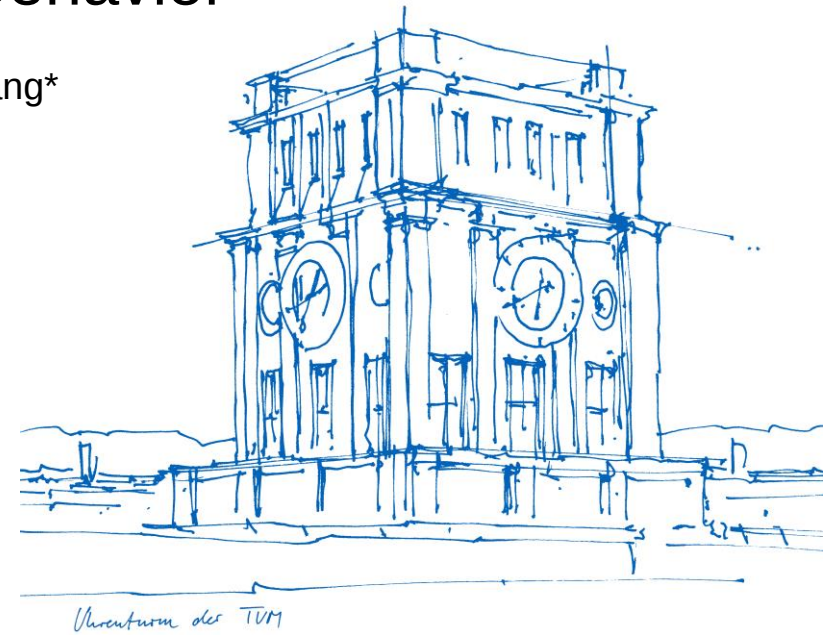
Identifying Origin/Destination Hotspots in Floating Car Data for Visual Analysis of Traveling Behavior

Mathias Jahnke*, Linfang Ding*,**, Katre Karja*, Shirui Wang*

* Technical University of Munich
Chair of Cartography

** University of Augsburg
Group of Applied Geoinformatics



Motivation

Understanding taxi mobility has significant social economic impacts on the urban areas.

- Intelligent transportation systems .g. taxi dispatching/recommendation system.
- Urban planning e.g. discovering functional regions.
- Human mobility pattern e.g. intra urban human mobility.

Detect popular places => high density of origin/destination hotspots on street level.
Identify places/venues close to the detected spots.
Explore the temporal pattern on the spots on district and street level.
⇒ Spatiotemporal visualization methods

Data

A one week subset of Shanghai‘s taxi floating car data.

Attributes: location (lat, lon), car id, timestamp, car status.
The car status indicates whether a taxi cab is occupied or not.

Origin/destination

Extracting origin and destination points.

The order of change in car status (1, 0) indicates a origin or a destination

- ⇒ 0 to 1 = origin (pick up)
- ⇒ 1 to 0 = destination (drop off)

The visual results of a pre-clustering based on different time intervals (1h to 4h) revealed a interval of 3h for aggregating origin or destination points.

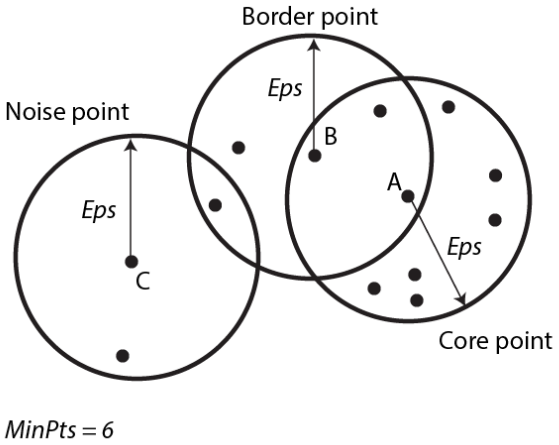
Too short intervals results only in very few and sparse point groups.
Too long intervals results in too condensed point groups.

