Intelligent Traffic Decision Analysis System Based on Big Data Mining

Yachao Jia, Xin Wang
Northwestern Polytechnical University Mingde College ,710124

Abstract: The urban intelligent traffic decision analysis system is a system designed to solve urban road traffic problems, improve traffic safety and comfort, and improve the level of traffic services. The application of big data mining technology can improve the ability of various types of road traffic information analysis and processing, expand the scope of application, and provide decision-making reference for the application of intelligent transportation systems, so as to ensure the smoothness and safety of roads. This article will introduce the concept of big data mining technology and intelligent transportation system, and analyze the architecture of urban traffic intelligent system based on big data technology and the design and implementation of mining system.

1. Introduction
The development trend of urban intelligent traffic decision-making system is intelligence perception, real-time processing, information fusion and public service. The means of traffic management is also becoming more abundant. The service industry provided for public travel is increasingly diversified. With the development of electronic information technology, especially after the rapid development of cloud computing and the Internet of Things, big data mining has begun to be applied in the field of urban intelligent transportation. On the one hand, it can use the characteristics of real-time, concurrency, scalability, and security of big data to optimize the collection and analysis of traffic data; on the other hand, big data can be applied to the management of mass traffic data to achieve the service goals of urban road traffic. To give full play to the value of urban intelligent transportation system it will contribute to improving the orderliness of traffic system operation, improving the efficiency of traffic operations, reducing safety accidents and pollution probability, optimizing traffic organization methods, and improving the level of transportation services.

2. Big Data Mining Technology Overview
Big data mining technology refers to extracting valuable information or knowledge from a massive volume of data, and then providing it to users through the form of services. Compared with traditional data mining techniques, the purpose of both is to obtain valuable information. However, the technical development background, processing objects, and the depth and breadth of mining differ between these two aspects [1].

2.1 Technological Development Background
With the development of science and technology and the increase in demand for valuable information, both traditional data mining and big data mining have achieved certain development. However, the background of the generation and development of traditional data mining is in the Internet era and database era. The amount of data to be processed does not have a large volume of data in the era of big
data mining technology, and the complexity of data information does not have the complexity of data in the era of big data mining. The background of the emergence and development of big data mining technology is cloud computing, Internet of Things, and mobile Internet. This technology is based on the characteristics of big data and the problems faced by current systems. It is based on the integration of cloud computing and related technologies. To achieve data mining, development and application are still in the stage of continuous exploration and development.

2.2 Processing Object
Big data mining technology and traditional mining technology also have different processing objects. This is mainly related to the background of the two technologies. The traditional data processing based on the development of the Internet only deals with the data generated by a certain range of information management systems. Although it also contains active data generated by users, it is still mainly structured by passively generated structured data. The big data mining technology is born in the background of cloud computing, the Internet of Things, and the mobile Internet. Therefore, the data onto its processing objects includes the data onto the information management system and the data generated by users in the web system, and also includes the generated simulation data automatically formed by sensing system. In other words, large data mining technology needs to deal with a large volume of objects, complex data types, more extensive and comprehensive collection of data, and more timely and rapid data processing. However, big data mining technology does not require high accuracy in data mining, so there will be problems of increased data redundancy and uncertainty.

2.3 Excavation
The breadth and depth of traditional data mining and big data mining technologies are not the same when analyzing and processing data. With the increase of data volume and complexity of data types, coupled with the development of data of different structures and patterns, only the integration of related technologies based on cloud computing through big data mining technology can analyze and process these complex data. The scope of its data processing is wider, and it is deeper in data mining analysis. However, traditional data mining can only deal with small-scale data information, the scope of processing is limited, and the type of analysis data is relatively simple. In addition, the traditional data mining is limited by its own platform system, and its calculation and mining algorithms are not very scalable. Therefore, it is difficult to handle multi-source heterogeneous information and it also lacks the timeliness.

