



# Measuring the Impact of Public Transit on the Transmission of Epidemics

Yuan Bai<sup>1</sup>, Qiuyang Huang<sup>2</sup>(✉), and Zhanwei Du<sup>1</sup>

<sup>1</sup> Department of Integrative Biology, University of Texas at Austin,  
Austin 78705, USA

<sup>2</sup> College of Computer Science and Technology, Jilin University,  
Changchun 130117, China  
[huangqy17@mails.jlu.edu.cn](mailto:huangqy17@mails.jlu.edu.cn)

**Abstract.** In many developing countries, public transit plays an important role in daily life. However, few existing methods have considered the influence of public transit in their models. In this work, we present a dual-perspective view of the epidemic spreading process of the individual that involves both contamination in places (such as work places and homes) and public transit (such as buses and trains). In more detail, we consider a group of individuals who travel to some places using public transit, and introduce public transit into the epidemic spreading process. Our simulation results suggest that individuals with a high public transit trip contribution rate will increase the volume of infectious people when an infectious disease outbreak occurs by affecting the social network through the public transit trip contribution rate.

**Keywords:** Community structure · Subway system · Early prediction

## 1 Introduction

Epidemiological research is complex and involves aspects such as public policies [1]. There are an increasing number of applications and tools that support real-world decisions in public health epidemiology. It is critical to use mathematical models to analyze epidemic spreading in public health epidemiology, to help overcome the problems of sparse observations, inference with missing data, and so on [1, 2]. These models can describe the spreading mechanisms of viruses, quantify the interventions' effects, and identify key factors related to the course of an outbreak [2]. Infectious disease spreading models are powerful tools for controlling the development of infectious diseases [2]. Mathematical and computational models have proven useful when addressing the 2014 Ebola outbreak, the 2009 H1N1 outbreak, and so on.

Consider the difference between urban transport in China and other Western countries; people would like to choose public transit for their daily travel needs. In contrast, in some cities in China, such as Beijing and Shanghai, the average public transit trip contribution rate can be as high as 40%. However,

individual-based models are important tools for studying the transmission mechanism of pandemic influenza [3]. With regards to individual-based models, many of these models ignore the possible risk inherent in commuting. For example, a heterogeneous graph modeling method was used by Dongmin et al. to describe the dynamic process of influenza virus transmission using clinical data [4]. The VirusTracker app simulates the spread of a virus and highlights the critical role of vaccinations in combating a disease outbreak [5]. Dynamic Behavior Visualizer is a tool for visualizing people’s dynamic behaviors and movements in a disaster [6]. However, these aforementioned models primarily assume that people are infected at a fixed location. Public transit, as a high-risk place where different crowds contact each other, is often neglected. This overlook can be understood, particularly in some developed countries (such as the United States) where almost ninety percent of the people travel by private vehicles and few travel by public transit or other modes of public transportation [7].

These methods cannot describe the infection risk due to crowded public transit, such as that in large cities in China. Thus, we can see that the intervention of public transit can affect the trend of epidemic diseases to some extent and enlarge the transmissibility threshold to some extent (see the analysis of the transmissibility threshold in the Methods section). In this regard, a new model is needed that considers urban public transit. In this respect, our work may provide a means for modeling the impact of public transit trips and for estimating the effectiveness of infection controls during public transit trips.

## 2 Materials and Methods

We use the data of a resident in Blacksburg from <http://ndssl.vbi.vt.edu/synthetic-data/> to construct the social network. We know when and where the resident is for one day. Moreover, the public transit information data can be obtained from <http://www.gtfs-data-exchange.com/agency/blacksburg-transit/>. With the public transit information, we can schedule people’s daily trips with the help of Google Maps. The resident information includes resident identification, start time of resident activity, resident location, the duration in this location, and so on. The public transit information includes public transit stops, trips, stop times, routes, and so on. Then, we propose a novel modeling framework for describing the dynamic process of individuals transmitting influenza virus by integrating public transit. In this model, we use the chain-binomial model in susceptible-exposed-infected-recovered (SEIR) models based on [8] to simulate the disease transmission process in public transit.

### 2.1 Methods

We use a general computational approach for networked epidemiology based on [10], which can generate a social contact network of the region under consideration. Three main steps are involved in the process of constructing synthetic populations [10]: Step 1 constructs an artificial population with

open-source databases. Step 2 connects daily activities to individuals for each household through daily surveys [10,11]. Step 3 assigns public-transit-based activities between the two activity locations. Demographics and home locations are considered here. The social contact graph is constructed with activity information for each person. Google Maps can be used here to compute the public-transit-based activities between the two activity locations.

