Design and Application of Enterprise Intelligent Marketing System Under the Background of Big Data

Benyu Zhao[1,a], Qian Chen[1,b], Yan Hu[1,c*]

40684646@qq.com[a], 842134437@qq.com[b], *295738740@qq.com[c]

Chongqing College of Architecture and Technology, Chongqing, China[1]

Abstract: In order to promote the information economy process of small and medium-sized enterprises in China, promote the prosperity and development of market economy, and improve the problems of poor information management conditions of traditional enterprises and weak internet background marketing ability, this paper establishes an intelligent marketing application software by combining big data technology. The system combines hadoop platform to collect, clean, calculate and process data, and uses javaweb technology to realize data visualization. The FP-growth association rule algorithm can effectively help enterprise managers to understand customers' needs and consumption preferences, intelligently analyze the most suitable marketing data reports for consumers, and achieve marketing accuracy by means of mass data Internet platform and interactive media, thus further improving the marketing effect.

Keywords: Hadoop, Intelligent Marketing, Big Data, System Design.

1 INTRODUCTION

Nowadays, with the development and popularization of diversified terminals, a large amount of data is generated in the operation of terminal equipment, so the current era is in the era of big data with data explosion. All walks of life need to keep pace with the times and use big data tools to improve their competitiveness. The traditional marketing method is to transmit a large amount of advertising information to network users, but it is easy to lead to the problem of mismatch between advertising information and audience, which brings a lot of negative effects on the marketing products of enterprises and wastes marketing resources to some extent. [10] However, the marketing method using big data technology can collect and sort out a large amount of effective data based on consumers and conduct in-depth analysis, and can draw data reports such as consumer behavior and consumer preferences. Through these intelligent data reports, the marketing executives of enterprises can integrate the consumer groups with relevant commodity attributes, and then establish user groups to realize targeted marketing. According to the above analysis, the author of this paper thinks that the enterprise intelligent marketing system based on big data technology should be developed. The system adopts data mining technology, data warehouse technology and other big data technologies to realize the integration and analysis of business data, and then helps enterprises to make relevant marketing decisions. The development of this system can provide a strong support for the current enterprises to realize information construction. [8]

BDEDM 2023, January 06-08, Changsha, People's Republic of China

DOI 10.4108/eai.6-1-2023.2330361

2 KEY TECHNOLOGIES

2.1 Hadoop Processing Platform

Hadoop platform is the most widely used big data processing platform in the current era, and it has many advantages. Hadoop has good scalability and convenient maintenance; The distributed characteristics make it highly fault-tolerant and highly reliable; Hadoop is a platform made up of multiple components, which has better performance than other platforms. Hadoop was developed in 2004 by Doug Cutting. According to the academic contents of three documents published by Google, its author developed and created them in java language, and chose to publish them in an open source way under Apache software. Hadoop platform is a development and application ecosystem that can support data-intensive applications, and its component team is growing with time. The most important components are distributed file system HDFS and parallel programming model MapReduce. HDFS is responsible for the distributed storage of massive data, while mapreduce is to realize centralized parallel computing of distributed data, and the two complement each other.

As a distributed file storage system, HDFS is a key part of big data technology. HDFS adopts a typical master-slave structure design. The master node names the information file of the slave node, and cuts the information data into multiple blocks, which are saved by different datanode slave nodes. The client sends a data request to the master node namenode, which requests block data from each slave node cluster, and then the slave node sends information to the client to read the returned result data. Hadoop ecosystem is shown in Figure 1. Besides Hadoop and mapreduce, there are many subprojects such as Ambari, Hive, HBase, Zookeeper, Flume and Mahout. With the cooperation of multiple components and clear division of labor, even inexperienced developers can use the advantages of clusters to deal with big data conveniently and quickly. [2]

