关于春节返程人口流动对新型冠状病毒肺炎(COVID-19)疫情影响的讨论

石耀霖^{1,2}, 程惠红^{1,2*}, 任天翔³, 黄禄渊⁴

1. 中国科学院计算地球动力学重点实验室, 北京 100049;

2. 中国科学院大学, 北京 100049;

3. 中国地质科学院, 北京 100037;

4. 中国地震局地壳应力研究所, 北京 100085

* 联系人, E-mail: chenghuihong@ucas.ac.cn

当 2020 年春运拉开帷幕时, 预期春运从 1 月 10 日~2 月 18 日, 在 40 d 时间里, 中国有 30 亿人次出行(http://www.gov.cn/xinwen/2020-01/10/content_5468027.htm). 虽然新型冠状病毒肺炎(以下简称新冠肺炎, COVID-19)的暴发导致今年春运 40 d 客流量同比下降 45%, 但武汉采取封城措施前, 已经有 500 万人离开武汉, 造成了新冠肺炎全国性的扩散. 春运返程预期有农民工 2.2 亿人次以上, 学生约 1 亿人次 (http://www.xinhuanet.com/politics/2020-02/16/c_1125580511.htm). 若不管控, 春节返程人口流动是否会造成新冠肺炎感染的回升呢? 这是人们普遍关心的问题.

大家都关注新冠肺炎每天全国新增确诊病人的数量, 但在传染病传播研究中, 研究人员还需要关注每日的新发病病人数量, 以及每日的新增被感染, 处于潜伏期尚未发病的潜在病人数量.

要获得每天发病病人数量, 原则上只要问诊时了解确诊病人的发病时间, 进行统计即可. 但是, 鉴于病人从发病、到就诊、再到确诊需要一定的时间(目前一般平均 5 d), 这样最近 10 d 以内发病的病人还没有全部确诊, 需要等今后 10 d 陆续统计的数据才能得到. 在实际操作中, 特别在每天全国确诊人数高达数千人的情况下, 往往难以及时获得每个患者的发病时间并完成数据的集成统计.

为此, 本文试图采用排队论对数据进行处理. 排队论关注等待时间、排队长度等数量指标的统计规律, 根据这些规律来改进服务系统的结构、优化服务系统, 以最经济的费用最好地满足对象的需求, 在医药卫生体系内也有这类应用^[1-3]. 在排队论中的 Erlang 概率分布可用于确定每日确诊病人中, 从发病到确诊已经等待了多少天的概率.

Erlang 分布的概率密度函数为

$$f_{X_k}(x) = \begin{cases} \frac{\lambda^k x^{k-1} e^{-\lambda x}}{(k-1)!}, & k=1, 2, 3, \dots; x \geq 0, \\ 0, & x < 0, \end{cases} \quad (1)$$

石耀霖 中国科学院大学地球与行星科学学院教授, 中国科学院院士. 长期从事地球动力学基础研究工作. 2003 年, 提出“SARS 流行病传播的系统动力学概率模型及应用”课题, 把地球动力学研究中的一种思想方法应用于 SARS 的传染病学研究, 成功建立了 SARS 传播概率模型.

程惠红 中国科学院大学地球与行星科学学院副教授. 从事计算地球动力学研究工作, 致力于地震孕育及同震形变等高性能并行计算方面的研究工作.

式中, k 为阶数, 影响着概率密度函数的形态; λ 为速率参数, 它的倒数 μ 被称为均值参数, $k\mu$ 为均值的期望. 实际上, Erlang 分布是伽马分布(gamma distribution)的一个特例, 即阶数 k 为整数的特殊情况. 排队论中两个顺序独立事件之间时间间隔为 x 的发生概率遵从二阶 Erlang 分布. Erlang 分布被广泛地应用于排队论中, 例如电话拨号的等候时间^[4], 医学上被用于预测诊断为充血性心力衰竭患者的再住院频率和时间^[5], 管理上优化护士与病人的比率^[6]等问题上.