Clustering

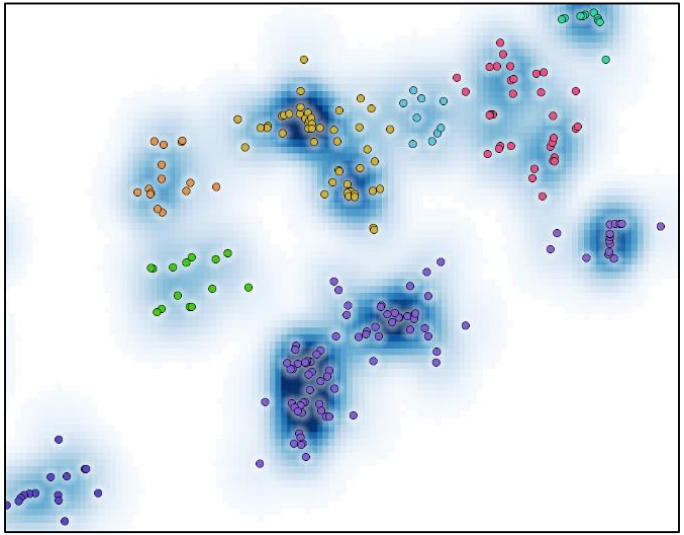
Finding pattern in a large collection of data by grouping data items based on their similarities in data space.
Geographic or non geographic.

In this case the DBSCAN clustering was used to find accumulations of origin or destinations points
⇒ o/d hotspots

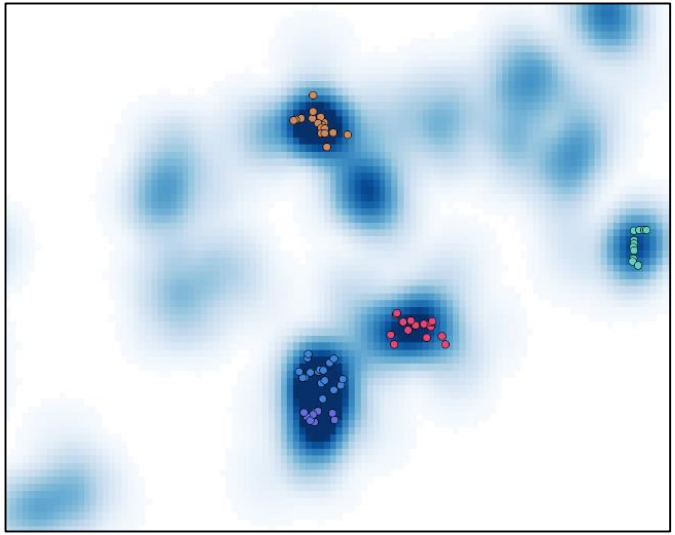
Concentrating on areas with high densities of o/d hotpots.
Varying number of clusters.



Clustering

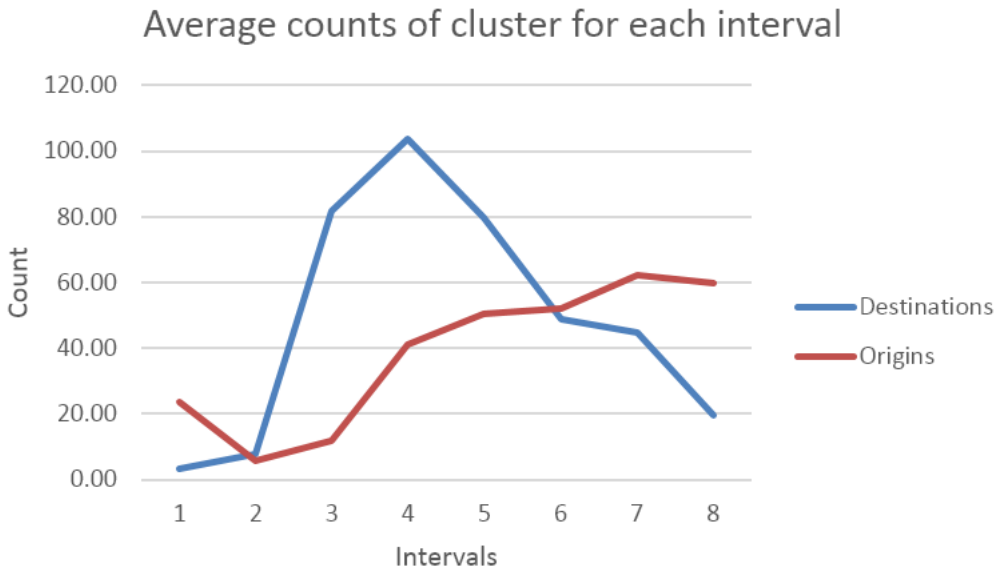


Eps = 200m, minPts = 10



Eps = 25m, minPts = 6

Clustering results



Semantic information

The semantic information was derived from OSM data based on the layers *buildings* and *landuse* plus additional expert knowledge.

