

## Research Article

# Analysis of the Impact of Big Data on E-Commerce in Cloud Computing Environment

Rongrui Yu <sup>1,2</sup>, Chunqiong Wu <sup>1,2</sup>, Bingwen Yan,<sup>1,2</sup> Baoqin Yu,<sup>1,2</sup> Xiukao Zhou,<sup>1,2</sup> Yanliang Yu,<sup>1,2</sup> and Na Chen<sup>2,3</sup>

<sup>1</sup>Business College, Yango University, Fuzhou City, Fujian Province 350015, China

<sup>2</sup>Big Data Business Intelligence Engineering Research Center, Fujian University, Fuzhou City, Fujian Province 350015, China

<sup>3</sup>Software Engineering Institute, Lanzhou Institute Technology, Lanzhou City, Gansu Province 730050, China

Correspondence should be addressed to Rongrui Yu; rryu@ygu.edu.cn

Received 23 April 2021; Revised 12 May 2021; Accepted 19 May 2021; Published 27 May 2021

Academic Editor: Zhihan Lv

Copyright © 2021 Rongrui Yu et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This article starts with the analysis of the existing electronic commerce system, summarizes its characteristics, and analyzes and solves its existing problems. Firstly, the characteristics of the relational database My Structured Query Language (MySQL) and the distributed database HBase are analyzed, their respective advantages and disadvantages are summarized, and the advantages and disadvantages of each are taken into account when storing data. My SQL is used to store structured business data in the system, while HBase is used to store unstructured data such as pictures. These two storage mechanisms together constitute a data storage subsystem. Secondly, considering the large amount of data in the e-commerce system and the complex calculation of the data mining algorithm, this paper uses MapReduce to realize the parallelization of the data mining algorithm and builds a Hadoop-based commodity recommendation subsystem on this basis. We use JavaEE technology to design a full-featured web mall system. Finally, based on the impact of cloud computing, mobile e-commerce is analyzed, including relevant theories, service mode, architecture, core technology, and the application in e-commerce, which can realize e-commerce precision marketing, find the optimal path of logistics, and take effective security measures to avoid transaction risks. This method can avoid the disadvantages of the traditional e-commerce, where large-scale data cannot be processed in a timely manner, realize the value of mining data behind, and realize the precision marketing of e-commerce enterprises.

## 1. Introduction

The development of information technology is an important characteristic of our society. This field has brought great application value to China's economic development and greatly highlighted the diversity of social development. With the continuous innovation of China's technological development, the field of e-commerce also begins to face the problem of innovation and development. Although e-commerce is an emerging industry formed by continuous development in recent years, the development of this field is bound to undergo new changes due to technological innovation. Among them, big data technology [1], cloud computing [2, 3], and other high and new technologies play a huge role in promoting this content. With the help of the

information collection and analysis capabilities of big data technology, it can bring real-time and dynamic service experience to the operation of e-commerce; that is, it can have a favorable control over the information, so as to better conduct e-commerce related businesses. E-commerce is the product of the era of big data. With the development of information technology, the traditional business model is facing unprecedented challenges. The emergence of e-commerce makes up for the deficiency of the traditional business model. Under the background of the information age, the e-commerce model focuses on the construction of the information network. From the perspective of cloud computing, big data technology and cloud platform have been comprehensively applied. By integrating the advantages of various technologies, a more powerful marketing

platform can be built to highlight the overall advantages of the Internet in today's era [4]. For marketing platforms, the overall advantage of data information lies in analysis and screening. Considering the characteristics of big data technology and cloud computing technology, it is necessary to analyze the connotation of these technologies as a whole, so as to ensure the rationality and perfection of cognition and the best application effect.

