Vaite: a Visualization-Assisted Interactive Big Urban Trajectory Data Exploration System

Chuang Yang

2019.09.19

Outline

- **Introduction**
- System Architecture
  - Three Layer Architecture
- Case Study
  - Data Model
  - Three Real World Cases
- Summary

Introduction --- Background

Urban Trajectory Data

- **<longitude, latitude >, timestamp, trip attributes**
- Heterogeneous: Taxi, Bus, Pedestrian
- Massive
- Valuable



Applications

- Congestion detection [BigData 2017]
- Taxi communities, outlier identification [TKDE 2017, WWW 2017]
- Billboard placement location [KDD 2018]

Chuang Yang, Yilan Zhang, Bo Tang, Min Zhu. Vaite: a Visualization-Assisted Interactive Big Urban Trajectory Data Exploration System. IEEE 35th International Conference on Data Engineering(ICDE), Macau, China, April 2019.

Introduction --- Motivation

General urban trajectory data analysis

- Particular analysis procedures for specific trajectory analysis tasks.
- e.g., Traffic Congestion Prediction [BigData 2017],
- Finding significant places [TVCG 2013].
- SQL-like query language
- **Professional Knowledge: Required.**

City manager explore the potential issues of the city ?

Taxi company explore the interesting facts among taxi drivers ?

Chuang Yang, Yilan Zhang, Bo Tang, Min Zhu. Vaite: a Visualization-Assisted Interactive Big Urban Trajectory Data Exploration System. IEEE 35th International Conference on Data Engineering(ICDE), Macau, China, April 2019.

Introduction --- Motivation

Interactive data exploration analysis

- Interactively pose ad-hoc visualization queries to identify potential relationships or glean insights.
- **Professional Knowledge: Non-required.**

Chuang Yang, Yilan Zhang, Bo Tang, Min Zhu. Vaite: a Visualization-Assisted Interactive Big Urban Trajectory Data Exploration System. IEEE 35th International Conference on Data Engineering(ICDE), Macau, China, April 2019.

Introduction --- Challenges

- **User-friendly Exploratory Analysis**
  a) issue an ad-hoc query should be convenience
  b) the corresponding results should be easy to interpret.