11500 buildings and 2700 records with different landuse.
7000 not meaningful or unclassified objects were not taken into account.

Categorizing the objects into hotel, commercial, office, exhibition center, restaurant, public and business.

At least 3700 objects for adding labels to every cluster.

The semantic information and the object were combined based on a distance threshold.

Application

APP DEVELOPMENT

Square

Crossfilter

Fast Multidimensional Filtering for Coordinated Views

Leaflet

mongoDB

Flask

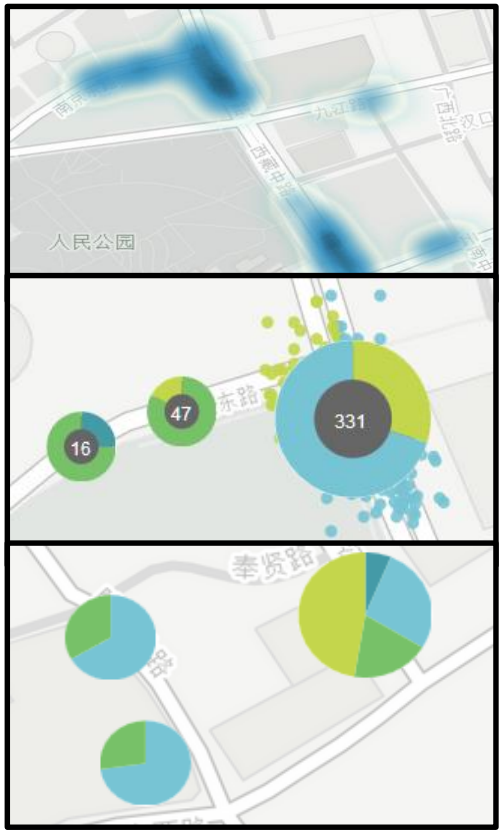
web development,
one drop at a time

Applications map-views

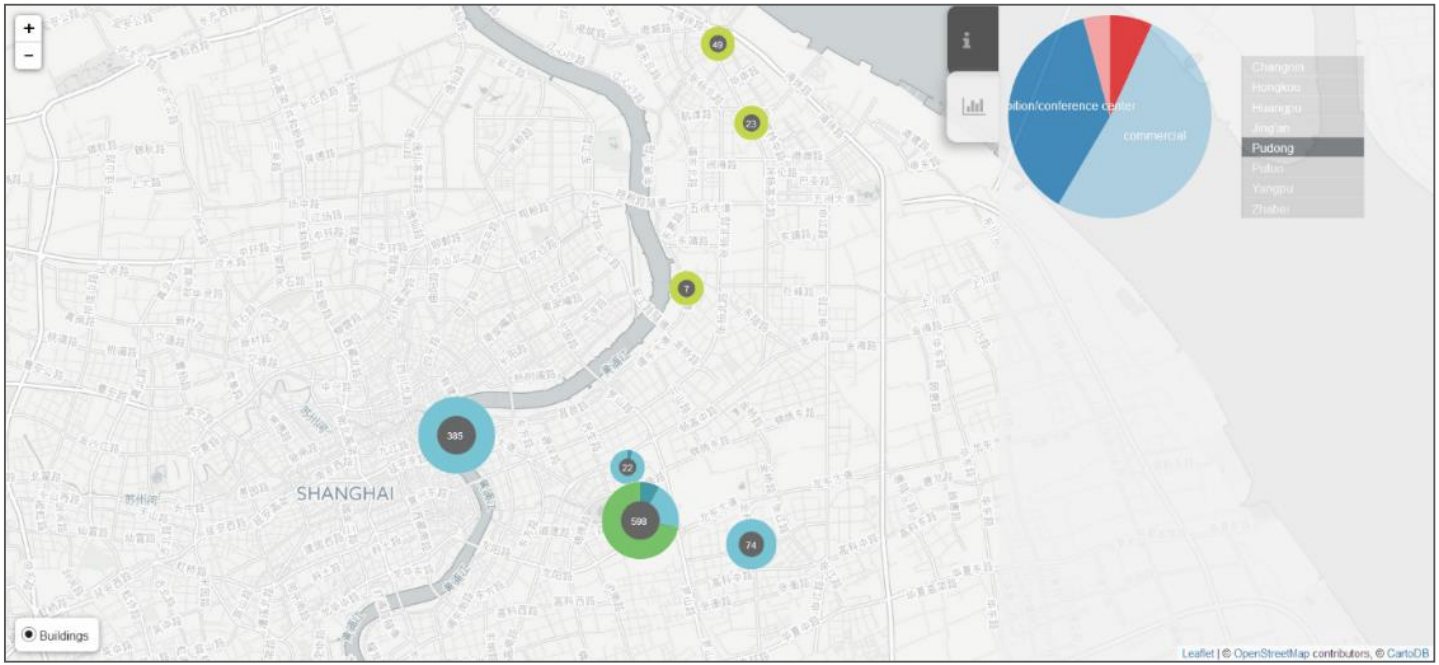
Heat layer => locations of hotspots

Donut chart => the distribution of tagged items

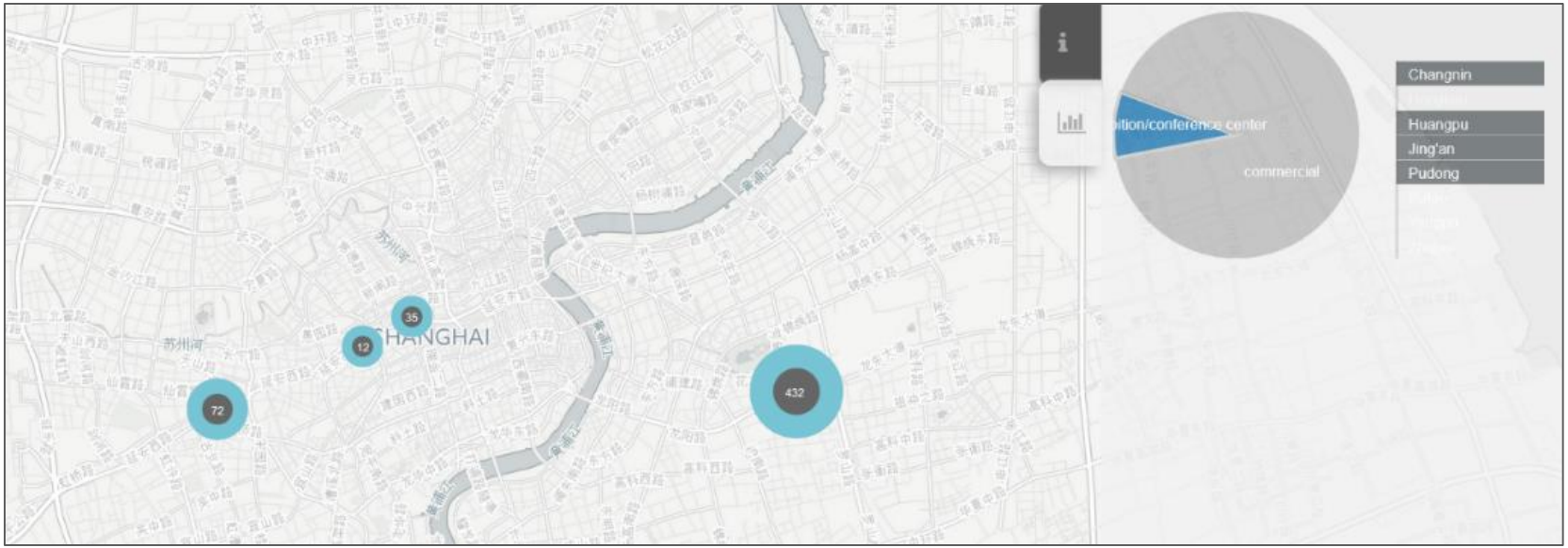
Pie charts => the distribution of occurring time intervals for each hotspot



