Linking Group-level Mobility Modeling and Disaster Relief using Social Media Mining

Lu Cheng (1213398522) Jiayong Mo(1213555809) Linzhen Luo(1213315543)

Understanding human mobility is of great importance to disaster relief. While there has been fruitful research on modeling human mobility using tracking data (e.g., GPS traces), the recent growth of geo-tagged social media (GeoSM) brings new opportunities to this. In this project, our team propose to apply a group-level mobility modeling method using both GeoSM data. Our insight is that the GeoSM data usually contains multiple user groups, where the users within the same group share significant movement regularity. Therefore, we have two different tasks to tackle: **(1) user grouping**: softly group the users in U such that the members in the same group have similar moving behaviors; and **(2) mobility modeling**: for each user group, discover the latent states that reflect the users’ activities from a multidimensional (where-what-when) view, and find out how the users move between those latent states. However, according to research, only 2% of the social media data are geo-tagged and lots of the text extracted are Non-English speakers, which means we will lose most of information of data we crawled using TweetTracker[1]. Therefore, to enrich our data set, we plan to use Machine Translation and Location Prediction based on existing geo-tagged data and users’ social ties.

## Goal

1. Develop a robust group-level human mobility model with a multidimensional structure, jointly consider location, time and text of each tweet posted by users.
2. Infers people's ne-grained location, even when they keep their data private and we can only access the location of their friends.
3. Explore the Hurricane Harvey data set we crawled from TweetTracker, a system that is specifically created to crawl tweets based on the keywords and locations specified.

## Research Questions

1. What are the intrinsic states underlying people’s movements under disaster response?
2. How do people move sequentially between those latent states?
3. How location estimation can help improve human mobility model development?
4. How the roles of people evolve over time, i.e., Pre-disaster, disaster and Post-disaster.

## Approaches

There are mainly two methods we are going to use: Hidden Markov Models and Dynamic Bayesian Networks.

**Ensembled HMM:** This is a group-level HMM, which generates an ensemble of Hidden Markov Models

(HMMs) to characterize group-level movement regularity. For each group, we have a well-trained HMM model;

**Dynamic Bayesian networks (DBNs)**: DBNs are generative probabilistic graphical models of sequential data. Nodes in the graph represent random variables and edges represent conditional dependencies. In a typical setting, a subset of the random variables is observed, while the others are hidden and their values have to be inferred. A DBN is composed of slices, each of which is a time interval in our case.