*Individual-Based Model.* We embedded an individual-based model in the PTF model to denote the virus transmission mechanism by considering the travel of an individual using public transit. The individual-based model of an epidemic features a dynamic process, including factors such as individuals visiting places, visiting public transits, and daily infection transmission.

To model this process, we defined the following functions. The subject's infection risk at location  $p$  is defined as:

$$\tau(i, j, p) = \beta \cdot c \cdot t(i, j) \quad (1)$$

where  $\beta$  indicates the infection rate.  $t(i, j)$  represents the contact time between individual  $i$  and individual  $j$  at location  $p$ .  $c$  represents the contact rate and is set as two different values according to the current location, which belongs to the public transit set  $S_{\text{public transits}}$  or the place set  $S_{\text{places}}$ . It is assumed that if someone remained near a symptomatic patient for more than  $h$  hours in the public transit, then the infection rate for this person is 100% of  $c_p$ , where  $a$  is the infectiousness at a certain location. If the person remained at some location for  $h$  hours, then the probability decreases proportionally to the duration.

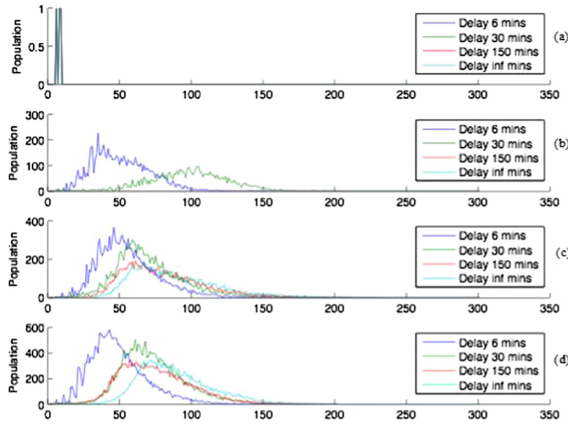
$$c = \begin{cases} \alpha \cdot c_p & \text{if } p \in S_{\text{public transits}} \\ c_p & \text{if } p \in S_{\text{places}} \end{cases} \quad (2)$$

The infection force of susceptible individual  $i$  caused by the infected neighbors  $j(j = 1, \dots, S_i)$  at location  $p$ :  $\lambda(i, p) = 1 - \prod_{S_i}^{j=1} (1 - \tau(i, j, p))$

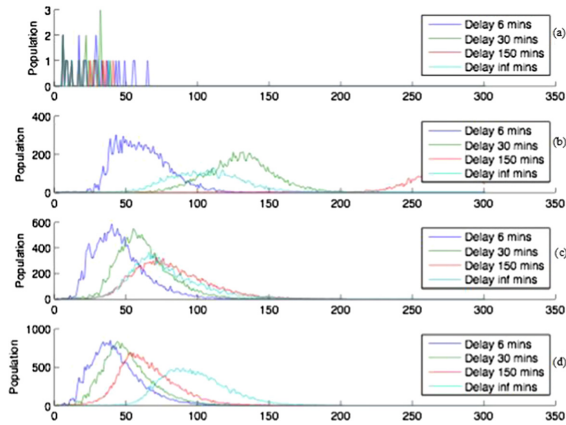
where  $\tau(i, j, p)$  is the infectivity of infected contact  $j$  at location  $p$ , capturing the probability of infected individual  $j$  infecting others.

### 3 Results

Our PTF considers the process. To measure the effect of an individual visiting public transit during the epidemic, the following experiments are conducted by setting different public transit trip contribution rates in the different basic reproductive number  $R_0$  (see Figs. 1, 2 and 3). From these three figures, we can see that as the number of people who take public transit is reduced, less people will be infected. In each figure, a higher  $R_0$  corresponds to a larger infected population. In each subplot, a larger time threshold  $h$  corresponds to a lower infected population. Moreover, when only 10% of people take public transit, the infected population is almost the same for the different  $h$  values (30, 150 and  $\infty$ ), as shown in Fig. 1. This means that when there are only a few people who

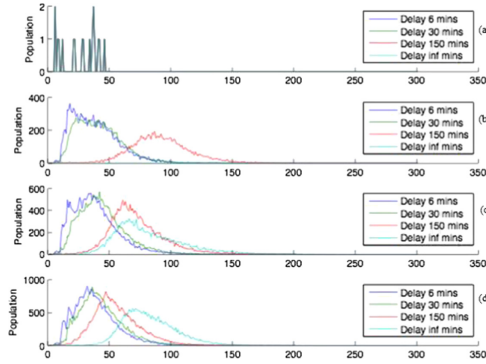


**Fig. 1.** Simulation of the influenza epidemic curve with 50% PTTCR. The number of people who take public transit is reduced by 50%. Additionally, seven scenarios were simulated for this intervention. In the four scenarios, the time threshold  $h$  (see Methods) in public transit is increased by 6 min, 30 min, 150 min and an infinite number of minutes, and the basic reproductive number  $R_0$  is increased by (a)  $R_0 = 0.9$ , (b)  $R_0 = 2.1$ , (c)  $R_0 = 4.0$ , and (d)  $R_0 = 4.5$



