Transit Data Analysis
Part B: Usage Pattern of Bike-sharing Service

Students: Siyue Yang(CityU), Danielle Stacy(UA),Jing Guo
Mentors: Candace Brakewood, Cheng Liu, Kwai Wong

Introduction

Bike-sharing System, as one of the burgeoning transportation services that have spread out worldwide in the recent years, is providing various information of human mobility patterns. This value can be further investigated when combined with some emerging smartphone applications which provides a platform for users to book sharing bikes. By using the back end data from smartphone applications, this project aims to investigate the usage features of bike-sharing services in both spatial and temporal scales.

Two datasets are involved when conducting the data analysis on users. One is from the Transit App, widely used in New York City, Chicago and approximately 125 other cities, which provides real-time transit vehicle information, trip planning and shared mobility services. The other one is from Divvy, a bike sharing system in Chicago. When utilizing Transit App to book Divvy sharing bikes, the users will be directed to Divvy website to complete the order. Both of the datasets contain information of booking orders in 2016, including the booking time and bike locations.

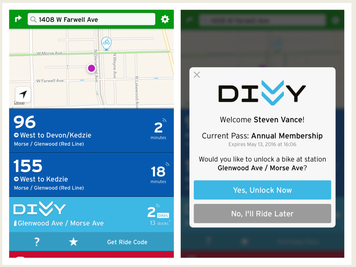

Figure1: Transit App Screen-shots

Figure2: Transit data distribution

Research Questions

Inured to travel times in the highly congested areas and public health, the bike share has continued its brisk growth, with 35 million trips taken in 2017, 25% more than in 2016. With its sustained development and various benefits to the society, we aim to provide a statistical analysis of bike-sharing usage, help bike-sharing operators to better understanding the behavior of users and manage the bikes.

The usage patterns include the frequency of people using the system, the difference between weekdays and weekends and ways that people using the bike share, such as some users may unlock the bikes consecutively to avoid paying.

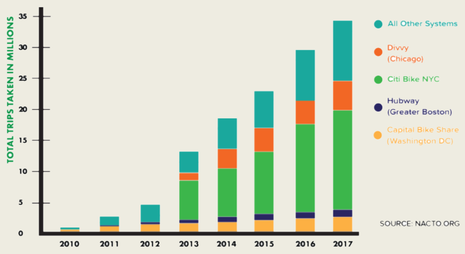

Figure 3: Growing ridership of bike-sharing system

Travel patterns of time

Firstly, the analytical hierarchy process was used to investigate the time usage patterns. By the data from the Divvy company, it was found that sharing-bike system was mostly used at 17pm of a day and in September of a year. The discovery showed that the frequency of bike share usage was related to the temperature and the daily activity.

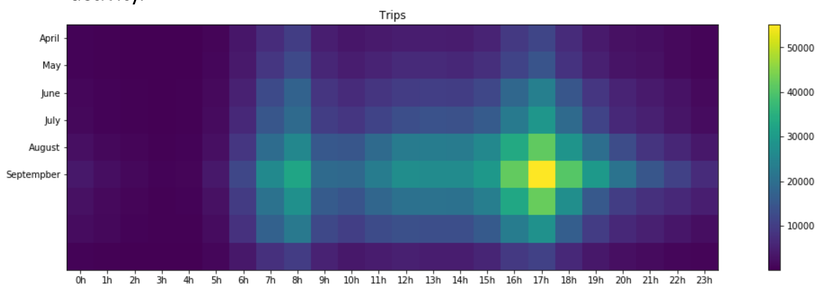

Figure 4: Frequency of bike share usage by analytical hierarchy process

Secondly, the trends in bike-share utilization over time was investigated. The use of Transit app, shown in blue bars, represented approximately 9% of all bike-sharing trips. Since both of the trends processed similar peaking patterns, the data from Transit App could be representative of the whole dataset.

Two peaks were observed during the weekdays: one was in the morning, most likely corresponding to morning commuting trips; a higher peak was in the evening, perhaps related to a combination of evening commuting trips and leisure activities. The lack of shower access at work could be a barrier to bike-sharing in the morning. There was also a smaller peak observed in the afternoon. During weekends, ridership in Chicago seemed to be more regular across time, probably corresponding to leisure or recreational trips.

