Visual Analysis Research of Traffic Jam Based on Flow Data

Wei Tian[1,2], Jinming Zhang[1,2(✉)], and Jialin Ma[1,2]

[1] Information Engineering University, Zhengzhou 450052, China
1158529851@qq.com
[2] Institute of Remote Sensing and Digital Earth,
Chinese Academy of Sciences, Beijing 100101, China

Abstract. With the acceleration of urbanization, road congestion is getting worse. To quickly determine the traffic jam, taking the example of Jiashan traffic data collected by ground sense coil. The paper firstly cleans raw data, eliminates error data, supplements missing data and processes redundant data, which provides reliable data for visual analysis; then using different visual components select congested roads, and collaboratively interacts three visual component which including heat map, bar graph and chord diagram, visually expresses road congestion from macroscopic to microcosmic and judges the position of the traffic jam, and provides reliable route for public trip.

Keywords: Ground sense coil traffic data · Road congestion · Visual analysis · Heat map · Chord diagram

1 Introduction

With the rapid development of society, the acceleration of urbanization, increasing the population and scale of cities, there are different levels of traffic congestion in most large and medium-sized cities, urban problem has became one of the bottlenecks of economic development [1]. Traffic congestion caused huge economic losses, including time loss, fuel loss and environmental costs. The data released by Chinese Ministry of Transportation shows that economic losses caused by traffic congestion are in 20% of the disposable income of the urban population, equaling 5–8% of annual gross domestic product (GDP) loss about up to 250 billion RMB a year, whereas loss in Beijing is 700 billion RMB one year (which more than 80% is the congestion time loss) [2]. Thus, the demand of community traffic control, vehicle navigation and personal travel path planning is growing, an urgent need for real-time traffic information to guide travelers, vehicles and transportation planning needs of individuals travel routes also have the accelerating growth trend.

Due to traffic congestion complexity, dynamic, and irrelevance reasons, it is difficult to predict traffic jams. Using prior knowledge and experience, combined with computer visualization technology, as well as the prediction model to analyze traffic congestion [3]. The paper uses traffic flow data (induction coil, microwave sensors) to analyze the traffic jam by visibility analysis, using the visual components judges congestion roads, which makes users acquire traffic congestion through the visual

H. Yuan et al. (Eds.): GRMSE 2016, Part I, CCIS 698, pp. 284–292, 2017.
DOI: 10.1007/978-981-10-3966-9_32

interface and make decisions according to their own needs. For example, taxi drivers want to know the current traffic conditions, or which way is saving time and nearest from the current position to destination; travelers need to choose transport routes considering traffic congestion.

In this paper, we mainly study the visual expression and visual analysis by the ground sensing coil flow data. Using visual component selects congested roads. Heat map is used to display the distribution of ground sense coil flow, from which can directly observe what roads congested and which road unimpeded; histogram, specific display according to flow value recorded by various ground sense coil; chord diagram, show the gridlocked traffic, which makes users understand the interrelation of each section. Visualizing the flow data of the ground sensing coil to analyze the road congestion in Jiashan, judge the congested roads, provide reliable information for public travel.

2 Related Work

In the process of traffic flow data visual analysis, including traffic data cleaning, data visual analysis two parts.

2.1 Traffic Flow Data Cleaning

Traffic flow data use toroidal coil detector to collect the data, the loop is a wide range of traffic sensor in urban roads, can acquire real-time of flow, speed, time share, travel direction and so on, are applied in traffic information collection and traffic state identification, traffic event monitoring system. Traffic flow data cleaning mainly includes data filtering, data recovery, outliers elimination, etc. Chu et al., used dynamic traffic data of Shanghai outer ring to clean noise data by the maximum possible value filling the missing data and moving average method [4]; Lu et al. came up with anomaly detection method based on curve fitting according to traffic/time occupancy rate of the inverted "V" shaped curve model, which can effectively identify the traffic flow data of the abnormal data [5]; Zhang et al. according to the relativity traffic flow theory of flaw detector intersection traffic of prediction [6], is a kind of effective data recovery method.