At present, with the rapid development of Internet technology, users can retrieve a large amount of data information on the network, and this data information is called big data. Big data technology belongs to the product of The Times development, and it can be in a fixed period of time; regarding the conventional software information and a variety of technical information collection and processing, only the use of more powerful technology, to complete the specific data collection and analysis, the processing of large data capacity is relatively strong and can accommodate record amounts of data. In the context of entering the data era, big data technology has gradually received attention. Big data covers a wide range of information, not limited to a certain field, and big data information includes all kinds of data in all fields of society and people's lives. And if the collection and processing of big data becomes the key to social development at this stage, big data processing is the process of processing massive data, selecting valuable data for investment and creating more value. At present, there are many kinds of big data processing technologies. Cloud computing and distributed databases are the main processing technologies of big data at present [5]. Enterprises also rely on all kinds of big data processing technology to collect more useful data for enterprises and promote their development. Big data has the following characteristics [6]:

- (1) It has a huge amount of data
- (2) Diversified types of big data and traditional data tend to be of a single type, most of which are original data types
- (3) Big data has fast processing capacity
- (4) Big data has low value density

However, the original data contains a lot of useless information, so the value density of such data is not high. It needs to be processed by big data processing technology. As mentioned above, big data has corresponding characteristics, but it also has its limitations. The existing data processing technology cannot meet the needs of life and production.

El-Sayed et al. [7] proposed that the development of big data technology brought new development space for e-commerce, and new sales model began to become the mainstream direction of e-commerce development. Wu and Lin [8] mainly introduced the introduction of big data and big data processing, analyzed the main characteristics and significance of big data, discussed the impact of big data processing on e-commerce, and put forward feasible suggestions for the development of e-commerce in the era of big data. Talib and Alomary [9] firstly expounded the actual use of big data technology in e-commerce system and analyzed

the actual use situation. Secondly, they focused on analyzing the problems existing in the use of big data technology in e-commerce system. Finally, they proposed feasible solutions based on the practical problems. In the association mining marketing of big data commodities, it is found that if the big data marketing model can be reasonably applied in the field of e-commerce, it can not only enhance the correlation between commodities, but also ensure the scope of big data mining to be more extensive. Balachandran and Prasad [10] studied the incremental assessment and management of security risks of e-commerce system in big data platform, preliminarily constructed a dynamic wind direction assessment framework, focused on the analysis of frequent threats and vulnerabilities, constantly enriched the information of the key-value pairs of threats and vulnerabilities, and supplemented the corresponding security measures. To solve the problem of dynamic risk assessment in cloud environment. Song et al. [11], from the perspective of the originated development of electronic commerce, combined the characteristics of cloud and big data technology and advantages, from the perspective of the construction and application of e-commerce cloud platform, and analyzed how to rely on the cloud and big data technology, build a faster, more accurate, more secure e-commerce cloud platform, with the rapid development of power electronic commerce mode. Behl et al. [12] focused on the key issues of efficient processing of large-scale transaction data in e-commerce. Juncai and Shao [13] built the mobile e-commerce system model based on cloud computing, proposed the Hadoop open-source cloud computing platform, designed the mobile e-commerce system, and clarified the key technical points such as the interaction between client and server. Oliveira et al. [14] discussed the precision marketing strategies adopted by some brands in actual marketing activities and summarized marketing experience. Based on the analysis of the relatively perfect precision marketing mode at present, this paper explores the inspiration of precision marketing strategy.

This paper first describes the actual use of big data technology in the electronic commerce system and analyzes the actual use of the situation, then focuses on the analysis of the use of big data technology in the electronic commerce system, and finally combines the actual problems and puts forward feasible solutions. Big data technology is being used more and more frequently in business activities. As the future is only a form of business activities, e-commerce should make more active use of big data technology to analyze and study market product positioning, customer portrait service, and market development through big data technology. Through analyzing the massive data existing in the network, the management of the e-commerce platform or enterprise can be provided with decision-making basis to improve the competitiveness of the e-commerce platform or enterprise. This paper combines the advantages of cloud computing for data storage and processing and applies it to data analysis and processing in the mobile e-commerce system, so as to avoid the disadvantages that large-scale data of traditional e-commerce cannot be processed in a timely manner. It can be used to dig out the value behind the data,

make efficient use of the data and information, and provide a sentence for e-commerce enterprises to do a good job in accurate marketing.