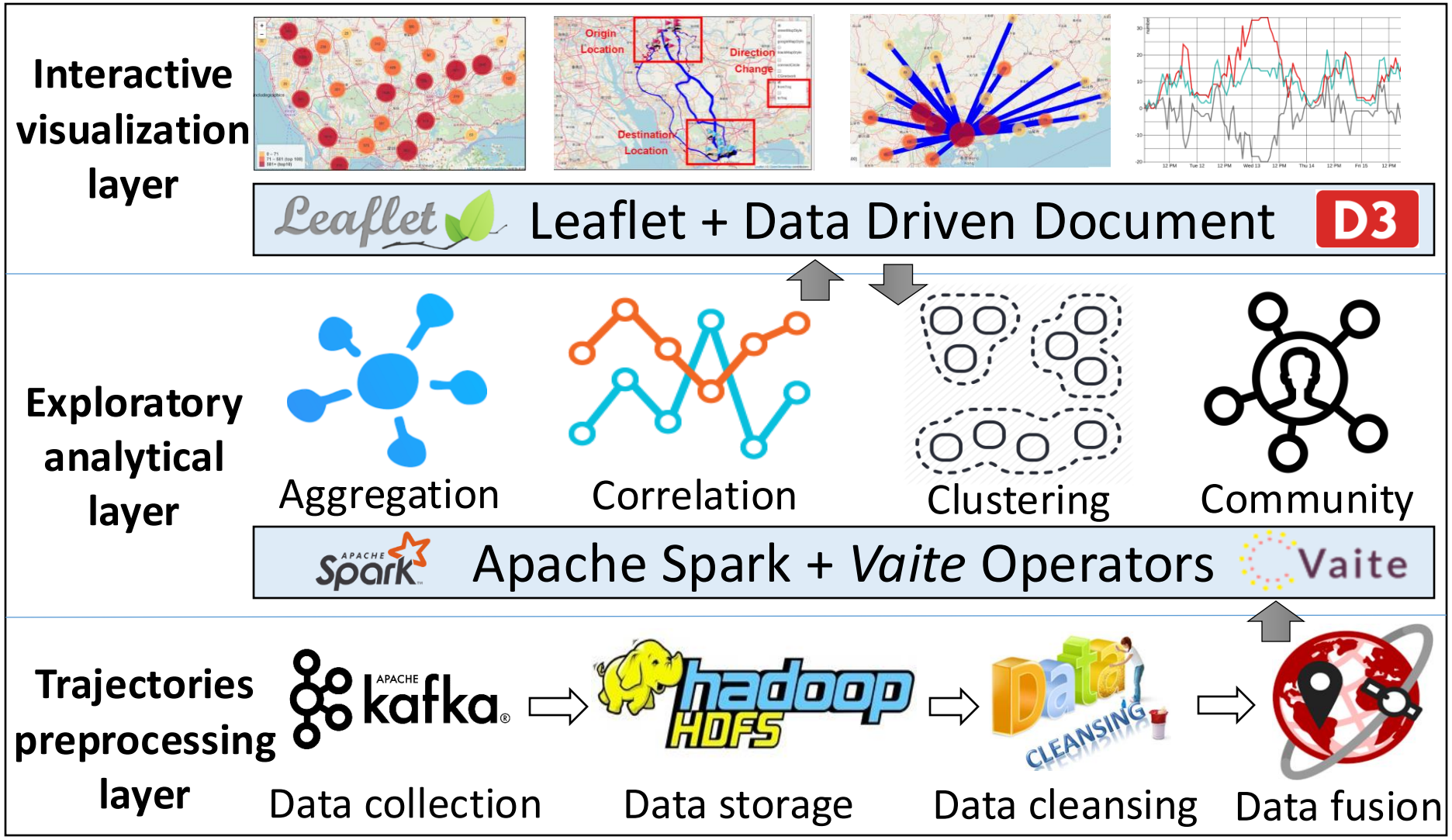
- **Incremental and Interactive Visualization**
  - **The underlying analysis algorithms and visualization techniques to be scalable.**

Chuang Yang, Yilan Zhang, Bo Tang, Min Zhu. Vaite: a Visualization-Assisted Interactive Big Urban Trajectory Data Exploration System. IEEE 35th International Conference on Data Engineering(ICDE), Macau, China, April 2019.

