

Exploration of Real-Time Traffic Information System Based on Big Data Analysis

Yan Hou

Qilu Normal University, Jinan, Shandong, 250014, China.

122365871@qq.com

Abstract. In order to solve the problem brought about by traffic congestion to people's daily life, this experiment uses the effective combination of intelligent transportation and vehicle networking technology, and proposes the real-time traffic information system based on the vehicle networking end - tube - cloud architecture for mass real-time traffic data. First, the overall architecture of the system is designed and analyzed, and then the related technology of traffic data perception and acquisition is introduced. Secondly, the mode of real-time traffic information data acquisition based on floating car and the characteristics of the initial traffic data are introduced, and a suitable data storage format is designed in accordance with the characteristics of original traffic data collected. Finally, the test is carried out according to the spatiotemporal characteristics of traffic data. The results show that the algorithm used can effectively repair data, and the data repair performance is obviously better than time and space two data repair models. Through the relevant technology, the current road traffic conditions are obtained, and then the information is transferred to traffic management departments and travel personnel. In this way, it can provide the road condition for the travellers real-time and accurately, and help travellers reasonably choose the travel routes, thus alleviating traffic congestion. In a word, real-time traffic information can help traffic management departments to do a good job in traffic management.

Keywords: vehicle networking; big data; real-time road conditions; data repair.

1. Introduction

At present, with the growth of the economy, people's life has undergone tremendous changes, especially in the way of transportation, private cars have become one of the most important means of transportation, so the number of cars has increased rapidly. Car is originally used to facilitate people's travel, and save time on the road. But now, due to the rapid increase in the number of cars, the current road infrastructure has not been able to meet its needs, thus causing serious traffic congestion. Traffic congestion will not only impede economic development, but also cause waste of resources and environmental pollution. According to the Ministry of Environmental Protection, the amount of pollutants emitted by vehicles in congested conditions is 5-10 times that of normal driving. China will discharge more than 16.7 thousand tons of carbon dioxide and 95 thousand tons of nitrogen oxides and sulfur dioxide as well as other harmful substances every day due to traffic congestion.

2. State of the Art

At present, cars are basically equipped with vehicle navigation systems, and almost all of the electronic maps APP are installed on the phones, and these technologies are really convenient for our travel. But now, because of the technology or some algorithms, the information provided by the electronic map has a certain difference with the reality. Especially with the development of the technology of the vehicle networking and the big data technology, how to get the road information more accurately in real time is a hot spot in current studies.

Han et al. proposed a signal control priority passing method [1]. This method solved the problem of bus priority in urban roads. Rhee et al. put forward a method for bus lane reuse [2]. The method solved the problem of bus passing through urban public transport lanes, and allowed motor vehicles to enter the exclusive bus lane during the non-peak travel period. Rashid et al. successfully used the vehicle networking technology in urban transportation [3]. The method effectively utilized the information platform of the Internet of vehicles to serve the residents to travel, and solved the problem

of traffic congestion due to the lack of traditional traffic information. Fusco et al. applied big data in road traffic [4]. The method effectively utilized the big data in the field of road traffic, and solved the problem of traffic congestion caused by inaccurate information of residents' travel. Wang et al. put forward an intelligent transportation plan based on big data [5]. The scheme mainly described the intelligent transport system architecture and related technologies. Shafiei et al. proposed a traffic adaptive priority method based on vehicle networking [6]. The method effectively utilized the vehicle networking, and solved the problem of inaccurate collection of data in road traffic system.

To sum up, the above methods solve the problem of congestion in road traffic from different angles, but there are some limitations. A traffic condition plan based on big data of vehicle networking is proposed, and the urban traffic condition is studied by artificial intelligence technology.

The real-time road condition is a concept proposed in view of the traffic congestion at present. It is mainly used to reflect the current traffic status of the road and guide the most time saving travel route for pedestrians. It can relieve traffic congestion and avoid choosing the travel routes with traffic accidents in the course of travel. The real-time road condition technology has developed relatively well in Europe. As a new application in the vehicle networking, remote diagnosis has begun to enter the field of vision, and the combination of OBD (On-Board Diagnostics) and mobile APP has become the most common application form [7]. At present, all kinds of OBD terminals and mobile phone APP emerge as the times require. Now, the electronic map has been integrated into our life and become one of the necessary software in the mobile phone terminal.

