易感态的人群紧密连接在一起, 形成一个高度连接的易感簇团, 这个簇团对应于较低的流行病传播阈值——一旦有病毒入侵, 疾病很可能将大规模扩散. 因此, 自适应重连的局部效应会抑制流行病的传播, 而其拓扑效应又会促进疾病的传播. 这两个相反的作用会导致一系列复杂的动力学现象.

图2 不同重连概率下感染密度 i^* 随感染概率 p 的变化关系. (a)–(d) 分别对应 $w=0, 0.04, 0.2, 0.6$ 的情形. 其中实线代表理论结果, 圆圈代表模拟结果. 其他参数为网络大小 $N=10^5$, 边数 $K=10^6$, 恢复率 $r=0.002$. 引自文献^[42]

Figure 2 Infected density as a function of infection probability under different values of rewiring probability. (a)–(d) correspond to $w=0, 0.04, 0.2, 0.6$, respectively. The solid lines and circles correspond to the theoretical results and simulation results, respectively. Other parameters are $N=10^5$, $K=10^6$, $r=0.002$. Cited from ref. ^[42].

此外, 关于接触网络上的SIR模型(或其他一些流行病模型)的研究也相当丰富. 我们介绍的这些分析方法同样可以应用到这些流行病模型上来. 特别地, 针对SIR模型, 研究者还提出了边渗流、消息传递、边划分等多种方法, 深入地揭示了网络上SIR传播模型的内在机理. 文献[16]对这些基于接触网络模型的理论分析方法有比较全面的介绍.

4 基于集合种群模型的流行病传播

基于接触网络的流行病传播模型忽略了个体的运动性. 然而在实际情况下, 个体运动在流行病传播过程中扮演了举足轻重的角色. 在古代, 一些疾病的传染源往往是由商旅部队等从一个地方带到另一个地方, 从而引发在空间上的蔓延. 在现代, 随着飞机、火车等交通工具的发明, 个体可以毫无阻碍地进行长距离运动, 使得疾病的传播变得更为便利. 为了刻画个体的运动性, 研究者提出了另一种模型, 称为集合种群模型(或反应-扩散模型)^[29]. 在这个模型中, 整个人群按地理位置被划分为多个亚种群, 不同的亚种群之间通过人口的相互迁移发生联系, 从而形成一个亚种群网络, 如图3所示. 在这里, 网络中的每个节点代表一个区域, 理论上一个区域可以被无穷多的粒子所占据. 在同一个区域内的粒子之间会发生相互反应(按SIS感染机制或SIR感染机制等), 同时这些粒子会以一定的规则在不同的节点之间运动.

首先我们介绍集合种群网上的SIS模型^[29]. 这里考虑一种最简单的运动方式, 即个体在网络上做随机扩散. 考虑一集合种群网络, 其大小为 V , 度分布为 $P(k)$, 平均度为 $\langle k \rangle = \sum_k P(k)k$. 设节点 i 上的粒子数为 $n(i)$, 则网络中的总粒子数为 $N = \sum_{i=1}^V n(i)$. 对于任意度分布的网络, 一种较好的近似方法是基于度的平均场理论, 即认为网络中度为 k 的节点在统计意义上是等价的(包括其上的粒子数). 由此, 引入参量 ρ_k 表示度为 k 的节点上的平均粒子数:

$$\rho_k = \frac{1}{V_k} \sum_{i \in N(k)} n(i), \quad (17)$$

其中, $V_k = VP(k)$ 表示度为 k 的节点数, $N(k)$ 表示度为 k 的所有节点的集合. ρ_k 随时间的演化方程可写为

图3 (网络版彩图)集合种群模型示意图. (a) 从宏观水平来说, 该模型是由许多亚种群组成; (b) 从微观水平来说, 每个亚种群都包含了大量个体. 改编自文献[7]

Figure 3 (Color online) Schematic of meta-population model. (a) From macroscopic level, the model is composed of many subpopulations. (b) From microscopic level, each subpopulation is composed of many particles. Adapted from ref. [7].

$$\frac{d\rho_k}{dt} = -D_k \rho_k + k \sum_{k'} P(k' | k) D_{k'k} \rho_{k'}, \quad (18)$$

其中 D_k 表示度为 k 的节点上粒子的逃逸概率, $D_{k'k}$ 表示粒子从度为 k' 的节点沿连边 $k'k$ 跑到度为 k 的节点的概率, $\rho_{k'}$ 表示度为 k' 的节点上的粒子数, $P(k' | k)$ 是条件概率, 表示度为 k 的节点指向度为 k' 的节点的概率. 方程(18)等式右边第一项表示单位时间内从度为 k 的节点跑出去的粒子数, 第二项表示单位时间内从邻居节点跑进来的粒子数.

假设每个粒子的逃逸概率为一常数 r , 再加上粒子随机扩散的假定, 可得到 $D_{k,k} = r/k'$. 另外, 对于度无关联网络有 $P(k' | k) = k'P(k')/\langle k \rangle$. 将这些表达式代入方程(18), 可得到稳态时度为 k 的节点上的平均粒子数:

$$\rho_k = \frac{k}{\langle k \rangle} \rho, \quad (19)$$

其中 $\rho = N/V$ 表示每个节点上的平均粒子数. (19)式表明, 节点的度越大, 其被随机游走的粒子所访问到

的概率也越大.

在此框架下, 可引入SIS模型的感染机制. 在SIS模型中, 人群被分为两类: 感染者和易感者. 假设节点*i*上的感染态和易感态粒子数分别为 $n_i(i)$ 和 $n_s(i)$, 显然有 $n(i) = n_s(i) + n_i(i)$. 根据前面的思路, 引入两个新的参量:

$$\rho_{I,k} = \frac{1}{V_k} \sum_{i \in N(k)} n_i(i), \quad \rho_{S,k} = \frac{1}{V_k} \sum_{i \in N(k)} n_s(i) \quad (20)$$

分别表示度为*k*的节点上的平均感染态粒子数和平均易感态粒子数. 结合(17)式可知, $\rho_k = \rho_{I,k} + \rho_{S,k}$. 假设感染态粒子和易感态粒子的逃逸概率分别为 D_I 和 D_S . 由此可得到 $\rho_{I,k}$ 和 $\rho_{S,k}$ 随时间的演化方程:

$$\begin{aligned} d\rho_{I,k} / dt &= -\rho_{I,k} + (1 - D_I)[(1 - \mu)\rho_{I,k} + \beta\Gamma_k] \\ &\quad + k \sum_{k'} P(k' | k) \frac{1}{k'} D_I [(1 - \mu)\rho_{I,k'} + \beta\Gamma_{k'}], \\ d\rho_{S,k} / dt &= -\rho_{S,k} + (1 - D_S)[\mu\rho_{I,k} + \rho_{S,k} - \beta\Gamma_k] \\ &\quad + k \sum_{k'} P(k' | k) \frac{1}{k'} D_S [\mu\rho_{I,k'} + \rho_{S,k'} - \beta\Gamma_{k'}], \end{aligned} \quad (21)$$

其中 β 为感染概率, μ 为恢复概率, Γ_k 表示反应核, 具体形式与粒子间的接触模式有关. $\beta\Gamma_k$ 表示单位时间内度为*k*的节点上平均新感染了粒子数. 方程(21)第一式的前两项表示单位时间内度为*k*的节点上粒子经反应、扩散后导致的感染态粒子数的变化, 它源自该节点本身的贡献. 等式右边的最后一项表示从邻居节点跑进来的感染态粒子数.

为了求解方程(21), 需要知道 Γ_k 的具体表达式. 我们主要考虑两种典型的形式: $\Gamma_k = \rho_{I,k} \rho_{S,k}$ (类型I)和 $\Gamma_k = \rho_{I,k} \rho_{S,k} / \rho_k$ (类型II). 类型I表示节点上的每个粒子都可以跟同一节点上的其他粒子进行接触; 类型II表示每个粒子只能接触其中有限个其他粒子.

