Article

West Lake Tourist: A Visual Analysis System Based on Taxi Data

Yunliang Jiang [1], Junjie Cao [2], Yong Liu [3] and Jing Fan [3],*

[1] School of Information Engineering, Huzhou University, Huzhou 313000, China
[2] College of Mathematics and Computer Science, Zhejiang Normal University, Jinhua 321000, China
[3] Institute of Cyber Systems and Control, Zhejiang University, Hangzhou 310000, China
*   Correspondence: fanjing105@zju.edu.cn

check for updates

Abstract: Mining the mobile pattern of the urban population plays an important role in city construction, and visual analysis is a powerful technique in studying mobile patterns. In this paper, based on the taxi trajectory data in Hangzhou, we share our design for an interactive visual analytic system, which helps analyzers leverage their domain knowledge to gain insight into travel patterns, including travel time rules of tourists and the distribution rules of pick-up and drop-off locations. Besides, our system can present the dynamic travel process and the Point of Interest (POIs) information of the origin and the destination. A case study has been conducted, which verifies that our system can provide tools for urban managers or urban experts on the design of scenic spot open entrances and exits and travel route planning.

Keywords: visualization; system; taxi data; visual analysis

1. Introduction

In recent years, with the development of Internet of Things (IoT) technologies and the growing of urban populations, urban data, such as social entertainment data, trajectory data and financial data, is exponentially growing. Taxi GPS data is one type of trajectory data. Taxi data contains people's daily travel habits, such as time information, area information, router preference, speed, and other travel information. Thousands of taxis produce a lot of data, which is typically spatiotemporal data, every day. Therefore, with the help of visualization techniques, we can find the travel pattern of a large scale people based on taxi data, which provides decision-making assistance for city managers through visual analysis. For example, the scenic administrators can arrange their openings and closings more reasonably if they obtained the visit time rule of a tourist; traffic managers can take measures to avoid traffic congestion if they get the hot area of drop-off; in addition, tourism managers can also plan tourist routes reasonably. Moreover, Hangzhou is a famous tourist city, where taxis are the preferred transportation for foreign tourists, so taxi trajectory data contains great value.

In this paper, based on taxi trajectory data of Hangzhou, we provide a visual analysis system with the scenic West Lake as the target area. Through interactive analysis, managers can gradually deepen and explore the travel patterns of the target tourists, such as travel time patterns, travel area patterns, dynamic presentation of Origin and Destination (OD) travel process, and spatial semantic analysis combined with the Point of Interest (POI) data. Moreover, urban experts can also take advantage of their domain knowledge to explore by themselves. The contributions of our works are as follows:

(1) Find the travel patterns of tourists by visual analysis. Combined with the background of Hangzhou tourism city, through the design of the visualization system, the travel patterns specific to tourists can be explored by visual analysis. The tool can be directly used by experts in the city planning field.

(2) Dynamic presentation of OD travel process. OD travel process means the spatiotemporal trajectory evolution. The co-evolution demonstrates the travel patterns of the crowd. Therefore, our system demonstrates the entire process of a trip and uses the clustering algorithm to group the tourists. For these tourists, our system can offer both time-consuming comparisons and changes in spatial distance.

Next, we give related work in Section 2, while Section 3 introduces our dataset and high-level architecture of visual system. Section 4 introduces the interface design in detail, Section 5 is a smart tourism case analysis, and in Section 6, we summarize our work and introduce the next plan.

2. Related Works

2.1. Visualization of Selection and Query

In the visualization of data, visual selection is the fundamental action of visual analysis. Ahlberg et al. firstly proposed a dynamic query concept, and the query was generated by graphical components [1]. Heer et al. proposed an interactive relaxation query strategy [2]. Liu et al. contributed methods for interactive querying (e.g., brushing and linking) among binned plots through a combination of multivariate data tiles and parallel query processing [3]. Another noteworthy element is a system named Tableau mentioned by Polaris et al. [4]. The authors provided a visual mode for the user and, by dragging and dropping the column of the data table and putting it in the side bar, visualization results were displayed in the area to achieve interactive exploration of the data. Modoni et al. introduced the Knowledge Graph Semantic Framework, an interactive and explorative visual environment for the semantic search of information leveraging the knowledge Graph proposed by Google [5]. Modoni et al. introduced a dedicated software dedicated to the Factory telemetry visualization. This software comprises *x-y* charts, as well as scattered and waveform plotting, which permit to show, under various different views, the data telemetry acquired from the sensors [6]. In this paper, our system was designed with rectangular and circular frame selectors which are different from the aforementioned systems, as tourists are our visual selection target. Based on the frame selector, users can select target tourist data by visual frame and edit the selection frame at the same time.

2.2. Spatiotemporal Data Visualization

The interactive visual analysis can further release the object's movement pattern and the cluster category of the trajectory [7]. Schreck et al. proposed an interactive management and control framework [8], and an outstanding feature of this framework is that it can combine the automatic analysis of data with the monitoring of human experts, and users can control the map to obtain the desired clustering result. The OPTICS algorithm can be used to cluster traffic trajectory data [9], users can further define clustering results by visual frame selection, such as excluding one or partial subsets from a cluster to form a new cluster, or split a subcluster into smaller clusters, which are important in loop analysis tasks. Zeng provides a visual analysis system to study the network change patterns in mobile data [10], Interchange Circos Diagram (ICD) is designed to check network change patterns, and it supports visual exploration in three scales: cities, districts, and roads. At the city's scale, an ICD represents a portioned area, and in the scale of the road, an ICD represents the road junction. Guo designed a Triple Perspective Visual Trajectory Analytics (TripVista) to explore complex traffic trajectory data through interactive analysis [11]. The system consists of three windows, traffic flow, ThemeRiver plot and the parallel coordinate plot (PCP). The traffic window shows the spatial information, the theme river shows the direction information, and the PCP shows the statistics of the traffic data from multiple angles. According to these windows, the user can explore the rules and irregular traffic data autonomously. Taxi theme visual analysis is a visual analysis system designed by D. Chu [12], and this system is used to specifically find the movement pattern of taxi trajectory. D. Chu combines four windows: taxi theme map, theme route, street cloud, and PCP; and the main contribution of the system is the transformation of trajectories into documents, where latent Dirichlet

allocation (LDA) [13] is used to discover the hidden document semantic information to form a taxi topic, so that users can find interesting move patterns. In the visualization of two-dimensional spatiotemporal data, in order to unify temporal and spatial data into a graph, Space-Time-Cube (STC) is a widely used method [14]. In space-time cubes, 3D trajectory is mapped into a three-dimensional coordinate system, the *x*-axis and the *y*-axis are used to represent the longitude and the latitude, and the *z*-axis represents the time. In this way, the irregular trajectory in the spatial object is converted into a trajectory in the cube, which makes the rules of trajectory become traceable.

