

Design of intelligent transportation System based on Hadoop

Jianfeng Hu, Jinghai Yin, Zhendong Mu

Institute of Information Technology, Jiangxi University of Technology, Nanchang 330098, China

Keywords: Big Data; Intelligent transportation; Hadoop; MapReduce

Abstract: The rapid development of urban transport, while give people more and more convenient, have appeared in the urban traffic pressure is too large and cannot be efficient use of transport resources. In the face of very large scale travel, It has been the target of urban intelligent traffic system which can accurate and timely analysis of traffic status and traffic induced service for travelers to provide quality. This paper introduced the concept of "big data". The traffic data and to calculate is concentrated in the database. The use of graphs of the Hadoop distributed computing framework to flow processing of traffic data. The data processing platform packaging will be released to service into a whole and treated with distributed platform framework to ensure the timeliness and high fault tolerance.

Introduction

Car networking refers to through a variety of wireless communication technology, to realize information sharing of all the vehicle state information (including attribute information and static and dynamic information, etc.) and road traffic environment information (include the information of road infrastructure, traffic conditions, service information). According to the different functional requirements of all the vehicles running condition for effective supervision and comprehensive services, Car networking can be achieved with the car, car and road, the car and the exchange of information between people. It can help achieve "dialogue" between people, vehicle and road. Like the Internet connects every single computer, car networking can have independent car together. With development of advanced sensor technology, network technology, computing technology, control technology, intelligent technology, and car networking can achieved the comprehensive perception on the road and traffic, large scope, large capacity data between multiple systems of interaction. For every car traffic control all the way, on every road traffic control of time and space, give priority to in order to provide the traffic efficiency and traffic safety of network and application.

Intelligent traffic originated by the information technology such as simulation, real-time control and communication network to solve the urban traffic congestion to try. Traffic congestion has become due to the worldwide problem of urbanization and motorization, and cause the low efficiency of traffic facilities, air pollution and fuel consumption increase. The outstanding characteristic of intelligent transportation is information collection, processing, distribution, exchange, analysis, and used as the main line, the diversity of service for the traffic participants and managers.

With the popularization of vehicle-mounted mobile terminals, and other mobile devices, and application of sensor network technology, using mobile devices to access social network has gradually become the mainstream. At the same time, the use of GPS, mobile phone, mobile terminals provide activities such as public transportation card path information; can be predicted for some event or to infrastructure solutions such as rationality assessment. Mobile social network has brought a lot of research and application of new opportunities, such as location services, taxi abnormal trajectory recognition and detection, the optimal taxi passenger search strategy, etc.

This thesis is based on vehicle networking platform, combined with the increasingly mature mobile social network for urban intelligent transportation provides basic data and information services. Is

mainly the related applications for motorists, and between different manufacturers of equipment, cannot connect between the acquisition didn't get the Shared data and to provide information services, car network still has a lot of a relatively independent private network with each other. If build a unified and open car networking, if you have not only the same brand, between different brands of cyclists can also establish a repeat information sharing and communication between, at the same time, car networking information will also be able to provide car free family data and services, to bring people's transportation is convenient. The main three scenario of the application has the following:

(1) The bus stop, waiting for the car passenger login system, can check through the site bus routes, you can view all really moving vehicles on a route location, speed, and expected time of arrival in the site.

(2) The outer car free people want to go back, must find appropriate traffic tools, at that time he can have three options. The first is through the system search nearby friends, driving out contact if they can get on the ride back, relying on the mobile social network. The second is search in no-load condition near the taxi, contact the driver come here they have. And the third is driving a bus to bus station to check the data on the bus.

(3) The driver met traffic accident and traffic accident dispute, can take advantage of the system quickly contact in the nearby friends rushed to the scene, to help deal with related affairs. Leaving car owners also can monitor the movement of the vehicle and location information.

System Architecture

Enterprise Architect is a powerful tool for specifying, documenting and building your software and business process projects. Using Enterprise Architect's support for UML and its related standards, you can model new complex software and business systems, or visualize and maintain existing systems.

