



Design and Implementation of Smart Logistics Big Data Management Service Platform

Xianjuan Shang^(✉), En Guo, Yuxin Song, and Yusen Li

Shandong Xiehe University, Jinan 250109, Shandong, China
43026511@qq.com

Abstract. With the overall informatization level of the logistics industry increasing day by day, the logistics data information shows explosive growth. The traditional logistics information management system lacks corresponding processing methods for large-scale, real-time and heterogeneous data information, which affects the high-quality development of modern logistics industry. In this regard, this paper aims at improving the efficiency of logistics data information processing and application, and takes big data technology as the core to build a smart logistics big data management service platform, which sets a new paradigm for the development of logistics industry under the big data environment. Hadoop framework is used as the distributed data processing server in the platform to collect and store massive data, and Spark framework is used as the data analysis and processing engine. In addition, the platform will be integrated with Web2.0 technology to provide convenience in user login, data access and visual interaction. After simulation test, the platform has outstanding advantages in data collection, data storage and real-time calculation, which realizes the application of logistics data value and makes a positive attempt to promote the construction of smart logistics mode.

Keywords: big data technology · Smart logistics · Manage service applications · Hadoop · data system

1 Introduction

Under the current background, the new generation of digital information technology represented by big data, Internet of Things and cloud computing is accelerating the innovative integration with the logistics industry, which has pushed the logistics industry into a period of great change from traditional logistics to modern logistics. [1] As an important content and symbol of modern logistics, the informatization degree of logistics industry can intuitively reflect the changes of business chain, management chain and value chain, and also clarify the future development trend of logistics industry. In the process of logistics informatization development, the volume of logistics data information is constantly expanding, and the corresponding data processing methods are gradually developing from stand-alone database software programs to networking and intelligence. Especially under the influence of big data technology, the sources of logistics data information are broader and more diverse, the data types are diverse

and mixed, and the update speed is fast and unstable, which puts a severe test on the traditional logistics information analysis and processing mode. In addition, the overall development level of logistics enterprises is uneven, the function of logistics information management system is relatively fixed, and it lacks control over the whole process of logistics industry, and it can not play the value behind logistics data information. [2] In view of this, this paper believes that in the big data environment, logistics enterprises need to upgrade the existing logistics information management system according to the current practical application requirements. With the application advantages of big data technology, network information technology and computer software technology, this paper builds a smart logistics big data management service platform. It realizes the rapid collection, distributed storage, classification, tracking, identification, query, analysis and visualization of massive logistics information, and completes the transformation from the traditional logistics system to a systematic and automated logistics information system [3].

2 Development Process

First of all, for the platform's big data infrastructure system, it is mainly built by Hadoop framework. According to the function and performance requirements of the platform, Hadoop framework will be deployed in clusters. Hadoop cluster consists of three nodes, named as Master1, Slave1 and Slave2 respectively, in which Master stands for master node and Slave is slave node. [4] The underlying operating system of each node is Linux Ubuntu 16.04, which is matched with JDK version 1.8.0_111. Hadoop chooses v2.6.1, which is installed in each node respectively, and the components such as Yarn, HDFS, Zookeeper, HBase, Kafka, Flume, Sqoop are also deployed in each node to form the basis of platform logistics data information collection, transmission, storage and management. [5] In addition, the big data basic system will adopt MapReduce distributed computing system to realize offline data processing, and combine with real-time dynamic data processing framework Spark Streaming to improve the processing efficiency of logistics information. Secondly, the design of the service management system of the platform will mainly rely on the Javaweb technology system. [6] Java is selected as the basic development environment of the service management system, MyEclipse V 2022 as the integration tool, Tomcat 8.0 as the Web server and MySQL 5.7 as the database server. Complete the configuration of Tomcat in the Preference option under MyEclipse. Then, based on the Spring architecture, the integration and encapsulation of the whole system are completed. [7] Through the introduction of the above key technical theories, the overall environment of system development, the configuration of related software and tools are determined, and the technical feasibility of the overall project of smart logistics big data management service platform is also clarified.

3 Functional Implementation

3.1 Real-Time Monitoring of the Vehicle Position

Under this function module, users can intuitively see the location information of vehicles currently operating online on the platform after logging in to the platform. It mainly includes vehicle state, coordinate position, speed, direction and historical trajectory.

Table 1. Real-time monitoring information of the vehicle position

No.	Vehicle ID	Longitude	Latitude	Speed per hour	Running state	Time	Monitoring and determination
1	0053	103.31	111.69	70.6km/h	Run	12–15 18:07	Normal
2	0193	105.87	114.33	0.0km/h	Stop	12–15 19:26	Normal
3	0268	-	-	-	Lose contact	-	Abnormal

The platform will aggregate data sources through the flume-ng-sdk component in the vehicle-mounted GPS navigation system, and package and send the corresponding vehicle position information to Kafka message middleware, which will be split according to the data information type and flow to the Spark Streaming real-time computing framework to complete data processing. [8] Table 1 shows the platform simulation test data, in which the vehicle states are running, stopped, damaged, lost, etc. The position coordinates are all double data between 0 and 120, and the speed is floating-point value. During the test, the data transmission rate of GPS navigation system is 3 times per second, and the GPS navigation system of each car is in a separate thread to avoid mutual influence.

