



电子科技大学学报

Journal of University of Electronic Science and Technology of China

ISSN 1001-0548, CN 51-1207/T

《电子科技大学学报》网络首发论文

题目: 基于时变参数-SIR 模型的 2019-nCoV 疫情评估和预测
作者: 喻孜, 张贵清, 刘庆珍, 吕忠全
收稿日期: 2020-02-02
网络首发日期: 2020-02-10
引用格式: 喻孜, 张贵清, 刘庆珍, 吕忠全. 基于时变参数-SIR 模型的 2019-nCoV 疫情评估和预测[J/OL]. 电子科技大学学报.
<http://kns.cnki.net/kcms/detail/51.1207.t.20200207.1546.002.html>



网络首发: 在编辑部工作流程中, 稿件从录用到出版要经历录用定稿、排版定稿、整期汇编定稿等阶段。录用定稿指内容已经确定, 且通过同行评议、主编终审同意刊用的稿件。排版定稿指录用定稿按照期刊特定版式 (包括网络呈现版式) 排版后的稿件, 可暂不确定出版年、卷、期和页码。整期汇编定稿指出版年、卷、期、页码均已确定的印刷或数字出版的整期汇编稿件。录用定稿网络首发稿件内容必须符合《出版管理条例》和《期刊出版管理规定》的有关规定; 学术研究成果具有创新性、科学性和先进性, 符合编辑部对刊文的录用要求, 不存在学术不端行为及其他侵权行为; 稿件内容应基本符合国家有关书刊编辑、出版的技术标准, 正确使用和统一规范语言文字、符号、数字、外文字母、法定计量单位及地图标注等。为确保录用定稿网络首发的严肃性, 录用定稿一经发布, 不得修改论文题目、作者、机构名称和学术内容, 只可基于编辑规范进行少量文字的修改。

出版确认: 纸质期刊编辑部通过与《中国学术期刊 (光盘版)》电子杂志社有限公司签约, 在《中国学术期刊 (网络版)》出版传播平台上创办与纸质期刊内容一致的网络版, 以单篇或整期出版形式, 在印刷出版之前刊发论文的录用定稿、排版定稿、整期汇编定稿。因为《中国学术期刊 (网络版)》是国家新闻出版广电总局批准的网络连续型出版物 (ISSN 2096-4188, CN 11-6037/Z), 所以签约期刊的网络版上网络首发论文视为正式出版。



基于时变参数-SIR 模型的 2019-nCoV 疫情评估和预测

喻 孜¹, 张贵清^{2*}, 刘庆珍³, 吕忠全¹

(1. 南京林业大学理学院 南京 210037; 2. 天津科技大学理学院 天津 300457; 3. 中国医学科学院血液病医院 天津 300020)

【摘要】文章基于 2019-nCoV 疫情发展到 2 月 1 日所呈现的特点, 对 SIR 模型进行了修正, 使用易感再生数、当日感染率和潜伏感染率来求解病毒演化动力学方程, 研究了感染人数的变化趋势, 并分析了政府行政行为对趋势变化产生的影响。结果表明, 从 2020 年 1 月 24 日后, 政府的行政行为有效降低了病毒蔓延趋势。与 1 月 24 日之前呈现的趋势相比, 截至 2 月 1 日, 实际感染人数较原趋势预估人数下降了超 1/2。易感再生数、当日再生数和潜伏再生数都大幅度降低。基于目前的趋势, 对易感再生数、当日感染概率、潜伏感染概率随时间的变化进行了分析, 利用时变参数对疫情发展进行了预测。结果表明在 2 月 9 日左右, 疫情发展会达到高峰, 随后确诊人数将出现下降。

关键词 2019-nCoV; SIR 模型; 疫情评估; 预测

中图分类号 O414; TP391

文献标志码 A

doi:10.12178/1001-0548.2020027

The Outbreak Assessment and Prediction of 2019-nCoV Based on Time-Varying SIR Model

YU Zi¹, ZHANG Gui-qing^{2*}, LIU Qing-zhen³, and LÜ Zhong-quan¹

(1. College of Science, Nanjing Forestry University Nanjing 210037; 2. College of Science, Tianjin University of Science & Technology Tianjin 300457;