2.2 Visual Analysis of Traffic Flow Data

The visual analysis of road traffic data can be used to demonstrate the characteristics of spatial and temporal changes of traffic data, assist managers to understand and analyze the operation situation of urban traffic, and provide decision support for managers and users. In recent years, the research on the visual analysis of road traffic is increasing. Jiang et al. using taxi origin and destination point data, designed a Web version's taxi O/D visual analysis system data based on B/S architecture, and a set of explore cognitive rules of taxi O/D data analysis method [7]; Wang et al. make a thorough study about visual analysis of trajectory data, and divide the visual analysis of trajectory data

into three types: direct visualization, aggregation visualization, characteristics visualization [8]; Li et al. made massive data and statistical results highly graphical based on the traditional simple chart, showed information and rules more in a shorter period of time [9]; He used the arrow diagram and flow chart to express visualization of road traffic data [10]; Andrienko et al. [11] established a framework of interactive visual interface, which can effectively support the understanding of mobile behavior and mobile mode analysis of human; TripVista interactive visual analysis system researched by the Beijing University of visual development team where Wang [12] and others in, explore abnormal behavior through the research on the microcosmic traffic rules, and show traffic track in a variety of drawing way (direction icon theme River, with a timeline of the scatter diagram, show a high dimensional attribute information of the parallel coordinate plot).

3 The Data Processing

3.1 The Working Process

Visualization of flow data mainly includes two parts. The first part is data cleaning (processing redundant data, eliminating the error data, supplementing the missing data); the second part is visual analysis of traffic flow data, using different visualization techniques on the expression and analysis of Jiashan road congestion (Fig. 1).

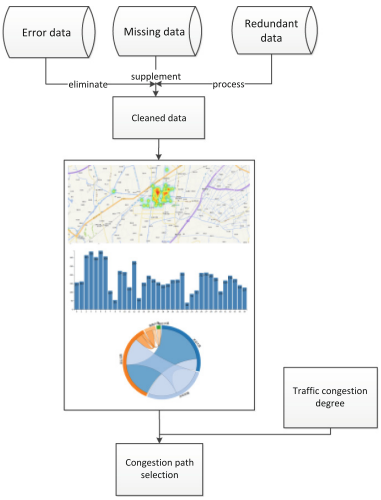

Fig. 1. Traffic data visual analysis work flow.

3.2 Select Data Sets

This paper chooses traffic data collected by the sense coil detector not floating car data, because sense coil detector technology is mature, easy to grasp; detection precision is high, can all-weather work; and low equipment prices. However, due to the limit of

coverage and sampling frequency floating car data is not every moment to meet the statistical requirements; in the city buildings, trees, and the viaduct caused blocking of inaccurate GPS positioning; GPS positioning error affects the floating car reliability and availability [1].

Take Jiashan sense coil flow data for sample, total 20 days data from 2015-9-20 to 2015-10-10, the sampling interval is about 30 s, and 36 sense coil detector which distributed in the city main roads, can monitor the city road better, and reflect the real-time traffic congestion. The main attributes of the flow data record of each ground sensor coil include equipment number, direction, number of the pile, average speed, time, average occupation rate, etc., and its properties are listed in Table 1. In this experiment used the device ID to represent selected sections (such as equipment number SBBH14 is Jinyang Road), use direction to judge the uplink and downlink (1 said uplink, 2 said downward), pile flow display number of vehicles through the sense of coil, the time is an important attribute of data traffic, which can determine the vehicle traffic laws through the different time periods and the existence of periodic.

Table 1. Attribute of flow data.

Data types	ID	Connotation
Flow data of ground sensor coil	***SBBH***	Equipment number
	FX	Orientation
	ZHLL	Number of the pile
	PJCS	Average speed
	QSSJ	Time
	PJZYL	Average occupation rate

3.3 Data Cleaning

Original data have many mistakes due to equipment, external environment and other factors, including three kinds of noise data: outliers, missing data, repeat the value in general [13]. By the analysis of the flow data, outliers accounted for 8% of these data, the data loss is 82%, the repeated number is 10%, so the data cleaning mainly for the lack of data processing.