据国家卫生健康委员会公布的数据, 全国患者从发病到确诊的时间间隔平均是 4.95 d(<https://m.chinanews.com/>

wap/detail/zw/gn/2020/02-17/9095025.shtml). 图1显示了在这次疫情初期北京的病人从发病到确诊等待时间的统计结果, 并与 $\mu=2.5$ 的二阶 Erlang 分布比较. 约 80% 以上的病人等待确诊时间在 9 d 以内.

原则上我们可以将每天确诊病人按此分布分配到前面若干天中, 得到每天的发病人数. 但湖北省的数据有两个特例: 一个是 1 月 27 日, 由于武汉条件改善使核酸检测时间缩短, 湖北省确诊病人数目从前一天的 371 人暴增到 1291 人, 这是由于平均等待时间缩短的结果, 并不是实际病人数目在那天突然增加; 另一个是 2 月 12 日湖北省确诊病人突然暴增达 14840 人, 该数据中只有 1508 人是基于

核酸检测, 13332 人是由于改变了确诊的认定标准, 把临床诊断确诊病人也纳入确诊统计. 这 13332 人也是把以前多日积累的病人一次计入, 他们的平均等待时间超过 5 d. 这两日的数据需要单独处理, 据文献[7]关于 2 月 12 日以前门诊确诊病人数目的统计, 等待时间定为 7 d, 将集中于那日的病人数目按 Erlang 概率密度分配到之前的几日. 图 2 是校正后的每天确诊人数(黑柱)和原始数据(绿柱)的比较. 可以看出消除了两个峰值的影响, 日最高确诊数出现在 2 月 7 日(第 69 日), 单日新增确诊 4000 人. Masè 等人^[8]在模拟数据延迟系列生成中也采用过 Erlang 分布. 本研究均以 2019 年 12 月 1 日作为计算时间起点^[9].

图1 COVID-19 疫情初期北京的病人从发病到确诊间隔天数的分布特征. 蓝柱为实际资料; 红柱为二阶 Erlang 分布理论值

Figure 1 Distribution of delay days interval between patient onset and diagnosis in the early days of COVID-19 in Beijing. The blue column is the actual data and the red column is the theoretical value of the second order Erlang distribution

图2 采用 Erlang 概率分布校正湖北省 COVID-19 每日确诊病人数目与未校正病人数目的比较, 且据校正的每日确诊人数推算每日发病人数和每日感染人数

Figure 2 Comparison of the daily confirmed number of patients before and after Erlang distribution correction in Hubei Province, and daily number of patients onset and infected from daily confirmed number of patient using Erlang distribution

在此基础上,再按照从发病到确诊平均等待时间为5 d,估算出每日发病人数.要想无遗漏地估算最近10余天,特别是最近几天的每天发病人数时,需要知道未来10余天的确诊人数,而这种资料目前无法得知.但是,我们可以对未来10余天确诊人数与现今持平、按一定速率比现今减少、或比现今增加做出不同程度乐观、悲观的多种可能估计.湖北省在严格管控下今后每天确诊人数在2月18日1693人的基础上,可按前10 d的平均速率每天确诊人数减少约100人估计.这些不同的估计仅仅影响大约2月9日以来的每日发病人数的估算,更早的结果不受影响.由图2可以看出,每日发病人数的高峰发生在1月31日(第62日),达3000人.

我们根据平均潜伏期(即从被感染到发病的平均等待时间)从每日发病人数估算出每日被感染人数. Backer等人^[10]对88例病例统计认为潜伏期为6.4 d, Li等人^[11]对425例统计认为潜伏期平均值为5.2 d, Guan等人^[12]对552例研究认为潜伏期中位数为3 d(遵从Erlang分布,相当于平均值为4.03 d).据试验两者引起的差别一般在10%以内,因此我们采用了介于6.4~4.0之间的5.2,结果见图2.可以清楚看出感染、发病、确诊的滞后特征,且每日被感染人数高峰发生在1月26日(第57日)左右,每日被感染人数高达约2800人.