3. Hematology Hospital, Chinese Academy of Medical Sciences Tianjin 300020)

Abstract The article is based on the early behavior of 2019-nCoV until February 1st. The SIR model was modified to solve the dynamic equation of virus evolution by using the number of susceptible persons, the probability of infection and the latent infection rate. The change trend of infected persons was studied, and the influence of government administrative actions on the trend was analyzed. The results showed that after January 24th, 2020, the administrative action of the government effectively reduced the spread trend of the virus and the number of infected people decrease obviously. The number of infections had reduced more than a half compared with number which according the trend of January 24th. The number of susceptible persons, the probability of infection and the latent infection rate were all greatly reduced. Based on the current trend and the optimal parameters, we fitted the trend equation and found that the outbreak will peak around February 9th and then decline.

Key words 2019-nCoV; the SIR model; outbreak assessment; prediction

自 2019 年 12 月在武汉发现第一批新型冠状病毒感染者以来, 该病毒在短时间内在武汉以及国内其他省份中快速传播, 并且在海外也很快出现了较多感染者。截止 2020 年 1 月 31 日晚 12 点, 全国确诊病患接近一万。其中 1 月 28 日—1 月 31 日的新增确诊人数约为 1 500 人、1 700 人、2 000 人和 2 102 人。政府也迅速采取了一系列措施来对抗疫情, 从 23 日开始, 地方政府陆续开始“封城”并通过

媒体实时发布确诊人数。专家通过网络及媒体传播防治疫情的注意事项。全社会自觉在家自我隔离, 戴口罩出门, 企业延期复工, 人员聚集场所关停。病毒的发展已经严重影响到了人们的日常工作与生活。2019-nCoV 疫情的发生和演变过程中, 逐渐开始蔓延恐慌情绪。全社会都在关注两个重要问题: 一是确诊人数将如何变化, 什么时间会下降? 二是迄今为止, 政府所采用的行政手段, 产生了多大的

收稿日期: 2020-02-02; 修回日期: 2020-02-04

基金项目: 国家自然科学基金(11247217); 天津科技大学青年教师创新基金(2014CXLG23)

作者简介: 喻孜(1981-), 男, 副教授, 主要从事理论物理方面的研究。

通信作者: 张贵清(1983-), 副教授, Email: nkzhangq@tust.edu.cn

积极作用? 回答这两个问题, 需要对疫情进行科学的预测和评估。

目前已经有大量工作对 2019-nCoV 疫情进行了研究。文献 [1] 根据最早 425 例确诊病例数据, 描述了病例特征, 并估计了关键流行病学延迟时间分布情况, 在病毒早期呈指数增长的初期, 估计了传染病倍增时间和基本再生数。文献 [2] 发现了第二代病例, 指出病毒会存在“人传人”, 对金银潭医院最开始收治患者时候的信息进行归纳研究, 指出非华南海鲜市场暴露的病例的存在。文献 [3] 基于 SEIR 模型, 根据 2020 年 1 月 25 日之前的发展趋势, 估算得到了 2019-nCoV 的再生数, 并对疫情发展进行了预测。还有很多其他研究^[4-9], 从不同角度对疫情的特点进行了评估。这些研究都对疫情的防治提供了重要的参考。在病毒传播研究中^[10-13], 由于 SIR 动力学模型能给出较为清晰的逻辑关系和准确的趋势预测, 因而被广泛采用。相比于之前的病毒传播事件, 2019-nCoV 疫情比较特殊。1) 相对于总的人口基数, 目前的治愈人数和死亡人数都较低, 因此出院人员 (Removed) 比率可以忽略不记。2) 从 2020 年 1 月 11 日, 官方开始报道确诊人数以来, 存在人口流动, 因此传统 SIR 模型认为总体环境是封闭 ($S+I+R=\text{常数}$) 并不符合现实。3) 由于染病确诊后将被隔离, 所以确诊病人并不形成新的感染。4) 此次病毒有较长的潜伏期, 会带来明显的延迟效应。5) 此次疫情突发, 武汉的诊断力量在初期存在明显的不足。综上所述, 为了更为准确地对疫情进行预测, 需要对 SIR 模型进行修正。本文根据实际情况, 对 SIR 模型进行了适当修正, 聚焦国内确诊人数的变化, 对趋势进行预测和判断, 并根据结果对现有政策进行评估。

