Jinjun Tang, Fang Liuc, Yinhai Wang, Hua Wang

Uncovering urban human mobility from large scale taxi GPS data

Summary

http://www.sciencedirect.com.zorac.aub.aau.dk/science/article/pii/S0378437115005853

Goal

Analyse and describe taxi mobility in Harabin

Data

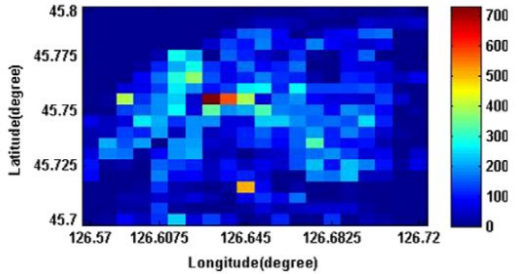
- 1100 taxis in Harabin (China) sampled every 30 secs (2880 samples of a taxi per day) during a half year

Table 1
Data sections of taxi GPS data in Harbin city.

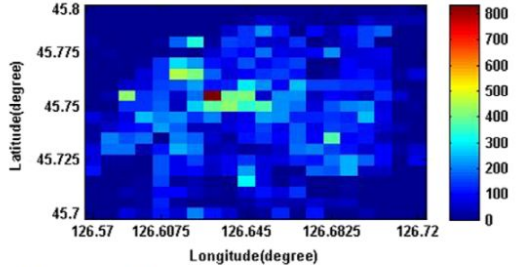
Taxi ID	Time	Latitude	Longitude	Speed	Orientation	Status
100300002	2012/8/1 6:59	45.738384	126.616920	35	109	0
100300002	2012/8/1 7:00	45.736588	126.614845	29	110	0
…	…	….	….	…	…	…
100300010	2012/8/1 11:08	45.757168	126.604280	40	8	1
100300010	2012/8/1 11:09	45.759000	126.605290	33	12	1
…	…	….	….	…	…	…

Location partition

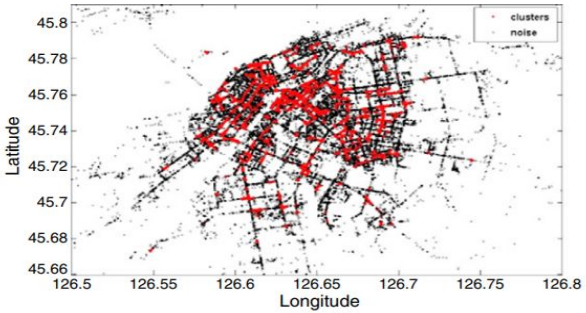
- Firstly 400 districts equal in size
- Then they used Density-Based Spatial Clustering of Applications with Noise (DBSCAN) algorithm to cluster pick-up and drop-off locations (red are clusters, black is noise)



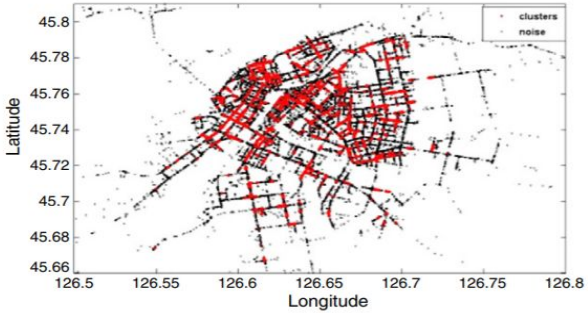
(a) Origins on weekday.



(c) Origins on weekend.



(a) Pick-up locations.

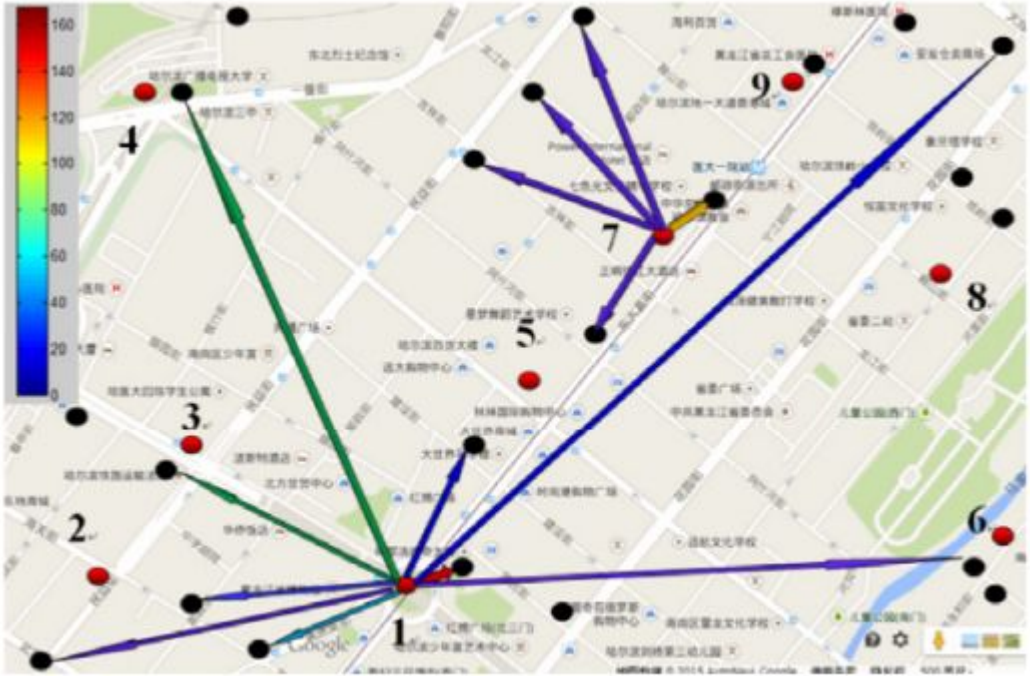


(b) Drop-off locations.