另外,由图2可以通过比较上升期和下降期不同阶段感染人数、发病人数和确诊人数,显示为什么回程春运不大可能造成疫情的大规模反弹.先看1月疫情发展期间,已确诊的病人数目只是冰山一角,已经感染未发病的潜伏期病人和已发病未就诊或未确诊的病人数远大于已确诊的病人数目.从第50~54天,即武汉封城前5 d,已经有

累计15000多人被感染而尚未发病,约3300人已经发病,但还没有被确诊.整个湖北在武汉封城前累计被感染人数估计超过32000人,按武汉离开与留守人口比率约为5/9考虑,并按离城人口中有30%流向湖北以外全国各地计算,有5500多个患者或潜伏患者流往全国各地.截至2月18日北京累计有约393名患者,大约一半与武汉有接触史,按此比例,全国湖北省以外其他省市目前确诊的12612名病人中,有5500名患者可能来自武汉的估计与实际情况大体吻合.最新统计结果表明,截至2月11日,全国除去湖北以外的患者中有武汉相关暴露史的人数达5900余人^[7].这也佐证了我们利用排队论Erlang分布对不规整原始数据进行校正处理的可信性.

据中国疾病预防控制中心新型冠状病毒肺炎应急响应机制流行病学组提供的截至2月11日的实际统计资料(图3),疫情从1月24日进入发病高峰期,2月1日为发病高峰,达3100余人,与我们估算1月31日为高峰、峰值约3000人很好地吻合.实际资料是1月24日报告确诊人数为500人,发病人数已经达到2600人,我们分别估计为370人和2200多人,也大致吻合.然而,我们采用的方法简单易行,仅仅利用每天公开发表的资料,迅速可以做出更新估计.对于最近10 d的每天发病人数统计,实际统计还需要等待10余天才能获得完备资料,而我们的方法可以对在乐观、悲观或一般的趋势预期下,获得有参考价值的估计.更重要的是,我们还可以估计出实践中难以完全获得的信息,例如每天新感染人数.在对潜伏期有一定了解的基础上,可以估算每天新感染人数,第57日(1月26日)感染人数达到峰值2800人.

在上述讨论中涉及的一个重要问题是未来10余天确

图3 中国疾病预防控制中心研究人员^[8]提供的截至2020年2月11日的实际统计资料.蓝色表示每日新发病人数,红色表示每日报告的确诊人数.由于许多12 d以前发病的病人还没有就诊或确诊,因此2月2日以来的发病人数尚有缺失,越临近2月12日缺失越严重,因此图中右端将图像模糊化表达.据中国疾病预防控制中心^[7]修改

Figure 3 The patient statistics provided by the researchers of China Center for Disease Control and Prevention (CDC)^[8] as of February 11, 2020. The blue column shows the daily number of new cases; the red column is the daily reported confirmed number of patient. Note that since many patients who had been onset 12 d go have not yet seen a doctor or been diagnosed, the number of onset patients has been missing since February 2. Moreover, the missing is more serious around February 12. Therefore, we blur the right of this figure. Modified according to CDC^[7]

诊人数尚属未知, 疫情的不同发展对本研究每天发病病人数目计算会有多大影响问题. 对此我们进行了敏感性检验, 未来 15 d 按确诊人数从第 80 天 1700 人基础起, 每天增加 100 人, 每天持平, 和依据传染动力学模型预测的因为防疫措施使传染数 R 降低, 在没有意外事件条件下持续减少的 3 种情况进行了比较, 结果如图 4. 虽然对最后 10 余天的数值有显著影响, 但对更早的峰值出现时间相差仅 1 d、峰值大小相差不到 10%. 图 4 也可以反过来解读, 与传染动力学模型相结合, 如果不能把传染数控制在红线对应的传染数之下, 疫情就会持续. 因此, 这对于已有措施是否足够, 要采取什么程度的防控措施, 可以提供参考.