1 模型

本文根据实际疫情对 SIR 模型进行了一些修正, 描述如下:

- 1) 由于不存在封闭情况, 考虑开放体系。
- 2) 新型冠状病毒的治愈人数和死亡人数相对较小, 因此只考虑两类人群, Susceptible(易感)、Infected(感染)。
- 3) 目前数据以天为单位发布, 因此不考虑连续变化情况, 只考虑离散的方程。
- 4) 当确诊后, 病人立即隔离, 因此不会作为新的感染源。因此感染源应该是第 $n+1$ 天和第 n 天诊断的病人差。假设平均染病者 (I) 会使 k 个人成为

易感者 (k 为易感再生数), 则每日新增的感染者为 $k(I_{n+1} - I_n)$ 。

5) 假设每天有 μ_1 的概率 (当日感染率) 易感者会成为感染者, 则当日新增病人为 $\mu_1 S_n$ 。其中 $k\mu_1$, 描述了当日病人再生数, 后文简称当日再生数。

6) 病情会有约 10 天的潜伏期, 在初期出现症状后, 很多暴露者才会去寻求诊断, 同时主要发病地武汉的初期诊断能力有限。这都会带来一个延迟效应。取 $\tau=4$, 代表平均延迟天数。在平均 4 天内的潜伏者会以 μ_2 的概率 (潜伏感染率) 被确诊为感染者。则每日潜伏感染人数为 $\mu_2 (S_n - S_{n-\tau})$ 。 $k\mu_2$ 描述了由于潜伏原因导致的病人再生数, 后文简称潜伏再生数。

基于上述考虑, 在每日模型迭代中, 各部分更新关系如下:

$$\begin{aligned} S_{n+1} - S_n &= k(I_{n+1} - I_n) - \mu_1 S_n - \mu_2 (S_n - S_{n-\tau}) \\ I_{n+1} - I_n &= \mu_1 S_n + \mu_2 (S_n - S_{n-\tau}) \end{aligned} \quad (1)$$

为了求解式 (1), 需要知道 (k, μ_1, μ_2) 3 个模型参数^[14-18], 以及 I 和 S 的初值。 $I(0)$ 通过官方数据给出 (数据来源于官方网站: <http://www.nhc.gov.cn/>), 剩下 3 个参数以及 $S(0)$ 则没有信息。文章先给定 3 个参数和 $S(0)$ 的初值, 求解方程, 将得出的解与真实数据进行比较, 通过梯度下降法寻找方差的最小值, 从而确定模型参数。由于梯度下降法对初值极为敏感, 因此, 文章还根据经验在一定初值范围内进行多次搜索, 保证得到的是最优解。

模型参数 (k, μ_1, μ_2) 是根据一个阶段的数据求出, 因而能够反应病毒在该阶段的传播趋势。通过截取不同阶段的数据来求解不同阶段的 (k, μ_1, μ_2) , 就能分析病毒演化趋势是否在外力作用下发生改变。最后再通过模型参数 (k, μ_1, μ_2) 的变化规律, 对未来进行预测。

官方从 2020 年 1 月 11 日开始公布数据, 前 4 天均无确诊人数变化。本文选择 2020 年 1 月 15 日作为数据研究的起始点。然后变化数据的结束点, 来截取不同时间段的数据进行求解。为了描述方便, 后文对时间段使用结束日期的英文简写来标识。例如 1 月 15 日至 1 月 24 日这个时间段, 简称为 Jan.24th 时间段, 1 月 15 日至 1 月 28 日这个时间段, 简称为 Jan.28th 时间段。Jan.24th 时间段是文章研究的第一个时间段, 此后每增加 1 天构成一个新的时间段。这是基于以下两点考虑: 1) 从 Jan.24th 时间段开始, 已经积攒了足够多的数据 (10 天的数据); 2) 政府从 1 月 11 日开始公布确诊

人数,采取措施,趋势形成也需要一个时间,并且病毒的最长潜伏期约为14天。需要指出,在理想状况下, $S(0)$ 应该是一个确定值,然而,在开放体系下,通过现有的数据无法评估 $S(0)$ 的真实值。因此只能通过式(1)的解与数据拟合来反向求解 $S(0)$ 。这会造成不同阶段的 $S(0)$ 略有不同。然而,由于本文采用了大范围梯度下降搜索,因此每一组参数都是在与现有数据拟合最优情况下得到的,因此文章求解得到的 (k, μ_1, μ_2) 能体现数据背后隐藏的病毒演化趋势。