The disadvantage of the stack cube is that it is static. Essentially, it uses space to exchange time, however, space-time cube cannot find the information released during the time change. So, taking into account the needs of the analyzer, we present the time and space through dynamic demonstration. The time evolution and spatial evolution are presented to users at the same time, and the trajectory is drawn with different colors for subsequent analysis by users.

2.3. POI Data Analysis

The development of a mobile internet makes it possible to provide richer mining based on a combination of multi-source data. Jiang et al. constructed population flow patterns of more than 300 cities and regions in China using the big data provided from Baidu LBS open platform [15]. Huang conducted Weibo public opinion analysis through Weibo text dataset and visualized the microblog public opinion profile image contained in the dataset [16]. Liu Bomin selected the Baidu Heatmap data of the waterfront area along the Qinhuai River in Nanjing under the dynamic background provided by Baidu Thermal Map [17]. Through GIS analysis and process, the Qinhuai Riverside was summarized from three parts of the spatiotemporal type. It provides an evidence for the spatial planning, design, and vitality improvement of the Qinhuai River. In view of this, West Lake, as a waterfront space and as our research area, also provides evidence for Hangzhou city managers on space planning and design. Meanwhile, we used the POI data provided by the Gaode API to assist us. By exploring the clustered central area, the numbers of typical POIs in the area demonstrate the social function.

3. Dataset Introduction and System Architecture

Next, we will introduce our data and target area in Section 3.1. In Section 3.2, we will introduce the process of data cleaning, the architecture of our visual system, and the Redis cache technology which was applied in our system.

3.1. Data Introduction

3.1.1. Target Area

Hangzhou, located in the southeast region of China, is the capital of Zhejiang province and one of the most famous tourist cities of China. Considering the character of HangZhou, we draw a polygon area around the West Lake, which is a famous scenic spot as the target area. The red dot enclosed area marked by the black circle in Section 4 shows our target area.

3.1.2. Taxi Data

The taxi GPS dataset comes from the taxi company of Hangzhou and spans seven consecutive days, from 1st November 2014 to 7th November 2014. Within the target area, the daily record number is 953,165, 848,780, 907,115, 912,148, 9,096,031, 918,853, 969,141, and the data size is 8 GB. Each record consists of taxi_id, gps_x, gps_y, state, speed, and date_time. The value of attribute state is 0 or 1, indicating whether or not there are passengers in the taxi. A vehicle can be uniquely identified with its taxi_id.

3.1.3. POI Data

We apply for having the Gaode map developer obtain the POI data online through the Gaode API. Four kinds of POIs (schools, shopping malls, residents, enterprises), respectively, represent four

functional areas: cultural educational areas, commercial areas, residential areas, and industrial areas. The number of POI stands for the spatial semantic information.

3.2. Data Cleaning and System Achitecture

3.2.1. Data Cleaning

Source dataset inevitably has some dirty data. For example, if all records of a taxi is all 0 or 1, we must filter them.

3.2.2. System Architecture

We can get the architecture of our system from Figure 1. The web of our visual system is programed by HTML, CSS, and JavaScript language combined with the D3.js and Leaflet.js libraries. The service is programmed in Python language, combined with the Flask framework to respond to requests. We used MongoDB as our database. MongoDB is a database that is good at storing JSON format data. The JSON format uses {key, value}, a common format for current data exchanges, and our data is, therefore, organized in JSON format. Mysql, which is another popular database, can also be used to store data, but the data format is not JSON; however, when the amount of data is too large, its performance is not as good as MongoDB. We added a layer of caching with Redis to speed up data response, which can improve users experience. Redis is also a key-value cache database, where data is stored in memory and speeds up access, which has a wide applications in the cache databases.
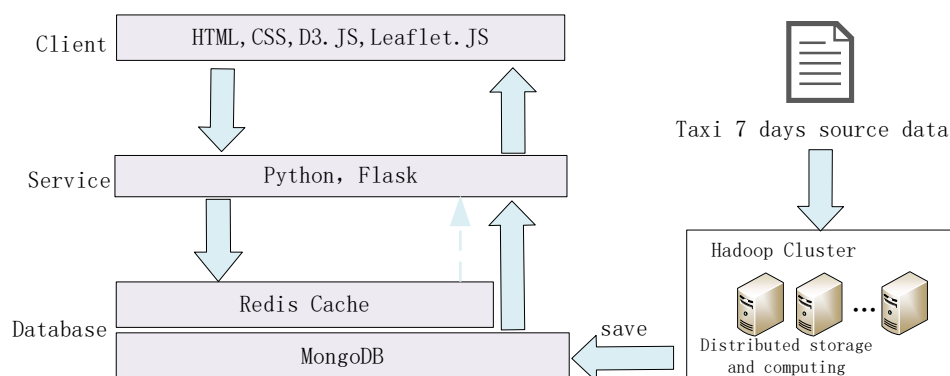

Figure 1. System architecture.

3.2.3. Redis Cache Technology

Cache technology is a method used by the server to speed up data access requests. The data in the cache will not be requested again from the server and will be directly returned to web from cached of Redis. The pseudocode of data cache is described as follows:

```
if key in Redis:
        value = GET key
else:
        access MongoDB;
        value = GET key
        access Redis;
        SET key value
return value;
```

The user will request similar data in the interactive analysis, so we made a layer of Redis cache. For example, in the OD module exploration, the user will select the tourists with rectangle after requesting the number of pick-up and drop-off according to the time form request, and the data cached

in Redis is sent to the server along with the coordinates of the rectangle. There are two advantages: one is the data will be quickly responded to, which improves the user experience, and another advantage is that it reduces server calculation resource waste. The Redis cache flow graph is showed in Figure 2.
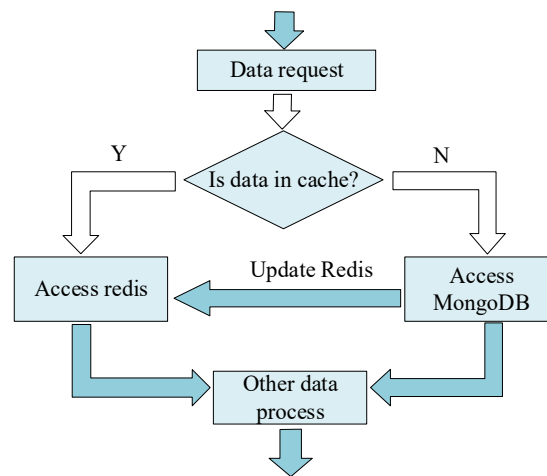

Figure 2. Data cache flow graph.