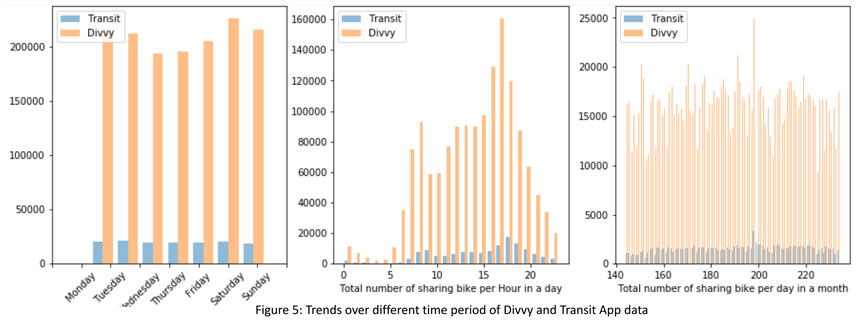

Figure 5: Trends over different time period of Divvy and Transit App data

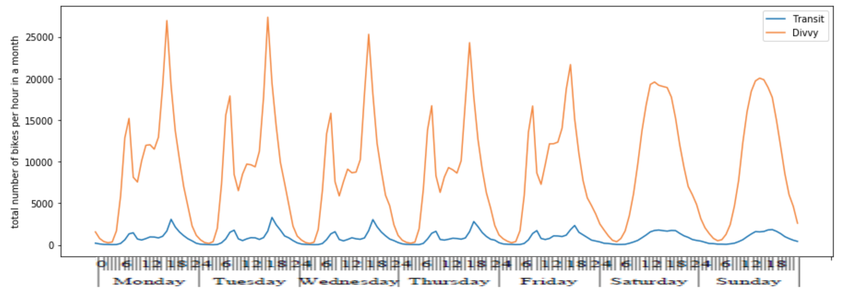

Figure 6: Total number of unlocked bikes per hour in a week

Preliminary Classification of users

Since the data from Transit App can be a representative sample of the whole bike-sharing orders as discussed in the last part, the records were used to classify the users into different groups. The k-means clustering method was utilized and it grouped the users mainly into four groups. The details of clustering results was shown below:

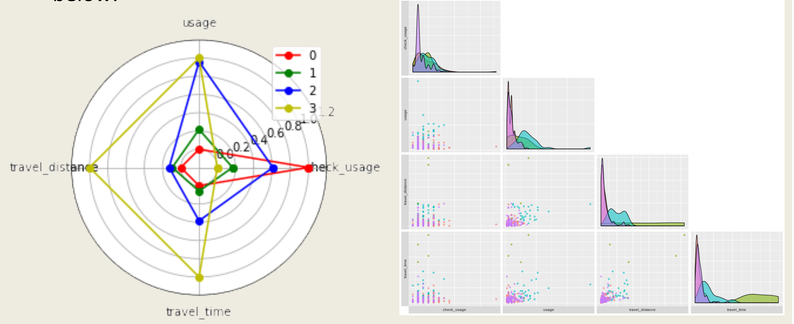

Figure 7: The characteristics of each of the groups

According to the results of clustering analysis, 45% of the users are potential tourists, who use seldom use Transit App nor booking sharing services. 36% of people use Transit App more but occasionally book the sharing bikes. 16% of them do short-distance trips by bikes and 2% of the sharing bike users may travel in a long distance.

Future Work

Previously, only four variables were used to classify the users. More variables should be considered for clustering analysis, such as number of times using service in the morning and afternoon, number of times using service in the weekdays and weekends. It may help indicate the commuters and tourists.

There was an Interesting phenomenon that some users may unlock different bikes consecutively in a period. One explanation is that people could avoid paying fees by doing so. More efforts should be put into the investigation of consecutive unlocks in the next stage.

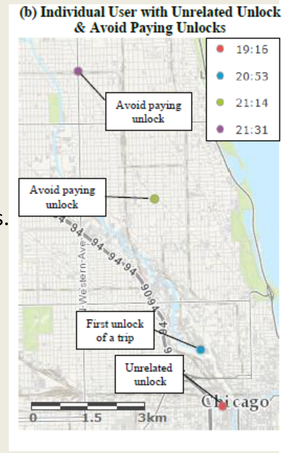

Figure 8: Examples of consecutive unlock patterns

Acknowledgements

This project was sponsored by the National Science Foundation through Research Experience for Undergraduates (REU) award, with additional support from the Joint Institute of Computational Sciences at University of Tennessee Knoxville. This project used allocations from the Extreme Science and Engineering Discovery Environment (XSEDE), which is supported by the National Science Foundation.

In addition, the computing work was also performed on technical workstations donated by the BP High Performance Computing Team.