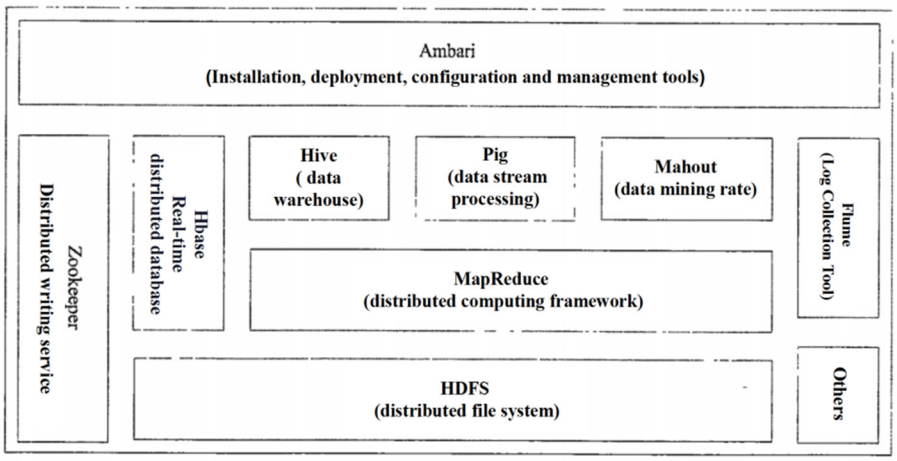

Figure 1: Structural diagram of the Hadoop platform

2.2 FP-Growth Algorithm

FP-growth algorithm is one of the most classic algorithms in association rule mining. The role of association rule algorithm is to find the relationship between massive data. Compared with other association rule algorithms, FP-growth algorithm has the advantage that it can realize frequent pattern data mining without generating candidate patterns. Next, this article will introduce the operation of FP-growth algorithm. First, we need to traverse the data packet to get the first frequent itemsets, and then we need to traverse the data packet twice and get the support count according to the first data. After that, the records in the database can be sorted and counted. In the sorting process, the infrequent items need to be filtered and cleared away. Finally, the algorithm composition can be iterated according to the mining condition FP-tree to obtain the frequent item set. Suppose we get the frequent itemsets (f:3,c:4,a:2,b:3,m:4,p:3), randomly expand the screening and delete the infrequent items in the data set to establish the memory sorting tree FP-tree. First, select the header table items, take P node as an example, and traverse all the paths involved in P node, so as to obtain {f,a,b,n:3}{cp:2} two-day search result paths, which are also the conditional patterns of P. In this random mode, FP-tree mining algorithm and construction algorithm are operated, and frequent itemsets with prefix P can be obtained without generating new paths after continuous iteration. [9]

However, considering that the traditional FP-growth algorithm model has some shortcomings in clear classification, this paper improves the algorithm on this basis. When considering the first frequent itemsets, the category attribute data is no longer taken into account. Instead, the object of counting operation is changed to the relative support number. The implementation process of the improved FP-growth algorithm model is shown in Figure 2.

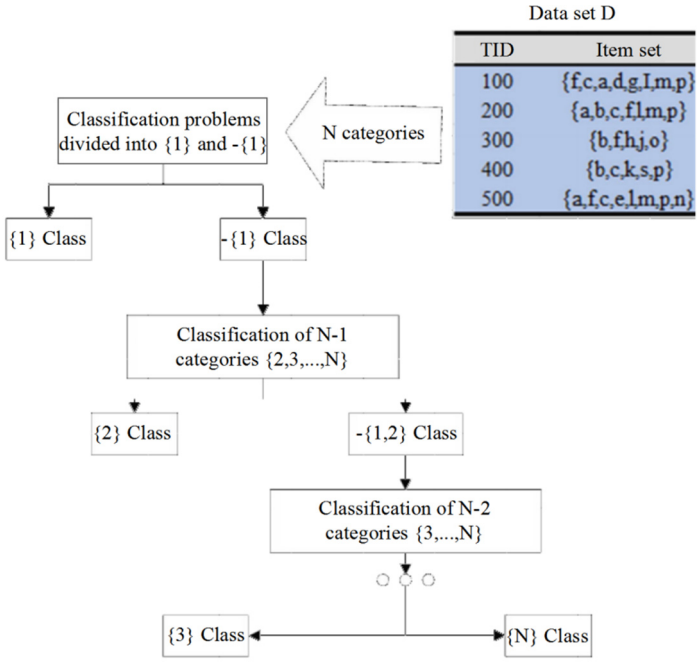

Figure 2: Implementation of improved FP-growth algorithm

The tree established by the improved algorithm model is named FP-treeR, and the improved algorithm support and confidence formulas are shown in formula (1) and formula (2). [4]

$$\text{support} = \frac{\text{supprot\_count}+|\text{R\_support\_count}|}{2\text{N}} \tag{1}$$

$$\text{confidence} = \frac{\text{supprot\_count}+|\text{R\_support\_count}|}{2\text{supprot\_count}} \tag{2}$$

2.3 Development environment

In this paper, the author briefly introduces the related technologies of platform development and use. The enterprise intelligent marketing system uses Hadoop as a big data server cluster to process data and store it in MySQL database, and uses JavaWeb technology to develop the corresponding application platform.