既然用 Erlang 概率分布计算发病人数和感染人数可信, 据此可探讨春节前疫情暴发阶段和现在疫情下降和趋于尾声阶段的差异. 图 5 是每天累计感染人数、累计发病人数、累计校正的确诊人数和原始报告确诊人数随时间变

化图, 可以看出它们之间存在滞后的关系. 图 6 则是图 5 中每天累计感染人数和累计发病人数的差值, 反映该日已感染但尚未发病的人数, 如果他们没有被发现和监控, 则发病以后会成为移动的感染源, 成为最危险难防的人群, 而在疫情开始阶段他们的相对数量也最多; 反之在疫情被控制的后期, 这类人的数目会显著地减少. 图 6 也给出累计发病人数和累计确诊人数的差值, 反映已经发病有症状有传染性但还没有被确诊的人数, 如果他们还没有被作为疑似患者隔离, 而是未就诊或居家休养, 则有很大传染别人的危险.

是否会发生大规模跨地域感染取决于两个因素: 一是有多少病人或已感染尚未发病的感染者, 他们数量越多则危险越大; 二是他们是否得到了监控, 如果没有得到监控而自由流动则扩散危险大. 图 5 和 6 显示出暴发阶段的特点是已感染未发病人员和已发病未确诊人员的数目, 可

图 4 假设未来 15 d 确诊人数变化对计算的每天发病人数曲线影响的敏感性试验

Figure 4 Sensitivity test of the impact of the confirmed number of patients in the next 15 d on the daily number of patients onset

图 5 每日累计病人数随时间的变化. 感染人数(蓝色)、发病人数(红色)、校正的确诊人数(黑色)和原始报告确诊人数(黄色)

Figure 5 Daily cumulative numbers vary with time. Blue presents the daily cumulative number of infected patients; red means the daily cumulative number of patients onset; black shows the daily cumulative confirmed number of patients after correction and yellow is the original daily cumulative confirmed number

图6 累计感染人数与累计发病人数差值(蓝色)随时间变化,反映每天已感染但尚未发病的人数;累计发病人数和累计确诊人数的差值(绿色为与已校正的确诊人数差值,红色为与原始报告确诊人数差值),反映每天已发病但尚未就诊确诊的病人数

Figure 6 The difference (blue) between the daily cumulative number of patients infected and onset over time, which reflecting the number of patients who have been infected but have not yet been diagnosed. And the difference between the daily cumulative number of patients onset and confirmed (green is the difference between the daily cumulative number of corrected confirmed patients, red is the difference between the daily cumulative number of confirmed people who have been diagnosed with the original report), which reflecting the number of patients who have been onset but have not yet been diagnosed

以远远大于已经确诊的病人数目。以1月23日封城时为例,虽然正式宣布的当日新增仅105例,累计确诊人数只有549人,医学观察也只有2776人(http://wjw.hubei.gov.cn/bmdt/ztzl/fkxxgzbdgrfyyq/xxfb/202001/t20200124_2014808.shtml),但图6显示那时累计感染但尚未发病的人数约达13500人,发病而尚未确诊的人数约达18000人。这30000多人基本不受管控,这是春节前大规模人口流动造成COVID-19在全国大规模扩散的原因。

反之,在疫情已经进入下降期时,即使看起来每天确诊人数可能还比较高,但新感染但尚未发病和已发病尚未确诊的人员数目比新增确诊人员数目还要少。虽然湖北在2月18日确诊病人数仍停留在1693人的高水平(http://wjw.hubei.gov.cn/fbjd/dtyw/202002/t20200219_2129954.shtml),但图6显示截至2月18日累计发病尚未确诊的人数仅有1200人,累计感染尚未发病的人数仅为3700人。考虑到当时在管控之下还有疑似病例3462人、医学观察68345人,所以这4900人逃脱管控的可能性很小。截至2月29日,当天新确诊病人570人(http://wjw.hubei.gov.cn/bmdt/ztzl/fkxxgzbdgrfyyq/xxfb/202003/t20200301_2164974.shtml)。图6显示累计感染尚未发病的人数仅有约110人,这些天累计发病尚未确诊的人数也仅为770人,考虑到当时处于管控之下的还有疑似病人646人、医学观察35766人。这不到900人完全逃脱管控的可能性就更小。病人或被感染未发病的人员数量低,而且绝大多数已经处于管控之下,这是春节后人口流动不会造成大规模反弹的原因。