Now, China is stepping up research on the vehicle networking, striving to make it infiltrate into every sector of the industry related to vehicle. The vehicle networking can change and integrate the rear-service industry chain, and can rebuild a car service industry. The vehicle networking can also transform the vehicle insurance into UBI (Utilization-based Insurance), which can change the management mode of the government and enterprise fleet. Moreover, it can change and create new vehicle finance and vehicle management modes, and create a completely new logistics information market [8]. In the future, the vehicle networking will definitely make our life more intelligent, safe and convenient.

3. Methodology

3.1 Vehicle Networking

The vehicle networking uses sensors, communication networks and other technologies to realize the network interconnection and information communication between people and vehicles, between vehicles, and between vehicles and roads. Intelligent traffic management is realized by intelligent management and control of people, vehicle, and roads. Traffic congestion time and crowding times are reduced, and the macro concept of Internet of things technology is embodied in the present situation.

Vehicle networking technology can provide good technical support for real-time traffic information system, and the vehicle networking is also a powerful support for building intelligent transportation system. However, there are few introductions to the application of vehicle networking technology in real-time road condition information system. Therefore, the real-time road condition information system architecture based on vehicle networking will be introduced briefly. The overall architecture of the system is shown in Figure 1.

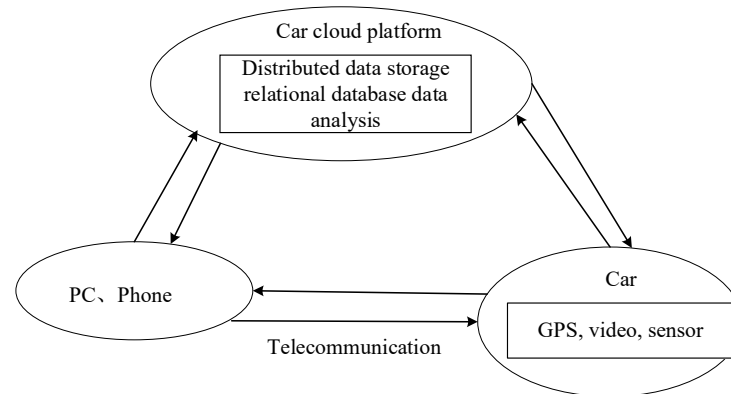


Figure 1. Overall Architecture of Real-Time Road Condition Information System Based on Vehicle Networking

3.2 GPS Positioning Technology

GPS is the global positioning system, which can be divided into two types of services according to the nature of application: military services and civilian services. The GPS system can provide us with location, moving speed and direction, time information and so on. Nowadays, our daily life is inseparable from GPS positioning. For example, vehicle navigation systems are installed on cars, which can provide navigation for users. The GPS system mainly consists of three parts: GPS space constellation, ground monitoring, and user equipment.

3.3 Vehicle Networking Big Data

With the rapid development of communication network technology and sensing technology, the extensive application of vehicle networking in the field of traffic information services has accumulated a large amount of data in the field of transportation. The vehicle networking big data effectively solved the problem of inaccurate traffic condition information and improved the accuracy of traffic condition prediction. The big data technology of vehicle networking has three characteristics: multi-source, real-time and correlation.

First, multi-source. The vehicle networking big data not only comes from sensors such as vehicle-mounted cameras, GPS, and road cameras, but also comes from the weather bureau and road construction providers. Therefore, the wider the source of vehicle networking big data is, the more data will be collected, the more accurate the analysis result is, and the more accurate the traffic prediction will be.

Second, real-time. The system realizes data transmission and information dissemination through 3G/4G mobile network and optical fibre Internet, so as to ensure the real-time of data transmission, thus improving the reliability of traffic road condition.