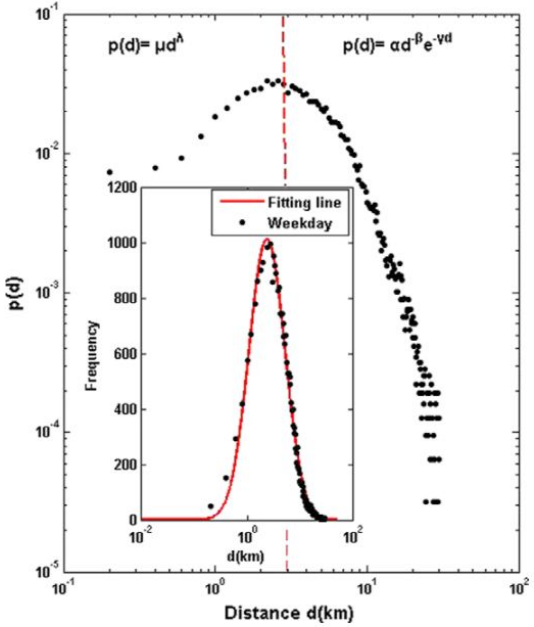
Analyses

- Taxi drivers’ choice of pick-up clusters (Huff model)
- Distance distribution
- Travel time distribution
- Average speed distribution
- Trip distribution based on entropy-maximizing model (https://en.wikipedia.org/wiki/Trip_distribution#Entropy_analysis)

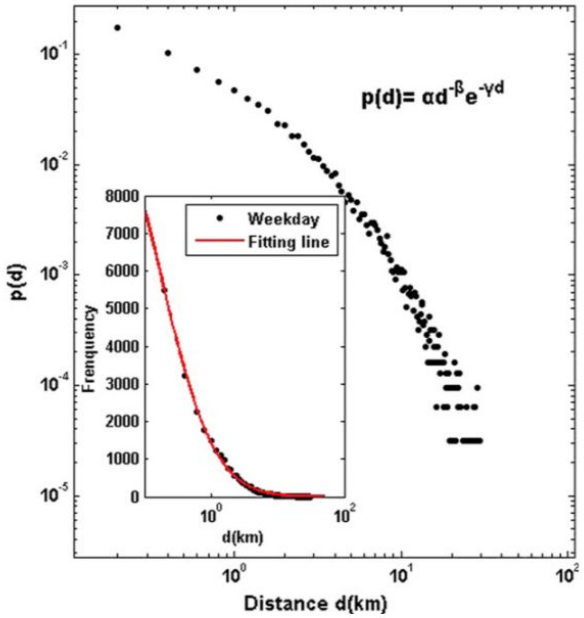
Choice of pick-up clusters



Distance distribution

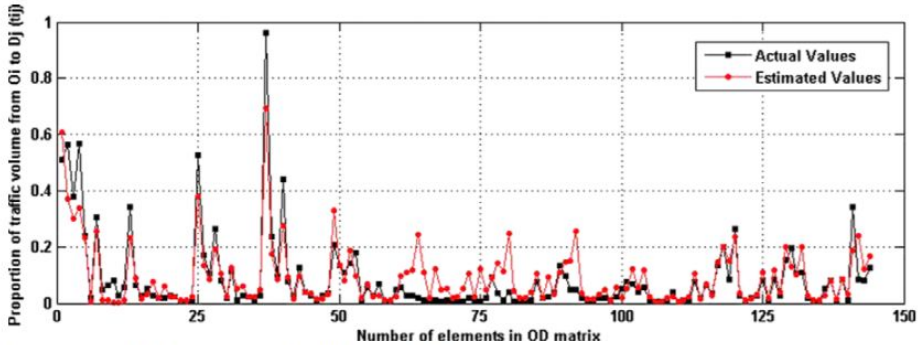


a) Occupied trips.

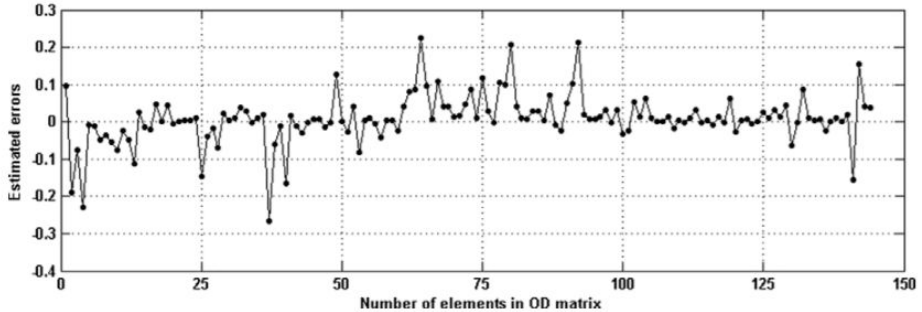


(b) Non-occupied trips.

Trip distribution



(a) Comparison between estimated and observed values.



(b) Estimation errors.