Research on the application of big data in intelligent transportation system

Shengneng Hu¹, Wei Liu²

¹ School of Civil Engineering & Communication, North China University of Water Resources and Electric Power, Zhengzhou 450011, China

² China communications materials & equipment company limited, Beijing 10010, China

Abstract

Intelligent traffic management system is the development trend of the traffic system. It is also an effective means to ease traffic and improve traffic efficiency. The application of the big data cloud platform in the intelligent transportation system is analyzed. At the same time, using big data analysis technology to build a big data analysis platform of intelligent transportation, the intelligent traffic data analysis platform is decomposed into infrastructure layer, data analysis layer and terminal publishing layer; The characteristics of the cloud platform are analyzed, and the specific technology implementation of the large data cloud platform is studied. The platform can solve the problem of storage, analysis and multi terminal distribution of mass data, provide traffic information services to traffic management departments and the public, It is a useful attempt to apply advanced information technology to the transportation industry.

Keywords: Big Data, Intelligent Transportation, Large Data Storage, Cloud Computing.

1. INTRODUCTION

At present, with the rapid development of the city, people are more and more requirement on transportation, facing the normalization problem, such as city traffic congestion, traffic safety, traffic organization and so on, the traditional way of thinking has been unable to solve these problems (Cao, 2015). With the rapid development of science and technology such as geographic information, communication, sensor and computer technology, the new thinking mode of intelligent transportation has gradually become a practical application from concept conception, which makes traffic management more efficient, information-based and extensive (Qiu, 2013). Big data is a disruptive technological change, following the cloud computing, Internet of things, which has the characteristics of large data volume, large data type, low value density, high commercial value and fast processing speed. In the field of transportation, the massive data include all kinds of traffic monitoring, data of services and applications, such as roads, waterways, passenger stations and ports and other video monitoring data, traffic flow detection of city road and highway, meteorological data, urban public transport, taxi and passenger vehicle satellite positioning data, and highway and waterway toll data, etc., these types of traffic data are numerous and huge.

Through market research and analysis, there are some systematic products for traffic management both at home and abroad, but there are still some problems, such as single system function, lack of integration and backward technology, and that is mainly reflected in the construction of application system distributed mass data, lack of effective integration of traffic data, low utilization rate, the data value cannot be brought into full play, and limited, traffic information dissemination is difficult to timely access to traffic warning etc. (Zhou et al., 2013; Liu, 2011; Xu, 2012). With the development of information technology, traffic departments urgently need a more advanced intelligent data analysis method, in order to carry on the efficient, real-time analysis to the massive transportation industry’s data, and to provide real-time and accurate traffic information service for travelers, so as to provide reference for traffic management departments to deal with unexpected accidents and illegal traffic behavior. With the emergence of mass traffic data, the technology of large data analysis brings new opportunities for the development of intelligent transportation. Large data technology storage capacity and computing power will be more reasonable configuration of traffic resources, more effectively support transportation planning, management, operation, service and security, and provide new ideas and means for public safety and social management.
2. BIG DATA ANALYSIS CLOUD PLATFORM APPLICATION ADVANTAGES

1) Improving the efficiency of traffic operation

Data technology can improve traffic operation efficiency, road network capacity and facilities frequency, and control traffic demand. Traffic improvement involves a large amount of engineering, and large volume characteristics of large data help to solve this dilemma. The real time of big data makes traffic running more reasonable. When static idle data is processed and need to be used, it can be used intelligently. Large data technology has higher prediction ability, it can reduce the probability of false positives and missed reports, and provide real-time monitoring for the dynamic nature of traffic.

2) Improving the level of traffic safety

The real-time and predictability of large data technology can help improve the data processing ability of traffic safety system, such as joint roadside inspection vehicle trajectory detector, big data technology rapid integration of various sensor data, construct the security model after a comprehensive analysis of the traffic safety, which can effectively reduce the possibility of traffic accident. In the field of emergency rescue, with its fast response time and comprehensive decision model, large data can provide auxiliary function for emergency decision-making command, improve emergency rescue capability, reduce casualties and property losses.

3) Providing environmental monitoring methods

Large data technology plays an important role in reducing road traffic congestion and reducing the impact of automobile transportation on the environment. Through the establishment of regional traffic emission monitoring and forecasting model to share traffic operation and environmental data, and establish a traffic operation and environment data sharing test system. Large data technology can effectively analyze the impact of traffic on the environment, while analyzing historical data. The technology can also provide the basis for decision making of traffic signal intelligent control to reduce traffic delay and reduce emissions (Yang, 2015). A low emission traffic signal control prototype system and a vehicle emission environment impact simulation system are established.