Outline

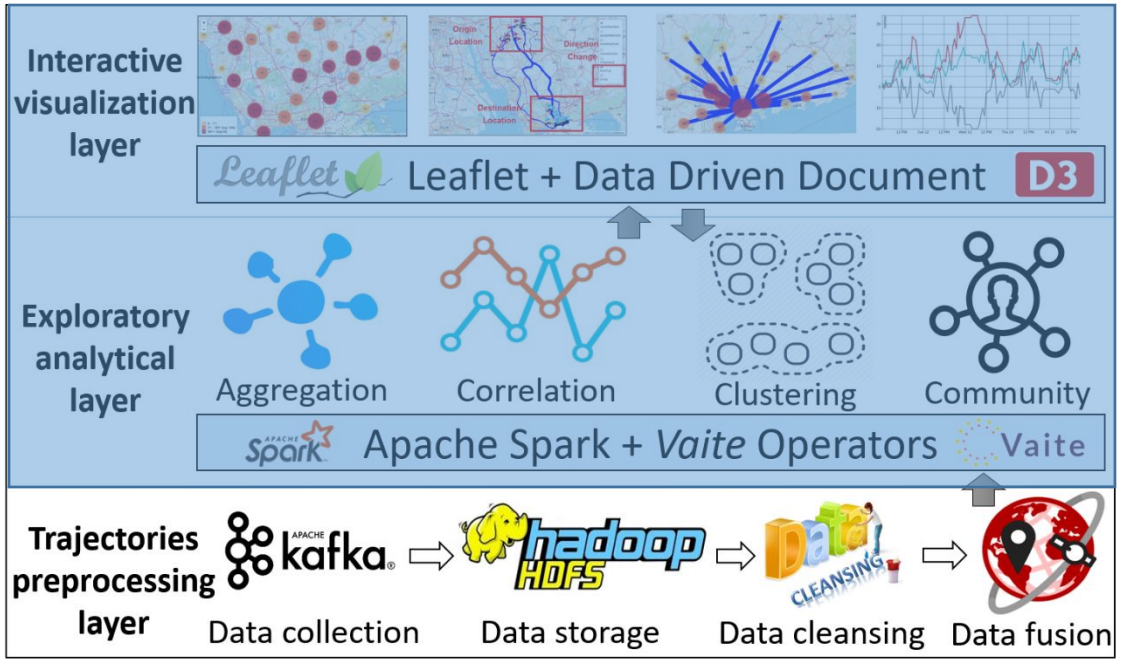
- Introduction
- **System Architecture**
  - Three Layer Architecture
- Case Study
  - Data Model
  - Three Real World Cases
- Summary

System Overview



Chuang Yang, Yilan Zhang, Bo Tang, Min Zhu. Vaite: a Visualization-Assisted Interactive Big Urban Trajectory Data Exploration System. IEEE 35th International Conference on Data Engineering(ICDE), Macau, China, April 2019.

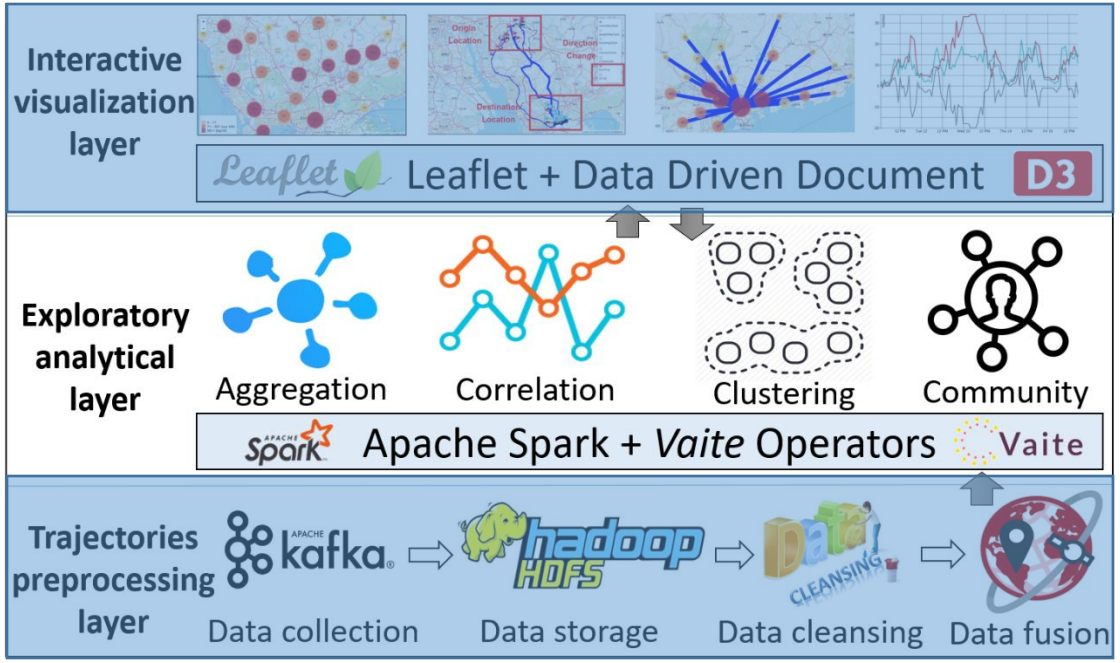
Vaite Framework --- Data preprocessing layer



- Collecting urban trajectories.
- Storing data on HDFS
- Matching road network.
- Cleaning the data errors
  - E.g., value missing, wrong data records

Chuang Yang, Yilan Zhang, Bo Tang, Min Zhu. Vaite: a Visualization-Assisted Interactive Big Urban Trajectory Data Exploration System. IEEE 35th International Conference on Data Engineering(ICDE), Macau, China, April 2019.