考虑一种简单的情形, 即 $D_I = D_S = 1$, 同时假设网络是度无关的. 此时, 对于第一种反应类型(类型I)有稳态解:

$$\rho_I = \rho - \frac{\mu \langle k \rangle^2}{\beta \langle k^2 \rangle}.$$

由此可知, 稳态时系统中存在感染态粒子的条件是 $\rho > \mu \langle k \rangle^2 / \beta \langle k^2 \rangle$. 在这种情况下, 种群网络的拓扑结构对流行病的传播有着重要的影响. 特别地, 对于

度分布满足 $P(k) \sim k^{-\gamma}$ 且 $\gamma \leq 3$ 的异质网络, 在热力学极限下 $\langle k \rangle^2$ 趋于无穷, 此时流行病的传播阈值 $\rho_c = 0$. 图4(a)展示了相应的模拟结果. 这里采用的是度分布满足 $P(k) \sim k^{-3}$ 的UCM网络(无关联配置模型), 其中 $\beta = 0.1$, $\mu = 0.2$. 当网络大小 $V = 1000$ 和 $V = 10000$ 时, $\langle k \rangle^2 / \langle k^2 \rangle$ 的值分别为0.7和0.49, 由此可得 ρ_c 分别为1.4和0.98. 模拟结果很好地验证了理论预期.

对于第二种反应类型(类型II), 有稳态解:

$$\rho_I = \rho(1 - \mu / \beta).$$

此时流行病的暴发条件为 $\beta / \mu > 1$, 与网络的拓扑结构无关. 图4(b)显示了相应的模拟结果, 与理论解符合得较好. 之所以这样, 是因为在第二种反应类型中, 节点*i*上粒子间的相互反应项中多了一个缩放因子 $n(i)$, 即除以节点上的总粒子数. 这样一来, 无论节点上有多少个粒子, 对于每个粒子来说都没有区别, 因为它只能接触有限个其他粒子(如只能接触一个粒子). 此时流行病的传播主要取决于 β 和 μ 之间的平衡, 而与节点的属性无关.

接下来考虑集合种群网上的SIR模型^[57]. 对于这种情形, 人们更关心种群网络中的每个节点上是否会暴发流行病. 考虑一个亚种群(节点)*j*, 其上的总粒子数为 $N_j = S_j + I_j + R_j$, 其中 S_j , I_j , R_j 分别表示易感态、感染态

图4 (网络版彩图)(a) 第一种反应类型下相对感染密度 ρ_I/ρ 随粒子密度 ρ 的变化, 其中方框和圆圈分别代表网络大小为 $V=1000$ 和 $V=10000$. (b) 第二种反应类型下相对感染密度 ρ_I/ρ 随 β/μ 的变化关系

Figure 4 (Color online) (a) The relative infected density ρ_I/ρ as a function of particle density ρ for type-I reaction process, where the squares and circles represent the network size $V=1000$ and $V=10000$, respectively. (b) The relative infected density ρ_I/ρ as a function of β/μ for type-II reaction process.

和免疫态的粒子数. 设每个易感态的粒子以比率 $\beta I_i/N_i$ 被感染, 同时处于感染态的粒子以比率 μ 进入免疫态. 由于每个节点内的粒子是均匀混合的, 因此可用基本再生数 $\lambda = \beta/\mu$ 来判断该节点上是否发生流行病的暴发($\lambda > 1$ 时能暴发). 假设每个粒子在网络上随机游走, 其跳跃率为 p , 且每个亚种群的基本再生数 $\lambda > 1$ (即满足流行病暴发的条件), 则可得到如下两个极端情况: 当 $p=0$, 若初始时刻某一节点上暴发疾病, 则该病毒会被限制在该节点上, 网络中的其他节点均不会受到影响; 当 $p \rightarrow 1$, 粒子不断地从一个节点跳到另一个节点, 此时整个系统等价于一个均匀混合的单一种群. 由于 $\lambda > 1$, 则流行病必然会在整个种群网络上暴发. 因此, 可能存在一个阈值 p_c , 当 $p > p_c$ 时, 在热力学极限下, 网络上最终会有有限比例的节点暴发流行病; 反之, 该比例趋于零.

为了刻画这种情况, 引入一个新的参量 λ^* 表示一个感染的亚种群能够感染的平均新亚种群数. 与粒子均匀混合情形下定义的 λ 类似, λ^* 在种群层面上的传播能力: 当 $\lambda^* > 1$, 流行病能在种群网络上全局传播. 为了求得 λ^* 的值, 考虑一度分布为 $P(k)$ 的种群网络. 设粒子运动达到稳态后度为 k 的节点上的粒子数为 ρ_k . 假设病毒在种群层面上的传播类似于树状分支过程, 定义第0代度为 k 且被感染的亚种群数为 D_k^0 (即初始时刻内部暴发流行病的节点数), 第1代度为 k 且被感染的亚种群数为 D_k^1 (每个感染亚种群会向其邻居节点输送染病粒子), 以此类推, 第 n 代被感染的度为 k 的亚种群数为 D_k^n . 由此可得到如下递归方程:

$$D_k^n = \sum_{k'} D_{k'}^{n-1} (k' - 1) P(k | k') \left[1 - \frac{D_k^{n-1}}{N_k} \right] \left[1 - \left(\frac{1}{\lambda} \right)^{\omega_{k,k}} \right], \quad (22)$$

其中 $P(k | k')$ 为条件概率, $(k' - 1)P(k | k')$ 表示 $(k' - 1)$ 个邻居节点中度为 k 的平均节点数, $(1 - D_k^{n-1}/N_k)$ 表示这些节点在第 $n-1$ 代尚未被感染的概率, $[1 - (1/\lambda)^{\omega_{k,k}}]$ 表示节点在接受 $\omega_{k,k}$ 个感染粒子后可能暴发的概率. 最后一项中 $p_{\text{ex}} \equiv 1/\lambda$ 表示一个感染态个体入侵一个易感种群后病毒消亡的概率, 所以当入侵感染态粒子数为 $\omega_{k,k}$ 时, 不发生流行病暴发的概率为 $(1/\lambda)^{\omega_{k,k}}$.

$\omega_{k,k}$ 的表达式可根据以下计算获得. 考虑一粒子

数为 ρ_k 的种群. 在流行病暴发期间, 总感染态粒子数记为 $\alpha \rho_k$, 其中 α 与具体的流行病模型及相应的参数有关. 粒子处在感染态的平均时间为 μ^{-1} , 在这段时间内, 感染态的粒子以比率 $D_{kk'}$ 跳到度为 k' 的邻居节点上. 由此可得到局部流行病暴发期间, 从 k 跳跃到 k' 的平均染病粒子数: $\omega_{k'k} = D_{kk'} \alpha \rho_k \mu^{-1}$. 假设粒子在网络上随机游走, 则有 $D_{kk'} = p/k$, 其中 p 为粒子的逃逸概率. 另外, 由式(19)可知稳态时 $\rho_k = kp / \langle k \rangle$. 最终可得到 $\omega_{k'k} = p \alpha \rho \mu^{-1} / \langle k \rangle$.

假设我们只关心每个亚种群处在疾病暴发阈值附近的情形, 即 $\lambda - 1 \ll 1$; 同时假设初始阶段 $D_k^{n-1}/N_k \ll 1$, 且网络是度无关联的. 根据这些条件, 方程(22)可重新写为

$$\Theta^n = (\lambda - 1) \frac{\langle k^2 \rangle - \langle k \rangle}{\langle k \rangle^2} \frac{p \alpha \rho}{\mu} \Theta^{n-1}, \quad (23)$$

其中 $\Theta^n = \sum_k D_k^n (k' - 1)$. 由(23)式可知, 流行病暴发的条件为

$$\lambda^* = (\lambda - 1) \frac{\langle k^2 \rangle - \langle k \rangle}{\langle k \rangle^2} \frac{p \alpha \rho}{\mu} > 1. \quad (24)$$

根据上述不等式, 可以计算粒子的临界跳跃比率 p_c : 当 $p > p_c$, 流行病能在集合种群网上全局传播. 对于SIR模型, 当 $\lambda \rightarrow 1$, 有 $\alpha \approx 2(\lambda - 1)/\lambda^2$ [7,57]. 将此代入(24)式可得

$$p_c \rho = \frac{\langle k \rangle^2}{\langle k^2 \rangle - \langle k \rangle} \frac{\mu \lambda^2}{2(\lambda - 1)^2}. \quad (25)$$

(25)式表明, λ 越趋近于1, 粒子的临界跳跃比率越大. 模拟结果验证了这一结论, 如图5所示.

前面的研究都是假定粒子在种群网上做随机运动. 而在实际情况中, 个体的运动往往具有目的性 [35,58,59]. 基于此, Tang等人 [35]研究了集合种群网上粒子按最短路径旅行的情形. 它与粒子随机扩散有着本质的区别: 对于随机扩散, 粒子当前所在节点的所有邻居节点都有相同的概率在下一步被该粒子访问; 而对于目的旅行, 只有处在最短路径上的那些节点才能被访问. 他们的结果表明, 相比于随机扩散, 粒子按最短路径旅行会加速流行病的传播. 另外的一些实证研

图5 系统中感染粒子数比例(Global Attack Rate)随基本再生数 λ 和粒子逃逸概率 p 的关系. 改编自文献[57]

Figure 5 The global attack rate as a function of the basic reproduction number λ and the jumping probability p . Adapted from ref. [57].