2 模型参数的影响

2.1 易感再生数、当日感染率、潜伏感染率的影响。

图1, 2, 3给出了不同时间段的拟合效果,从图中发现模型的计算结果和实际公布的数据吻合较好,说明模型是合理的。同时也发现易感再生数、当日感染率、潜伏感染率参数也确实发生了变化,并且是在不断减小,这和政府的决策是有直接关系的。随着政府的强制行政行为、民众对病毒的认识加深,更多的民众选择佩戴口罩、防护设备以及选择居家少外出,以上原因直接导致了易感再生数、当日感染率、潜伏感染率的下降。为了更直观地看到趋势,图4给出了3个时间段确诊人数随时间变化的趋势图。从图中可以清楚发现,疾病的传播趋势确实是在减缓的,如果按照Jan.24th时间段的疫情发展趋势,那么在1月31日的确诊人数将会达到近40 000,同样如果按照Jan.28th时间段的疫情发展趋势,那么1月31日的确诊人数将会达到15 000左右,明显高于实际疫情的发展情况。所以图4表明政府的行政行为和民众的居家防护对于疫情的控制确实起到了很明显的作(确诊人数按照趋势变化减少了将近1/2)。鉴于此,我们同样提醒大

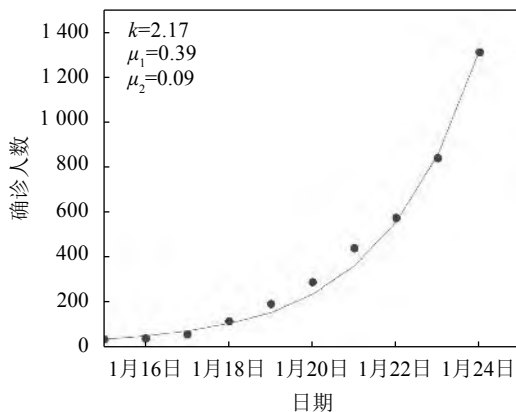


图1 Jan.24th时间段的拟合效果图,其中曲线是模型计算结果,圆点代表官方统计数据。

家能不外出就不外出,若必须外出,请一定做好有效的防护。

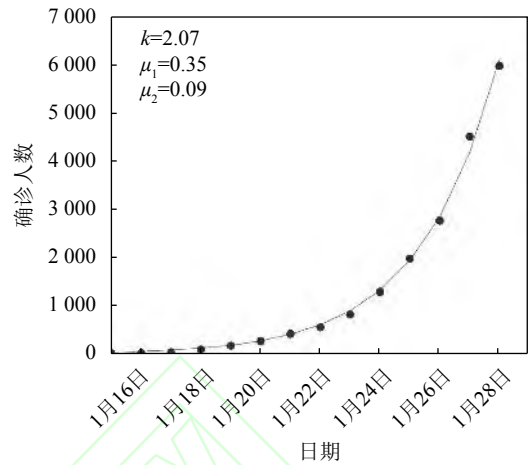


图2 Jan.28th时间段的拟合效果图,其中曲线是模型计算结果,圆点代表官方统计数据。

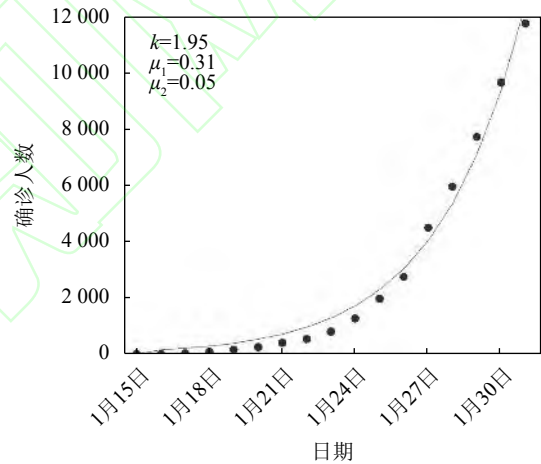


图3 Jan.31th时间段的拟合效果图,其中曲线是模型计算结果,圆点代表官方统计数据。