以上是对湖北省的分析,并没有对全国各省份逐一进行类似分析。但2月18日当天其他省市确诊人数仅为湖北省的约3.4%(58/1693),2月29日仅为湖北省的0.5%(3/570),按此比例其他省份这段期间全部染病未发作和发病未确诊人数在百人以下或更少,而且这类人很多已经在管控监管之下,不会像1月份以前那样随意流动。因此,绝对不会再发生类似春节前大量感染者从武汉流入各地的情景。这些定量数据表明,应该对不同地区分类管理,分批次稀释返程人群密度,在交通运输过程中注意卫生防疫,在出发地做好体温监测,对流动人口登记或通过大数据手段掌握其流动过程,便于在一旦出现病例时追溯监控,在接收地做好社区或单位的接收和隔离监管。只要保证万一发现偶发病例,能够追踪其路径和查寻其路上交通工具和密切接触者人群,坚持谨慎科学的管理,错峰返程运作一定会成功。春节后回程人口流动产生疫情反弹的可能性很小。当然要实现不反弹的前提是思想不能懈怠、必须维持的基本措施不能放松。

另外值得注意的是,境外国家(地区)新冠肺炎疫情正处在发展期,而且其中有些病人传染链并未被掌握,参照本研究揭示的武汉初期的一个重要特征是,对于疫情暴发阶段传染链未被识别的病人群,发病而未确诊的人数可能远大于确诊人数,而已被感染尚未发病的人数又可能大于发病人数。因此在防范人口流动造成疫情反弹时,来自国外的疫情传播风险已远大于国内节后农民工和学生返程的风险,防止境外传入成为新的防控重点。

致谢 感谢国家自然科学基金(40344007)资助;感谢三位匿名审稿人对文章提出的意见。

推荐阅读文献

- 1 Fomundam S, Herrmann J W. A Survey of Queuing Theory Applications in Healthcare, Technical Report. The Institute for Systems Research, 2007. 24
- 2 Zhao S J, Yang Z, Zhong Y J. BPR of medical workflow based on queuing theory in hospital (in Chinese). *Chin Hosp Manag*, 2003, 23: 10–12 [赵树进, 杨哲, 钟拥军. 基于排队论的医疗工作流程重组. *中国医院管理*, 2003, 23: 10–12]
- 3 Peng Y C, Dong S B, Chang W H. Measuring the efficiency of the outpatient process with queuing theory model (in Chinese). *Chin J Hosp Admin*, 2005, 21: 806–809 [彭迎春, 董斯彬, 常文虎. 运用排队论模型测量医院门诊流程效率. *中华医院管理杂志*, 2005, 21: 806–809]
- 4 Lin Y B, Chlamtac I. Effects of Erlangcall holding times on PCS call completion. *IEEE Trans Veh Technol*, 1999, 48: 815–823
- 5 Bardhan I, Oh J H, Zheng Z, et al. Predictive analytics for readmission of patients with congestive heart failure. *Inform Syst Res*, 2015, 26: 19–39
- 6 de Véricourt F, Jennings O B. Nurse-to-patient ratios in hospital staffing: A queueing perspective. ESMT Working Paper No. 08-005, 2008
- 7 The Novel Coronavirus Pneumonia Emergency Response Epidemiology Team. Vital Surveillances: The Epidemiological Characteristics of an Outbreak of 2019 Novel Coronavirus Diseases (COVID-19)—China, 2020. *China CDC Weekly*, 2020, 2: 113–122
- 8 Masè M, Faes L, Nollo G, et al. Determination of synchronization of electrical activity in the heart by Shannon entropy measure. In: Amsterdam M, ed. *Recent Advances in Multidisciplinary Applied Physics*. Amsterdam: Elsevier, 2005. 235–239
- 9 Huang C L, Wang Y M, Li X W, et al. Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. [https://doi.org/10.1016/S0140-6736\(20\)30183-5](https://doi.org/10.1016/S0140-6736(20)30183-5)
- 10 Backer J A, Klinkenberg D, Wallinga J. Incubation period of 2019 novel coronavirus (2019-nCoV) infections among travellers from Wuhan, China, 20–28 January 2020. *Eurosurveillance*, 2020, 25: 2000062
- 11 Li Q, Guan X, Wu P, et al. Early transmission dynamics in Wuhan, China, of novel coronavirus-infected pneumonia. *N Engl J Med*, 2020, 382: 1199–1207
- 12 Guan W, Ni Z, Hu Y, et al. Clinical characteristics of 2019 novel coronavirus infection in China. *N Engl J Med*, 2020, doi: 10.1056/NEJMoa2002032