究指出, 真实社会系统中人类的空间运动具有高度的周期性和可预测性^[60–63]. 例如, 绝大多数人经常在少数几个地点之间来回运动, 偶尔会访问一些新的地点. 为了反映这种运动的记忆性^[64–68], Balcan和Vespignani^[64]提出了一种新的粒子运动模式, 即来回往返运动来研究集合种群网上的流行病传播. 他们的模型很简单: 对于一个集合种群网络, 假设节点 i 上的每个粒子会以比率 σ_{ij} 迁移到 i 的某一邻居节点 j 上; 同时这些发生位移的粒子会以比率 τ^{-1} (τ 为平均停留时间)返回到它们的初始节点. 这样, 节点 i 上的粒子数 N_i 可分为两部分: N_{ii} 和 N_{ji} , 分别表示来自节点 i 且处在 i 节点的粒子数和来自节点 j 但处于 i 节点的粒子数. 在来回往返运动模式中, 参数 σ_{ij} 和 τ 决定了整个传播动力学行为. 例如, 当 $\sigma_{ij} \rightarrow 0$, 病毒只能被囚困在局域节点, 无法引发全局暴发; 另外, 如果 $\tau \rightarrow 0$, 感染态粒子待在邻居节点的时间极短, 没有充分的时间与易感人群接触, 病毒同样很难在其中传播开来. 以SIR模型为例, 可以引用前面(随机扩散)介绍的思路进行解析求解. 结果表明, 只有当粒子的迁移比率和返回比率都达到一定的阈值, 流行病才可能在整个种群网络上传播开来. 相比于随机扩散的情形, 该阈值以一种更复杂的方式依赖于网络的拓扑结构(牵涉到度分布的1–4阶矩).

集合种群模型还在其他方面得到了不同的扩展. 例如, 考虑到现实生活中某些场合(如音乐会、购物商场等)会聚集大量的人群, 导致局域密度极大. Tang等人^[69]将此类比于统计物理中的凝聚现象, 研究了粒子发生零区域相互作用凝聚时对流行病传播的影响. Ruan等人^[34]在反应-扩散模型中加入了旅途过程中的

感染反应. 在经典的模型中, 个体从一个节点跳到另一个邻居节点是在瞬间完成的, 忽略了旅途过程带来的效应. 而经验事实告诉我们, 旅途中的感染反应也是极有可能的. 为此, 他们提出了“旅行-感染”模型, 发现旅途感染确实会加剧疾病的传播. 此外, 他们还研究了双层网络(公路网和高铁网)上的集合种群模型, 发现站点的存在会极大地促进流行病的传播^[33]. Wang等人^[70]考虑了地域相关的异质性接触模式对流行病传播的影响, 发现这种异质性接触会显著降低流行病的传播阈值.

在集合种群模型中, 一般只考虑种群层面的网络结构, 而认为在每个亚种群内部, 粒子之间是均匀混合的. 然而这只是一粗略的近似. 事实上, 在区域内部, 人群也是一个接触社会网络. Mata等人^[71]扩展了这一模型, 他们将每个亚种群看成是一个均质的网络, 粒子的扩散过程看作是不同网络(相互连接的亚种群)之间粒子的随机交换. Ruan等人^[72]进一步发展了这一想法, 他们提出了一个简单的模型, 这个模型同时考虑了亚种群内部的网络结构以及粒子在不同亚种群之间做来回往返运动的特点. 作为第一步尝试, 他们仅考虑两个相互耦合的亚种群—— A 和 B , 其中每个亚种群都是一个ER随机网络. 网络 $A(B)$ 中的所有节点(代表粒子)之间以概率 $p_A(p_B)$ 随机相连. 为了方便起见, 假设这两个网络的大小是相等的, 即 $N_A=N_B=N$, 但连边密度不一样, 即 $p_A \neq p_B$. 这是因为考虑到不同区域由于文化、经济等差异导致人群联系紧密程度不一样. 在两个种群之间引入“潜在”连边(见图6中的红色虚线): 网络 $A(B)$ 中的每个节点以概率 $p_{AB}(p_{BA})$ 与网络 $B(A)$ 中的每个节点相连. 在 $N_A=N_B$ 的假设下, 显然有 $p_{AB}=p_{BA}$. 这些连边表示两个个体之间可能是朋友关系, 但是由于他们处在不同的种群, 无法发生物理上的接触, 因此通常情况下对流行病的传播没有影响. 但当其中一个区域的个体访问另一个区域时, 这些潜在连边就会被激活, 此时它们能传播相应的疾病.

为了刻画粒子在两个种群之间的来回往返运动, 将每个接触网络映射到另一个亚种群, 并使相应的节点为空, 如图6所示. 空节点不代表粒子, 或表示不被粒子所占据. 只有当空节点变成实的, 或被粒子占据了, 它们才有意义. 假设处在 $A(B)$ 网络的粒子(实节点)会以概率 σ 跃迁到它的映射节点位置(空节点, 处在网

图 6 (网络版彩图)模型示意图. 其中蓝色和黄色圆圈分别代表亚种群A和B的居民, 红色曲线代表两个种群之间的潜在连边. 每个种群中的粒子以概率 σ 跳到另一个种群, 同时移位的粒子以速率 τ 返回到它们的初始位置. 引自文献[72]

Figure 6 (Color online) Schematic illustration of the model. The blue and yellow nodes are the residents of subpopulations A and B, respectively. The dotted red lines denote the potential links between sub-populations. Individuals in each subpopulation travel to the other subpopulation with a per capita diffusion rate σ and the displaced individuals return to their initial residence with rate τ . Cited from ref. [72].

络 $B(A)$), 同时这些已经发生位移的粒子会以速率 τ 返回到它们的初始位置. 这里值得注意的是, 粒子在同一时间只能占据一个节点, 一旦它离开初始位置, 原来所在的节点就会变空, 而映射的节点会变实. 这一过程生动地模拟了这样一种情况: 当个体从一个区域旅行到另一个区域时, 他会暂时切断原先的朋友接触网络, 而在新的区域建立起一些新的联系.

通过考虑SIS模型, 可以从数值模拟和理论分析上得到一些有意思的结果. 图7展示了每个亚种群中稳态感染密度随感染概率 β 和粒子跳跃概率 σ 的关系, 其中红线是理论结果. 可以看出, 跳跃概率对流行病的暴发阈值有着非单调的影响: β_c 随着 σ 的增加先增加后减小. 这表明个体旅行对流行病的传播有着一种复杂的作用. 一方面, 通过输入和输出感染粒子, 粒子迁移可以直接导致不同亚种群中感染密度的变化; 另一方面, 粒子的跳跃还会引发网络结构的变化, 例如跳出去的粒子会使原先占据的节点变空, 等价于这些节点暂时从网络中被移除了. 这会进一步影响流行病在亚种群中的传播.

图 7 (网络版彩图)稳态时每个亚种群中的感染密度随感染概率 β 及粒子跳跃概率 σ 的关系. 这里红色实线为理论结果. 参数为 $N=5000$, $p_A=0.002$, $p_B=0.004$, $P_{AB}=0.06$, $\tau=1$ 以及 $\mu=1$, 初始感染密度为0.5%. 引自文献[72]

Figure 7 (Color online) The infected density at the steady state in each subpopulation as a function of infection probability β and traveling rate σ . The red lines are the theoretical results. The parameters are chosen as $N=5000$, $p_A=0.002$, $p_B=0.004$, $P_{AB}=0.06$, $\tau=1$ and $\mu=1$, the initial infected density is 0.5%. Cited from ref. [72].