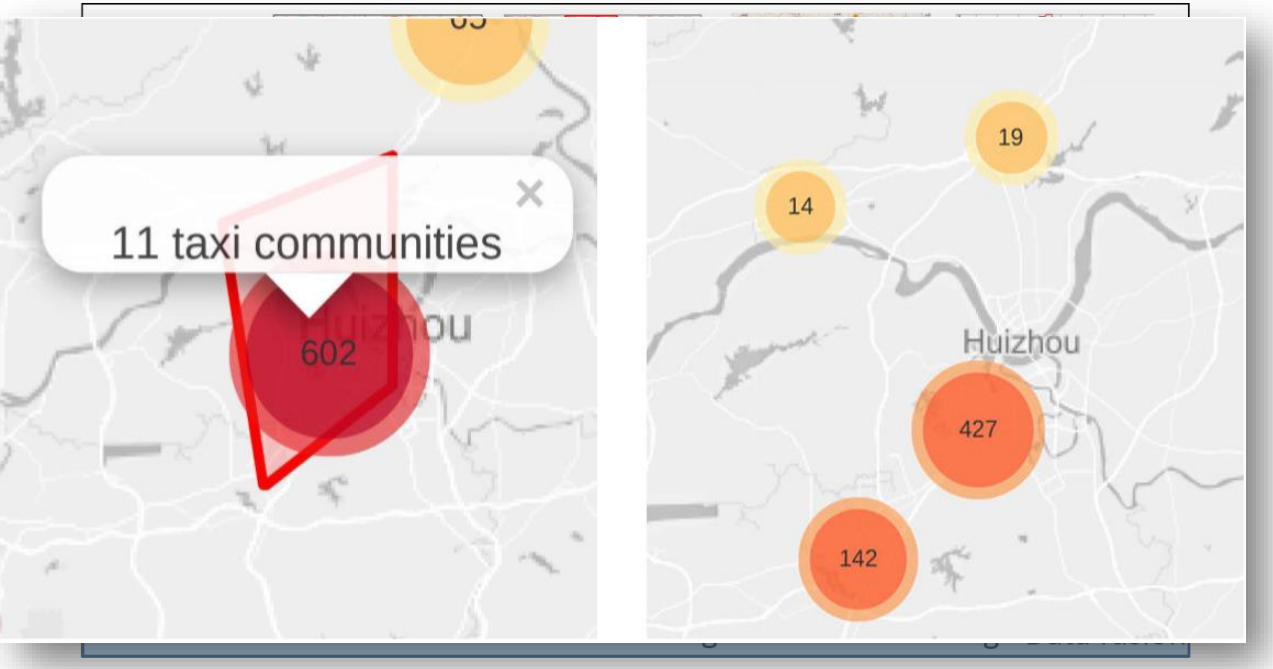
Vaite Framework --- Exploratory analytical layer



- Core Computation Layer
- A suite of atomic operators:
  - *event extractor, sub-traj extractor…*
- A set of analysis models:
  - *clustering, correlation…*

Chuang Yang, Yilan Zhang, Bo Tang, Min Zhu. Vaite: a Visualization-Assisted Interactive Big Urban Trajectory Data Exploration System. IEEE 35th International Conference on Data Engineering(ICDE), Macau, China, April 2019.