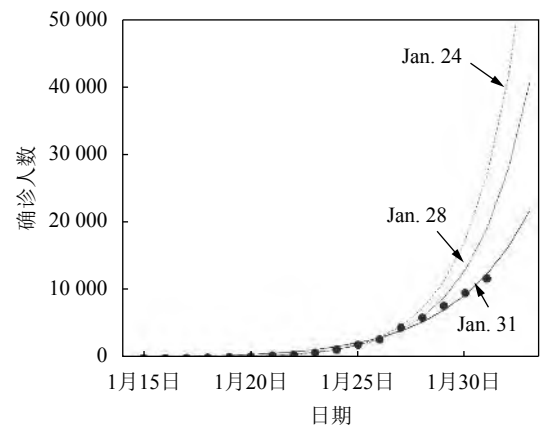


图4 不同时间段确诊人数的走势图,圆点代表官方统计数据。

2.2 易感再生数、当日再生数、潜伏再生数的变化趋势

为了更清楚地展示从初始阶段(Jan.24th时间

段)开始,趋势的变化情况,本文从第一个阶段(1月15日-1月24日)开始,每增加一天构成一个新的时间段进行求解,得到了易感再生数、当日感染再生数和潜伏感染再生数随时间的演化。结果如图5,图6,图7所示。图中的横坐标代表了时间段结束日期。从图中,可以清楚地看到3个参数都有下降趋势。同时,趋势每天都在改变。说明政府行为和民众的防护意识确实产生了强力效果。需要注意,再生数在1月27日、28日两天会有波动,1月27日,多地医疗工作者驰援武汉,使得检测能力得到大幅提升,所以导致27、28日的确诊人数大幅上升出现了明显的和前期不一样的易感再生数、当日再生数和潜伏再生数。这些参数在27、28日的波动,也从侧面证明了本文模型能够敏感地体现外力对趋势的改变。另一方面,也说明参数

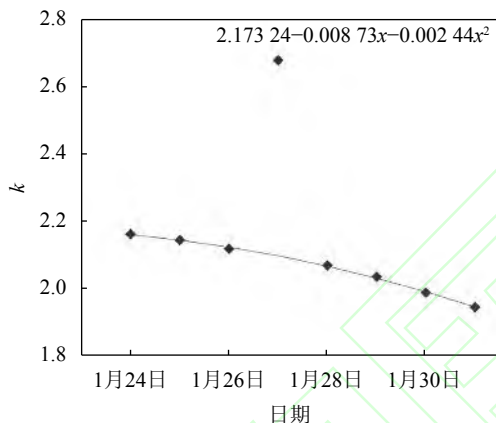


图5 易感再生数随着时间的演化情况。图中横坐标为阶段的截至日期,选择不同阶段进行拟合,反应趋势的变化。去除了27日的波动较大点。图中x的值取时间段结束日期与1月23日间隔的天数(例如图中1月24日对应 $x=1$,1月25日对应 $x=2$)。

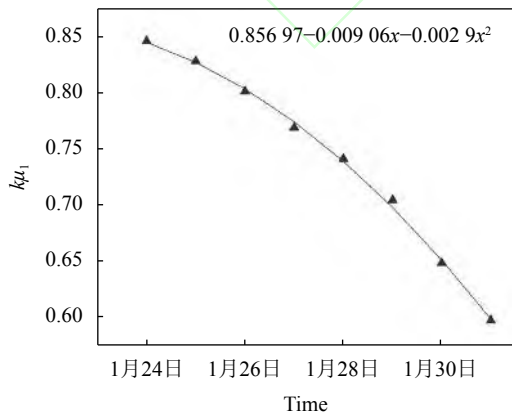


图6 当日再生数随着时间的演化情况。图中横坐标为时间段的截至日期,选择不同阶段进行拟合,反应趋势的变化。图中x的值取时间段结束日期与1月23日间隔的天数(例如1月24日对应 $x=1$,1月25日对应 $x=2$)。

