An Analysis of Beijing HFMD Patients Mobility Pattern during Seeking Treatment

Zhaonan Zhu, Zhidong Cao, and Daniel Dajun Zeng

The State Key Lab of Intelligent Control and Management of Complex Systems
Institute of Automation, Chinese Academy of Sciences, Beijing, China
{zhaonan.zhu,zhidong.cao,dajun.zeng}@ia.ac.cn

Abstract. Previous epidemiological researches have studied HFMD transmission pattern, but the study on the patients mobility pattern while seeking treatment is absent. In this paper, we present a statistical analysis of the spatial-temporal pattern of the Beijing HFMD patients mobility based on a complete works dataset containing over 40,000 cases in 2010. The main findings are as follows: (1) the patents incline to take their sick children to hospitals on weekdays, especially on Monday; (2) the patients living far from city center are more likely to get to hospitals in the afternoon, while the patients in the downtown in the morning; (3) patients (or their parents) either select the nearby hospitals, or the Class-A hospitals in the downtown. Furthermore, we employ a gravity model to describe the spatial mobility pattern of patients. The experiment results show a good fit.

Keywords: Hand-foot-mouth disease, mobility pattern, spatial-temporal analysis, gravity model.

1 Introduction

Hand-foot-mouth disease (HFMD), caused by Coxsackie A16 virus, can transmit among 0 to 5-year-old children through close contact, tableware and human excrement. Despite the fact that the mortality of HFMD is not much high (for example, in 2010, 19 out of 45409 cases died, so that is 0.022%), this disease brings pain to sick children and their families. Additionally, the number of HFMD patients is increasing in recent years (18445 cases in 2008, 24483 in 2009 and 45409 in 2010, from Beijing CDC), and the seasonal outbreaks of HFMD make the government suffer a high cost in preventing and controlling the transmission.

Over the past years, many researchers have conducted related research ranging from molecular biology to epidemiology. Cao et al. [1] gave an empirical analysis of the 2010 Beijing HFMD epidemiological data, revealing some spatial-temporal patterns of HFMD transmission. Chen et al. [2] analyzed epidemiologic features of Taiwan HFMD epidemic.

In the city epidemic, infected people from all over the city seek medical service. Large numbers of HFMD patients from all districts to all hospitals may cause potential risk of cross-transmission, for the patients may have close contact with the others

D. Zeng et al. (Eds.): ICSH 2013, LNCS 8040, pp. 24–30, 2013.

in the crowded public places such as buses, subway stations and hospitals. And for hospitals and public health department, the macro spatial-temporal law of the patients' mobility is very important for epidemiological survey and epidemic controlling.

Previous study mainly focused on the epidemiology mechanism, but few work studied the patients' mobility pattern when they are seeking treatment. Some human mobility-related work [4] proposed some basic spatial-temporal analysis method applied in the field of commuting data analysis, social media[5] and infectious disease[3].

The remainder of this paper is structured as follows. We begin with an introduction to the data used in this study. Then we report a spatial-temporal analysis of the data. Finally, we adopt an unconstrained gravity model to fit the commuting data, and the results analysis are also shown.

2 Data

2.1 Data Format

The 2010 Beijing HFMD case dataset provided by Beijing Center for Disease Control and Prevention (Beijing CDC) consists of detailed information of all 45409 cases that occurred in 2010. The format of the data is as follows: {caseID, gender, date of birth, profession, residential address, hospital, date of onset, date and time of visiting hospital}.

2.2 Data Preprocessing

As the raw data was filled by doctors at first, some cases' information is absent or fallacious. The cases without detailed address or the addresses of which do not belong to Beijing are removed.

We use standard place database to get all the administrative district where each case lives, and locate all the 44607 valid cases with the latitude and longitude information (accurate to the street office, which is the basic residential administrative unit in China, and there are over 340 street offices in Beijing) and all the 195 hospitals on the map of Beijing. All the 44607 cases live in 312 basic administrative units.

3 Spatial-Temporal Analysis

3.1 Wheres of the Cases

We locate all the valid cases with the geographic coordinates on the map. By generating the heat map, it is easy to see the spatial features of the cases.

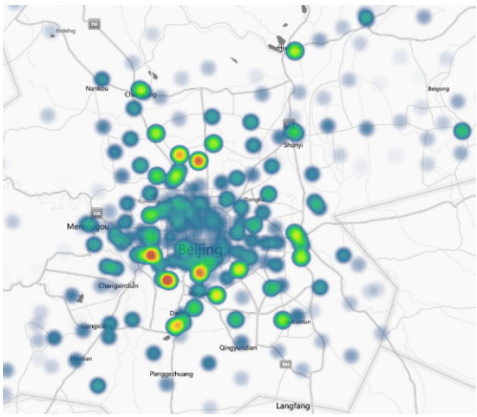

Fig. 1. The heat map of Beijing 2010 HFMD epidemic

We can see from figure 1 that the heat points are mainly located in the junction area of urban and suburban districts. It is because in those area the people commute often from suburban to urban districts. Frequent flow of people may result in more infected cases.

