Introduction for the literature review

1. Background

The spread of an infectious disease through a population is an inherently spatial process. Infection happens frequently recent years, especially at high densely-populated regions. Many researchers try to reveal about the spatial distribution of infection based on the mobility data. Mobility data is a dataset, what describes the mobility behaviour of people in a particular area, such as the journey-to-work data in a country. Many researches have done research at national or international scale, and also others focus on infection that may occur within individual buildings such as school or hospitals. Mobility data used by these kinds of researches is comparably easy to analyse. There are just a few at metropolitan scale(Moss, Naghizade, Tomko & Geard, 2019). Naturally, how to gain this kind of data is the most important and difficult part in this kind of researches. Yu Zheng et al. proposed the whole process of trajectory data mining and pointed out the possible algorithms used in each section in this process. But there are still some problems cannot be solved: (1) the selection of research scale (e.g. at national scale), if the scale is too small, the possibility that the analysis can be affected by report bias is increased. (2) some important data is hard to collect, such as those characterise private features. (3) it is difficult to determine whether a person is infected or not, and it also involves privacy problems.

1. Outline the question

This research will focus on using existing methodologies and algorithms to mine new trajectory dataset at urban scale that consider the factors that may affect the chance of infecting, such as the length for a person staying at one place and so on.

1. The step to solve the problem

Firstly, I will read literature involves methodologies (such as Mean (or Median) Filter which is used to remove outlier, or how to convert a sequence of raw latitude/longitude coordinates to a sequence of road segments) and algorithms used to mine trajectory dataset (such as “Trajectory Data Mining: An Overview” by Yu Zheng et al. and what it refers), and some papers reveal about the distribution of infection at global, urban and smaller scale. Then, I will learn how to apply the methodologies and algorithms to raw data. Some knowledge of building model based on the dataset is also needed.

1. Reference

Moss, R., Naghizade, E., Tomko, M., & Geard, N. (2019). What can urban mobility data reveal about the spatial distribution of infection in a single city?. BMC Public Health, 19(1). doi: 10.1186/s12889-019-6968-x

Zheng, Y. (2015). Trajectory Data Mining. ACM Transactions On Intelligent Systems And Technology, 6(3), 1-41. doi: 10.1145/2743025

Rolls, D., Geard, N., Warr, D., Nathan, P., Robins, G., & Pattison, P. et al. (2015). Social encounter profiles of greater Melbourne residents, by location – a telephone survey. BMC Infectious Diseases, 15(1). doi: 10.1186/s12879-015-1237-9

Eames, K., Bansal, S., Frost, S., & Riley, S. (2015). Six challenges in measuring contact networks for use in modelling. Epidemics, 10, 72-77. doi: 10.1016/j.epidem.2014.08.006