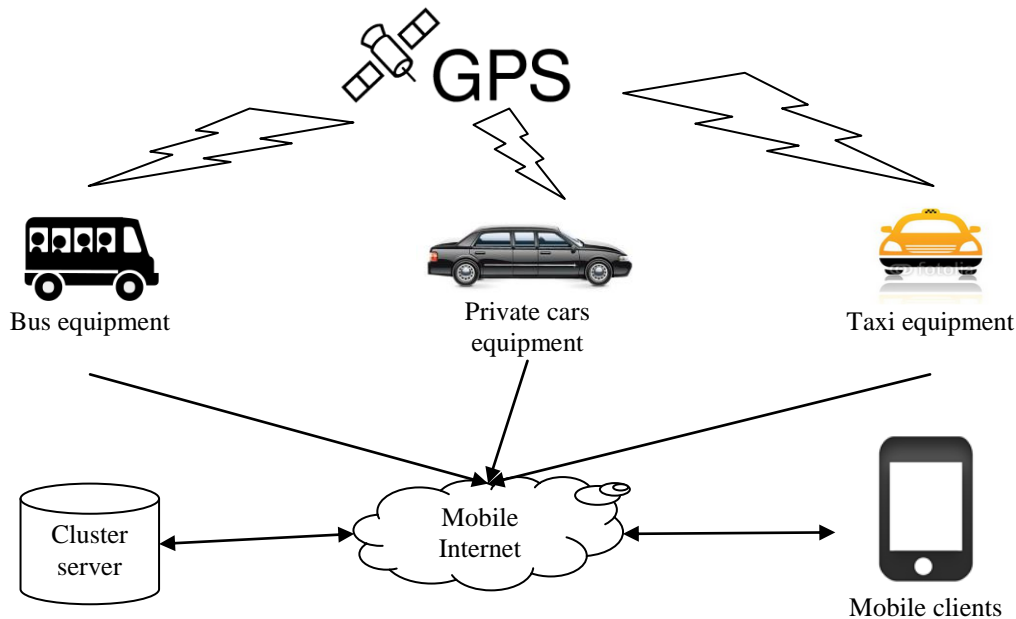


Figure 1. System structure diagram

As it can be seen from Figure 1, all kinds of vehicles will be obtained from GPS satellite positioning information through the mobile Internet upload to server, all kinds of mobile client is also through the mobile Internet access to relevant information of the vehicle from the server. Finally realize if you have

these services main foothold is interconnected. It is through the integration of vehicle, road, all kinds of information and services, in the end is a (the car, and focus on the car) to provide services, therefore, can provide information and if you have access to Internet service is not only a vehicle terminal, but all access to mobile Internet terminals, tablets, smart phones are also you connected terminals. Existing technology and the application of the mobile Internet can basically use in 1st of interconnected, including media entertainment, e-commerce, web application, information service, etc. Of course, compared with the existing mobile Internet, it has two key features: one is related to the vehicle and road, the second is the position information as a key element.

System Analysis

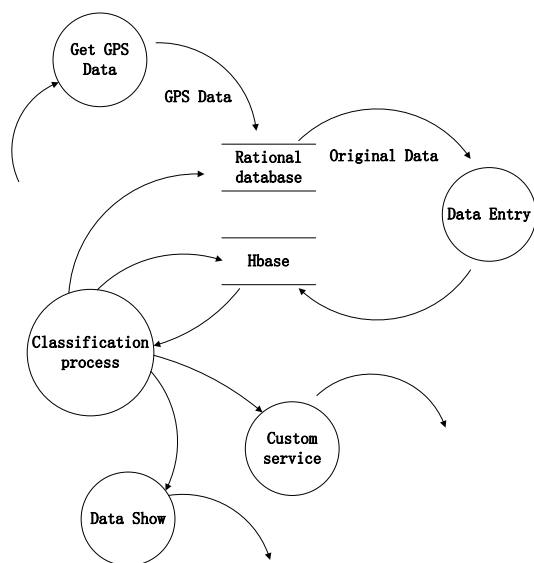


Figure 2. Data flow diagram

As shown in figure 2, on-board equipment through the GPS satellite positioning information, and then the positioning data, through the mobile Internet to upload to server, server load first temporarily stored in the platform of raw data, using the data from table schema designed in advance, to load data into Hbase distributed database, to facilitate the next step of calculation. Visit Hbase data query module, the use of graphs the parallel computing model, write custom data processing function of the Map function and a Reduce function, to store the results of calculation to the Hbase, and to deal with before and after processing the data query in data query module. The final state of the key nodes of data processing data display to the user, and can according to different application requirement for data, the custom data format, the module for the calculation of the second module plays a leading role. In the process of data of the whole flow, storage and graphs of Hbase framework calculation is one of the key steps, on the basis of fully understand the characteristics of the data, implement multi-node distributed processing technology, and ensure the consistency of the reliability of the platform and data storage.

System Design

Graphs programming model is easy to use, easy to implement related applications, although there are various different transaction business logic, but they are the principle is the same. Loading from HDFS graphs using TextInputFormat keys to refer to the Map function mapping.