5 流行病传播的预测

研究流行病传播的一个重要目标是进行预测, 这在现实生活中具有重要意义^[73]. 当流行病暴发时, 及时预测接下来疾病可能到达的地点及时间, 可以让相关机构提前进行部署, 从而有效地控制疾病的肆意传播. 这方面的研究主要集中在全球性的流行病传播^[38-41], 可看作是集合种群模型的一个应用(基于实际数据). 通过结合全球航空网络(这种快捷的长距离运动已成为全球性疾病传播的一种主要渠道)^[73,74], 科学家对一些全球性的流行病进行实时预测已成为可能. Hufnagel等人^[39]首先分析了全球性流行病传播的可预测性. 按照集合种群模型的框架, 他们的模型分为两部分: 局域感染动力学(SIR模型)和全球旅行动力学. 在模型中他们充分考虑了随机性因素的影响. 例如对于感染动力学, 考虑的是随机朗之万方程:

$$\begin{aligned} \frac{ds}{dt} &= -\alpha sj + \frac{1}{\sqrt{N}} \sqrt{\alpha sj} \zeta_1(t), \\ \frac{dj}{dt} &= \alpha sj - \beta j - \frac{1}{\sqrt{N}} \sqrt{\alpha sj} \zeta_1(t) + \frac{1}{\sqrt{N}} \sqrt{\beta j} \zeta_2(t), \end{aligned} \quad (26)$$

其中 s 和 j 分别表示易感态和感染态的人口密度, α 为感染率, β 为恢复率, $\zeta_1(t)$ 和 $\zeta_2(t)$ 分别表示两个独立的高

斯白噪声. 值得注意的是, 当总人口数 $N \rightarrow \infty$, 方程(26)回到经典的确定性方程(1). 而对于有限的 N , 涨落项在流行病暴发初期就变得不可忽略. 对于旅行动力学, 作者分析了全球最大的500个机场的所有民用航班, 由此构建了一个带权重的全球航空网络, 其中连边权重代表线路的乘客容量. 个体在不同区域之间的运动可由转移率矩阵 γ_{ij} 刻画, 它正比于连边的权重. 引入向量 $\chi = \{S_i, I_i, \dots, S_M, I_M\}$, 其中 S_i, I_i 分别表示区域 i 中的易感态和感染态人数. 这一向量定义了系统的随机状态, 通过考虑感染过程和运动过程导致的感染态人数的变化, 可得到关于概率密度函数 $p(\chi; t)$ 的主方程.

值得注意的是, 由于引入了随机因素, 流行病的预测变得很困难. 即使初始条件一样, 由于涨落的影响, 最终的结果可能会很不一样. 但是将上述模型应用到 SARS 传播案例时(SARS 病毒的基本再生数较小, 且相应的感染人数也不多), 预测结果却出乎意料的准确. 通过模拟 Hufnagel 等人^[39]发现, 网络的异质性, 尤其是航空网络上个体转移率 γ_{ij} 分布的高度异质性, 会使得流行病的可预测性变得更好. 之后, Colizza 等人^[38]发现航空网络对流行病传播的预测有着一种更为复杂和微妙的影响. 一方面, 航空网络上节点之间的巨大连接数量导致病毒可以沿大量不同的可能路径进行传播, 使得不同的流行病暴发实现之间可能存在很大的差异, 即便它们具有相同的初始条件. 从这个角度来说, 航空网络的拓扑结构会导致流行病传播的预测性变差. 另一方面, 航空网络上交通流量的分布是高度异质的, 这使得疾病的扩散通道特定化, 从而提高了其可预测性.

最近, Brockmann 和 Helbing^[40]基于全球航空网络提出了一个新的概念——有效距离, 从而使得全球性流行病传播的预测变得更为直观、简洁. 我们知道, 在古代, 流行病在空间上的传播是逐渐蔓延的, 然而随着现代交通工具(如飞机)的出现, 流行病在全世界范围内的传播变得扑朔迷离: 有时候距离暴发源邻近的城市安然无恙, 而远在千里之外的城市却被感染了. 这使得流行病传播的时空斑图似乎毫无规律可循, 如图8下半部分所示.

根据航空网络数据, 定义转移矩阵 $\mathbf{P}_{mn}$ 为从节点 n 到节点 m 的乘客比例: $\mathbf{P}_{mn} = F_{mn} / F_n$, 其中 F_{mn} 表示从节

图8 (网络版彩图)流行病传播的时空图, 其中上半部分是基于有效距离, 下半部分是基于地理距离. 引自文献[40]

Figure 8 (Color online) The spatiotemporal pattern of an epidemic, where the upper part is depicted based on the effective distance, and the lower part is based on the geographic distance. Cited from ref. [40].

点 n 到 m 的(每日)乘客数量, $F_n = \sum_m F_{mn}$. 考虑到尽管疾病可沿多条不同的路径在任意两个节点之间传播, 然而整个感染过程是由最可能的一些轨迹(可通过转移矩阵得到)所决定的. 根据这一假设, 可定义从节点 n 到与之相连的节点 m 的有效距离为

$$d_{mn} = (1 - \log P_{mn}) \geq 1. \quad (27)$$

(27)式表明, 如果从节点 n 到 m 的客流量很小, 则 $n \rightarrow m$ 的有效距离会很大. 值得注意的是, 有效距离是不对称的, 即 $d_{mn} \neq d_{nm}$.

利用这一概念将网络上的节点重新进行排列, 可得到一幅清晰的流行病传播图像: 从初始暴发源开始, 疾病的传播呈现出类似于水波的扩散方式, 如图8上半部分所示. 设疾病到达某一城市的时间为 T_a , 则有

$$T_a = \frac{D_{\text{eff}}(P)}{V_{\text{eff}}(\beta, R_0, \gamma, \varepsilon)}, \quad (28)$$

其中 $V_{\text{eff}}(\beta, R_0, \gamma, \varepsilon)$ 为疾病传播的有效速度, 它由疾病的相关参数决定, 而与网络和初始暴发位置无关; $D_{\text{eff}}(P)$ 为距离源点的有效距离, 它只与网络的结构有关, 可通过 d_{mn} 计算获得. 假设疾病的初始暴发位置在节点 i 上, 则可根据(28)式计算该疾病到达任意节点 j 和 l 的时间之比:

$$\frac{T_a(j/i)}{T_a(l/i)} = \frac{D_{\text{eff}}(j/i)}{D_{\text{eff}}(l/i)}. \quad (29)$$

上式意味着, 当流行病暴发时, 只要知道源点到达某一城市的时间, 就可以根据有效距离(可根据航空网络的数据计算获得)立刻预测出该疾病到达任意其他城市的绝对时间, 哪怕我们对于该疾病本身的性质一无所知.

随着计算机科学技术的发展, 基于大规模数据集的流行病学建模的研究取得了巨大进展(例如前面介绍的这些工作). 基于此, 研究者建议使用这些模型作为全球性流行病传播的实时预测工具. 由此, GLEAM(Global Epidemic and Mobility)计算模型应运而生^[75]. GLEAM的基本框架如图9所示^[76]. 它需要输入三部分数据: 人口数据、人群移动数据和流行病学模型. 中间是基于个体水平的模拟过程, 包括种群内的感染过程和种群间的运动过程. 最后输出的是感染病例数等相关结果. GLEAM计算模型在实际运用中取得了很好的效果^[75]. 例如, 对于2009年H1N1病毒的传播, GLEAM预测在北半球感染的高峰时期是在11月, 而通常流感样病毒流行的典型高峰期是在1月或2月. 事实证明了预测的准确性. 此外, GLEAM还可以用来预测估计一些干预措施带来的影响, 这为相关疾病控制机构提供了宝贵的信息.

6 流行病传播的控制

研究流行病传播的最终目的是为了能够控制它的传播^[77]. 制定相应的疫苗或抗病毒药物是一项极为重要的任务. 然而, 疫苗数量往往是有限的, 如何选择关键的个体对他们进行接种疫苗, 从而最大限度地抑制流行病的传播是一个值得深入思考的问题. 研究发现有效的免疫策略取决于网络结构.

对于均匀随机网络, 考虑SIS模型, 其流行病传播

图 9 (网络版彩图)GLEAM计算模型的框架图, 主要分为三部分: 输入、模型及输出. 引自文献^[76]

Figure 9 (Color online) The framework of GLEAM, which is divided into three parts: the input, the model and the output. Cited from ref. ^[76].

的动力学方程由方程(4)所描述. 此时一种简单有效的方法是进行随机免疫, 即每个节点以相同的概率被选为免疫对象^[78]. 设 g 为网络中被免疫的节点比例. 在这种情况下, 可等价认为方程(4)中的 β 值变为 $\beta(1-g)$. 由此, 可得到稳态时:

$$\begin{aligned} \rho(\infty) &= 0, & \text{if } g > g_c, \\ \rho(\infty) &\sim \beta - \beta_c, & \text{if } g \leq g_c, \end{aligned} \quad (30)$$

其中 $g_c = (\beta - \beta_c) / \beta$, 代表免疫节点比例的临界值. 当免疫节点比例 $g > g_c$, 流行病就无法在网络中暴发.