2.4 Big Data Mining Architecture
According to the above analysis of big data mining theory and data warehouse design theory, a big data mining architecture that integrates multiple computing modes and storage modes can be constructed under the big data environment, that is, the ITS data mining platform framework. In this architecture, functions can be divided into data layer, algorithm layer, analysis layer, and application layer. First of all, the data layer provides rich resources and dynamic support for the analysis and processing of data and stored data by the big data mining technology, which is supported by the platform, and uses the related technologies and processing tools based on cloud computing to analyze and process complex and large data. In order to build a resource-rich clouds environment. In addition to providing resources such as data, hardware, and software to the outside world in the cloud environment, it can also provide power support for data preprocessing, data analysis, and mining of big data mining. The algorithm layers refer to the analysis of data onto the cloud platform which is based on user characteristics and actual needs, using data mining and analysis tools that has high storage and analysis capabilities. The application refers to the result of data processing after the analysis and processing using the big data mining technology on the cloud platform, and provides the user with the help of visual technology and other technical service forms.
3. Implementation of Intelligent Traffic Data Mining System

The intelligent traffic data mining system has the characteristics of high reliability, high extensibility, and high efficiency. The computation mode is mainly batch processing and stream processing. The structure can be divided into three layers: data source, big data mining platform, and user layer. During the construction of the system, traditional databases and processing tools, graph parallel computing, and in-memory computing are also integrated into the platform. The functions are rich and varied. Relying on supporting technologies, complex data can be processed in real time and quickly. In the following, the internal working flow will be introduced to the hardware and software environment of the system, system data preprocessing, system data storage, system data calculation and analysis, and system data display to analyze the data processing method and features of intelligent traffic data mining system based on big data mining.

3.1 System Hardware and Software Environment and System Design

The hardware and software environment of the intelligent traffic data mining system includes several subsystems such as data area, traffic application area, communication area, GIS area and terminal area. A database server is arranged in the data area, a traffic application server is arranged in the traffic application area, a GIS server is arranged in the GIS area, a communication server is arranged in the communication area, and a plurality of operation terminals are arranged in the terminal area. The big data mining theory is used to build the intelligent traffic big data mining system. The system has comprehensive features. The system design includes a cube customization module, an OLAP analysis mode selection module, a graph display module, and a data mining model selection module. The cube customization module can set the dimensions of the cube and then import the required analysis dimensions or filter out unnecessary dimensions; the OLAP analysis selection module is actually a tool that enables users to use multidimensional views for analysis; The function of the display module is that the user can set the type, vertical and horizontal coordinates of the chart, set the font, name, size, and legend of the chart, and display the chart results; the function of the data mining model selection module is based on the required analysis of data-reactor characteristics and analysis goals selection and determination of mining models [3].

3.2 Data Preprocessing

Traditional data mining in data preprocessing adopts the method of pre-existing mode and data, that is, using processing tools to perform queries and updates in a given mode, and then preprocessing static data to protect data and it has the characteristics of protecting the integrity and accuracy and the high precision. Big data mining technology adopts the pre-processing mode of pre-existing data, and adjusts the mode as the data changes without specific modes. When preprocessing data, big data mining technology uses traditional preprocessing technology, stream processing technology, multi-modal entity recognition, and remote automatic acquisition and fusion to improve parallel computing, iterative calculations, and data consolidation and sharing in preprocessing and other capabilities. However, the pretreatment of big data focuses on the correlation between data and does not pay attention to the causal relationship between data. Moreover, the emphasis on real-time performance when processing data does not focus on completeness and accuracy, so the quality of data preprocessing results is not high.

3.3 Data Storage

Traditional data mining uses data in line storage to store static and deterministic structured data in multidimensional data models or entities and links. The storage is passive and random, and its flexibility and scalability are poor. On the one hand the big data mining technology in the storage of traffic information data, in addition to database storage also includes distributed storage, on the other hand it can store a wide range of data types, structured data, there are semi-structured and unstructured the data, the storage method is mainly column storage and rank hybrid storage. The storage mode is flexible and simple and highly scalable.
3.4 Data Calculation and Analysis

Traditional data mining mainly concentrates on batch processing during data calculation and analysis. However, big data mining technology integrates multiple computing models and data processing tools to perform distributed parallel processing on big data. When dealing with complex, diverse, and massive volumes of data, traditional data mining cannot be automatically analyzed in depth, and it is difficult to express complex analysis models. Big data mining can solve poor scalability of the analysis tools and clouds in traditional data mining. The weak platform analysis function improves data parallel computing and analysis capabilities.

3.5 Data Display

Traditional data mining mainly uses text, reports, and a few visual graphs to reflect the model's performance, performance, and mining information during data presentation. However, it is only suitable for data with small data and simple relationships, for multidimensional, massive, and dynamic data. The data can only be displayed through big data mining. Big data mining visually presents a large amount of complex big data in the form of images and animations through the visualization of human-computer interaction, and then uses automated analysis tools for analysis and mining to help users further understand the data. At present, the representative visualization technologies include the cosmic planet map and tag cloud[4]. However, in order to realize more efficient visual analysis, the timeliness of visualization technology, load balancing, and node communication need further research.