Vaite Framework --- Interactive visualization layer



- Trajectory exploratory visualization tools:
  - Spatial bubble view
    - design for spatial aggregation task.
    - supports a set of aggregation operators,
    - e.g., *Min*, *Max*, *Avg*, *Sum*
  - Trajectory-map view
  - ……
- Interactive action:
  - Convert keyboard/mouse input to analysis operator and pass them to computation engine.

Chuang Yang, Yilan Zhang, Bo Tang, Min Zhu. Vaite: a Visualization-Assisted Interactive Big Urban Trajectory Data Exploration System. IEEE 35th International Conference on Data Engineering(ICDE), Macau, China, April 2019.

Outline

- Introduction
- System Architecture
  - Three Layer Architecture
- **Case Study**
  - Data Model
  - Three Real World Cases
- Summary

Case Study --- Data Model

- **Data Source:**
  - Shenzhen Taxi Trajectories Data
- **Data Size:**
  - nearly 6GB/day, nearly 80,000,000 GPS Records.
  - overall Data: one month data
  - in use: one week data(2016_01_25 ~ 2016_01_31)
    - Size: 41.8 G
    - Records: 562,476,072
- **Data Characteristics:**
  - Massive: growing day by day.
  - Noisy: e.g., Value Missing, Wrong Records, GPS Location Floating

```
val schema = StructType(
  StructField("TerminalNumber",StringType,nullable=true)::
    StructField("Longitude",DoubleType,nullable=true)::
    StructField("Latitude",DoubleType,nullable=true)::
    StructField("GPStime",TimestampType,nullable=true)::
    StructField("DeviceNumber",IntegerType,nullable=true)::
    StructField("V",IntegerType,nullable=true)::
    StructField("Direction",IntegerType,nullable=true)::
    StructField("P_State",BooleanType,nullable=true)::
    StructField("WarnCode",IntegerType,nullable=true)::
    StructField("SimId",IntegerType,nullable=true)::
    StructField("S_State",BooleanType,nullable=true)::
    StructField("Color",StringType,nullable=true)::Nil
)
```

```
粤B642ZB,113.960197,22.572617,2016-01-25 00:00:07,1457167,0,315,1,,,0,蓝色,
粤B8NN15,113.807899,22.626801,2016-01-24 23:59:57,1398983,0,0,0,,,0,蓝色,
粤B8MT13,114.035599,22.559999,2016-01-24 23:59:43,1382069,0,0,0,,,0,蓝色,
粤B4V4P7,113.856598,22.585800,2016-01-25 00:00:08,1569738,0,0,0,,,0,蓝色,
粤B8V7V9,114.137581,22.609249,2016-01-25 00:00:06,1571029,0,90,0,,,1,蓝色,
粤B7705D,113.984703,22.694599,2016-01-24 23:59:57,1145153,17,0,0,,,0,蓝色,
粤B3K0P6,114.209114,22.732332,2016-01-25 00:00:06,1589342,82,225,0,,,0,蓝色,
粤B0P4A0,114.186798,22.641899,2016-01-24 23:59:42,1653325,26,0,0,,,0,蓝色,
粤B6VU46,114.131065,22.562017,2016-01-24 23:59:53,1422883,38,225,0,,,0,蓝色,
粤B3WK76,113.950600,22.546700,2016-01-24 23:59:49,1427958,0,0,0,,,0,蓝色,
粤B4HR57,113.814468,22.715549,2016-01-25 00:00:05,1372065,39,290,0,,,1,蓝色,
粤BW2555,114.304047,22.727768,2016-01-25 00:00:05,1453336,0,90,0,,,0,黄色,
粤B1HL13,113.920799,22.672501,2016-01-24 23:59:57,1370266,40,0,0,,,0,蓝色,
粤B1P1R2,113.914497,22.528601,2016-01-25 00:00:06,1653419,0,0,0,,,0,蓝色,
粤BN43F1,114.119316,22.597668,2016-01-25 00:00:07,1511529,23,315,0,,,0,蓝色,
粤B3WK76,113.950600,22.546700,2016-01-24 23:59:48,1427958,0,0,0,,,0,蓝色,
粤B2H5T2,114.064468,22.537867,2016-01-25 00:00:07,1575015,33,225,0,,,0,蓝色,
粤BR45W0,114.114250,22.546949,2016-01-25 00:00:04,1531606,24,283,0,,,1,蓝色,
粤B4NJ18,113.821899,22.650999,2016-01-24 23:59:57,1395340,9,256,0,,,0,蓝色,
```