对于异质随机网络(其流行病传播动力学方程见方程(6)), 若仍然采用随机免疫的方法则效果甚微. 沿用前面的思路, 用 $\beta \rightarrow \beta(1-g)$ 进行替换, 通过求解方程(6)可得到

$$1 - g_c = \frac{1}{\beta} \frac{\langle k \rangle}{\langle k^2 \rangle}. \quad (31)$$

上式意味着, 当 $\langle k^2 \rangle \rightarrow \infty$, $g_c \rightarrow 1$, 即网络中几乎所有的节点都被免疫, 才能抑制流行病的暴发.

对于这种情形, 一种有效的免疫策略是采取基于度值的目标免疫, 即免疫网络中度大的节点^[78]. 首先, 对网络中的节点按度值大小从高到低进行排序, 然后选取 gN (N 为网络大小)个排名靠前的节点进行免疫. 当网络很大时, 可引入阈值上限 $k_t(g)$, 使得满足 $k > k_t(g)$ 的节点均被免疫, 相应的免疫节点比例为 $g = \sum_{k > k_t} P(k)$, 其中 $P(k)$ 为网络的度分布. 由此可知, 对于度小于 k_t 的所有节点的平均度和方差分别为 $\langle k \rangle_t = \sum_m^{k_t} m P(k)$ 和 $\langle k^2 \rangle_t = \sum_m^{k_t} m^2 P(k)$, 其中 m 表示最小度值. 任取一条连边它指向免疫节点的概率为

$$p(g) = \frac{\sum_{k > k_t} k P(k)}{\sum_k k P(k)}. \quad (32)$$

考虑到被免疫的节点及其连边等价于从网络中被移除了, 移除之后的节点满足度分布:

$$P_g(k) = \sum_{q \geq k} P(q) \binom{q}{k} (1 - p(g))^k p(g)^{q-k}. \quad (33)$$

这一新的度分布对应的平均值和方差分别为

$$\langle k \rangle_g = \langle k \rangle_t (1 - p),$$

$$\langle k^2 \rangle_g = \langle k^2 \rangle_t (1-p)^2 + \langle k \rangle_t p(1-p).$$

当不存在免疫节点时, 异质随机网上流行病暴发的阈值为 $\beta_c = \langle k \rangle / \langle k^2 \rangle$. 当网络中 $k > k_t(g)$ 的节点均被免疫时, 则有

$$\frac{\langle k^2 \rangle_{g_c}}{\langle k \rangle_{g_c}} \equiv \frac{\langle k^2 \rangle_t}{\langle k \rangle_t} [1 - p(g_c)] + p(g_c) = \beta^{-1}. \quad (34)$$

上式确定了免疫节点比例的临界值. 特别地, 对于 BA 网络, 通过近似计算可得到

$$g_c \simeq \exp(-2 / m\beta). \quad (35)$$

以上结果表明, 基于度值的目标免疫对于 BA (或无标度) 网络十分有效, 其免疫阈值在大部分的 β 取值范围内都是很小的, 如图 10 所示.

然而, 目标免疫在实际应用中存在很大的局限性, 因为这一策略要求知道每个节点的度值, 即网络的全局信息, 这在真实情况中是很难获取的. 为了克服这一困难, Cohen 等人^[79]提出了著名的“熟人免疫”策略: 首先从网络中随机选取一部分节点放入集合 A , 对于集合 A 中的每个节点, 随机选取该节点的一个邻居节点, 将它们放入集合 B , 然后将 B 中的所有节点都进行免疫. 这一策略只需要知道网络的局部信息就能达到比随机免疫更好的效果. 这是因为, 在网络中任取一条连边它指向度大的节点的概率更大. 但是熟人免疫策略的效果并不是最优的, 为此, Gallos 等人^[80]改进了这一方法: 首先随机选取一部分个体, 然后通过询问每个个体其邻居节点中度值可能比它自己大或大于某一给定阈值的节点, 并将其进行免疫. 实验发现, 对于具有胖尾分布的网络, 该方法具有极好的免疫效果.

对于社区网络, 基于度值的目标免疫策略和熟人免疫策略可能都会失效. 此时, 一种有效的方法是基于节点介数的目标免疫策略. 社区网络往往由多个不同模块组成, 每个模块内部节点之间的连接稠密, 而不同模块之间的连接比较稀疏^[81]. 通常连接两个不同模块的连边的两端节点被称为桥节点, 它们一般具有很高的介数值^[82]. 对这些节点进行免疫可以有效地阻止疾病从一个社区扩散到另一个社区. 另一种方法是基于随机行走的节点中心性^[83], 它定义为一个随机行走者在任意两节点 m 和 n 之间随机运动时经过某一节点的频次. 实验结果表明, 基于随机行走的中心性策略

图 10 BA 网络上目标免疫(方框)与随机免疫(圆圈)的效果. 这里 ρ_g/ρ_0 表示相对感染密度, g 表示免疫比例. 改编自文献 [78]

Figure 10 The immunization effects of targeted and uniform immunization in the BA network, where ρ_g/ρ_0 is the reduced infected density, and g is the fraction of immune nodes. Adapted from ref. [78].

在绝大多数情况下比基于介数的策略效果要更好.

以上这两种方法同样需要知道网络的全局信息. Salathé 等人^[84]针对社区网络提出了一种基于局域信息的免疫策略(类似于熟人免疫策略)——Community-Bridge-Find (CBF). 该算法利用邻居节点的结构信息, 通过自避免随机游走的方法来识别社区网络中的桥节点, 然后对其进行免疫. 与熟人免疫策略相比, CBF 策略在绝大多数情况下表现得更好, 但在社区结构很强时, 熟人免疫策略的效果要更好. 这是因为当网络的社区性很强时, 病毒很可能被限制在一个局域社区内. 在这种情况下, CBF 免疫策略并不是一个有效的方法, 因为它主要是用来寻找桥节点. 相反, 熟人免疫策略可以找到该社区内度大的节点并对其进行免疫, 从而有效地抑制疾病在该区域内的传播. CBF 免疫策略算法的一个改进版本是 Bridge-Hub-Detector (BHD)^[85], 它同时考虑了桥节点的度异质性. 在一些真实的社区网络中, BHD 算法的免疫效果往往更理想.

上述这些免疫策略的思想还被应用到许多不同的网络结构上; 同时, 一些新的免疫策略方法也被相继提出^[77,86-91]. 然而这些策略都暗含了同一个假设, 即免疫是强制性的. 在现实生活中, 个体是否采取接种免疫会受到很多因素的影响, 事实上在很多情况下取决于个体的主观意愿. Ruan 等人^[92]基于个体具有危机意识的假设, 提出了一个信息驱动接种的流行病传播模型. 其

研究表明, 当信息在网络中充分传播时, 少量的接种人数就能达到控制流行病传播的效果. Funk等人^[93]考虑了防范信息与疾病同时传播的情形. 一方面, 疾病的存在会诱导防范信息的产生, 而防范信息的传播会触发个体产生相应的行为反应, 导致其易感性发生改变, 进一步影响疾病的传播. Wang等人^[94]进一步研究了双层耦合网络上信息传播与流行病传播相互影响的情况, 结果表明, 信息的传播会抑制流行病的暴发.

另一种更真实的情况是考虑博弈过程^[95,96]. 在决定接种疫苗前, 人们往往会权衡利弊, 考虑各种可能的因素. 这是因为接种疫苗需要付出一定的“代价”, 如时间和金钱的花费、副作用等; 而不接种就面临着被感染的风险. 这是一个典型的博弈过程. Zhang等人^[97]研究了无标度网络上基于博弈理论的自愿接种行为. 考虑SIS模型, 如果一个易感态节点有 k_{inf} 个感染邻居, 则该节点被感染的概率为 $\lambda = 1 - (1 - \beta)^{k_{\text{inf}}}$, 其中 β 为一个易感态个体碰到一个感染态个体被感染的概率. 假设实际的接种行为主要受两个因素的影响: 被感染的风险和接种带来的风险及花费等, 则可引入两个效益函数:

$$\begin{aligned} p_n &= c_1 \lambda / \mu, \\ p_v &= c_2, \end{aligned}$$

分别表示个体采取不接种和接种两种行为时所对应的成本, 其中 c_1 与感染风险有关, c_2 与接种风险或花费有关. 上式表明, 一个易感态节点不接种的代价与它被感染的概率 λ 以及处于感染态的平均时间 $1/\mu$ (μ 为恢复概率)成正比. 当 $p_n > p_v$, 个体就会采取接种行为. 研究表明, 在流行病暴发初期, BA网上的个体比ER网上的更倾向于接种, 所以更容易得到控制. 这是因为, 根据个体接种的条件可知(代入相应的表达式): $p_{\text{inf}} \geq [\log_{1-\beta}(c_1 - c_2) / c_1]$ ([]表示取整), 意味着只有当个体的邻居节点中感染人数达到一定的阈值, 该个体才可能接种. 显然, BA网上的hub节点更容易满足这一条件, 所以流行病的传播能被有效地抑制. 然而, 这个模型假设个体是完全理性的. 在现实情况中, 个体所做的决策是非完全理性的. Fu等人^[98]从演化博弈的角度研究了模仿接种行为对疾病传播的影响. 在他们的模型中, 每个节点(在当前季节末期)有一个收益值, 然后个体会从邻居节点中随机挑选一个节点进行模仿(决定是否下个季节接种). 研究发现, 当个体更倾向于

模仿那些成功者(或搭便车者, 他们没采取接种但也没被感染, 所以收益值最高)的策略时, 接种水平会降低. 目前, 将博弈理论与疾病传播动力学相结合也是研究复杂网络上流行病传播的一个主要方向之一.