3. PLATFORM ARCHITECTURE

Intelligent traffic big data analysis cloud platform consists of basic service layer, data analysis layer and terminal publishing layer. As shown in figure 1.

![Intelligent transportation large data analysis cloud platform overall architecture](image)

Figure 1. Intelligent transportation large data analysis cloud platform overall architecture
3.1 Basic service layer

The basic service layer is the basis of the data analysis layer and terminal distribution layer. The main purpose is to use cloud computing technology to integrate data from different systems, to analyze and store heterogeneous data, and to have security and stability.

This layer provides data and performance guarantees for the data analysis layer and the terminal distribution layer, as a collection of multiple traffic application system, data update frequency is very fast, including insert, delete, add, reorder and so on, in the case of dynamic updates, it is important to ensure the security, stability and accuracy of the data.

3.2 Data analysis layer

The data analysis layer should combine the basic services layer to provide actual requirements for data and traffic management, the use of big data and data mining technology, to help the traffic subject decision-making and judgment, which need to have the accuracy of data, analysis of real-time and functional diversity of three characteristics.

The accuracy of data requires data mining model to be fully and rigorously validated, and the model is evaluated accurately on a regular basis; Analysis of the real-time requirements of the technical platform can provide users with the latest analysis of the conclusions, make effective decisions based on the latest conclusions, this requires large data technology to quickly store, process and operate (Floridi, 2012; Jeffrey, 2016). The diversity of functions requires the platform to have multidimensional analysis capabilities. On the data time scale, the technical platform needs to analyze the daily, weekly, and grade data; In the audience, the platform needs to meet the different needs of the government departments, management departments and the general public; in the data model, the platform should be able to provide the status quo analysis and trend prediction.

3.3 Terminal publishing layer

The main function of the terminal distribution layer is to store the analysis results from the data analysis layer into the cloud, according to different terminal requests, different analysis results are provided and displayed, and the terminal distribution layer is easy to use. This layer is directly oriented to application objects, and the services provided must be easily understood and used by users. The interface is friendly and easy to operate, and it can accurately find the functional entrance corresponding to the requirements.

4. TECHNICAL REALIZATION

Platform specific technical realization idea is using a large data processing and storage technology of traffic data, combined with the big data and cloud computing technology, The use of expert mathematical model for mass traffic data for multi-dimensional analysis and mining, and through cloud publishing services, the analysis results will be communicated to all types of terminals, improve people's perception of the condition of the road bridge and traffic conditions, it enables traffic participants to complete traffic evaluation and decision-making quickly, comprehensively and accurately, and realizes intelligent traffic management.

4.1 Technical realization of basic service layer

This layer is mainly implemented by cloud computing. Cloud computing refers to the Internet based super computing paradigm, which combines large amounts of information and processor resources stored on personal computers, mobile phones, and other devices to work together. This is a new approach to sharing infrastructure, it can connect huge pool of systems to provide various services.

This paper uses OpenStack to build cloud computing platform, transforms all kinds of traffic IT infrastructure into equipment resource service, and uses Hbase to realize fast data access.

4.1.1 Traffic data source integration

Cloud computing has realized the integration of distributed system and heterogeneous data, and effectively organize traffic infrastructure information, traffic object information and vehicle information in different systems, and then using a large data processing technology, to solve the integration of massive data storage and read. As shown in figure 2.
4.1.2 Cloud platform infrastructure services

Cloud platform infrastructure services is the use of OpenStack to build a cloud computing platform, the transportation of various IT infrastructure into equipment resource services, so as to realize the on-demand service of infrastructure cloud, resource pooling, rapid expansion and service scalability (Dhruba, 2010; Konstantin, 2010). The overall architecture of infrastructure cloud services is divided into four layers: resource pool layer, virtualization layer, management layer and service layer, as shown in Figure 3.

In Figure 3, the resource pool is at the bottom of the architecture, including all the physical devices in the infrastructure cloud services, such as hardware servers, network devices, storage devices, and other hardware devices. The resources in the resource pool are not individual physical device individuals, nor are scattered physical devices, but are all focused on a pool in the image, forming a centralized pool of resources. The virtualization layer is used to select and package resources from a pool of resource pools in accordance with user or business requirements to form computing resources of different sizes. The management function is to carry out maintenance and management of the unified resource pool layer, including the collection of information resources, to understand the running state of each resource and performance, and the resources are packaged with different virtualization technologies, and ensure the availability of computing resources and load balancing after packaging. The main function of the service layer is to provide an interface for users, and to provide users with the ability to use management, virtualization, and resource pools.