Many reasons lead to data loss, such as the detector scan frequency is not fixed, the transmission line fault and the vehicle is too dense when the detector can't detect vehicles, etc. The processing methods aiming at data missing are removing the data object, interpolation computing missing value, neglecting missing value, using the probability model to estimate the missing value, using the historical data and so on [14]. In this paper, we first make a visual expression of the data (Fig. 2), judge whether the data is missing, and recover the missing data by using historical data [15]. Because people travel is similar in time, the traffic flow data is periodic in different days, using the same time history data to recovery missing data. As follows:

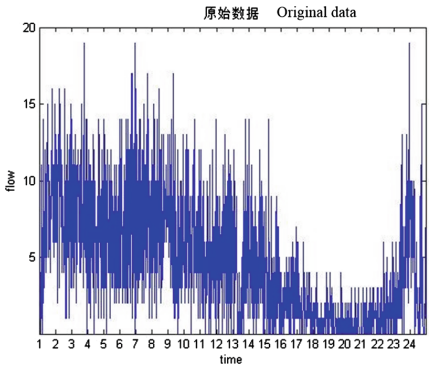

Fig. 2. The original data visualization of line chart.

$$\hat{x}_i = Hx_i \tag{1}$$

In this formula: x_i—estimated value
Hx_i—history data

Using historical data to recovery the missing data, 95% of the missing data can be recovered (Fig. 3), which meets the data requirements better, reflects trends in data change, provides reliable data for subsequent visual analysis.

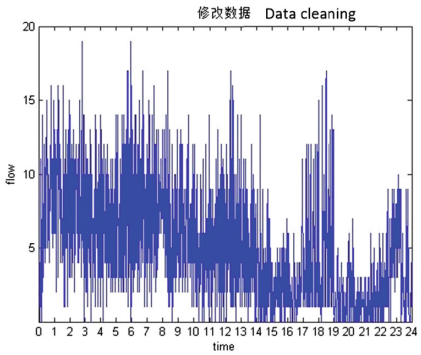

Fig. 3. The visual line chart after data cleaning.

4 Visual Design and Analysis

4.1 Congestion Path Selection

When analyzing Jiashan road congestion data, we need to choose the congestion path, determine which path is clear, which road is congested firstly, then make a visual analysis of the congested road. In this paper, the method of multi view collaborative visualization is used to choose the congested road section.

Multi view collaborative visual expression is used to choose the congestion section, and the congestion section is judged by different visual expression methods. The interactive analysis of the visual components is completed as shown in Fig. 4. Using the flow data acquired by ground sense coil to describe, the visual component 1 which is flow data measured by induction coil is used for the heat map expression, display the flow data of each sense coil collection to highlight form, red is represent for vehicle intensive, green is vehicle through the darker sparse, said the deeper the color the larger the flow, the color is more shallow said the flux is small; the visual component 2 histogram which display specific flow of the sense coil in a time, make up the visual component 1. Component 1 can only judge the range, but cannot determine the specific road congestion, where the component 2 can clearly observe the specific traffic flow at a time and the specific number of the flow. Comprehensive visual components 1 and 2, the congested road is Tangong north road and wood Avenue direction; visual component 3 is displaying the interrelationship diagram of Jiashan County main road congestion by the chordal graph, through the visual component 1 and 2 to determine the main road congestion is wood Avenue, Tangong north road, outer ring road, by the visual component 3 observation to the trend of congested road traffic, such as most traffic of Tangong north road towards the direction of wood Avenue.

Fig. 4. Traffic data visualization synergy. (Color figure online)

4.2 Traffic Jam Analysis Validation

From above visual components analyzes Jiashan main congestion line is Muye Avenue, outer ring Road, Tangong north road. Due to the existed research on the data flow processing traffic congestion rarely exist, so this paper calculates road congestion degree according to the flow, which judges whether the road congested, and makes traffic data quantify, in order to verify the correctness of the above conclusion of visual analysis.

In the process of deciding the traffic congestion threshold, we cannot to measure congestion just by a simple flow for that the number of vehicles can't discern the road congestion and recording of small flow may be less in the period of driving and is not caused by congestion. Zheng Shujian and others' article described the traffic congestion measured by congestion degree in Japan, defined as the ratio of the evaluation

benchmark amount among actual traffic volume and 24 h or 12 h a day in a road [16], such as the formula (2).

$$DC = (\omega \times Q)/C \tag{2}$$

Type: ω as the weight coefficient, $\omega = 1 - \alpha/100 + \alpha \times \beta/100$; α as large vehicles mixed rate; β for the standard vehicle equivalent coefficients of large vehicles; Q for 12 h day traffic volume; C as the evaluation base 12 h traffic volume, benchmark traffic can be gained by the level of planning and design traffic capacity, peak rate, in the same direction rate. According to the actual situation of Jiashan (no large vehicle traffic in the urban area), the paper improves the formula (3), as follows:

$$DC = Q/C \tag{3}$$

Type: Q as 1 h traffic volume; C as evaluation benchmark 1 h traffic.