3.2 Whens of Visiting Hospitals

HFMD patients are mostly 0 to 5-year-old children, so they are largely likely taken to hospitals by their adult custodians. When custodians consider when to go, they may account for both the date and the distance. We compare the different distributions of different groups of patients.

First, we compare the patients visiting hospitals on weekends and on weekdays.

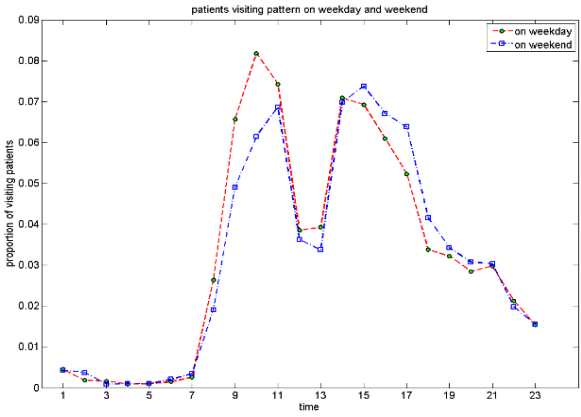

Fig. 2. The number of patients seeking treatment at different time on weekdays and weekends

Figure 2 illustrates that the patients visiting hospitals on weekdays are more willing to go in the morning, while those who seek treatment on weekends are more likely to go in the afternoon. This is because i) people tend to sleep later on weekend; ii) the doctor's duty time on weekend is shorter, and professor medicals are not always on duty on weekend, which results in less medical service seekers who long for better treatment lead by professor medicals;

Second, we compare the patients with a range of distance from living place to hospitals. We divide all the cases into 8 groups according to the distance. They are above 100 kilometers, 80-100 km, 50-80 km, 30-50 km, 15-30 km, 5-15 km, 2-5 km and below 2 km. The distributions are presented in Figure 3.

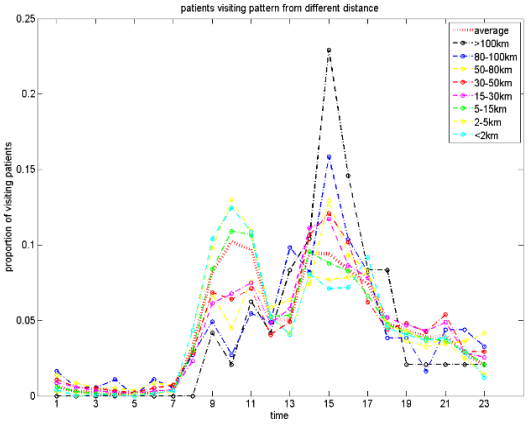

Fig. 3. The number of patients seeking treatment at different time over different distance

From the figure above we can conclude that the farther the patients are from living places to destined hospitals, the later the patient flow peaks. As the figure shows, when the distance is more than 100 kilometers, most patients come to see doctors at 15:00; however, those who are very close to hospitals (less than 2 kilometers) are most likely to visit hospitals at 10:00. This is because the patients living far from hospitals would take a long time on the way. And 10:00 and 15:00 are both the time at which doctors' workload gets the highest. This also explained why the peaks occur at 10:00 and 15:00.

4 Exploring Spatial Mobility Pattern Using Gravity Model

We consider the probability p_{ij} that a given case from area i visiting the hospital j. It is intuitive that a sick kid's parents should take some factors into consideration when they decide which hospital they would send their child to. First is the level of the hospital. A professional class-A pediatric hospital is much better than a local clinic. Second is the distance. Those too far from home are not good choice, because of the

traffic cost and the time delay. These two basic factors impact parents' choice, hence patients mobility pattern.

Gravity model was first proposed by Casey in 1955[4]. It has been used to forecast trade between countries and traffic flow between provinces. It is also widely used in describing the movements of people over distance [3]. The basic idea of gravity model is to simulate the flow among two places as the gravitational force between two celestial body. The mass stands for some kind of resource that the places hold. The basic form of gravity model is:

$$p_{ij} = \alpha \frac{Q_i Q_j}{d_{ij}^2}$$

Where p_{ij} stands for the amount of the flow , and Q_i, Q_j stand for the population of district i and district j. d_{ij} stands for the distance between district i and j. And α is a scale factor.