Chuang Yang, Yilan Zhang, Bo Tang, Min Zhu. Vaite: a Visualization-Assisted Interactive Big Urban Trajectory Data Exploration System. IEEE 35th International Conference on Data Engineering(ICDE), Macau, China, April 2019.

Case Study --- Taxi Community Identification

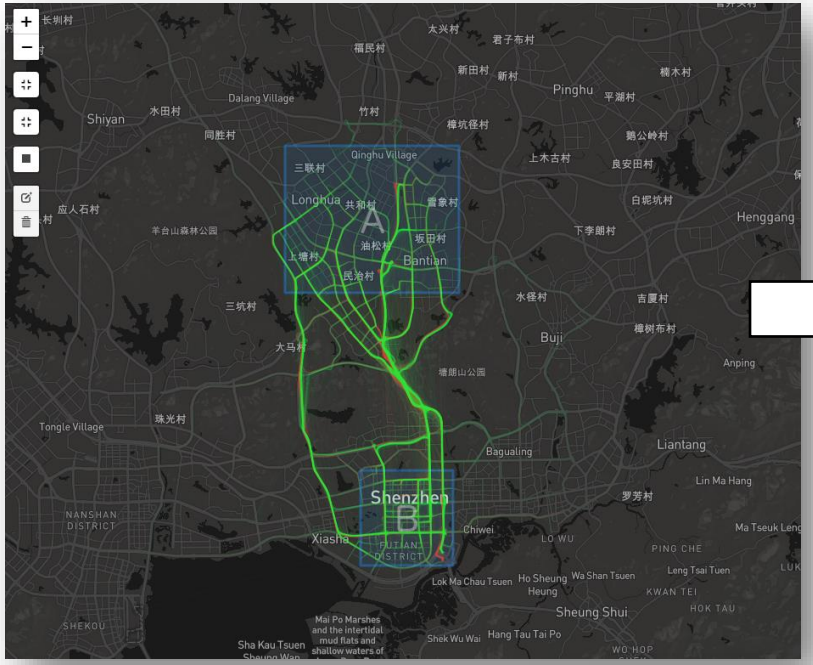
a) Extracting the pick-up events by **event extractor** and Visualizing them by **heat-map**

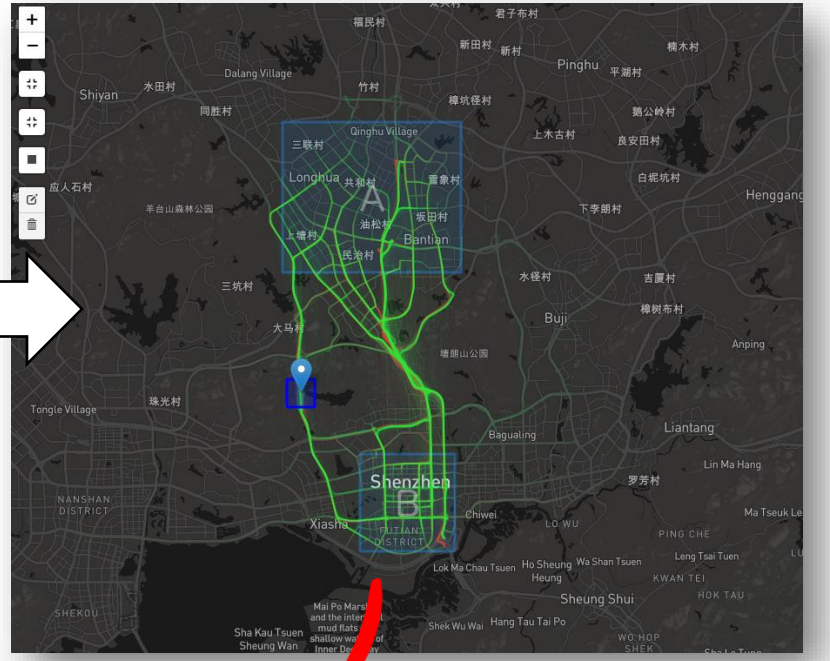
d) Applying **aggregation analysis model** and identifying taxis with top-10 pick-up events