3.2 Driving Route Planning

When the platform performs the function of driving route planning, it needs to obtain the information of logistics orders and transport vehicles through the data interface, and call the map information data through the resource access interface. Under the Spark computing engine, the A Star algorithm is called to calculate the order information, vehicle information and map information to obtain the shortest distance route. As shown in formula 1, it is a distance formula for calculating the longitude and latitude of two points on the map, in which the latitude and longitude of point A are (lat_a, lon_a) , the latitude and longitude of point B are (lat_b, lon_b) , m is the difference between the latitudes of two points, n is the difference between the longitudes of two points, S is the distance between two points, and R stands for radius of the earth. [9] After the operation of A STAR algorithm is completed, Spark will return the final result to HDFS for saving, and support the Web Server to call related results and return them to the client in the form of pictures, as shown in Fig. 1.

$$S = 2 \arcsin \sqrt{\sin^2 \frac{m}{2} + \cos(lat_a) \times \cos(lat_b) \times \sin^2 \frac{n}{2}} \times R \quad (1)$$

3.3 Demand Forecast

Under this function module, users can make reasonable decisions on short-term business needs and long-term development planning in the future. The platform can support many

Table 2. Comparison of the prediction algorithm model effects

Judgement index	LSTM algorithm	ELM algorithm
RMSE	274.59	251.19
MAE	223.06	193.46
Consumed time (s)	176.55	1.76

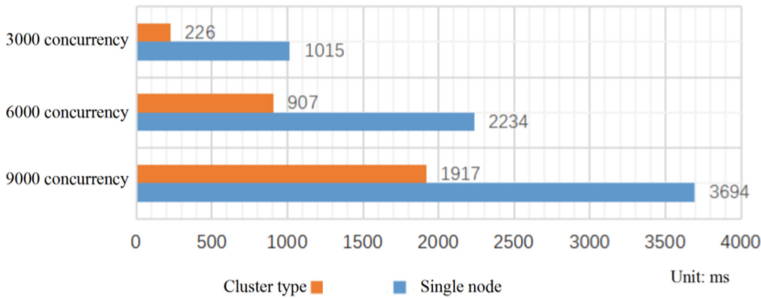


Fig. 1. Comparison of the platform operation efficiency

forecasting models, such as time series forecasting algorithm (LSTM) and extreme learning algorithm (ELM), to analyze and forecast the historical logistics data information, and compare and analyze the forecasting models of these algorithms with the indexes of standard error, average absolute error and running time. [10] The simulation test shows that the performance of the ELM algorithm model is better than that of the LSTM algorithm model in three criteria, which can adapt to the diverse application scenarios of the platform and meet the functional design requirements of the platform for demand forecasting. The results are shown in Table 2.

In addition, after the overall design and construction of the platform is completed, the running stability and concurrency support of the platform are simulated. The experiment consists of three groups of Hadoop single node and cluster node running time comparison under different concurrent processing conditions. The test results are shown in Fig. 1. The results show that the platform can improve the data processing efficiency by 40.8% by using Hadoop cluster, and has good high concurrent processing ability and load balancing ability.

4 Conclusion

This paper aims at improving the efficiency of logistics data information processing and application, and takes big data technology as the core to build a smart logistics big data management service platform, which sets a new paradigm for the development of logistics industry under the big data environment. The platform realizes the rapid collection, distributed storage, classification, tracking, identification, query, analysis and

visualization of massive logistics information, completes the transformation from traditional logistics system to systematic and automated logistics information system, and promotes the digitalization and intelligent construction of logistics industry.

References

1. Huang Bin. Research on the Intelligent Transformation Path of Traditional Logistics Industry in the Era of Big Data[J]. Journal of Technical Economics & Management. 2021.12.
2. Lu Minglu. Research on Optimization Strategy of Smart Logistics in the Era of Big Data[J]. Industrial Innovation Research. 2020.11.
3. Liao Sheng. Construction and Operation of Smart Logistics Business System under the Background of Big Data[J]. China Science and Technology Information. 2022.11.
4. Shi Fangxia, Gao Yi. Application Analysis of Hadoop Big Data Technology[J]. Modern Electronics Technique. 2021.09.
5. Xu Wenpeng, Li Shengguang, etc. Analysis of Big Data Platform Based on Hadoop Framework[J]. China Security & Protection. 2020.04.
6. Shi Feng. Development and Application of JavaWeb Based on MVC Pattern[J]. Electronic Technique. 2021.05.
7. Huang Junyong. Research and Design Based on Spring Framework[J]. Computer Knowledge and Technology. 2018.01.
8. Xie Yanqing. Research on the Application of Spark Streaming in Real-time Computing[J]. Computer Knowledge and Technology. 2018.09.
9. Yang Mingliang, Li Ning. Improved A* Algorithm for Mobile Robot Path Planning[J]. Mechanical Science and Technology for Aerospace Engineering. 2022.03.
10. Yu Qiong, Tian Xian. Nonlinear Time Series Prediction Algorithm Based on Combined Model[J]. Computer Engineering and Science. 2021.10.

Open Access This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial 4.0 International License (<http://creativecommons.org/licenses/by-nc/4.0/>), which permits any noncommercial use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

