13th COTA International Conference of Transportation Professionals (CICTP 2013)

An insight into traffic safety management system platform based on cloud computing

Xian Xiaobing[a], Bao Chao[b], Chen Feng[c]

[ac]*Informatization Office, Changshu Institute of Technology, Changshu, Jiangsu Province, China ,215500*
[b]*School of Computer Science &Engineering,Changshu Institute of Technology, Changshu, Jiangsu Province, China ,215006*

Abstract

With the recent development of information technology and the increasing road traffic collisions, better gathering, analyzing and tackling useful information considering traffic safety has become one of most pressing concerns for local traffic authorities. The ultimate goal is to ensure sound traffic management and quick access to updated road traffic information for city travelers. Therefore, we attempt to design a traffic safety information platform based on cloud computing technology, which serves as an excellent channel for communication between traffic authority, staff and the public. This platform relies on data gathering, analyzing, mining model, readily adapted to present traffic conditions and reliable data from traffic accidents in Beijing. The platform is fully integrated with corresponding variables such as people, vehicles, environment, weather and road conditions.

1. Introduction

Due to accelerating urbanization and increasing traffic flow, traffic injuries tend to pose a serious threat to roadway safety. According to the statistics published by World Health Organization (WHO), annual road traffic deaths and injuries around the world has reached 1.2 million and 50 million respectively, which often resulted in permanent damages. China has witnessed around 500, 000 annual traffic accidents averages and more than 100, 000 death toll closely associated with road accidents in the past few years unfortunate to stand second to none in the world. There is no doubt road traffic accidents have became one of the leading causes of death(Shi & Li, 2011) and are expected to rank from the ninth in 1990 up to the sixth in 2020. To sum up, road traffic injuries have became a searching concern facing our world.

With the recent development of information technology and the ever-increasing threat to life from road traffic collisions, it has become one of most pressing concerns for local traffic authorities to better gather, analyze and tackle useful information considering traffic safety with a view to ensuring sound traffic management and for city travelers to get quick access to updated road traffic information in the hope of guiding their travel behaviors as well (Zeng & Yan, 2011). Therefore, we attempt to design a traffic safety information management platform based on cloud computing, data warehousing and data mining, which serves as an excellent channel for communication between traffic authority, staff and the public. It aims to take full advantage of road safety information database in attempt to generalize, judge and predict various safety information such as traffic accident black-spots, vehicle types liable to serious accidents and weather based on already occurred accidents on one hand and gather, tackle and share detailed traffic conditions without delay such as road condition, potential risks, accidents, disasters and weather on the other hand (Zheng, 2012).

* Xian XiaoBing.Tel.8613814995953.

 Email: 82850836@qq.com

2. Related technologies and concepts

2.1. Cloud Computing

Cloud Computing can be used as a platform and one type of application program abreast of the latest computing development such as distributed computing, parallel computing and grid computing.

2.2. Hadoop

As infrastructure of distributed system, Hadoop is developed by Apache Foundation. Users can explore the distributed programs and make full use of cluster of high-speed operation and storage without knowing details about the distributed architecture substrates (White, 2011).

Figure 1.Hadoop component

HBase	Pig	Hive	Mahout	Avro
Map/Readuce		HDFS		ZooKeeper
Hadoop Common				

As can be seen from the picture, it contains a lot of components. We would here introduce only a few main modules such as Map/Reduce and HDFS aiming at tackling mass data.

2.3. Hive and Mahout

Hive is a data warehouse platform based on Hadoop. By utilizing Hive we can easily realize data extraction, transformation and loading. Furthermore, it's possible to transform the program from QL to corresponding Mapreduce and base them on Hadoop by HQL. Mahout is a distributed framework of machine learning and data mining, providing some classical algorithm in the extensible field of machine learning. It attempts to help developers create intelligent application in easier and friendlier way. Mahout contains many functions such as clustering, classification, recommend filtration, frequent subitem mining. In addition, Mahout can effectively spread to the cloud by means of Apache Hadoop library (About Mahout).

3. Problems and construction mode

3.1 Problems

Public traffic information service has been in the limelight worldwide. However, there are still a lot of weaknesses in this area:
- Limited information service ability. At present most information services rely on government and information gathering, tackling, and releasing are independent to some extent. What's more, information content seems to be simplistic and impractical.
- Drawback of service content, approach, quality and range. The present service content mainly touches upon travel information and service approach is considered to be outdated and thus leave much room for improvement.
- Lack of constantly updated traffic safety information. At present time, the relevant traffic safety information is incomplete for information service provider and it's really tough gathering, storing and processing mass information.

Given problems above, we try to establish a traffic safety information platform based on cloud computing for local traffic authorities and the public through establishing traffic safety knowledge system peculiar to Beijing. With the platform any users can connect cloud service platform with their own terminal equipment and keep them well informed of any traffic safety information they are concerned about. At the same time we strive for accurate analysis of these information and recommend to those people sharing the same interest targeted at promoting the sharing and transmitting of traffic safety information and knowledge in society at large and further establishing traffic safety information

platform based on cloud computing .

3.2 The main research content

During the implementation of the project, we bring many research approaches and experiment means into focus, that is, cloud computing, data warehousing, data mining. We attempt to design a traffic safety information management platform based on cloud computing, data warehousing and data mining, which serves as an excellent channel for communication between traffic authority, staff and the public (Gray, 2011). It aims to take full advantage of road safety information database in attempt to generalize, judge and predict various safety information such as traffic accident black-spots, vehicle types liable to serious accidents and weather based on already occurred accidents on one hand and gather, tackle and share detailed traffic conditions without delay such as road condition, potential risks, accidents, disasters and weather on the other hand.