Visualization

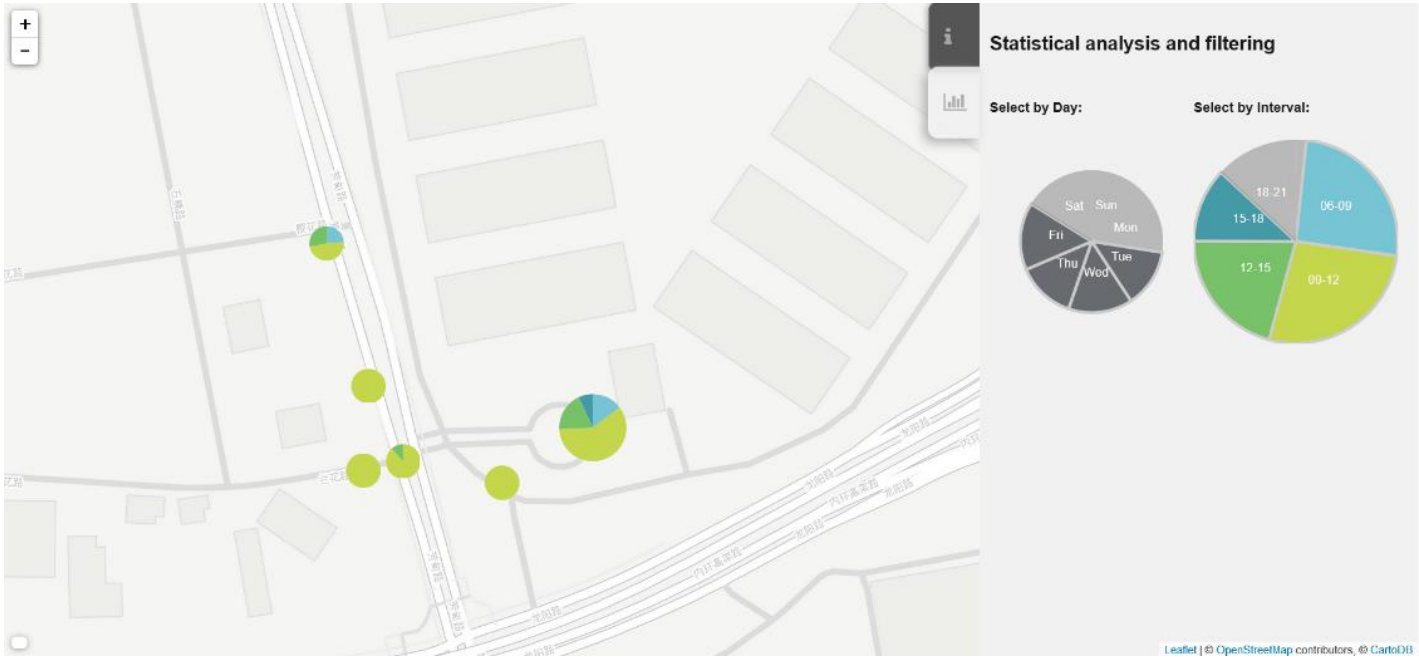


Visualization



Visualization

Visualization



Conclusion

Processing of floating car data to extract o/d points, densities

- Filter, cluster, visualization

The extracted o/d points were combined with sematic information from OSM and local experts.

This was brought to a web based geo-visual application to gain deeper insights.

- Map view and different charts/diagrams to visual the time dependent information.

Outlook

Improvements on the visualization:

- more interactivity,
- 3d views.

How to make semantic data integration more certain (one to many instead of one to one, include a probability value).
Increase the used data subset.

Acknowledgment

We would like to kindly thank Prof. Chun Liu from Tongji University for sharing the Floating Car Data with us.

Thanks!

Mathias Jahnke
mathias.jahnke@tum.de

Technical University of Munich
Chair of Cartography