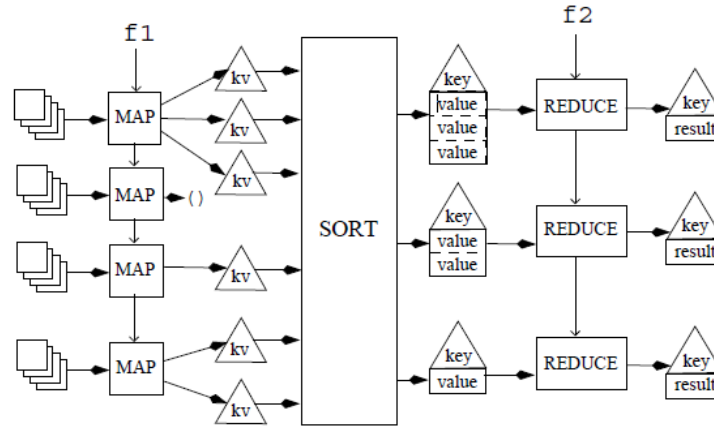


Figure 3. MapReduce flow chart

Hadoop contains a Java graphs API. It can be used to output key/value pair, call context object. The write method can collect new key values to the output results. In addition, it is also used to update count and status information. This API in org. Apache is Hadoop. The Map function by the Mapper class implements, the latter declares a Map method, the input for the one row of a text file, output intermediate key/value pair <license plate number, point information>. The Map according to the program to customize the mapping mode, from a text files to extract the useful information in a line of data mapping into the intermediate key/value pair. Graphs provide partition interface, the default partitioning tool is Hash Partitioned, it in the middle of the key value for the Hash operation after the key values of evenly distributed to Reduce, Reduce each process a partition. Among the key value output to a local buffer after the partition, when the buffer is overflow is filled to disk. Overflow wrote at the same time, according to the key value for the sorting, or digital order since the childhood in alphabetical order. And sorted key/value pair will merge on the disk, the key value of the same key value to merge together.

Merged in the middle of the key from the Map in the HTTP way to Reduce end. In Reduce side key corresponding to the same value will be passed an iterator in order to Reduce method. (key, values Iterator) after processing to produce the final result. Reduce function on every element in the list one by one to perform a function, the end result is done through the handling of all run Reducer operating results are classified and explained. Within this transaction is to retrieve one day after a two level of vehicle information, Reduce function between input key/value pair <license plate number, point and information of its statute, results of key-value pairs 2 > < license plate number, point 1 + point>. It will be the value in the list of the number of point to a queue, two traverse queues if matching level 1 and level 2, the output. Finally, the write method to collect the new key values to the output is called the context object.

Create jobs in the run function class, which defines the Job execution specification, we are using it to control the operation of the whole operation. Then you need to specify the type of map and reduce, control output keys to type, the path of the input and output data..

Discussion

This paper designed the intelligent transportation system based on Hadoop, Hbase to distributed database with distributed data storage function, using the graphs distributed computing framework provides efficient distributed processing technology, build on a lot of cheap computer cluster, the cluster service from the original data entry to business application output as a "cloud", realize the automatic transfer of data. At present, the platform is still in further development, has a good scalability and efficient data parallel processing ability.

Acknowledgment

This work was supported by Jiangxi province department of science and technology support project [20142bbe50030] and Natural Sciences Project of Jiangxi Science and Technology Department [20122BAB201049]. The authors are grateful for the anonymous reviewers who made constructive comments.

References

- [1] Pavlo A, Paulson E, Rasin A. A comparison of approaches to large-scale data analysis. Proc. of the SIGMOD. Rhode Island: ACM Press, 2009, 165—178.
- [2] Chu CT, Kim SK, Lin YA, Yu YY. Map-Reduce for machine learning on multicore. In: Seholkopf B, Platt JC, Hoffnlan T, eds. Proc. of the NIPS. Vancouver: MIT Press, 2006, 281-288.
- [3] Panda B, Herbaeh JS, Basu S, Bayardo RJ. PLANET: Massively Parallel learning of tree Ensembles With MapReduce. PVLDB, 2009, 2(2): 1426-1437.
- [4] Lin J, Sehatz M. Design Patterns for efficient graphal gorithms in MapReduce. In: Rao B, Krishna Puram B, Tomkins A, Yang Q, eds. Proc. Of the KDD. Washington: ACM Press, 2010. 78-85.
- [5] Chuck Lam. Hadoop in Action. Publisher: Manning Publications. December 6, 2010.
- [6] Jimmy Lin, Chris Dyer. Data-Intensive Text Processing With MapReduce. Publisher: Morgan and ClayPool Publishers, April 11, 2010 P34-P35..