

病毒如何传播之不确定性未来

(中国科学院理化技术研究所 戴 闻 编译自 Susanna Manrubia. *Physics*, November 16, 2020)

长期以来,天文学家精确的预测令世人惊讶。预言3000年4月26日,日全食可观察带,将顺着南美洲、大西洋和北非一线延伸,这一预测没有人反对。但试图描述病原体未来传播的科学家,必须面对非常大的不确定性。研究人员既缺乏关于传染病如何传播的可靠数据,感染能力的微小差异又可以成倍地扩大病毒的传播,从而难以预测各类流行病的严重程度。来自伊利诺伊大学的Nigel Goldenfeld和他的同事,最近以伊利诺伊州为例,开发了一个模型,评估非药物干预时COVID-19病毒传播的短期影响。研究者还探索和量化了模型的局限性,阐明了它无法预测的场景。他们使用该模型,向公民、临床医生和决策者,作出了示范性展示。

用于预测COVID-19扩散的绝大多数模型是所谓的隔离间模型,它将群体中的个体分组,纳入隔离间。最常用的SIR隔离间是:易感(S)、感染(I)和移除(R),R表示死亡或已康复的群体。其他隔离间可以包括无症状的、行动受限的(非感染个体)、检疫隔离、住院和接种疫苗者。更复杂的COVID-19模型包括:



“隔离间”模型将个体分组到一个个隔离间,如住院组(左)、受限组(中)和感染组(右),然后根据人在隔离间之间的移动方式,预测病毒将如何传播

在SIR动力学之上的附加规则,这些规则可以解释隔离间的地理依赖定义,如个人的流动性等。

广泛使用SIR模型来研究COVID-19传播路径的一个令人信服的原因是:这些模型在复制过去的过程时,证明了惊人的有效性。但是,再现已经发生的事情并不意味着模型可以准确地预测未来——为了使模型适合过去的的数据而获得的参数值,并不一定适合未来数据。在这方面,SIR模型有一些局限性,这些局限性来自:一是模型无法代表现实世界的复杂、分布式和异构性质;二是缺乏准确的数据来训练模型。虽然这两个问题可以被潜在解决,但第三个问题不能解决:使用SIR模型进行的预测,对分配给模型的参数初始值的微小变化特别敏感。这种敏感性不可避免地将预测能力限制在不远的将来。还有第四个关键角色,个人和团体对COVID-19扩散的反应差别。

Goldenfeld和合作者开始构建自己的模型时,非常清楚SIR模型的优点和局限性。他们为预测COVID-19在伊利诺伊州的传播,进行了量身定做。他们的模型是一个所谓的感染年龄变体,它“记住”一个人何时到达一个隔离间,然后使用这个时间来计算经过一段时间延迟后移动到下一个

隔离间的可能性。这一特性在大多数SIR模型中没有。

为了获得模型参数(共有22个参数),该小组采用了数值“马尔可夫链蒙特卡罗”方法。这个方法可以评估给定参数集与输入数据的兼容性。为了训练模型,该团队输入了来自伊利诺伊州公共卫生部的数据,其中包括COVID引起死亡的每日数量和COVID阳性患者占用重症监护床位的每日数量。使用这些可靠的数据为预测增加了鲁棒性。

他们预测,如果不迅速实施居家命令,感染数量将呈指数增长,尽管第一次开始研究该模型时,伊利诺伊州只有19例。预测结果强烈表明,有必要发布这样一个命令,这就是伊利诺伊州州长在阅读了他们的研究结果后所做的决定。SIR模型预测的一个必要条件是规则保持不变,与模型包含的规则相匹配,并且个人都遵循规则。正如《纽约时报》报道的那样,当COVID-19阳性的大学生去参加聚会和上课,而不是像预设的那样在宿舍隔离时,Goldenfeld等人所做的预测就失败了。

大流行建模的有效性时间尺度,仍然是一个悬而未决的问题。我们需要改进病毒建模基础设施,跟踪病毒感染的全球数据收集系统,并实时运行模型。这些模型必须考虑到一个群体改变其行为以应对危机的能力,以及病原体的特定生物学特性。

· 124 ·