固定不变的动力学模型无法真实体现此次2019-nCoV疫情的变化情况。

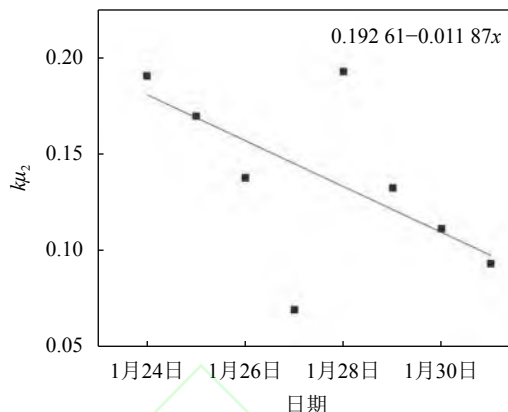


图7 潜伏再生数随着时间的演化情况。图中横坐标为阶段的截至日期,选择不同阶段进行拟合,反应趋势的变化。去除27、28日上下两个波动点后,发现直线拟合效果最佳,因此用直线拟合。图中x的值取时间段结束日期与1月23日间隔的天数(例如1月24日对应 $x=1$,1月25日对应 $x=2$)。

3 基于最优参数的疫情预测

为了结合现有趋势对未来进行预测,本文去除波动数据的影响对参数进行了拟合,从而得到了参数的变化趋势方程如下:

$$\begin{cases} k = 2.17324 - 0.00873x - 0.00244x^2 \\ k\mu_1 = 0.85687 - 0.00906x - 0.0029x^2 \\ k\mu_2 = 0.19261 - 0.0118x \end{cases} \quad (2)$$

式(2)中,x代表与1月23日间隔的天数,从1月24日开始取1。使参数按照式(2)的规律变化,代入式(1)中进行求解,预测1月31日以后的数据。预测图如图8所示。从图中可以看出26天后,也就是2月9日左右应该会出现确诊总数的峰值,随后确诊人数将会下降。

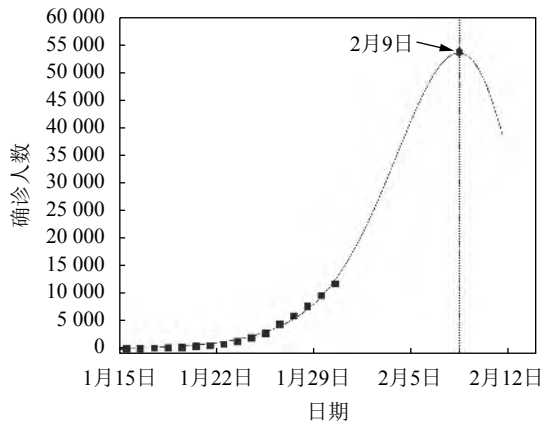


图8 图中给出了在模型最优参数条件下的模型估计值,从图中可以看出26天后,也就是2月9日左右应该会出现确诊总数的下降,同时确诊人数也达到最大的54000左右。

4 结束语

文章基于修正 SIR 模型研究了 2019-nCoV 疫情发展到 2020 年 2 月 1 日(政府隔天公布数据)为止的趋势和政策的效果。模型的参数根据实时数据拟合得到,因而修正过后的时变参数-SIR 模型结果能够反应不同阶段疫情的趋势变化。通过趋势变化能够解读政府行政行为产生的效果。结果表明,在 2020 年 1 月 27 日和 28 日,由于全国医疗系统驰援武汉使得确诊能力得到了加强,一些积压和潜伏病人得到收治,因此模型参数出现扰动。其他时间段,参数变化趋势明显。从 2020 年 1 月 24 日后,政府的行政行为有效地降低了病毒蔓延趋势,使得染病人数的变化趋势逐日下降。易感再生数、当日感染率和潜伏感染率都大幅度降低。基于目前的趋势,对疫情发展进行了预测。结果表明,在 2 月 9 日左右,疫情会达到高峰,随后确诊人数将会出现下降。

从动力学方程的解和实际数据的比较来看,文章提出的时变参数-SIR 模型能够及时反应现有趋势变化,甚至体现一些突发情况(1 月 27 日全国医疗系统驰援武汉)带来的影响。因此文章基于现有数据对趋势变化的分析结论应该是有效的。从预测角度来说,动力学预测方案本身就会存在一定的不确定性。然而,本文并没有采用固定参数去求解动力学方程,而是基于前期积累的趋势,采用动态参数求解,这在一定程度上减少了动力学方程的不确定性。因而,本文预测的趋势应该能够提供有益的参考价值。本文基于 2019-nCoV 病毒的流行特点,对 SIR 模型进行的理论修正,以及通过数据确定动态参数进行预测的方案,会对病毒的研究提供积极的理论参考。