Third, correlation. In the environment of vehicle networking big data, the focus is to find out the correlation between the road and the motor vehicles in the vehicle networking big data. At the same time, the causes of abnormal traffic conditions can also be analysed through correlation.

4. Results and Discussion

4.1 Collection of Real-Time Traffic Information

A floating car is a vehicle equipped with GPS and wireless communication equipment. As there are more and more applications related to intelligent traffic technology, floating car data will be an indispensable source of data for the applications in the future, and floating car data can be used to judge the threshold of traffic parameters on the road.

The GPS positioning system uses the satellite system to locate. The GPS device on the vehicle terminal will detect the time of satellite signal transmission in real time after receiving the satellite

signal. Thus, it can calculate the latitude and longitude of the location of the floating car and the instantaneous speed of the vehicle, store these data, and then process them with the difference correction data. The wireless communication module transfers these data obtained to the information processing platform, and it also receives the information transmitted from the information processing platform.

4.2 Cleaning of Original Traffic Data

The collected raw GPS data cannot be directly used for data mining analysis, because the original traffic data will have abnormal data, that is, "dirty data". The phenomenon of "dirty data" generally refers to noise, error, and loss in data. Therefore, in order to ensure the accuracy of the final analysis results, it is necessary to first clean the raw traffic data.

In order to ensure that traffic data can be better used for data mining, first of all, the data are formatted, so that the data have a unified storage format. After storing data in a specified storage format, data need to be cleaned. Generally, data cleaning can be divided into two kinds. One is to clean the signal which is interfered by noise. This cleaning is to eliminate noise, which belongs to the cleaning of normal data. The other one is to clean the abnormal raw data that are inconsistent, erroneous and repetitive, which belongs to the cleaning of abnormal data.

Table 1. Vehicle GPS Data Storage Format

Vehicle ID	Date	Sampling time	Longitude	Latitude	Speed(km/h)
540876	20150611	25787	118.944912	32.060386	0
540876	20150611	25800	118.944936	32.060596	5
540876	20150611	25812	118.945136	32.060600	9
540876	20150611	25842	118.946128	32.058830	36

4.3 Result Analysis

Due to the information acquisition equipment and the cause of the weather, the wrong data will exist in the traffic information collected. As shown in Table 2, it is part of the driving data of a vehicle obtained. Because of too many data, only a small part of the data is displayed. For the sake of simplicity, the table includes only the vehicle information number, the latitude and longitude of the vehicle, and the instantaneous speed of the vehicle.

Table 2. Raw GPS Data

Serial number	Longitude	Latitude	Speed(km/h)	Serial number	Longitude	Latitude	Speed(km/h)
1	118.801336	32.008336	16.78	11	118.801336	32.014504	24.46
2	118.801168	32.008476	18.05	12	118.801336	32.015400	6.7
3	118.801128	32.008584	18.76	13	118.801336	32.015584	27.32
4	118.801136	32.008660	19.17	14	118.801336	32.016052	25.76
5	118.801168	32.008786	19.78	15	118.801336	32.015968	28.58
6	118.801920	32.010536	21.88	16	118.801336	32.015316	29.47
7	118.802848	32.011818	-2.35	17	118.801336	32.015090	23.32
8	118.804376	32.013156	23.73	18	118.801336	32.014438	25.75
9	118.804640	32.013562	22.18	19	118.801336	32.013974	67.8
10	118.804920	32.014392	22.28	20	118.801336	32.013736	23.2

In the previous analysis, it can be seen that the speed of the floating car has a certain correlation in the adjacent time and adjacent space, and the correlation between these data varies with the change of time or space distance. As shown in Figure 2, it is the relationship between the traffic information data and time and road distance on a section of the road. It can be seen from (a) that the correlation between the data is getting smaller and smaller as the time goes on, and the same rules are presented in (b). The greater the distance between the two sections is, the smaller the correlation coefficient between the data on the section will be.