b) **Clustering** pick-up event by spatial features

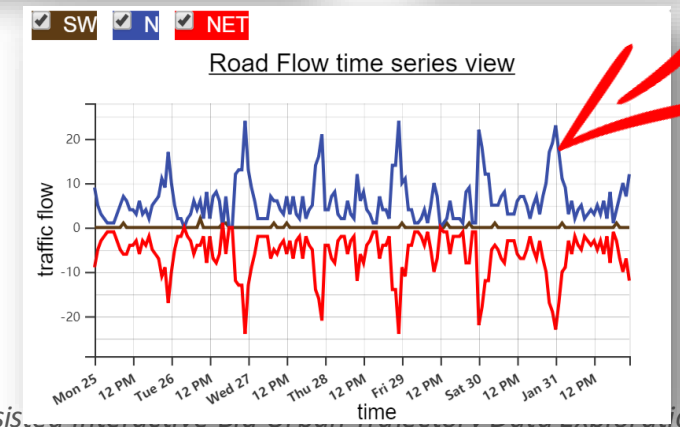
c) Aggregating clustering results and Visualizing them by **spatial bubble view**

Case Study --- Traffic Flow Exploration





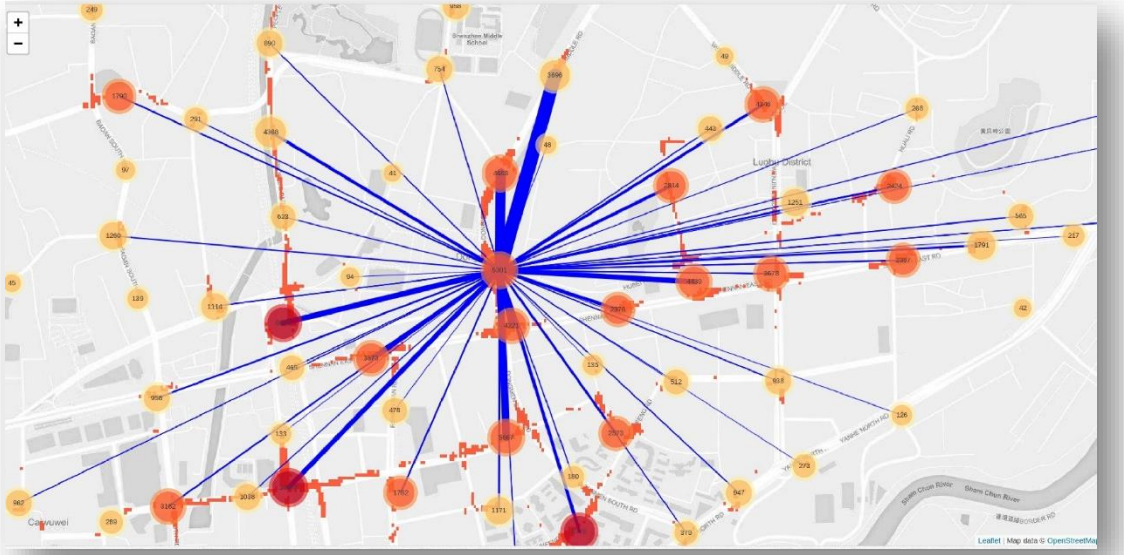
Step 1. Selecting two areas on the map (by mouse), extracting and visualizing all sub-trajectories between them (identifying popular paths).