3.6 System Design and Implementation

When designing an intelligent transportation system, it needs to meet the development needs of the transportation business, and also has certain degree of expansion and adaptability. The main application should be to achieve good results in the following aspects: First, GPS data, bayonet data, and aircraft can be used. The data of vehicle traffic flow detection and other data are combined with the road network model to calculate and analyze the traffic operation index through the spatial correlation analysis algorithm, so as to reflect the traffic conditions of the entire city traffic, the degree of congestion, etc., which can provide decisions for traffic planning and congestion management. Reference; Secondly, the use of video surveillance equipment to collect vehicle traffic information, through the establishment of prevention and control database for the dynamic tracking of the target vehicle, can achieve automatic alarm, identify abnormal driving behavior, traffic safety risk assessment and suspects early warning and other functions. Thirdly, statistics on the traffic flow of road sections can be calculated and forecasted through technologies such as intelligent video processing, neural networks, and association rules. Then, the dangers of each road section can be predicted based on the weather and the number of lanes, and the road sections where traffic accidents may occur are predicted. And alarms to improve the level of traffic management and reduce the chance of accidents. In addition, according to the data of traffic management business, the comprehensive processing of traffic flow, traffic accidents and other information is implemented, and the travel laws[5] are analyzed through the establishment of OD matrix model, thus providing supportive data for decision-making of traffic management organizations.

Using this system to analyze traffic flow, traffic conditions can be dynamically learned in real time. The implementation method is to sort and summarize the coil data which is collected in real time according to the analysis needs, and then form the real-time data of the corresponding road section. Thus, the traffic flow of a certain section can be predicted and the traffic flow state can be divided by the prediction model. For example, forecasting an elevated traffic flow from 7:00am on May 1, 2017, predicts the traffic flow at 7:10pm, and identifies traffic conditions. After that, the real traffic flow situation at 7:10pm and the predicted analysis results are compared with each other, so that short-term traffic flow prediction and classification results can be verified. The time for predicting the classification can then be extended, from 7:00am to 9:00am as the prediction phase, and the actual situation is compared with the predicted classification results, and finally the deviation rate is calculated. This method of using the big data mining system to predict and classify short-term traffic
conditions can improve the accuracy of the prediction. The output of the system can be used as an auxiliary basis for the release of traffic information. It can provide traffic managers with more accurate information on traffic conditions, so we need to improve the level of traffic management services and improve traffic conditions.

Table 1: Statistics of forecasted deviations

| Deviation rate (%) | Time |
|--------------------|------|
| 0                  | 7:00 |
| 2                  | 7:30 |
| 4                  | 8:00 |
| 6                  | 8:30 |
| 8                  | 9:00 |

4. Conclusion:
In short, big data mining technology is applied in the field of intelligent transportation. The flexible use of traffic big data can greatly change the development of transportation. This is because big data mining technology has the characteristics of real-time, distribution, high efficiency and predictability. Therefore, it is possible to analyze and deal with the massive traffic data, detect traffic anomalies in a timely manner and solve them, improve the efficiency of traffic management, and make traffic operations more reasonable. At the same time, it can also use big data to predict traffic conditions in a timely manner to reduce the probability of false positives and false negatives. Through the establishment of prediction models and real-time monitoring models, traffic information can be dynamically grasped in real time, improving the data processing capabilities of traffic safety systems and thus ensuring the safe driving.

References:
[1] Meng Xiaofeng, Ci Xiang. Big data management: concepts, technologies and challenges [J]. Computer Research and Development, 2013(01):146-169.
[2] Wu Weiqiang. Intelligent Traffic Decision Analysis System Based on Big Data Mining[J]. Mechatronic Engineering Technology, 2017(s2).
[3] Li Jianzhong, Liu Xianmin. An important aspect of big data: data availability [J]. Computer Research and Development, 2013(06): 1147-1162
[4] Tan Na, Liu Li. Big Data helps smart traffic development [J]. Traffic Construction and Management, 2015(z2): 235-237.
[5] Lu Biao, Li Yue, Zhang Wanli. Research and design of intelligent traffic data analysis platform system based on big data technology[J]. Journal of Hubei Institute of Science and Technology, 2016(05):6-9.