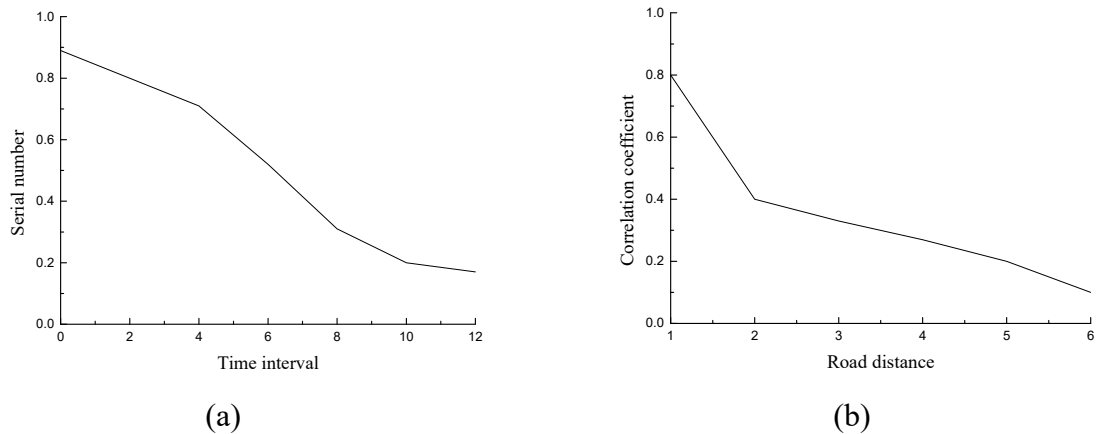


Figure 2. Correlation between Traffic Data and Time and Road Segment Distance

5. Conclusion

In the context of intelligent traffic and vehicle networking technology, a real-time traffic information system based on the vehicle networking end - tube - cloud architecture for mass real-time traffic data is proposed. First, the overall architecture of the system is designed and analyzed, and then the related technology of traffic data perception and acquisition is introduced. The original traffic information will have some exceptions, so it should be pretreated before the original data analysis, so as to improve the accuracy of data mining results. Secondly, several ways of real-time traffic information data collection based on floating car and the characteristics of the initial traffic data collected are introduced, and a suitable data storage format based on the characteristics of the original traffic data collected is designed. Thirdly, several rules for data cleaning are set up according to the previous research and experience. Finally, based on the spatial and temporal characteristics of traffic data, a dynamic adaptive data repair algorithm is proposed and simulated. The experimental results show that the algorithm can effectively repair the data, and the data repair performance is obviously better than that of the time and space two data repair models.

References

- [1]. Han, Youngjun, and Soyoung Ahn. Stochastic modeling of breakdown at freeway merge bottleneck and traffic control method using connected automated vehicle. *Transportation Research Part B: Methodological*, 2018,107, pp. 146-166.
- [2]. Rhee, Kyoung-Ah. Spatial regression analysis of traffic crashes in Seoul. *Accident Analysis & Prevention*, 2016, 91, pp. 190-199.
- [3]. Rashid, Bushra, and Mubashir Husain Rehmani. Applications of wireless sensor networks for urban areas: A survey. *Journal of Network and Computer Applications*, 2016, 60, pp. 192-219.
- [4]. Fusco, Gaetano, Chiara Colombaroni, and Natalia Isaenko. Short-term speed predictions exploiting big data on large urban road networks. *Transportation Research Part C: Emerging Technologies*, 2016, 73, pp. 183-201.
- [5]. Wang, Chao. Soft computing in big data intelligent transportation systems. *Applied Soft Computing*, 2016, 38, pp. 1099-1108.
- [6]. Shafiei, Sajjad, Ziyuan Gu, and Meead Saberi. Calibration and validation of a simulation-based dynamic traffic assignment model for a large-scale congested network. *Simulation Modelling Practice and Theory*, 2018, 86, pp. 169-186.

- [7]. Berecibar, M. Critical review of state of health estimation methods of Li-ion batteries for real applications. *Renewable and Sustainable Energy Reviews*, 2016, 56, pp. 572-587.
- [8]. Baecke, Philippe, and Lorenzo Bocca. The value of vehicle telematics data in insurance risk selection processes. *Decision Support Systems*, 2017, 98, pp. 69-79.