参 考 文 献

- [1] LI Qun, MED M, GUAN Xu-hua, et al. Early transmission dynamics in Wuhan, China, of novel coronavirus-infected pneumonia[J]. N Eng J Med. DOI: 10.1056/NEJMoa2001316.
- [2] CHAN J F W, YUAN Shuo-feng, KOK King-hang, et al. A familial cluster of pneumonia associated with the 2019 novel coronavirus indicating person-to-person transmission: A study of a family cluster[J]. The Lancet, Doi: 10.1016/S0140-6736(20)30154-9.
- [3] ZHOU T, LIU Q H, YANG Z, et al. Preliminary prediction of the basic reproduction number of 2019-nCoV.[EB/OL]. [2020-01-23]. arXiv: 2001.10530.
- [4] IMAI N, DORIGATTI I, CORI A, et al. Estimating the potential total number of novel coronavirus cases in Wuhan City, China[EB/OL]. [2020-01-23] <https://www.imperial.ac.uk/media/imperial-college/medicine/sph/ide/gida-fellowships/2019-nCoV-outbreak-report-22-01-2020.pdf>.
- [5] READ J M, JESSICA R B, CUMMINGS D A, et al. Novel coronavirus 2019-nCoV: Early estimation of epidemiological parameters and epidemic predictions[EB/OL]. [2020-01-23]. <https://doi.org/10.1101/2020.01.23.20018549>.
- [6] IMAI N, et al. Transmissibility of 2019-nCoV[EB/OL]. [2020-01-25]. <https://www.imperial.ac.uk/mrc-global-infectious-disease-analysis/news-wuhan-coronavirus/>.
- [7] 许小可. 基于英国帝国理工学院报告方法的武汉新型冠状病毒病例数量估计 [EB/OL]. [2020-01-29]. <https://mp.weixin.qq.com/s/zhjlj1Xsbi51MAYs5zvgw>.
- [8] 张江. 防控力度多大才能遏制疫情发展? 网络动力学推演给你答案 [EB/OL]. [2020-01-29]. https://mp.weixin.qq.com/s/cOR7vX7HDpze8bwjUWf_RQ.
- [9] 胡延庆. 复杂网络模型预测新冠肺炎感染人数将在最近 2 到 3 天内呈下降趋势![EB/OL]. [2020-01-30]. https://mp.weixin.qq.com/s/d4YcA50f9YM1tZEDRIuQ_Q.
- [10] 夏程遗, 孙世温, 刘忠信, 等. 考虑感染时延的局域世界复杂网络上疾病的传播行为[J]. 辽宁工程技术大学学报, 2009, 28(4): 615-618.
- [11] 李光正, 史定华. 复杂网络上类 SIRS 疾病传播行为分析[J]. 自然科学进展, 2006, 16(4): 508-512.
- [12] 杨慧, 唐明, 许伯铭. 自适应网络中的流行病传播动力学研究综述[J]. 复杂系统与复杂性科学, 2012, 9(4): 63-83.
- [13] Romualdo P S, Alessandro V. Epidemic spreading in scale-free networks[J]. Phys Rev Lett, 2001, 86(1): 3200-3203.
- [14] WANG Lin, LI Xiang, ZHANG Yi-qing, et al. Evolution of scaling emergence in large-scale spatial epidemic spreading[J]. PLoS One, 2011, 6(7): e21197.
- [15] WANG Lin, LI Xiang. Spatial epidemiology of networked metapopulation: An overview[J]. Chin Sci Bull, 2014, 59(28): 3511-3522.
- [16] WANG Zhen, Michael A A, WU Zhi-xi, et al. Coupled disease-behavior dynamics on complex networks: A review[J]. Phys Life Rev, 2015, 15: 1-29.
- [17] LI Chao, WANG Li, SUN Shi-wen, et al. Identification of influential spreaders based on classified neighbors in real-world complex networks[J]. Appl Math Comput, 2018, 320: 512-523.
- [18] CHEN Xiao-long, WANG Wei, CAI Shi-min, et al. Optimal resource diffusion for suppressing disease spreading in multiplex networks[J]. Stat Mech, 2018, 5: 053501.

编辑 蒋晓