According to the data volume and overall operation requirements of the system, this paper chooses to build a Hadoop3.3.1 cluster with three nodes, including a master node and two slave nodes. The main node is named namenode, and the secondary node is named datanode. Then, the distributed collaboration system zookeeper-3.4.1, distributed file system HDFS 2.6.5, flume1.9.0, Hive 0.13.1 and Hbase2.6.5 are installed and deployed in these three nodes synchronously, and the initial construction of hadoop cluster is completed. The cluster will be developed under Linux system. This paper selects Centos6.5 Server release version of Linux operating system.

The front-end development tool used in the JavaWeb application of this system is boomstrap+jquery, and the development language is JavaScript+HTML+CSS. The back-end Java development tool is IDEA 2021.1.3 (Ultimate Edition), the development environment is JDK 1.8, and the J2EE framework of Tomcat+Spring MVC+Spring+MyBatis is is used in the implementation of this system. The development language is Java, and MySQL 8.0.28 is selected to help manage data. Through the introduction of the above key technical theories, the overall environment of enterprise intelligent marketing system development, the configuration of related software and tools are determined, and the technical feasibility of the overall project is also clarified. [3]

3 REQUIREMENTS ANALYSIS

3.1 Functional requirements

The user group of the intelligent marketing system of enterprises is positioned in the marketing managers of enterprises. Firstly, the system should acquire and integrate the marketing data of the enterprise management system, and save these data in the system cluster, so it has the functions of data warehouse, search engine and classification. Secondly, the system should perform algorithm analysis on the data, and the system should have the function of issuing data analysis reports.

Finally, the data analysis report should have the function of data visualization in order to be more intuitive and concrete, and make it easier for users to understand. The data index system

in data warehouse includes customer attribute, customer consumption attribute and customer gender. The daily table update data of the system includes daily increment data and daily total data. Data labels include statistical labels, rule labels, mining labels, calculation labels, etc. The analysis includes commodity analysis, user analysis, traffic analysis and recommendation effect analysis. [1]

3.2 Overall design

Enterprise intelligent marketing system adopts B/S structure and is designed with multi-layer distributed idea. The overall design architecture diagram of the system is shown in Figure 3. The system is divided into customer layer, web layer, business layer and database layer. The client layer includes pages and application clients, and users access the system through PC user ports. The main function of Web layer is to collect and process requests from clients. The content of web layer includes javabean objects and JSP pages. The main function of the business layer is to handle business logic and realize the interface function between the client and the system server, and encapsulate a large number of functional components. The data layer is responsible for managing the database server, which is convenient for the system to immediately call the data resources. [5]

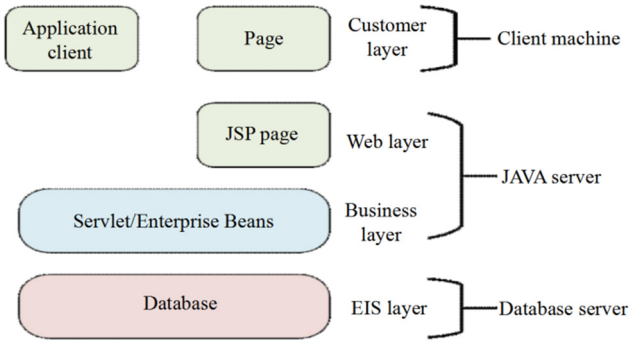

Figure 3: Overall architecture diagram of the system design

4 FUNCTIONAL IMPLEMENTATION

The main functions of enterprise intelligent marketing system can be divided into data warehouse, data classification, data query, data analysis and visual report. This section introduces the implementation logic of data analysis report viewing. When the user clicks on the front page of the data analysis function, selects the data range, and then clicks "Start Data Analysis" to start generating the data report. The back-end server will use the SummaryController to receive the get request sent by the front-end, and obtain all relevant data in MySQL, the database server, through the getDigest method. [6]

In order to improve the security of enterprise management information system, the system needs to judge the legality and permission category of system users. The user names and IP addresses of all relevant users will be saved in the background of the system database. Only those who meet the requirements of the database can log in normally. In the data retrieval function module, the user presentation layer of the system will display the retrieval conditions

related to the information categories queried by users, and the search engine of the system will make retrieval calls according to the classified keywords. Data access uses the encapsulated business data control component in the system, and after the results are found, they are returned to the presentation layer of the browser and fed back to the users. Among the data warehouse modules, two types of input modes are supported, and the automatic input mode supports the information input of other department information management systems of enterprises. The manual input method supports the salesman to manually input all kinds of sales order data of the enterprise into the background database of the system. The system will automatically enter the input information into the background database and establish statistical tables.

