Making Sense of Big Data from Ride-Sharing Services

Yanhua Li
Assistant Professor
Computer Science Department
Worcester Polytechnic Institute
Who am I?

Yanhua Li, PhD
Assistant Professor
Computer Science & Data Science
Worcester Polytechnic Institute (WPI)

PhD, Computer Science, U of Minnesota, 2013
PhD, Electrical Engineering, BUPT, 2009

Research Interests: Big data analytics, Smart Cities, Measurement, Spatio-temporal Data Mining

Industrial Experience: HUAWEI Noah’s Ark Lab, Mobike, DiDi
CityLines: A Hybrid Hub-and-Spoke System for Urban Transit Services

Guanxiong Liu, Yanhua Li, Zhi-Li Zhang, Jun Luo, Fan Zhang
WPI, CAS, Lenovo, UMN
Contact: yli15@wpi.edu
Today’s Urban Transit Services

Public Transits

- Affordable ride-sharing services

Private Transit

- Reduce the personal vehicle usage
Limitations of Today’s Public Transits

• Fixed Routes and Time Tables
  – Transit supply mis-match dynamic demands

• Large number of stops and transfers
  – Long travel time
Limitations of Today’s Private Transits

• **Expensive**
  – High operation cost,
  – Due to the exclusive service

• **Service delay**
  – On-demand services
  – Delay after the service request

• **Transit modes run independently**
  – Lack of inter-transit coordination
Future Urban Transit Services

Today’s Transits
• Private Transits
  – High Cost
  – Service delay
• Public Transits
  – Fixed routes
  – Fixed timetable
  – Long travel time
• No Inter-Transit Coordination

Future Smart Transit
• Dynamic services
  – Real time trip demands
• Short travel time
  – as private transits
• Low cost
  – as public transits

Private Transits: Point-to-point mode
Public Transits: fixed route mode
Hub-and-Spoke Transit Mode

- Traffic move along spokes connected via a few hubs
  - Less operation cost (than private), thus lower cost
  - Less stops/stations (than public), thus lower transit time

- A promising transit mode, and how to design it in urban areas?
CityLines Transit System

- **CityLines: a Hybrid Hub-and-Spoke Transit Mode**
  - point-to-point mode: high demand source-destination pairs
  - hub-and-spoke mode: low demand source-destination pairs

Private transit
Point-to-point model

Public transit
Fixed-route model

CityLines
Hybrid hub-and-spoke mode

Reduce routes, thus operation cost
Reduce stops/stations, thus travel time
Input Data Description

- Trip Demand Data (in Shenzhen):
  - **Source**: Taxi GPS, Bus, Subway Transactions
  - **Duration**: March 1st–30th, 2014.
  - **Size**: 19,428,453 trips in all transit modes
  - **Format**: Taxi ID, time, latitude, longitude, load

- Road Map, Subway Lines, and Bus routes:
CityLines Demo System

https://wpi.edu/~yli15/CityLines/
Planning Bike Lanes Based on Sharing-Bikes’ Trajectories

Jie Bao, Tianfu He, Sijie Ruan, Yanhua Li, Yu Zheng

Microsoft Research
Worcester Polytechnic Institute

ACM KDD 2017
Bike “Renaissance” in China

1980s  
Owning a bike

2017  
Sharing a bike
Effective Bike Lanes Promote Cycling Experiences

• Motivations to promote cycling
  - Reduce the air pollution  - Address the “last mile” in public transportation
  - Ease traffic congestions  - Advocate the healthier lifestyle
Why Sharing Bike (Mobike) Trajectories

• Large Number of Users
  – Better coverage of general bike users
    *1 Billion Users, 25 Million Daily Orders

• Station-Less Deployment
  – More realistic travel demands

• Detailed Trajectory Data
  – More information on mobility patterns

*http://www.hbspcar.com/1922.html
Mobike Data Description

- Dates from 2016/09/01 to 2016/09/30.
- 13,063 users, 3,971 bikes, and 230,303 trajectories.
Problem Settings

• Government Requirements
  – Budget Constraints
  – Best Serve the Public (coverage & continuity)
  – Construction/Management Convenience

Road Space & Money  User & Trip Coverage  K-Connected Components
Jinyun Road Station
Wanda Plaza
Existing Bike Lanes

Highly Populated Residential Area
On-going projects

• Dockless bike-sharing
  – Load balancing strategies
  – Incentive mechanisms

• Ride-hailing
  – Driver behavior analysis
  – Smart traffic management
Questions?