4. Visual Interface

The main interfaces of our visualization system are shown in Figure 3 and mainly consists of three modules. The first one is the basic data statistics module, as shown in Figure 3A, which shows some statistical results of the data in different time dimensions. The second one is the heatmap and scatter plot linkage module, as shown in Figure 3B. By clicking the scatter point, we can get information on both the travel space situation and numerical quantification. The third one is the OD trajectory visual analysis module, as shown in Figure 3C-1,C-2. By selecting the target traveler autonomously, we can get the distribution time and the travel trajectory of the target tourist. Combined with the POI information, the social functional of OD can be obtained.

Figure 3. Main interface of the visualization system: (**A**) Static data about pick-up and drop-off in different time dimensions; detailed description is in Section 4.1. (**B**) Heatmap and scatter plot linkage module, where we can obtain information on both visual perception and quantitative value. Origin and Destination (OD) travel visual analysis module, with the area surrounded by six red points in (**C-1**) our target area, the inner lake is West Lake, and (**C-2**) the result after a visual analysis.

4.1. Time Visual Analysis View

Before visual analysis, if users have a basic understanding about data, they can solve the problem more quickly. So, we designed a basic statistical module of data that we can use to get basic information on the data before expanding the specific visual analysis.

Figure 3A, upper left: This represents the total number of pick-up and drop-off per hour in the target area, which can be converted into a line graph through the widget.

Figure 3A, top right: This presents four different travel proportions: the morning, afternoon, evening, and midnight. The outer is drop-off, and the inner is pick-up. It is easier to find the crowd from a coarse-grained perspective.

Figure 3A, bottom left: This shows sum of trips per hour for seven days.

Figure 3A, bottom right: This shows different travel proportion in a week; proportion = the number of a day trips/the number of seven-day trips.

4.2. Heatmap and Scatter Plot Linkage View

The heatmap can more intuitively detect the abnormal flow in the area. The Itoh designed heatgrid presents all the heat points simultaneously [18]. However, there is no spatiotemporal, which is helpful to trajectory data analysis. Therefore, we designed heatmap and scatter plots at the same time to explore the travel changes in the target area in a wider spatial and temporal perspective. The heatmap expresses the change in space, and the scatter plot expresses the specific change of flow in time. An interactive exploration way, which complements the spatiotemporal perspective, has been given showed in Figure 4a,b.

Linkage Between Heatmap and Scatter Plot View

The 24 h is the *x*-axis of the scatter plot; that is, in the time axis, the *y*-axis represents the travel total numbers of this hour. Denote the travel amount as Q, and then the 24 h amount is (Q1, Q2, Q3 . . . Q23, Q24). The amount of an hour is represented by a scatter circle, and the base color is red. The Q is mapped to the radius R of the circle, and the larger Q is, the larger the R is. The color of each spatial point in the heatmap represents the congestion and patency of the trip. The larger the Q gets will change the color hue. Four kinds of colors represent different levels of Q. Besides, interactive operation is provided. When the date is selected, the 24 h travel scatter plot is first rendered. When the mouse clicks on a red scatter, the hour's heatmap will appear on the map, and the mouse will display the Q. This will help users analyze the data from two perspectives: visual perception and quantitative value.

(**a**)   (**b**)
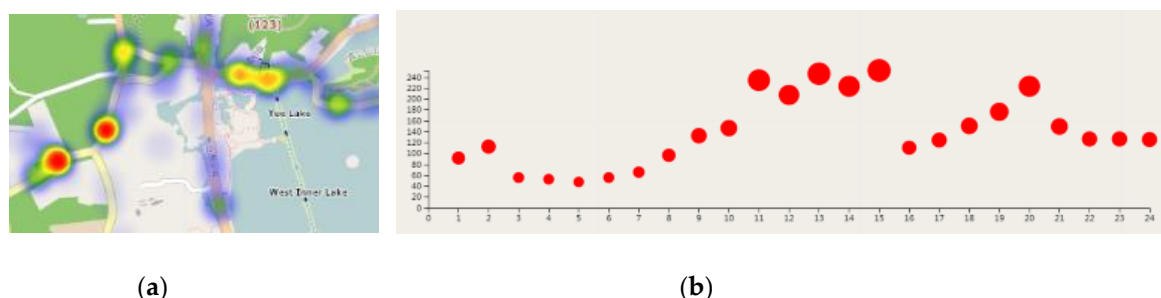
Figure 4. Visualization system partial views: (**a**) Heatmap is designed to show the temporal information. (**b**) *x*-axis stands for 24 h, and *y*-axis stands for the total numbers of corresponding hour. Click the circle, and the number will float above it; the heatmap will be shown on the map at the same time.

4.3. OD Travel Visual Analysis View

4.3.1. OD Visual Analysis Process

The user can freely explore the travel trajectory in the target area; the steps are as follows:

Step 1 Select the time period *t*, and generate the pick-up point or the drop-off point in the target area.

Step 2 Click the rectangle button to select or edit the area. Click on the cluster button to cluster the coordinates of origin or destination of the target tourist using the *K*-Means algorithm. The clustering attribute is the coordinates of the pick-up or the drop-off. The clustering center will be displayed as a bigger marker on the map.

Step 3 Click the OD button, and the markers which stand for the tourist will move, leaving a travel trajectory on the map.

Step 4 Clicking the cluster center will make a time bar view pop up.

Step 5 Clicking on the POI button will make the system calculate the amounts of POI (school, enterprise, residential, shopping) as a radius of 500 m nearby the cluster center, as well as display the results in the fireworks view.

4.3.2. Visual Selection and OD Trajectory View

Visual selection is an important part of visual analysis. We have designed rectangle and circle as selectors. After selection, it can be edited, changed, and deleted. Meanwhile, the button for adding markers on the map is a aid, as shown in Figure 5a.