最后值得一提的是, 一些非药物的干预措施也能在一定程度上控制流行病的传播, 例如关闭学校、自发回避行为(倾向于待在家里)等^[99,100]. 这方面的研究需要更多实际数据的支持.

7 总结

关于复杂网络上流行病传播的研究在近20年来已取得了巨大进展. 一方面, 在数学建模方法上, 从最初的单层、静态网络到多层、时变网络等, 已越来越接近真实的世界. 与此同时, 理论解析方法也在不断发展, 用以揭示新模型的本质. 另一方面, 随着计算机科学技术的迅速发展, 大量真实数据的获取变得越来越便捷, 这将流行病传播的研究带入了一个新的阶段. 将实证数据与理论模型相结合, 可以用来分析、预测真实世界中的流行病传播过程, 这为相关机构部署、防御流行病的大规模扩散提供了坚实的理论依据. 这些数据的获取, 同时也带来了一些新的挑战. 例如, Scarpino等人^[101]发现在美国国内的17次流行病暴发实证数据、25年的州内流感数据以及波多黎各19年的登革热病毒数据中存在暴发式的相变过程. 流行病传播的这种一级相变(传统的流行病传播模型只能发生二级相变)最近几年引起了研究者的广泛关注, 科学家已提出了多种不同的模型来解释这种现象^[102-105]. Zheng等人^[49]考察了香港的流行病数据, 发现香港的流感暴发并没有严格的季节性, 而是具有非周期性的特征(这一特征同样在美国纽约与巴尔的摩的麻疹数据中被发现), 由此提出了一种新的SIRS模型来解释这种非周期性. Liu等人^[106]根据以上数据进一步发现了复发式流行病的迟滞现象.

实证数据的获取给流行病传播的建模研究提出了挑战, 而新模型的提出又给数学分析方法提出了挑战, 这一过程从未停止. 真实世界中流行病的传播是一个非常复杂的过程. 它牵涉到不同网络之间的相互作用(例如接触网络和在线社交网络等)^[107-109], 也可能牵涉到不同病毒之间的相互耦合^[104,109-111], 同时还涉及个体的心理活动和认知^[7]等. 所以无论是在实证数据分

析(特别是多层网络的实际数据)还是理论分析(特别是多种动力学耦合)方面, 都还存在着很多严峻的挑战, 有待进一步的研究. 总而言之, 复杂网络上流行病传

播的研究是一个充满活力的研究领域, 其也在其他领域, 特别是社会传播领域得到了广泛的应用^[112–115]. 我们相信, 目前的研究进展甚至才刚刚开始.

参考文献

- 1 Anderson R M, May R M. *Infectious Diseases of Humans*. Oxford: Oxford University Press, 1992
- 2 Yang H L. *The Great Plague in History* (in Chinese). Beijing: Chinese Development Press, 2007 [杨红林. 历史上的大瘟疫. 北京: 中国发展出版社, 2007]
- 3 McNeill W H. *Plagues and Peoples* (in Chinese). Beijing: China Environmental Science Press, 2010 [威廉·H. 麦克尼尔. 瘟疫与人. 北京: 中国环境科学出版社, 2010]
- 4 Chen L S, Chen J. *Nonlinear Biodynamical Systems* (in Chinese). Beijing: Science Press, 1993 [陈兰荪, 陈键. 非线性生物动力系统. 北京: 科学出版社, 1993]
- 5 Duncan J W, Steven S H. Collective dynamics of “small-world” networks. *Nature*, 1998, 393: 440–442
- 6 Barabasi A. Emergence of scaling in random networks. *Science*, 1999, 286: 509–512
- 7 Pastor-Satorras R, Castellano C, Van Mieghem P, et al. Epidemic processes in complex networks. *Rev Mod Phys*, 2015, 87: 925–979, arXiv: [1408.2701](#)
- 8 Newman M E J. The structure and function of complex networks. *SIAM Rev*, 2003, 45: 167–256
- 9 Boccaletti S, Latora V, Moreno Y, et al. Complex networks: Structure and dynamics. *Phys Rep*, 2006, 424: 175–308
- 10 Fang J Q, Wang X F, Zheng Z G, et al. *New interdisciplinary science: Network science* (in Chinese). Prog Phys, 2007, 27: 239–343 [方锦清, 汪小帆, 郑志刚, 等. 一门崭新的交叉科学: 网络科学. 物理学进展, 2007, 27: 239–343]
- 11 He D R, Liu Z H, Wang B H. *Complex Systems and Complex Networks* (in Chinese). Beijing: Higher Education Press, 2009 [何大韧, 刘宗华, 汪秉宏. 复杂系统与复杂网络. 北京: 高等教育出版社, 2009]
- 12 Albert R, Barabási A L. Statistical mechanics of complex networks. *Rev Mod Phys*, 2002, 74: 47–97
- 13 Cattuto C, Van den Broeck W, Barrat A, et al. Dynamics of person-to-person interactions from distributed RFID sensor networks. *PLoS ONE*, 2010, 5: e11596, arXiv: [1007.3680](#)
- 14 Isella L, Stehlé J, Barrat A, et al. What’s in a crowd? Analysis of face-to-face behavioral networks. *J Theor Biol*, 2011, 271: 166–180
- 15 Zhao K, Stehlé J, Bianconi G, et al. Social network dynamics of face-to-face interactions. *Phys Rev E*, 2011, 83: 056109, arXiv: [1102.2423](#)
- 16 Wang W, Tang M, Eugene Stanley H, et al. Unification of theoretical approaches for epidemic spreading on complex networks. *Rep Prog Phys*, 2017, 80: 036603, arXiv: [1612.04216](#)
- 17 Pastor-Satorras R, Vespignani A. Epidemic spreading in scale-free networks. *Phys Rev Lett*, 2001, 86: 3200–3203
- 18 Cai C R, Wu Z X, Chen M Z Q, et al. Solving the dynamic correlation problem of the susceptible-infected-susceptible model on networks. *Phys Rev Lett*, 2016, 116: 258301, arXiv: [1606.02037](#)
- 19 Granell C, Gómez S, Arenas A. Dynamical interplay between awareness and epidemic spreading in multiplex networks. *Phys Rev Lett*, 2013, 111: 128701, arXiv: [1306.4136](#)
- 20 Boguñá M, Castellano C, Pastor-Satorras R. Nature of the epidemic threshold for the susceptible-infected-susceptible dynamics in networks. *Phys Rev Lett*, 2013, 111: 068701, arXiv: [1305.4819](#)
- 21 Karrer B, Newman M E J. Message passing approach for general epidemic models. *Phys Rev E*, 2010, 82: 016101, arXiv: [1003.5673](#)
- 22 Shrestha M, Scarpino S V, Moore C. Message-passing approach for recurrent-state epidemic models on networks. *Phys Rev E*, 2015, 92: 022821, arXiv: [1505.02192](#)
- 23 Valdano E, Fiorentin M R, Poletto C, et al. Epidemic threshold in continuous-time evolving networks. *Phys Rev Lett*, 2018, 120: 068302, arXiv: [1706.05968](#)
- 24 Serrano M A, Boguñá M. Percolation and epidemic thresholds in clustered networks. *Phys Rev Lett*, 2006, 97: 088701
- 25 Schwarzkopf Y, Rákos A, Mukamel D. Epidemic spreading in evolving networks. *Phys Rev E*, 2010, 82: 036112, arXiv: [1001.3878](#)
- 26 Liu Z, Hu B. Epidemic spreading in community networks. *Europhys Lett*, 2005, 72: 315–321