Here we define the OD matrix $N_{i\times j}$. Each eliment n_{ij} stands for the number of patients living in district i going to hospital j (i = 1,2,…312; j = 1,2,…195). $R_i = \sum_j n_{ij}$ means the total number of the patients in district i, and $H_j = \sum_i n_{ij}$ means the total number of the patients visiting hospital j.

We consider an unconstrained gravity model:

$$p_{ij} = kQ_i^{\hat{a}} Q_j^{\hat{a}} f(d_{ij})$$

Where $p_{ij} = n_{ij}/R_i$, indicates the probability the patients from district i going to hospital j. The distance impedance function $f(d_{ij}) = d_{ij}^{\hat{a}}$. $Q_i = n_{ij}/H_j$, and $Q_j = H_j/\sum_i \sum_j n_{ij}$, which stands for the proportion that the number of patients visiting hospital j account for of the total number (that is 44607), explains the that the hospital level is one of the factors that the parents take into consideration when they choose hospitals, because high-level hospitals always attract larger number of patients for their better medical service. The data proves this. The top 6 hospitals receive 24075 cases, and that number is over the half.

By using a standard LSM method, we get the parameters as follows. $k = 5.105 \times 10^{-3}$, $\alpha = 0.9059$, $\beta = 0.6878$, $\gamma = -0.1997$. While the $R^2 = 0.6497$.

After simulation with the model, the residual is found to follow a normal distribution, which indicates that the residuals of each point is independent.

The simulation result is shown below. In Figure 4, we can find that most of patients are willing to choose the hospitals which are probably about 10 kilometers away from home. And a considerable number of patients would like to go more than 10 km but less than 40 km. Those who go too far away or too near from home are not the majority. In addition, out simulation could describe such a procedure that how patients (or their parents) choose the hospitals to go while seeking treatment. Distance and the level of the hospital are the two major factors that are taken into account.

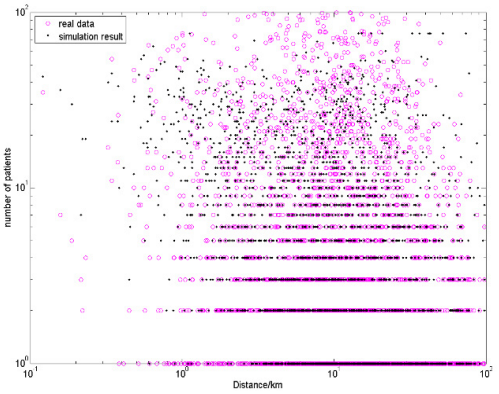

Fig. 4. The real data and the simulation result

5 Discussion

In this paper, we analyze the spatial-temporal pattern of HFMD patients seeking treatment, and an unconstrained gravity model is adopted to fit the mobility data. We find that the HFMD infected cases are mainly located in the junction area of urban and suburban districts. We also find that the temporal patterns of different group of patients seeking treatment are mostly different according to their living locations, distances to hospitals and the dates they go to see doctors. Finally, we use an unconstrained gravity model fit the mobility data. The coefficient of determination ($R^2 = 0.6497$) was not very satisfactory, it is partly because the granularity of the address information in the raw data is not fine. Over 40,000 cases located over Beijing are grouped into only 312 geographic unit, which may result in the general performance of the model.

Acknowledgments. This work is partially funded through NNSFC Grants #71025001, #71103180, #91124001, and #91024030, MOH Grants #2013ZX10004218 and #2012ZX10004801.

References

1. Cao, Z., Zeng, D., Wang, Q., Zheng, X., Wang, F.: An epidemiological analysis of the Beijing 2008 Hand-Foot-Mouth epidemic. Chinese Science Bulletin 55, 1142–1149 (2010)
2. Chen, K.T., Chang, H.L., Wang, S.T., et al.: Epidemiologic features of hand-foot-mouth disease and herpangina caused by enterovirus 71 in Taiwan, 1998-2005. Pediatrics 120, E244–E252 (2007)

3. Truscott, J., Ferguson, N.M.: Evaluating the Adequacy of Gravity Models as a Description of Human Mobility for Epidemic Modelling. PLoS Comput. Biol. 8, e1002699 (2012)
4. Casey, H.J.: Applications to traffic engineering of the law of retail gravitation. Traffic Quarterly IX(1), 23–35 (1995)
5. Cheng, Z., Caverleem, J., Lee, K., et al.: Exploring Millions of Footprints in Location Sharing Services. In: ICWSM 2011 (2011)