In this paper, the improved FP-growth algorithm is used to analyze the intelligent marketing data. The operation process is to first calculate the frequent data set of the project, then determine the association rules between consumers and customers to generate customer interest groups, and then get the recommendation rules of marketing products according to the preference list data of this group. The performance test is required for the intelligent marketing data analysis function of P-growth algorithm. Taking the marketing database of Company A as an example, we first select four data sets for functional testing. The data sets are shown in Table 1. [7]

Table 1: Test data set

Data set	Transaction item quantity	Data capacity
1	2000	30MB
2	5000	45MB
3	8000	100MB
4	15000	150MB
5	20000	200MB
6	80000	700MB

After that, we use Hadoop cluster to use the improved FP-growth algorithm to process data and control the same support variables. The data is processed in descending order of data capacity, and the support level is set to 15%. The data processing results are shown in Figure 4. By observing the data in the figure, we can see that the running time of the improved FP-growth algorithm studied in this paper is less than that of the traditional algorithm, thus achieving the purpose of improving the data processing efficiency.

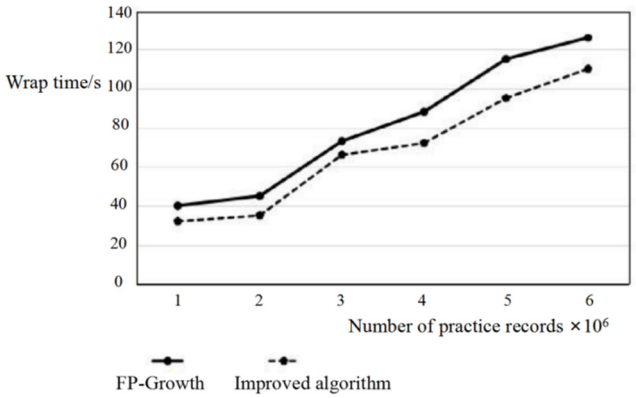

Figure 4: Comparison of running time before and after FP-Growth improvement

5 CONCLUSIONS

This paper studies the intelligent enterprise marketing system based on big data technology. The FP-growth association rule algorithm of this system can effectively help enterprise managers understand customers' needs and consumption preferences, and intelligently analyze the most suitable marketing data reports for consumers. Therefore, this research has certain application value. However, due to the author's limited ability and time, there are still some shortcomings to be improved. The intelligent marketing system designed in this paper can assist marketing leaders to make relevant decisions, but it can't directly realize the automatic intelligent marketing of the Internet platform, so there is still some room for development in intelligent marketing decision-making.

REFERENCES

[1] Du Fangwei. (2020) Construction of Intelligent Marketing Management System for Publishing Enterprises. Science-Technology & Publication.04.

[2] Fang Pengcheng. (2020) Design of Product Marketing System Based on Association Rule Algorithm.Zhejiang University of Technology.06.

[3] He Fang. (2021) Research on Enterprise Marketing Strategy in Big Data Era. China Market. 07.

[4] Hu Yuyang. (2021) Design and Development of Business AI Marketing System Based on Deep Learning. Wuhan Textile University.06.

[5] Li Jun, Zhang Yu. (2020) Construction of Enterprise Intelligent Marketing System Based on Full Chain Tracking, Platform Linkage and Emotional Interaction Mechanism. Journal of Green Science and Technology.05.

[6] Li Xinghao. (2021) Precision Marketing Mode of E-commerce Enterprises Based on Big Data. Cloud Times.04.

[7] Wang Yating. (2020) Design and Implementation of Precision Marketing Management System Based on Big Data. Nanjing University of Posts and Telecommunications.04.

[8] Xue Yu. (2021) Research on Company A's Precision Marketing Scheme. Shandong Normal University.10.
[9] Zhang Xuan. (2021) Design and Implementation of Retail Marketing System Based on Micro-service Architecture. Xidian University.06.
[10] Zeng Jingwen. (2022) Transformation of Enterprise Marketing Mode under the Background of Big Data. Industrial Innovation Research.5.