- 27 Kuperman M, Abramson G. Small world effect in an epidemiological model. *Phys Rev Lett*, 2001, 86: 2909–2912
- 28 Stegehuis C, van der Hofstad R, van Leeuwen J S H. Epidemic spreading on complex networks with community structures. *Sci Rep*, 2016, 6: 29784, arXiv: [1611.06092](#)
- 29 Colizza V, Pastor-Satorras R, Vespignani A. Reaction-diffusion processes and metapopulation models in heterogeneous networks. *Nat Phys*, 2007, 3: 276–282
- 30 Meloni S, Perra N, Arenas A, et al. Modeling human mobility responses to the large-scale spreading of infectious diseases. *Sci Rep*, 2011, 1: 62
- 31 Balcan D, Vespignani A. Invasion threshold in structured populations with recurrent mobility patterns. *J Theor Biol*, 2012, 293: 87–100
- 32 Belik V, Geisel T, Brockmann D. Natural human mobility patterns and spatial spread of infectious diseases. *Phys Rev X*, 2011, 1: 011001, arXiv: [1103.6224](#)
- 33 Ruan Z, Hui P, Lin H, et al. Risks of an epidemic in a two-layered railway-local area traveling network. *Eur Phys J B*, 2013, 86: 13
- 34 Ruan Z, Tang M, Liu Z. How the contagion at links influences epidemic spreading. *Eur Phys J B*, 2013, 86: 149
- 35 Tang M, Liu Z, Li B. Epidemic spreading by objective traveling. *Europhys Lett*, 2009, 87: 18005
- 36 Wang L, Li X. Spatial epidemiology of networked metapopulation: An overview. *Chin Sci Bull*, 2014, 59: 3511–3522
- 37 Gómez-Gardeñes J, Soriano-Paños D, Arenas A. Critical regimes driven by recurrent mobility patterns of reaction-diffusion processes in networks. *Nat Phys*, 2018, 14: 391–395
- 38 Colizza V, Barrat A, Barthélemy M, et al. The role of the airline transportation network in the prediction and predictability of global epidemics. *Proc Natl Acad Sci USA*, 2006, 103: 2015–2020
- 39 Hufnagel L, Brockmann D, Geisel T. Forecast and control of epidemics in a globalized world. *Proc Natl Acad Sci USA*, 2003, 101: 15124–15129
- 40 Brockmann D, Helbing D. The hidden geometry of complex, network-driven contagion phenomena. *Science*, 2013, 342: 1337–1342
- 41 Balcan D, Colizza V, Gonçalves B, et al. Multiscale mobility networks and the spatial spreading of infectious diseases. *Proc Natl Acad Sci USA*, 2009, 106: 21484–21489
- 42 Gross T, D’Lima C J D, Blasius B. Epidemic dynamics on an adaptive network. *Phys Rev Lett*, 2006, 96: 208701
- 43 Demirel G, Barter E, Gross T. Dynamics of epidemic diseases on a growing adaptive network. *Sci Rep*, 2017, 7: 42352
- 44 Wang Z, Andrews M A, Wu Z X, et al. Coupled disease–behavior dynamics on complex networks: A review. *Phys Life Rev*, 2015, 15: 1–29
- 45 Zhao S, Bauch C T, He D. Strategic decision making about travel during disease outbreaks: A game theoretical approach. *J R Soc Interface*, 2018, 15: 20180515
- 46 Crépey P, Alvarez F P, Barthélemy M. Epidemic variability in complex networks. *Phys Rev E*, 2006, 73: 046131
- 47 Zhou T, Liu J G, Bai W J, et al. Behaviors of susceptible-infected epidemics on scale-free networks with identical infectivity. *Phys Rev E*, 2006, 74: 056109
- 48 Small M, Tse C K. Small world and scale free model of transmission of SARS. *Int J Bifurc Chaos*, 2005, 15: 1745–1755
- 49 Zheng M, Wang C, Zhou J, et al. Non-periodic outbreaks of recurrent epidemics and its network modelling. *Sci Rep*, 2015, 5: 16010
- 50 Zheng M, Zhao M, Min B, et al. Synchronized and mixed outbreaks of coupled recurrent epidemics. *Sci Rep*, 2017, 7: 2424, arXiv: [1610.02528](#)
- 51 Boguñá M, Pastor-Satorras R, Vespignani A. Absence of epidemic threshold in scale-free networks with degree correlations. *Phys Rev Lett*, 2003, 90: 028701
- 52 Chu X, Guan J, Zhang Z, et al. Epidemic spreading in weighted scale-free networks with community structure. *J Stat Mech*, 2009, 2009: P07043
- 53 Saumell-Mendiola A, Serrano M Á, Boguñá M. Epidemic spreading on interconnected networks. *Phys Rev E*, 2012, 86: 026106, arXiv: [1202.4087](#)
- 54 Liu S, Perra N, Karsai M, et al. Controlling contagion processes in activity driven networks. *Phys Rev Lett*, 2014, 112: 118702, arXiv: [1309.7031](#)
- 55 Eames K T D, Keeling M J. Modeling dynamic and network heterogeneities in the spread of sexually transmitted diseases. *Proc Natl Acad Sci USA*, 2002, 99: 13330–13335
- 56 Mata A S, Ferreira S C. Pair quenched mean-field theory for the susceptible-infected-susceptible model on complex networks. *Europhys Lett*, 2013, 103: 48003, arXiv: [1305.5153](#)
- 57 Colizza V, Vespignani A. Invasion threshold in heterogeneous metapopulation networks. *Phys Rev Lett*, 2007, 99: 148701, arXiv: [0802.3636](#)
- 58 Ruan Z, Wang C, Ming Hui P, et al. Integrated travel network model for studying epidemics: Interplay between journeys and epidemic. *Sci Rep*,

- 2015, 5: 11401
- 59 Zheng M, Ruan Z, Tang M, et al. Influence of periodic traffic congestion on epidemic spreading. *Int J Mod Phys C*, 2016, 27: 1650048
 - 60 Song C, Qu Z, Blumm N, et al. Limits of predictability in human mobility. *Science*, 2010, 327: 1018–1021
 - 61 Song C, Koren T, Wang P, et al. Modelling the scaling properties of human mobility. *Nat Phys*, 2010, 6: 818–823, arXiv: [1010.0436](#)
 - 62 Gonzalez M, Hidalgo C, Barabasi A. Understanding individual human mobility patterns. *Nature*, 2008, 453: 779–782, arXiv: [0806.1256](#)
 - 63 Yan X Y, Wang W X, Gao Z Y, et al. Universal model of individual and population mobility on diverse spatial scales. *Nat Commun*, 2017, 8: 1639
 - 64 Balcan D, Vespignani A. Phase transitions in contagion processes mediated by recurrent mobility patterns. *Nat Phys*, 2011, 7: 581–586, arXiv: [1103.4874](#)
 - 65 Soriano-Paños D, Lotero L, Arenas A, Gómez-Gardeñes J. Spreading processes in multiplex metapopulations containing different mobility networks. *Phys Rev X*, 2018, 8: 031039
 - 66 Poletto C, Tizzoni M, Colizza V. Heterogeneous length of stay of hosts' movements and spatial epidemic spread. *Sci Rep*, 2012, 2: 476, arXiv: [1207.4746](#)
 - 67 Belik V, Geisel T, Brockmann D. Recurrent host mobility in spatial epidemics: Beyond reaction-diffusion. *Eur Phys J B*, 2011, 84: 579–587, arXiv: [1106.3461](#)
 - 68 Poletto C, Tizzoni M, Colizza V. Human mobility and time spent at destination: Impact on spatial epidemic spreading. *J Theor Biol*, 2013, 338: 41–58
 - 69 Tang M, Liu L, Liu Z. Influence of dynamical condensation on epidemic spreading in scale-free networks. *Phys Rev E*, 2009, 79: 016108
 - 70 Wang L, Wang Z, Zhang Y, et al. How human location-specific contact patterns impact spatial transmission between populations? *Sci Rep*, 2013, 3: 1468
 - 71 Mata A S, Ferreira S C, Pastor-Satorras R. Effects of local population structure in a reaction-diffusion model of a contact process on metapopulation networks. *Phys Rev E*, 2013, 88: 042820, arXiv: [1305.4965](#)
 - 72 Ruan Z, Tang M, Gu C, et al. Epidemic spreading between two coupled subpopulations with inner structures. *Chaos*, 2017, 27: 103104
 - 73 Barrat A, Barthélemy M, Pastor-Satorras R, et al. The architecture of complex weighted networks. *Proc Natl Acad Sci USA*, 2004, 101: 3747–3752
 - 74 Guimerà R, Mossa S, Turttschi A, et al. The worldwide air transportation network: Anomalous centrality, community structure, and cities' global roles. *Proc Natl Acad Sci USA*, 2005, 102: 7794–7799
 - 75 Barabasi A L. *Network Science*. Cambridge: Cambridge University Press, 2015
 - 76 Tizzoni M, Bajardi P, Poletto C, et al. Real-time numerical forecast of global epidemic spreading: Case study of 2009 A/H1N1pdm. *BMC Med*, 2012, 10: 165
 - 77 Wang Z, Bauch C T, Bhattacharyya S, et al. Statistical physics of vaccination. *Phys Rep*, 2016, 664: 1–113, arXiv: [1608.09010](#)
 - 78 Pastor-Satorras R, Vespignani A. Immunization of complex networks. *Phys Rev E*, 2002, 65: 036104
 - 79 Cohen R, Havlin S, Ben-Avraham D. Efficient immunization strategies for computer networks and populations. *Phys Rev Lett*, 2003, 91: 247901
 - 80 Gallos L K, Liljeros F, Argyrakis P, et al. Improving immunization strategies. *Phys Rev E*, 2007, 75: 045104, arXiv: [0704.1589](#)
 - 81 Palla G, Derényi I, Farkas I, et al. Uncovering the overlapping community structure of complex networks in nature and society. *Nature*, 2005, 435: 814–818
 - 82 Morris M, Podhisita C, Wawer M J, et al. Bridge populations in the spread of HIV/AIDS in Thailand. *AIDS (London, England)*, 1996, 10: 1265–1271
 - 83 Newman M E J. A measure of betweenness centrality based on random walks. *Soc Networks*, 2005, 27: 39–54
 - 84 Salathé M, Jones J H. Dynamics and control of diseases in networks with community structure. *PLoS Comput Biol*, 2010, 6: e1000736
 - 85 Gong K, Tang M, Hui P M, et al. An efficient immunization strategy for community networks. *PLoS ONE*, 2013, 8: e83489
 - 86 Duan W, Chen Z, Liu Z, et al. Efficient target strategies for contagion in scale-free networks. *Phys Rev E*, 2005, 72: 026133
 - 87 Chen Y, Paul G, Havlin S, et al. Finding a better immunization strategy. *Phys Rev Lett*, 2008, 101: 058701
 - 88 Schneider C M, Mihaljev T, Herrmann H J. Inverse targeting—An effective immunization strategy. *Europhys Lett*, 2012, 98: 46002, arXiv: [1112.2957](#)