The OD trajectory is left by the target tourist after clustering. We use four colors (red, yellow, green and black) to represent different clusters, and we highlight the cluster center with larger marks. Different clusters and trajectories map different colors shown in Figure 5b.

(**a**)   (**b**)

Figure 5. (**a**) Edit the selection. (**b**) The trajectory after demonstrating dynamically.

4.3.3. Time Bar View

In Section 4.3.1, Step 2, four clusters in different color are generated. The time bar is designed according to the travel time of the clusters. Figure 6 represents the time distribution of some tourists. It was designed as a one-dimensional timeline, so that we can obtain the time information intuitively. Each colorful block represents a trip. Different colors represent the trips of different individuals.
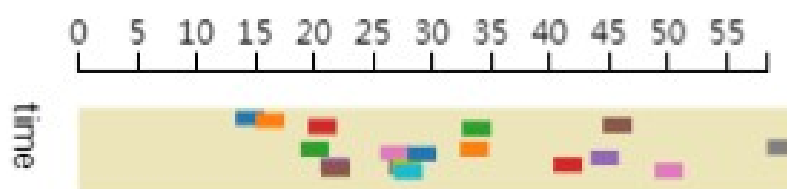

Figure 6. Time bar shows the time distribution of this cluster's tourists.

4.3.4. POI Fireworks View

In Step 5 of Section 4.3.1, according to the sum *N* of representative POI near the center of the cluster, a fireworks view is designed to display the results, so as to obtain social function information

of the area showed in Figure 7. Taking the yellow cluster as an example, N is the number of POI at a radius of 500 m around the cluster center. We select four representative POI (enterprise, residential, shopping, and school), to respectively represent four different functional areas (work area, living area, shopping area, and cultural and educational area). We then display the POI numbers in the form of arcs. Since the number of POIs is related to the selected area, the web interface is limited and the aesthetics are considered, so we apply *N*/20 as the number of arcs. Different kinds of POI are mapped with different colors. Basically, the number of arcs represents the social functional area character. Because of the long term development of the city and the integration of functional areas, it is better to study an area in the multiply functional view.
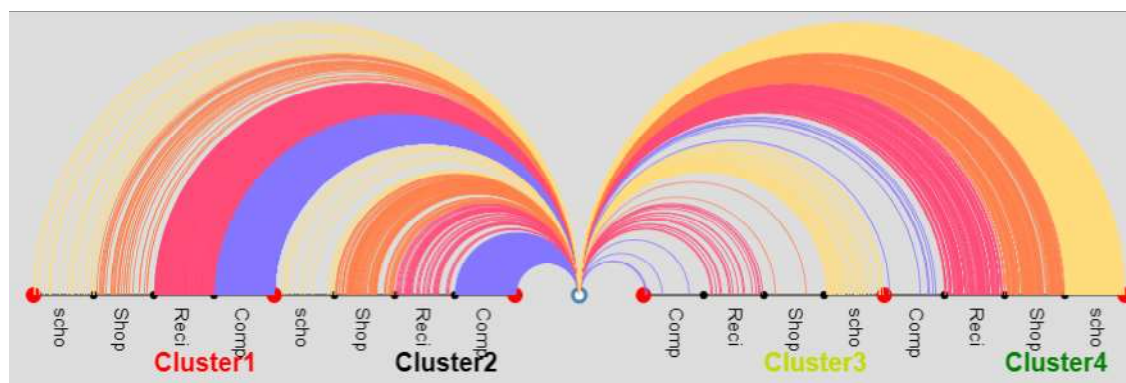
Figure 7. Point of Interest (POI) fireworks show the POI information of OD.

5. Visual Analysis Case

In this part, taking the 3rd day (Monday) as an example, we will show a complete layer-by-layer visual analysis process of the visualization system to obtain the behavior of tourists in West Lake. There are three parts: travel time, travel heat area, and travel OD analysis. We want to provide a migrated method on spatial-temporal data visual analysis process.

5.1. Travel Time Patterns

The first basic statistic module shows some basic data shown in Figure 8. The proportion of pick-up and drop-off each day are listed below. We can see that Monday has a higher proportion pick-up and drop-off, but how about every hour in detail a day?

drop-off: (12.86%, 14.01%, 14.23%, 9.48%, 16.03%, 19.26%, 14.13%);
pick-up: (9.54%, 15.95%, 16.43%, 17.42%, 15.35%, 19.71%, 5.6%).

