Analyzing the Riders behavior

Pramesh Kumar and Alireza Khani
Department of Civil, Environmental and Geo-Engineering
University of Minnesota
Outline

- Motivation
- Data
- Temporal Analysis
- Spatial Analysis
- Clustering
- Observations/Conclusion
- Ongoing Work
Motivation

- UPass sales and ridership increased in 2010 but decreasing since then
  - except a small increase after opening of Green Line.
- Difficult to analyze the U-Pass ridership trends
- What are changing shifts in the U-Pass use
  - by identifying the clusters of the UPass holders.
Data

- Automatic Fare Collection (AFC) data provides the passenger's entry or exit information from 2009 to 2016 which consists of:
  - Unique ID
  - Date
  - Time of the tag
  - Geographic Coordinates
  - Route and its Direction
  - Other information

- More than 37 million records, 35 GB. Big Data!
Temporal Analysis

- To analyze the variation of U-Pass ridership, no. of unique U-Pass and Average rides per unique card with
  - School Years
  - Season/Semester of the year
  - Time of the day
Ridership versus Semesters of different school years
No. of U-Pass versus different Semesters

Variation of No. of Unique Cards with respect to semesters of the University

Change in No. of Unique Cards with respect to semesters of the University
Average Rides/ Card versus different Semesters
Ridership versus time of the day
Weekdays

Upass Ridership variation with respect to the time of the day on Weekdays

Tag Frequency

0e+00 1e+05 2e+05 3e+05 4e+05

Time
00:30 01:30 02:30 03:30 04:30 05:30 06:30 07:30 08:30 09:30 10:30 11:30 12:30 13:30 14:30 15:30 16:30 17:30 18:30 19:30 20:30 21:30 22:30 23:30

Period
Sep09-Aug10
Sep10-Aug11
Sep11-Aug12
Sep12-Aug13
Sep13-Aug14
Sep14-Aug15
Sep15-Aug16
Spatial Analysis
First tag of the day assumed as home location
Mostly trips generate near U like Dinkytown, Como, Prospect Park, Marcy Holmes area
Other areas like Downtown, Uptown, Lyn Lake, Steven Square, Seward, Maple Grove, Airport and Mall of America.
In year 2012-13 areas like Roseville, Northeast Minneapolis and Burnsville
for the year 2014-15, areas near to the commuter rail stations like Ramsey and Anoka while other areas like Eagan, and Brooklyn Park were found to show a significant number of trips.
Clustering

- Frequency based Clustering
  - Frequency of tags per year
  - Frequency of card use per week
- Travel Pattern Clustering
- Regularity Clustering
Total U-Pass users versus no. of rides per year

Frequency based clusters of Upass Riders

Change in Frequency based clusters of Upass Riders in consecutive Years
Total U-Pass users versus No. of days U-Pass used per week
Travel Pattern of the Passengers

- Commuters show similar travel pattern both spatially and temporally.
- Research uses DBSCAN (Density-based spatial clustering of applications with noise) algorithm to cluster these trips.
- Define $\epsilon$ (radius required to search the neighborhood points) and $n$ (minimum number of points required to form a cluster).
Results
- Major Clusters Found
  - University Campus East bank + Dinky town
  - West bank + Seward + Phillips
  - St. Paul Campus + Como
  - Downtown
  - Uptown
  - Prospect park + Raymond
  - McMaster College + St. Catherine University

- Minor Clusters Found
  - Anoka
  - Brooklyn park
  - Coon Rapids
  - North town Transit Center
  - Louisiana Transit Center
  - Interstate 394 En & Gen Mills Blvd P&R
  - 95th Ave Park and Ride
clustering of UPass Data

Please enter Eps value
0.01

Please enter MinPts value
6

Please select the semester
Fall 2009

Please select the time range

| 00:00:00 | 01:00:00 |
| 02:00:00 | 03:00:00 |
| 04:00:00 | 05:00:00 |
| 06:00:00 | 07:00:00 |
| 08:00:00 | 09:00:00 |
| 10:00:00 | 11:00:00 |
| 12:00:00 | 13:00:00 |
| 14:00:00 | 15:00:00 |
| 16:00:00 | 17:00:00 |
| 18:00:00 | 19:00:00 |
| 20:00:00 | 21:00:00 |
Trip Chaining for Origin-Destination Estimation

- First tag of the day is assumed as the origin and can be interpreted as the housing location.
- For pay exit routes, the last tag of the day is considered as the origin of the first trip.
- Then next boarding can be assumed as the destination of current trip.
- The destination of the last trip of the day can be assumed as the origin of the next days.
- No other mode of transportation used in between the trips.
Regularity Pattern of the Passengers

- “Regularity” means how frequently transit riders make similar trips.
- The following parameters are used:
  - Number of days of travel in study period
  - Average number of trips per day
  - Number of similar boarding stops
  - Frequency of similar routes used to reach the destination
  - Frequency of similar departure time
- K-Means clustering algorithm was performed to prepare three types of clusters
Average number of trips taken in a day

Number of Similar Boarding Stops

Number of Similar Departure Times

Similar Boarding Stops

Similar Departure times
Observations/Conclusions.

- U-Pass ridership has decreasing trend.
- More decrease in Spring Semester
- No. of U-Pass used in 2015-16 is more
- Regular cluster found near university areas and also other areas
- Green Line has significant impact on ridership
Thank You

Metro Transit

Pramesh Kumar and Alireza Khani
Department of Civil, Environmental and Geo-Engineering
University of Minnesota
kumar372@umn.edu, akhani@umn.edu