- 89 Hinde J, Schwartz I B. Epidemic extinction and control in heterogeneous networks. *Phys Rev Lett*, 2016, 117: 028302, arXiv: [1604.07244](#)
- 90 Wang Z, Zhao D W, Wang L, et al. Immunity of multiplex networks via acquaintance vaccination. *Europhys Lett*, 2015, 112: 48002
- 91 Zhao D, Wang L, Li S, et al. Immunization of epidemics in multiplex networks. *PLoS ONE*, 2014, 9: e112018
- 92 Ruan Z, Tang M, Liu Z. Epidemic spreading with information-driven vaccination. *Phys Rev E*, 2012, 86: 036117
- 93 Funk S, Gilad E, Watkins C, et al. The spread of awareness and its impact on epidemic outbreaks. *Proc Natl Acad Sci USA*, 2009, 106: 6872–6877
- 94 Wang W, Liu Q H, Cai S M, et al. Suppressing disease spreading by using information diffusion on multiplex networks. *Sci Rep*, 2016, 6: 29259, arXiv: [1604.01644](#)
- 95 Bauch C T, Galvani A P, Earn D J D. Group interest versus self-interest in smallpox vaccination policy. *Proc Natl Acad Sci USA*, 2003, 100: 10564–10567
- 96 Bauch C T. Imitation dynamics predict vaccinating behaviour. *Proc R Soc B-Biol Sci*, 2005, 272: 1669–1675
- 97 Zhang H, Zhang J, Zhou C, et al. Hub nodes inhibit the outbreak of epidemic under voluntary vaccination. *New J Phys*, 2010, 12: 023015
- 98 Fu F, Rosenbloom D I, Wang L, et al. Imitation dynamics of vaccination behaviour on social networks. *Proc R Soc B*, 2011, 278: 42–49
- 99 Cauchemez S, Valleron A J, Boëlle P Y, et al. Estimating the impact of school closure on influenza transmission from sentinel data. *Nature*, 2008, 452: 750–754
- 100 Bayham J, Kuminoff N V, Gunn Q, et al. Measured voluntary avoidance behaviour during the 2009 A/H1N1 epidemic. *Proc R Soc B*, 2015, 282: 20150814
- 101 Scarpino S V, Allard A, Hébert-Dufresne L. The effect of a prudent adaptive behaviour on disease transmission. *Nat Phys*, 2016, 12: 1042–1046, arXiv: [1509.00801](#)
- 102 Chae H, Yook S H, Kim Y. Discontinuous phase transition in a core contact process on complex networks. *New J Phys*, 2015, 17: 023039
- 103 Böttcher L, Woolley-Meza O, Araújo N A M, et al. Disease-induced resource constraints can trigger explosive epidemics. *Sci Rep*, 2015, 5: 16571, arXiv: [1508.05484](#)
- 104 Gómez-Gardeñes J, Lotero L, Taraskin S N, et al. Explosive contagion in networks. *Sci Rep*, 2016, 6: 19767
- 105 Cai W, Chen L, Ghanbarnejad F, et al. Avalanche outbreaks emerging in cooperative contagions. *Nat Phys*, 2015, 11: 936–940
- 106 Liu H, Zheng M, Wu D, et al. Hysteresis loop of nonperiodic outbreaks of recurrent epidemics. *Phys Rev E*, 2016, 94: 062318
- 107 Wang W, Tang M, Yang H, et al. Asymmetrically interacting spreading dynamics on complex layered networks. *Sci Rep*, 2014, 4: 5097, arXiv: [1405.1905](#)
- 108 Liu Q H, Wang W, Tang M, et al. Impacts of complex behavioral responses on asymmetric interacting spreading dynamics in multiplex networks. *Sci Rep*, 2016, 6: 25617, arXiv: [1509.08183](#)
- 109 Newman M E J. Threshold effects for two pathogens spreading on a network. *Phys Rev Lett*, 2005, 95: 108701
- 110 Zhao Y, Zheng M, Liu Z. A unified framework of mutual influence between two pathogens in multiplex networks. *Chaos*, 2014, 24: 043129
- 111 Azimi-Tafreshi N. Cooperative epidemics on multiplex networks. *Phys Rev E*, 2016, 93: 042303, arXiv: [1511.03235](#)
- 112 Ruan Z, Iñiguez G, Karsai M, et al. Kinetics of social contagion. *Phys Rev Lett*, 2015, 115: 218702, arXiv: [1506.00251](#)
- 113 Ruan Z, Wang J, Xuan Q, et al. Information filtering by smart nodes in random networks. *Phys Rev E*, 2018, 98: 022308
- 114 Zhou J, Liu Z, Li B. Influence of network structure on rumor propagation. *Phys Lett A*, 2007, 368: 458–463
- 115 Zhang Z K, Liu C, Zhan X X, et al. Dynamics of information diffusion and its applications on complex networks. *Phys Rep*, 2016, 651: 1–34

Epidemic spreading in complex networks

RUAN ZhongYuan^{*}

College of Computer Science, Zhejiang University of Technology, Hangzhou 310023, China

Human history is accompanied by all kinds of epidemic diseases. Yet, the research progress in epidemic spreading is unexpectedly slow. In the last few decades, by applying the theory of complex networks, the study on epidemic spreading has obtained great success. In this review, we briefly introduce the important progresses on epidemic spreading in complex networks. We mainly focus on the following four aspects: (1) epidemic spreading based on contact network model; (2) epidemic spreading based on meta-population model; (3) the prediction of epidemic spreading; (4) the control of epidemic spreading. These achievements make us deeply understand the epidemic spreading in the real world, and make it possible for us to predict and control the real epidemics.

complex networks, epidemic spreading, contact network model, metapopulation model

PACS: 64.60.aq, 87.19.X-, 87.23.Ge

doi: [10.1360/SSPMA-2019-0128](https://doi.org/10.1360/SSPMA-2019-0128)