We can transform to the line chart using the top right corner widget. We can see, starting from 7 o'clock, the number of pick-up and drop-off starts to increase and the drop-off peaks at 10 o'clock, but the peak of pick-up is around 15 o'clock, indicating the different social attributes of the target area in different time periods. It is in the morning that people gathered at the West Lake, and it is in the afternoon that people keep away from the West Lake, which is also consistent with our common sense. There is a trough at 17 o'clock, and we can speculate that this phenomenon is caused by dinner. After that, the amount of travel resumes with its increasing paced, as people come out, and then decreases around 23 o'clock. This information about travel peaks and troughs is also reflected in the bottom left of Figure 3A. In terms of the overall trend, the total of drop-off is always greater than pick-up. This can also indicate that the taxi driver will not go to the congested place to pick up the passengers in the West Lake.
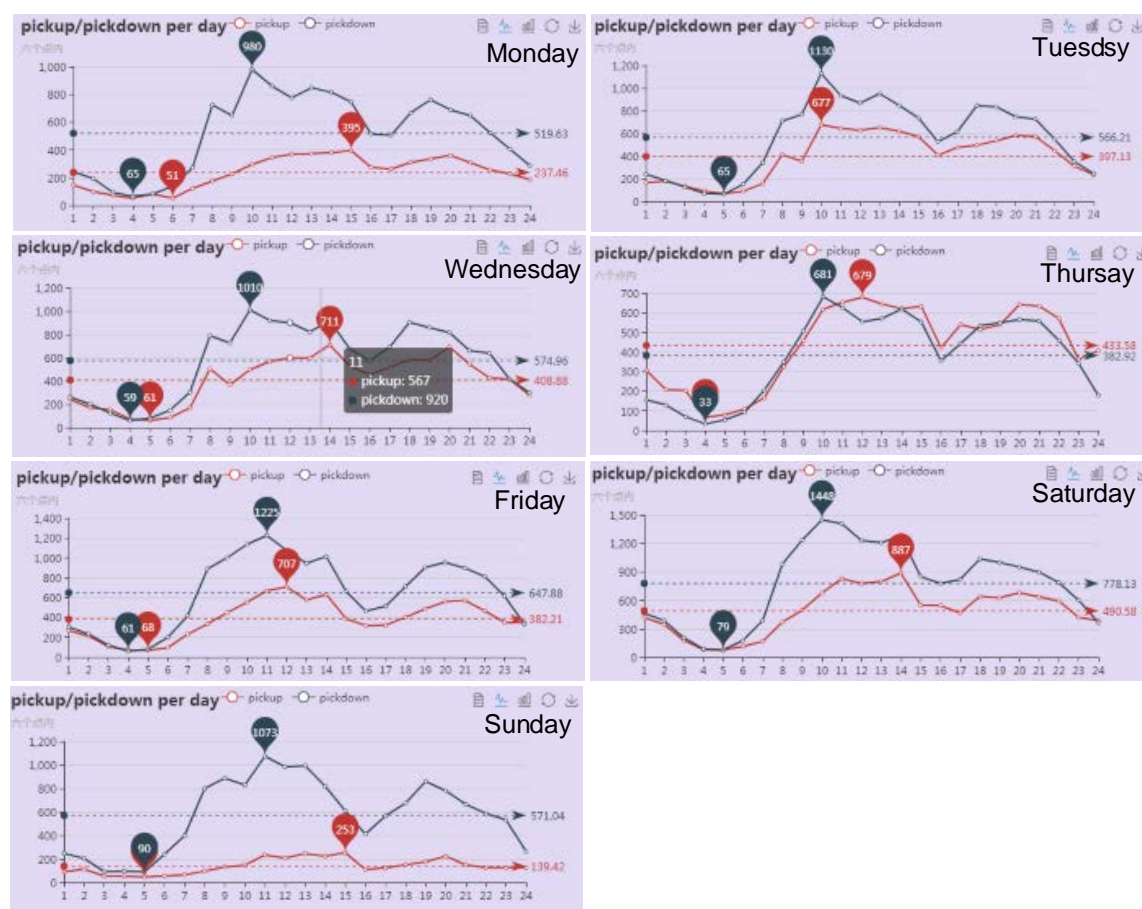
Figure 8. Distribution of travel times from Monday to Sunday.

5.2. Travel Regional Patterns

Next, if we know the preference area of pick-up or drop-off, we can provide evidence for traffic management, rather than just relying on experience of workers. Therefore, the second module heatmap is put to use.

We can see from Figure 9 that the hot spots are mainly located at the intersection of Nanshan Road and Beishan Street on Monday. The intersection with Nanshan Road is the second hot spot of pick-up, which reached its highest peak at 15:00 and then began to fall back. Until about 17 o'clock, it began to rise again. After about 20 o'clock, tourists began to leave.
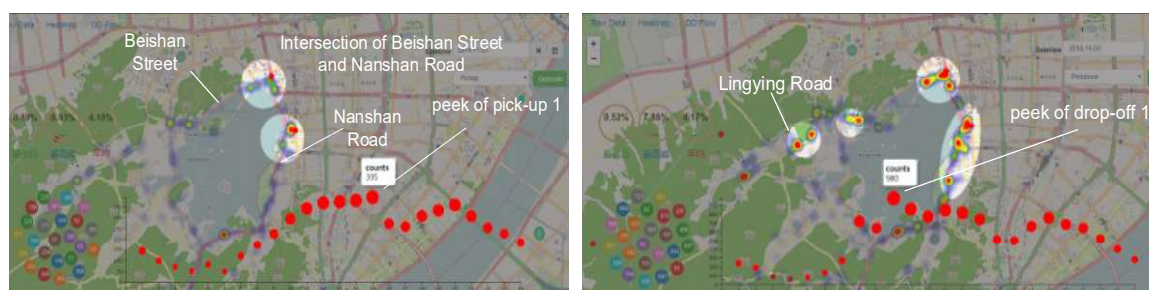

Figure 9. Heatmap images of pick-up and drop-off.

At about 20 o'clock in the evening, the number of pick-up reached another peak, mainly distributed at Nanshan Road and its intersection with Beishan Street. The main reason for the peak is that it this time was the end of the musical fountain show, and a large number of tourists were leaving. We

can also find that a lot of tourists appeared at the intersection of Nanshan Road and Beishan Street, and maybe these tourists from Nanshan Road came to take taxies. After 20 o'clock, the number of pick-ups for tourists gradually decreased.

We can get that the pick-up hot spot mainly distributed at the intersection of Nanshan Road and Beishan Street and the intersection of Nanshan Road and Jiefang Street. In terms of drop-off hotspot area, in addition to the area similar to pick-up, there are also many drop-off hotspots on Lingyin Road. Therefore, in addition to the area similar in pick-up, we should monitor these areas to prevent traffic congestion. Meanwhile, a city manager can broadcast this information to remind citizens of to avoid these hotspots.

The number of drop-off is much more active than that of the pick-up. Many tourists got off the taxi at the intersection of Lingzhu Road and Lingyin Road. We enlarged the map and found that the drop-off points were concentrated in one area, KFC. We can observe the change of the number of tourists one day in Figure 10. We can see that there are many tourists gathering during the day, even if the location of KFC is not in the central scenic area, so we can infer that the dining hall and rest facilities in the scenic area cannot satisfy the needs of tourists, and the manager should expand the catering and rest area.