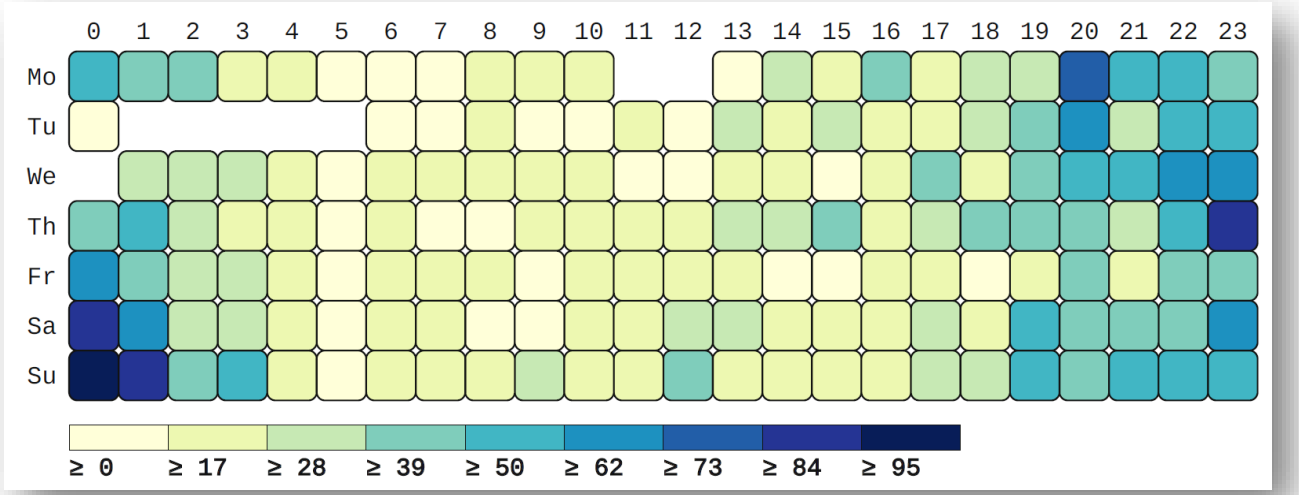


Step 2. Investigating one of popular path between two areas (by mouse), the traffic flow will be visualized.

Chuang Yang, Yilan Zhang, Bo Tang, Min Zhu. Vaite: a Visualization-Assisted Interactive Big Urban Trajectory Data Exploration System. IEEE 35th International Conference on Data Engineering(ICDE), Macau, China, April 2019.

Case Study --- Traffic Congestion Exploration





1. Finding traffic jams locations: **spatial bubble view**
2. Exploring traffic jams correlations: **star topology view**
3. Identifying the traffic jam trend: **pixel based view**

Chuang Yang, Yilan Zhang, Bo Tang, Min Zhu. Vaite: a Visualization-Assisted Interactive Big Urban Trajectory Data Exploration System. IEEE 35th International Conference on Data Engineering(ICDE), Macau, China, April 2019.

Outline

- Introduction
- System Architecture
  - Three Layer Architecture
- Case Study
  - Data Model
  - Three Real World Cases
- **Summary**

Summary

- **Interactively:** support ad-hoc exploratory analysis, near real-time computation
- **User friendly:** easy to use, results are easy to interpret
- **Extensible:** customize for users' analysis operator, easy to configure for specific analysis tasks

Chuang Yang, Yilan Zhang, Bo Tang, Min Zhu. Vaite: a Visualization-Assisted Interactive Big Urban Trajectory Data Exploration System. IEEE 35th International Conference on Data Engineering(ICDE), Macau, China, April 2019.

Thanks for your listening !
Q&A