Summary for “关于春节返程人口流动对新型冠状病毒肺炎(COVID-19)疫情影响的讨论”

Impact of returning population migration after the Chinese Spring Festival on the COVID-19 epidemic

Yaolin Shi^{1,2}, Huihong Cheng^{1,2*}, Tianxiang Ren³ & Luyuan Huang⁴

¹ Key Laboratory of Computational Geodynamics, Chinese Academy of Sciences, Beijing 100049, China;

² University of Chinese Academy of Sciences, Beijing 100049, China;

³ Chinese Academy of Geological Sciences, Beijing 100037, China;

⁴ Institute of Crustal Dynamics, Chinese Earthquake Administration, Beijing 100085, China

* Corresponding author, E-mail: chenghuihong@ucas.ac.cn

The outbreak of the novel coronavirus disease 2019 (COVID-19) and its spread throughout the China have caused a huge impact on China and the international community. And now it becomes a worldwide infectious disease which poses a major threat to the lives of people around the world. What is worth noting about China is five million people left Wuhan before the Spring Festival, which caused the nationwide spreading of COVID-19 epidemic. Then, it raises a question of concern, should the return of migrant workers and students after the Spring Festival cause an increase in the epidemic? In this study, we use the discrete stochastic model (DSM) to study the transmission dynamics of COVID-19. The DSM is different from the classical continuous variable ordinary differential equations, and has two characteristics. First, on account of few patients at the beginning of the COVID-19 and random fluctuations during the transmission process are prominent; the DSM can better reflect the initial transmission characteristics than the continuous variable deterministic ordinary differential equation model. Second, the DSM can easily track the changes in the epidemic situation, and well reflect the infectious rate varies with time due to different prevention and control measures, and then gradually estimate the development of the epidemic situation. Meanwhile, based on the facts that there are successive time lags among epidemic infection, symptom onset, and diagnosis confirmation, the Erlang probability density distribution, which is frequent used in the queuing theory, has been applied to the calculation of numbers of epidemic daily outbreaks and daily infections respectively from the corrected numbers of patients diagnosed and confirmed reported every day since the outbreak of the epidemic in Hubei Province. The number of symptom onset patients we calculated agrees well with recent statistics made by the Chinese Center for Disease Control and Prevention (CDC), showing the feasibility of our method. The calculation results indicate that in the rising stage of the epidemic, although the number of newly diagnostic confirmed patients reported before the day of lock down of Wuhan city was only 180 daily, the number of newly infected people may have reached about 2500, and the cumulative number of infected may reach 33000. More than 5500 people may have rushed out of Hubei Province uncontrollably, thus causing a national epidemic of COVID-19. However, it is in the epidemic decline stage now. Even if Hubei's daily confirmed diagnosis remains at a high level of more than 1700 people a day, the number of newly infected people may be less than 800 nowadays, and most of whom may already be concerned about quarantine. The number of infections in other provinces and cities is much lower than that in Hubei. Therefore, as long as the epidemic situation in different regions is distinguished, return trips could be arranged at reduced people density, attentive epidemic prevention of transportation, and well preparation at the receiving cities. Rebound of the epidemic is quite unlikely. Although, the possibility of an epidemic rebound is small, only in conditions of slack thinking and strict measures are carried out. In accordance with the transmission of the world epidemic, more attention must be paid to the inspection of the influx of foreign infected people.

COVID-19, returning population migration, queue theory, Erlang probability distribution

doi: 10.1360/TB-2020-0151