(**a**) 7 h   (**b**) 10 h

(**c**) 16 h   (**d**) 19 h
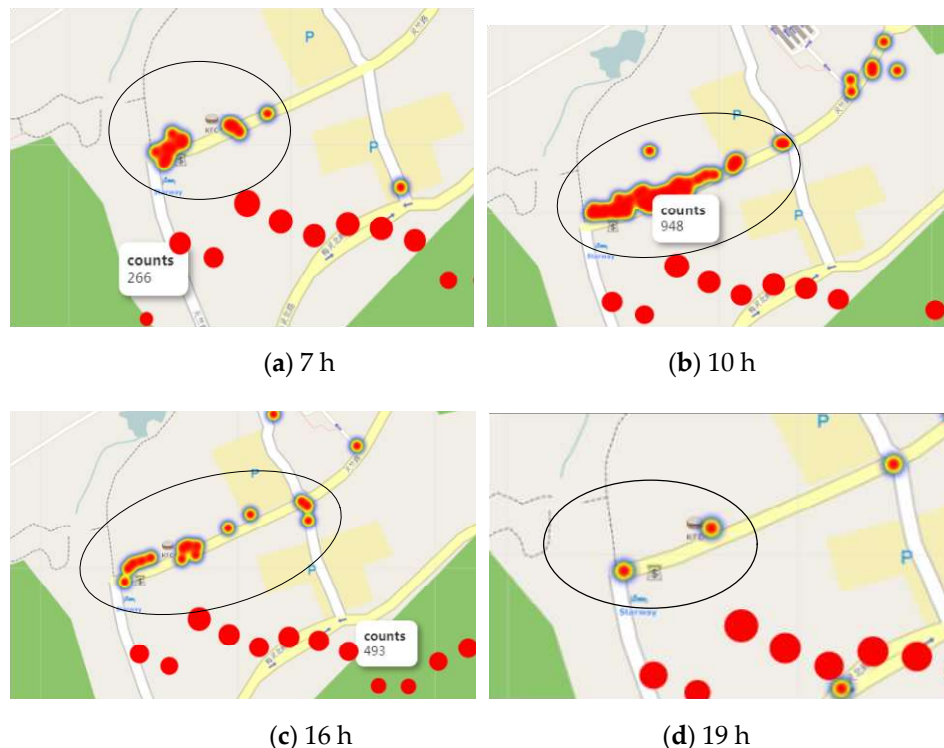
Figure 10. Heatmap at different time periods near KFC on Monday.

In the heatmap of Nanshan Road and Beishan Street, it can be seen from Figure 11a that the number of tourists that drop-off at the intersection of Broken Bridge and Nanshan Road and Beishan Street at 10:00 o'clock in the morning is more than the number of tourists that drop-off at the Jiefang Road intersection. But after 19:00 o'clock, the number of tourists who dropped off at Jiefang Road differs from the number of tourists who dropped off in the front of the Broken Bridge shown in Figure 11b. The difference between the two situations reflects the trend of tourists; during the day, people tend to go on the Broken Bridge for sightseeing, but more people will choose Jiefang Road as the drop-off point at night. Why? There are two possible destinations; one is a musical fountain and the other is for shopping, so we can conclude that people are more inclined to go shopping in the evening, at places such as at Intime Department Store.

The above analysis can draw some conclusions, such as tourists' sightseeing habits, in that people tend to go to the broken bridge to see the scenery during the day, but at night, people tend to go shopping, so the merchants can focus on sales at night. At the same time, since the three intersections are the hot areas for dropping off, the city management and security during the day and night should be strengthened at these intersections, so that traffic information can be broadcast by the car radio, if needed.

(**a**) 9:00 h   (**b**) 19:00 h
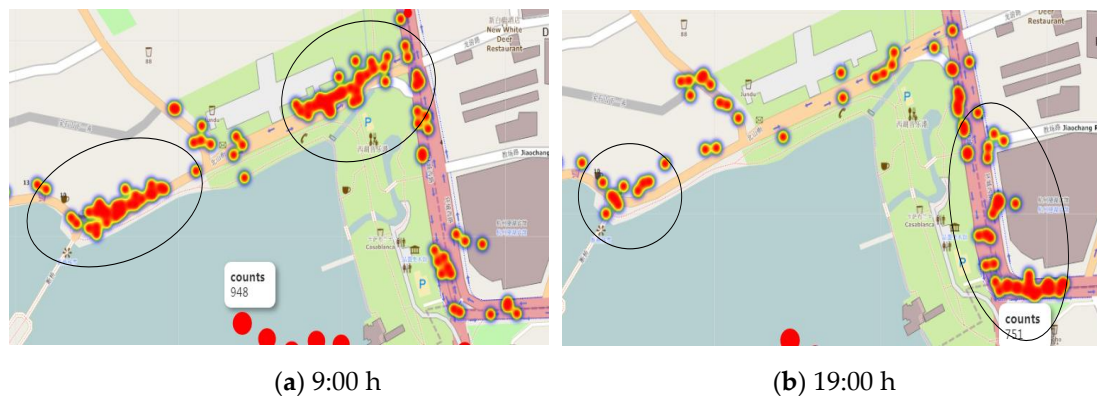
Figure 11. Analysis of the drop-off at the intersection of Nanshan Road and Beishan Street.

5.3. OD Spatiotemporal Information and POI Social Function Visualization

If we want to learn about which place some tourists come from, we can use the third module. We dynamically presented the travel process of the tourist painting the trajectory on the map, and visualized the travel time distribution of trips. As a coordinate can only get geographic information, we need to add the spatial semantics to the origin. Therefore, we requested the Gaode map API to learn the spatial semantics of area. The following four steps show a specific visual analysis process:

Step 1 Visual selection

Selecting the target tourist whom will be explored with the selector in the tool bar.

Step 2 Tourist clustering

Clicking the cluster button will cluster the destination area. As shown in Figure 12, the clustering results show four cluster centers centered on Tianmu Mountain, West Lake Cultural Square, Gongshu District, and Xianghu Lake. In the case of traveling on Monday, we can find that the yellow clusters centered on the West Lake Cultural Square cluster have the biggest number of tourists, indicating the main origin of the exploration area; the red clusters centered on the Tianmu Mountain represent the Xihu District, and the number of tourists in Gongshu District with Moganshan Road is the third.

Step 3 OD process displays dynamically

Clicking the OD button, markers standing for passengers will dynamically move and leave a trajectory. As shown in Figure 13, we can find the line of Lingyin Road and Beishan Street, represented by the red line, and the line of Nanshan Road, represented by the yellow line, is thicker than other routes, indicating that the travel volume of these road is larger. Upon clicking the cluster center marker of the yellow cluster, the time bar of the cluster will pop up, showing that the yellow cluster travel time period is mostly in 0 to 20 min, and the longest time cost is 40 min. Figure 14 shows the four clusters' travel time distributions, and it can clearly be seen that the people who are centered on the West Lake Cultural Square spend around 0 to 30 min, and the time cost centered on Tianmu Mountain is both 5 to 20 min and 35 to 45 min. The green cluster time cost of the Gongshu District, which is based on Moganshan Road, is 15 to 50 min, and there is only one route in black cluster. The time cost from Xiaoshan Airport is more than 50 min.