The key problem needs tackling:

There are three key technologies:

• Cloud storage of massive traffic data. Use large-scale distributed cloud storage structure to store all information including images, audio and video files with regard to traffic safety accidents and ensure quick and safe access to all information.

• Accurate analysis and recommendation of massive traffic data. Based on the processing platform of distributed mass data users can data mine and analyze traffic safety information after extracting, organizing, analyzing and processing structured mass information in order to recommend them to relevant users.

• Cloud terminal technology. Support various mobile terminal including mobile phones, pad and customized cloud terminal in diverse interactive forms so that users can use cloud services anytime and anywhere and share traffic safety information.

3.3. Architecture design of intelligent traffic safety information platform based on cloud computing.

Based on the awareness and understanding of above problem, we attempt to design the following system structure.

Figure 2.Architecture design of intelligent traffic safety information platform based on cloud computing
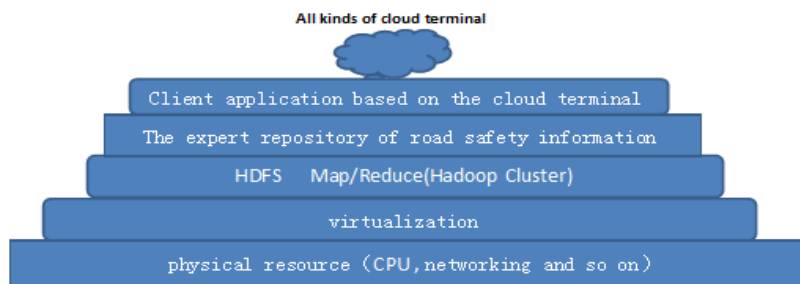

Figure 2 can be briefly explained as follows. Those mass data gathered by each branch system

provide support for dividing users in the field of traffic safety. In most cases drivers' random behaviors often result in irregular pattern, which is more appropriate for cluster analysis. Due to the large scale of user data, we decide to adopt Hadoop for parallel computing and distributed storage. Besides, to improve the utilization rate of resources we have realized the virtualization of server layer, built traffic information security knowledge based on all kinds of data mining algorithm with Mahout and design mobile terminal access for solving all kinds of cloud terminal access in the application layer. Only in this way can we ensure users get access to cloud computing traffic safety information platform which is based on Critrix XenDesktop via any tablet PC, smart phones, PC, Mac and thin clients.

3.4 Application case

Due to complicated factors in relation to traffic safety, it is not easy to reach a consensus on the definition and standard of traffic accident. More important, there is no widely accepted effective algorithm on account of incomplete statistical data so far. This paper attempts to confine influencing factors of road safety in a grey domain by adopting clustering algorithm (Jaworski, et al., 2011) and evaluate overall level of road traffic safety based on information screening, processing, extension and expansion. The procedures will be as follows:

To start with, we should be aware of the actual situation of data source, design target data and select desired attribute of influencing factors of traffic accidents after ETL processing based on Hadoop, which include data extraction and transformation. The final step is cluster analysis.

- Format conversion: Data format must be transformed into the input format that can be handled by clustering algorithm. What clustering algorithm can deal directly with the format is squencefile in Mahout. So we should write a Class which is used to convert the format into squencefile file: InputMapper and a Map function to implement InputMapper. Map function is defined as follows: Public void Map (LongWritable key, Text values, OutputCollector<Text, VectorWritable>Output, Reporter reporter)
- Invoke Apache Mahout
- Invoke parallel clustering algorithm of Mahout, which contains a couple of clustering algorithm. In this paper K - means clustering algorithm is a case in point.
- Obtain the clustering results of HDFS
- After finishing clustering algorithm, we still need analyze the clustering results. Namely, obtain the results directly in HDFS by Hive and extract them to local client for analysis.

4. Conclusion

Due to its potential theoretical and practical values in the field of traffic management, cloud computing has aroused wide concern among researchers. However, as we've explained, it is no easy to start from scratch and thus we call for further analysis and exploration.

Its application in the traffic field also depends on the development of related areas such as IT and logistics. That is to say, the development and application of cloud computing in various basic industries can also facilitate its development in traffic field.

References

Shi J., Li, X. (2011). Research on Traffic Information Cloud Computing and Its Application. *Journal of Transportation Systems Engineering and Information Technology*,(01):179-184.
Zeng, K., Yan, J. (2011). Cloud Computing and Its Application in Intelligent Transportation. *Modern Science &Technology of Telecommunication*,(05):45-51.
Zheng, X. (2012). Application of Cloud Computing in the Intelligent Transportation System in the Future. *Guangxi Journal of Light Industry*,(03):88-89.
White, T. (2011). Hadoop: The Definitive Guide. *O'Reilly Media*: 30-32
About Mahout . https://cwiki.apache.org/MAHOUT/quickstart.html s
Gray Evaluation Method Based on the Relative Merits of the Clustering of Road Traffic Safety. *Road Traffic & Safety*. 2006(11)
Jaworski, P., Edwards, T., Moore, J., Burnham, K. (2011). Cloud Computing Concept for Intelligent Transportation Systems. 2011 14th International IEEE Conference on Intelligent Transportation Systems. 2011: 931-936