[Figure 2. Traffic data integration sketch map]
4.2 Technical implementation of data analysis layer

Using data mining technology, according to the needs of traffic management, such as road network monitoring and control, public travel services, data analysis and so on, combined with the expert mathematical model, the massive data collected in state of highway bridge and traffic condition are processed and analyzed in real time, grasp the traffic status, interruption rate and congestion at anytime and anywhere, and realize a series of intelligent traffic behaviors, such as highway bridge security assessment, congestion warning, traffic guidance and so on.

4.2.1 Hadoop Technology

Hadoop is a software framework for distributed processing of large amounts of data. In reliability, it maintains multiple copies of work data, ensuring that nodes are re-distributed for failed nodes; in terms of speed, it works in a parallel manner, with parallel processing, processing speed, and processing of PB level data. These advantages of Hadoop make the platform more accurate and more efficient for traffic data processing.

4.2.2 Application of NoSQL in Intelligent Transportation

NoSQL is a non-relational database. Traffic data contains large amounts of pictures, video, audio and other data, the traditional relational database cannot be effectively stored and managed. Tables in NoSQL and relational databases are stored and formatted data; the structure is different, non-relational databases are stored on key assignments, its structure is not fixed, each tuple can have different fields, each tuple can add some of its own key assignments as needed, so that it will not be limited to the fixed structure of the traditional traffic database.
system, reducing the amount of time and space (Wang, 2010; George et al., 2016). Its advantage is that it can support the database, and can efficiently store and access the massive traffic data. It can make the database more scalable and more convenient. These characteristics are very beneficial to our handling of traffic data.

**Figure 5.** NoSQL and relational database combination sketch map

### 4.2.3 HBase Technology

Using HBase technology, traffic data can be stored in several data centers, what users see is virtual single data center, and each data center is backed up with each other. This enables the platform to have high throughput capability for massive traffic related data and non-relational data (such as traffic surveillance, video, audio, various application documents, etc.), fast access to data can be achieved.

**Figure 6.** Intelligent transportation cloud platform — HBase Technology

### 4.2.4 Data analysis technique

Data analysis can dig out hidden information and rules from huge amounts of data. Through the analysis of the existing traffic data, not only can the traffic status be fully understood, but also the future traffic situation can be predicted. The correct method of data analysis can reduce the probability of false positives and missed reports of traffic accidents, so as to assist traffic decision-making and improve traffic operation efficiency.
4.3 Technology implementation of terminal distribution layer

After analyzing the mass data of highway bridge status and traffic situation, the traffic status can be carried out through cloud computing service for multi terminal and multi type publishing services. Support PC, Pad, smart phones (iOS/Android) and other terminals; release information, including traffic status, traffic replacement, auxiliary decision-making data, charts and other types; The information audience including industry executives and citizens, making the industry management personnel can quickly evaluate and deal with all kinds of traffic emergencies, the public is informed of the safety status of the road bridge and the real-time state of the traffic, so as to improve the level of traffic management and service satisfaction.

The information is released in real time by cloud services via smart mobile-phones and PC terminals, the service cloud platform release, can let the public share of traffic and environmental data at any time, so as to make the reasonable travel plans, ensure public security, reduce road traffic congestion, reduce the impact on the environment of automobile transportation.
5. CONCLUSION

This paper uses cloud computing, big data, data mining and other advanced technologies, cloud platform for large data analysis of intelligent traffic is constructed, and the three-layer architecture of platform is expounded, the characteristics of each architecture layer should be defined, and the technical implementation of each layer will be introduced. The platform can solve the problem of massive data storage, analysis and multi terminal distribution, and provide a theoretical basis for the construction of large data cloud platform for intelligent transportation.

1) Cloud computing technology can realize the integration of distributed and heterogeneous data, and organize the infrastructure information, traffic object information and vehicle information in different systems effectively. The use of large data processing technology to solve the integration of massive data access timeliness.

2) The parallel data mining and service based on cloud computing can help the traffic department to discover useful information from the data, describe and predict the traffic situation, and play the potential value of the data. Using large data technology, it can optimize data processing and analysis process, and improve data analysis ability.

3) Based on the multi-dimensional traffic information generated by data mining, using the cloud publishing platform, through the PC, mobile phones and other terminals to receive equipment, users can understand the traffic information in real-time. This platform can guarantee the compatibility of terminal, the usability of function, the real time of data, the universality of region, the accuracy of information, and provide personalized traffic information service for users.

ACKNOWLEDGEMENTS

Thanks to China communications materials & equipment company limited. for providing relevant research materials and data. In the course of the project, colleagues from China communications materials & equipment company limited. have offered a lot of help. Thank you for your help.

REFERENCES

Qiu W.Y. (2013). Cloud data platform technology for intelligent traffic data analysis. China traffic information, 8(10), 106-110