Step 4 POI semantic analysis

In the deeper exploration of the spatial semantics of the OD, we combined the urban POI. The generated POI fireworks view is shown in Figure 15. We can see the West Lake Cultural Square represented by the yellow arcs, which is the most densely curved, covers different functions, shopping malls, and schools as the first, and residence and enterprises is second. In the Gongshu District, represented by the green arcs, the POI numbers of residence and the school are more than the shopping malls and enterprises. The West Lake District, represented by the red arcs, are mainly residential areas.
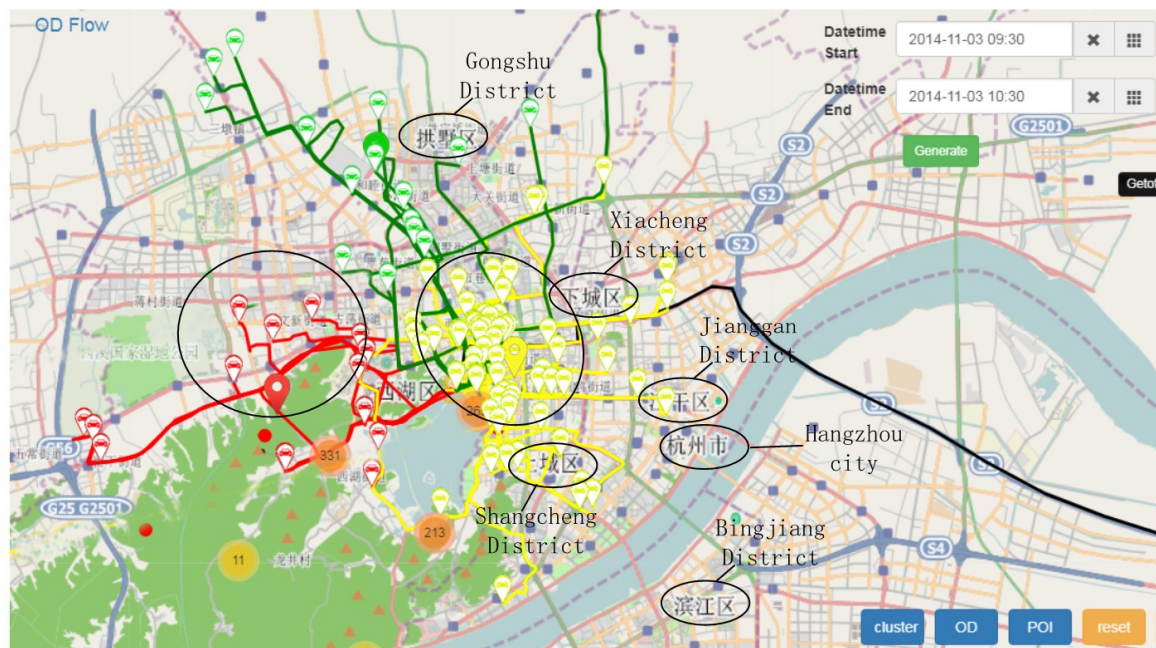

Figure 12. Cluster result: Four different color clusters showing the distribution of tourist origin from 9:30 to 10:30 a.m.
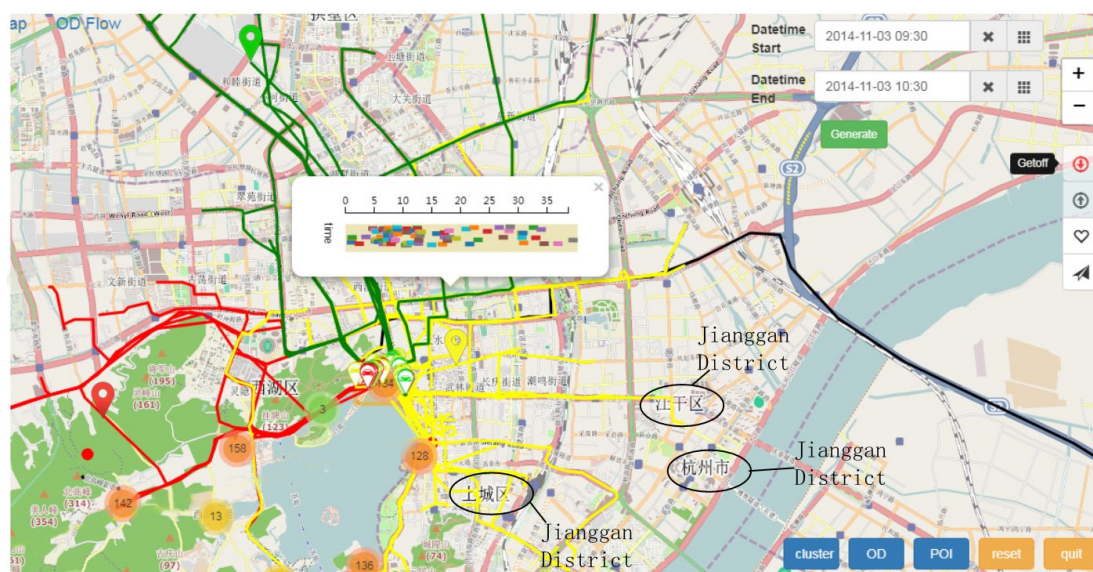
Figure 13. Travel trajectory: Trajectory left can show the main travel routes, and the time bar shows the distribution of travel time about this cluster; when you click the cluster center, it will pop up.
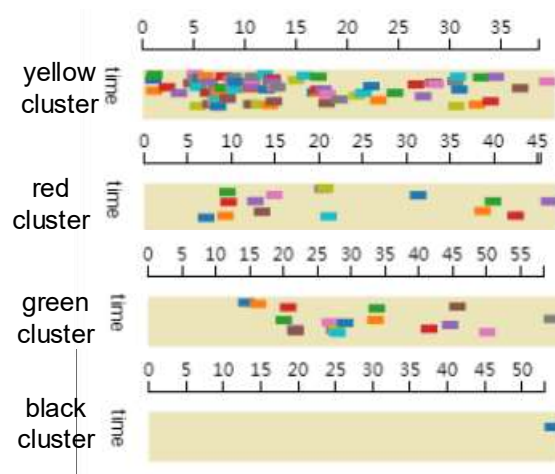
Figure 14. Travel time distribution of four clusters: Yellow cluster is less than 35 min mostly, and it is dense. Red cluster is in 5 min to 45 min, it is sparse. Green cluster is in 15 min to 55 min. There is only one in black cluster that exists, and its time cost is more than 50 min.
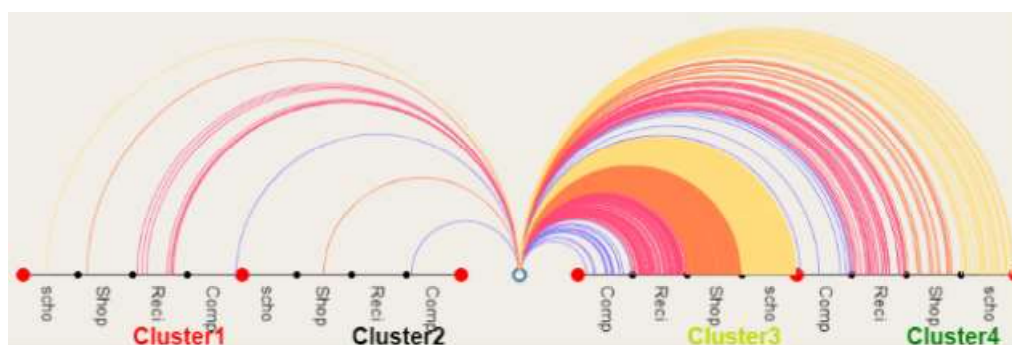
Figure 15. POI information fireworks: Near the Yellow cluster, POI is dense, residence, shopping mall and school is more, and company is few. This indicates that this is a mixture district of business and residence. Green amount is second, and the POI function is similar. Red is occupied by residence. Black is so sparse that we cannot obtain useful information.

