**Spatial and Temporal Characteristics of Population Movement in Central Business District Using Cellular Data, a Case Study of Shanghai**

STAT 5544 Spatial Statistics Project Proposal

Danni Lu & Wenyu Gao

**Background**

The rising traffic congestion at rush hours is a nightmare for every growing metropolitan. Along with the issue, Central Business Districts (CBD) are becoming controversial as they are top areas that have been bringing tremendous pressure to transportation system due to their high traffic attraction. To reduce periodical traffic congestion and balance commute traffic is a long term strategy. The very first step to this strategy should be figuring out where the traffic pressure origin and how intense it is. Identifying urban area with most traffic attraction, exploring where the population come from and analyzing spatial and temporal characteristics of population movement are essential for traffic dispersion and alleviating rush hour pressure. As a result, we are going to study the spatial and temporal behavior of population movement in Shanghai.

Nowadays, deep penetration of mobile phones provides us detailed travel information with high accuracy and precise timing. Thus, our study is based on cellular data from people concentrating at CBDs in Shanghai.

**Data description**

We collected geodetic position information of all users from one mobile carrier in Shanghai for a continuous 24 hour period. The dataset includes coded mobile phone ID, date, time, latitude and longitude.

When a mobile phone is in service, it connects with the nearest cell tower every one second for regular tracking. The Mobile Carrier system communicate with the mobile device through cell tower and record the latitude and longitude of cell tower. In addition to regular tracking, the system also keep the record of mobile device when one of the following events happen:

(a). mobile phone users send or receive massages;

(b). mobile phone users make a phone call or receive a phone call;

(c). mobile phone users travel from area related to one cell tower to area related to another cell tower.

The data we collected are records that filtered out regular tracking, without the loss of the information related to user travel from one place to another.

In total, we collected data of more than 1.1 billion records. These records are from 37450 different cell tower all over Shanghai with average density 5.91 cell tower per kilometer square.

**Research Plan**

We will start from exploratory analysis for point patterns to achieve a general idea of how population moves with time and location. Next we are going to fit the models. Spatial-temporal models, hierarchical models and Bayesian models are possible candidates and some other models may also be considered. Model selection techniques are used to select the best fitted and predicted models. Afterwards, we will do statistical inferences on the best fitted model and finally we are going to predict the spatial and temporal characteristics of population movements with the best predicted model.