**Fig. 2.** Simulation of the influenza epidemic curve with 30% PTTCR. The number of people who take public transit is reduced by 30%. Additionally, seven scenarios were simulated for this intervention. In the four scenarios, the time threshold  $h$  (see Methods) in public transit is increased by 6 min, 30 min, 150 min and an infinite number of minutes, and the basic reproductive number  $R_0$  is increased by (a)  $R_0 = 0.9$ , (b)  $R_0 = 2.1$ , (c)  $R_0 = 4.0$ , and (d)  $R_0 = 4.5$

take public transit, the intervention performance of public transit cannot play a large role in the epidemic process. Regardless of how much effort is expended here, we cannot obtain a substantial improvement.



**Fig. 3.** Simulation of the influenza epidemic curve with 0% PTTCR. The number of people who take public transit is reduced by 0%. Additionally, seven scenarios were simulated for this intervention. In the four scenarios, the time threshold  $h$  (see Methods) in public transit is increased by 6 min, 30 min, 150 min and an infinite number of minutes, and the basic reproductive number  $R_0$  is increased by (a)  $R_0 = 0.9$ , (b)  $R_0 = 2.1$ , (c)  $R_0 = 4.0$ , and (d)  $R_0 = 4.5$

Rather, when approximately 20% of people take public transit, the infected population is different for the different  $h$  values (30, 150 and inf), as shown in Fig. 3. This means that when there are many people who take public transit, the intervention performance of public transit will play an important role in the epidemic process. We can obtain a substantial improvement if we can control the spreading in public transit. School closure is a common intervention in an epidemic.

### 4 Conclusion

Our simulation results suggest that individuals with high public transit trip contribution rates will increase the number of infectious people when there is an infectious disease outbreak, similar to the school closure intervention, guiding similar research in other fields. We conclude that the public transit trip contribution rates will have an impact on the process of the spread of an infectious disease because they can affect the social network. In this respect, our work provides a means for modeling the impact of public transit trips and for estimating the effectiveness of infection controls for public transit trips.

**Acknowledgments.** This work was supported by the Open Fund of Key Laboratory of Urban Land Resources Monitoring and Simulation, MNR (No. KF-2019-04-034).

## References

1. Marathe, M., Vullikanti, A.K.S.: Computational epidemiology. *Commun. ACM* **56**(7), 88–96 (2013)
2. Lofgren, E.T., Halloran, M.E., Rivers, C.M., et al.: Opinion: mathematical models: A key tool for outbreak response. *Proc. Natl. Acad. Sci.* **111**(51), 18095–18096 (2014)
3. Ohkusa, Y., Sugawara, T.: Application of an individual-based model with real data for transportation mode and location to pandemic influenza. *J. Infect. Chemother.* **13**(6), 380–389 (2007)
4. Guo, D., Li, K.C., Peters, T.R., et al.: Multi-scale modeling for the transmission of influenza and the evaluation of interventions toward it. *Sci. Rep.* **5** (2015)
5. VirusTracker. <http://ndssl.vbi.vt.edu/apps/virustracker/>
6. DBV. <http://ndssl.vbi.vt.edu/apps/dbv/>
7. Santos, A., McGuckin, N., Nakamoto, H.Y., et al.: Summary of Travel Trends: 2009 National Household Travel Survey (2011)
8. Hladish, T., Melamud, E., Barrera, L.A., et al.: EpiFire: an open source C++ library and application for contact network epidemiology. *BMC Bioinf.* **13**(1), 76 (2012)
9. Swarup, S., Eubank, S.G., Marathe, M.V.: Computational epidemiology as a challenge domain for multiagent systems. In: *Proceedings of the 2014 International Conference on Autonomous Agents and Multi-agent Systems*. International Foundation for Autonomous Agents and Multiagent Systems, pp. 1173–1176 (2014)
10. Eubank, S., Guclu, H., Kumar, V.S.A., et al.: Modelling disease outbreaks in realistic urban social networks. *Nature* **429**(6988), 180–184 (2004)
11. Barrett, C.L., Beckman, R.J., Khan, M., et al.: Generation and analysis of large synthetic social contact networks. In: *Winter Simulation Conference*. Winter Simulation Conference, pp. 1003–1014 (2009)