The relationship between the congestion degree and operation level of road network, is that congestion degree by how much to open or congestion, as shown in Table 2:

Table 2. Relationship between the congestion degree and road network.

Congestion degree DC	1 h operation level of road network
DC < 1	Unimpeded
1 ≤ DC < 1.75	Congestion gradually increasing
DC ≥ 1.75	Chronic congestion

Statistical calculation by formula (3), traffic situation of Jiashan Muye Avenue, Zhujiang Road – Lingshan Road in 2015-9-28 9:00–10:00 as shown in Fig. 5, of which Tangong north Road to Muye Avenue is 75% to the congestion state, where Tangong

Fig. 5. Traffic route choice.

north Road - Zhujiang Road - Lingshan Road is smooth. The result of verifying the traffic flow data visualization by congestion degree quantization is correct.

5 Conclusion

The paper directs at the road congestion matter, taking flow data collected from Jiashan for example, cleans the original flow data, determines the time scale, visualizes and analyzes the flow data. Three interactive visualization components including heat map, bar histogram and chordal graphs, makes users analyze road congestion from macroscopic to microscopic, determine whether the road congested, and whether need to change the travel route. And quantifying the flow by calculating the traffic congestion degree, then judging the rode congestion and what extent of the congestion, which is a serious congestion or slight congestion. The visual analysis results are verified by calculating the congestion degree, which provides a reliable route for public travel.

Acknowledgment. This work is funded by the National Natural Science Foundation of China (No. 41371383).

References

1. Li, Q.: Traffic Geographic Information System Technology in the Development of Frontier. Science Press, Beijing (2012)
2. Beijing Traffic Development Research Center: The 2015 Beijing traffic development report. Traffic Transp. (3), 3–7 (2016)
3. Wang, Z., Lu, M., Yuan, X., et al.: Visual traffic jam analysis based on trajectory data. IEEE Trans. Vis. Comput. Graph. **19**(12), 2159–2168 (2013)
4. Chu, H., Yang, X., Wu, Z.: Method analyzing and application of dynamic traffic data pre-processing. In: China Intelligent Transportation Annual Meeting (2005)
5. Lu, W., Shang, N., Qin, M.: A traffic data preprocessing method based on curve fitting exception detection. J. Comput. Res. Dev. **43**(s3), 642–646 (2006)
6. Zhang, H., Yang, Z., Li, Y.: Study on forecast of traffic volume at non-detector intersections. J. Highw. Transp. Res. Dev. **19**(1), 91–95 (2002)
7. Jiang, X., Zheng, C., Jiang, L.: Visual analysis of large taxi origin-destination data. J. Comput. Aided Des. Comput. Graph. **27**(10), 1907–1917 (2015)
8. Wang, Z., Yuan, X.: Visual analysis of trajectory data. J. Comput. Aided Des. Comput. Graph. **27**(1), 9–25 (2015)
9. Li, W., Zhou, F., Zhu, W.: The rail transit network traffic data visualization research. Chin. Railw. **05**(2), 94–98 (2015)
10. He, X.: Visual analytics of road traffic with large scale taxi GPS data. Zhejiang University of Technology (2014)
11. Andrienko, G., Andrienko, N., Wrobel, S.: Visual analytics tools for analysis of movement data. ACM Sigkdd Explor. Newsl. **9**(2), 38–46 (2007)
12. Guo, H., Wang, Z., Yu, B., et al.: TripVista: triple perspective visual trajectory analytics and its application on microscopic traffic data at a road intersection. In: IEEE Pacific Visualization Symposium, pp. 163–170. IEEE Computer Society (2011)

13. Chen, W., Shen, Z., Tao, Y.: Data Visualization. Electronic Industry Press, Beijing (2013)
14. Wang, M.J., Pan, Q.M., Chen, W.: Survey of visualization data cleaning. J. Image Graph. **20**(4), 0468–0482 (2015)
15. Jin, S.: Research on preprocessing methods of loop detector data. Jilin University (2007)
16. Zheng, S., Yang, J.: Traffic congestion evaluation index calculation methods at home and abroad. Highw. Automot. Appl. **04**(1), 57–61 (2014)