6. Conclusions

In this paper, we proposed a visual analysis system based on taxi trajectory data of Hangzhou. Our system integrates time static view, the heatmap and scatter linkage view, and the OD travel visual analysis view, which explored the travel pattern of tourist in West Lake. Besides, we conducted a case study to verify that our system can provide tools for urban managers on the design of a scenic spot's open entrance and exit, as well as travel route planning.

In the next step, our target area will be expanded to Xixi Wetland to further discover the pattern of the people in the scenic spot under the tourist city.

References

1. Ahlberg, C.; Williamson, C.; Shneiderman, B. Dynamic queries for information exploration: An implementation and evaluation. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, Monterey, CA, USA, 3–7 May 1992.

2. Heer, J.; Agrawala, M.; Willett, W. Generalized selection via interactive query relaxation. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, Florence, Italy, 5–10 April 2008.

3. Liu, Z.; Jiang, B.; Heer, J. imMens: Real-time Visual Querying of Big Data. In *Computer Graphics Forum*; Blackwell Publishing Ltd.: Oxford, UK, 2013.

4. Peng, Y. A System for Query, Analysis and Visualization of a Multi-Dimensional Relational Database. *IEEE Trans. Vis. Comput. Graph.* **2002**, *8*, 52–65.

5. Modoni, G.E.; Sacco, M.; Terkaj, W. A semantic framework for graph-based enterprise search. *Appl. Comput. Sci.* **2014**, *10*, 66–74.

6. Modoni, G.E.; Sacco, M.; Terkaj, W. A telemetry-driven approach to simulate data-intensive manufacturing processes. *Procedia CIRP* **2016**, *57*, 281–285. [CrossRef]

7. Andrienko, G.; Andrienko, N.; Rinzivillo, S.; Nanni, M.; Pedreschi, D.; Giannotti, F. Interactive visual clustering of large collections of trajectories. In Proceedings of the 2009 IEEE Symposium on Visual Analytics Science and Technology, Atlantic City, NJ, USA, 12–13 October 2009.

8. Schreck, T.; Bernard, J.; Von Landesberger, T.; Kohlhammer, J. Visual Cluster Analysis of Trajectory Data with Interactive Kohonen Maps. *Inf. Vis.* **2009**, *8*, 14–29. [CrossRef]

9. Rinzivillo, S.; Pedreschi, D.; Nanni, M.; Giannotti, F.; Andrienko, N.; Andrienko, G. Visually driven analysis of movement data by progressive clustering. *Inf. Vis.* **2008**, *7*, 225–239. [CrossRef]

10. Zeng, W.; Fu, C.W.; Arisona, S.M.; Qu, H. Visualizing interchange patterns in massive movement data. In *Computer Graphics Forum*; Blackwell Publishing Ltd.: Oxford, UK, 2013.

11. Guo, H.; Wang, Z.; Yu, B.; Zhao, H.; Yuan, X. Tripvista: Triple perspective visual trajectory analytics and its application on microscopic traffic data at a road intersection. In Proceedings of the 2011 IEEE Pacific Visualization Symposium, Hong Kong, China, 1–4 March 2011.

12. Chu, D.; Sheets, D.A.; Zhao, Y.; Wu, Y.; Yang, J.; Zheng, M.; Chen, G. Visualizing hidden themes of taxi movement with semantic transformation. In Proceedings of the 2014 IEEE Pacific Visualization Symposium, Yokohama, Japan, 4–7 March 2014.

13. Blei, D.M.; Ng, A.; Jordan, M. Latent dirichlet allocation. Journal of machine Learning research (3). *J. Mach. Learn. Res.* **2003**, *3*, 993–1022.

14. Tominski, C.; Schumann, H.; Andrienko, G.; Andrienko, N. Stacking-Based Visualization of Trajectory Attribute Data. *IEEE Trans. Vis. Comput. Graph.* **2012**, *18*, 2565–2574. [CrossRef] [PubMed]

15. Jiang, X.; Zheng, C.; Tian, Y.; Liang, R. Large-scale taxi O/D visual analytics for understanding metropolitan human movement patterns. *J. Vis.* **2015**, *18*, 185–200. [CrossRef]

16. Huang, G. Research and Implementation of Weibo Public Opinion Visualization System. Ph.D. Thesis, Jiangsu University, Zhenjiang, China, 2016.

17. Liu Band Zhang, L. "Development in the water"-the city's waterfront era is coming. In Proceedings of the 2017 China Urban Planning Annual Meeting, Dongguan, China, 18–20 November 2017.

18. Tomita, Y.; Kitsuregawa, M.; Itoh, M.; Yokoyama, D.; Toyoda, M.; Kawamura, S. Visual Exploration of Changes in Passenger Flows and Tweets on Mega-City Metro Network. *IEEE Trans. Big Data* **2016**, *2*, 1.