## 2. E-Commerce Cloud Platform Based on Big Data

The rise of cloud computing and big data platformization has changed the traditional operation mode of e-commerce. Many small- and medium-sized e-commerce enterprises quickly build e-commerce platforms by renting various services from cloud service and data center suppliers. The powerful computing capacity of cloud computing can meet the customized needs of all kinds of e-commerce users and, at the same time, better promote the collaboration and sharing among e-commerce enterprises.

*2.1. System Requirement Analysis and Overall Architecture.* To build an e-commerce cloud platform, first of all, it is necessary to define the system design requirements and conduct demand analysis. The system design should meet the following requirements:

- (1) To complete the installation and deployment of Hadoop platform and HBase distributed database, realize the efficient communication and information sharing among all nodes in the distributed network and ensure the healthy and stable operation of the whole system.
- (2) A combination of relational database My SQL and distributed database HBase is used to store data in the system. Structured data is stored in My SQL, while semistructured or unstructured media file data such as pictures and Flash are stored in HBase, through the combination of the two to achieve complementary advantages.
- (3) Realize the parallelization of data mining algorithms based on Hadoop platform to ensure that the system can achieve the goal of efficient data analysis in the environment of massive data.
- (4) Complete the design of the key business processes in the Web mall system, so that each module of the mall system can realize its own independent functions according to the requirements, and ensure that each module can be perfectly connected.

In this paper, the design and development of electronic commerce system based on cloud computing technology should meet the demand of the following functions:

- (1) The front desk receptionist shopping system user shopping system mainly realized the function of online shopping and contained the user login/register function, commodity information query and display function, classification of goods query function, the function of order with online shopping, and online payment function.
- (2) Background management system to achieve the management and maintenance of goods, orders,

users, and other information mainly includes rights management function, commodity management function, order management function, and user management function.

- (3) The data storage subsystem includes My SQL database and HBase-based mass picture storage system, which are used to store structured and unstructured data, respectively.
- (4) The recommendation functions realized by the commodity recommendation subsystem are all based on the analysis of users' historical data.

With the increase of transactions, the scale of historical data is bound to be extremely large. Therefore, the recommendation subsystem is required to run efficiently and stably in the environment of massive data. The transaction business process of an e-commerce system based on big data is shown in Figure 1.

This e-commerce system is composed of cloud service layer, virtual resource layer, and physical resource layer. Its network architecture is shown in Figure 2. Cloud computing service types can be generally divided into three levels, namely, IaaS, PaaS, and SaaS. These three levels of the overall architecture of the cloud computing technology level may include some of the virtualization technology and applications, automated deployment, and distributed computing technology, such as the technology architecture of advantage that can show extremely good ability of parallel computing and massive scalability and flexibility. The virtual resource layer can realize resource integration in the cloud computing environment and form one or more logical virtual resource pools, which share computing, storage, and network resources [15]. This layer can delegate control over the host (or cluster) resources and create and configure multiple resource pools as direct children of the host or cluster. The physical resource layer connects the cloud infrastructure, including servers, software resources, and application modules, to the network through the cloud terminal technology to provide hardware support for the application implementation of cloud services.

*2.2. Data Storage Subsystem Design.* This platform puts forward the design requirements for the data storage system. Image storage system is mainly used to store image data in the electronic commerce system, when the design requirements of HBase table structure on the premise meet the demand of system functions, which should not affect its performance, and when the image number is required after reaching a certain scale, it can still ensure that the image access time is within the acceptable waiting time range. HBase is a distributed database built on the distributed file system HDFS [16], and it has its own integrated cache, so HBase has the advantages of both massive data storage and real-time reading and writing. However, because HBase is a columnar nonrelational database, it does not have many characteristics of traditional relational databases. In order to meet the needs of the system for diversified operation of key data, as well as the requirements of mass image data storage

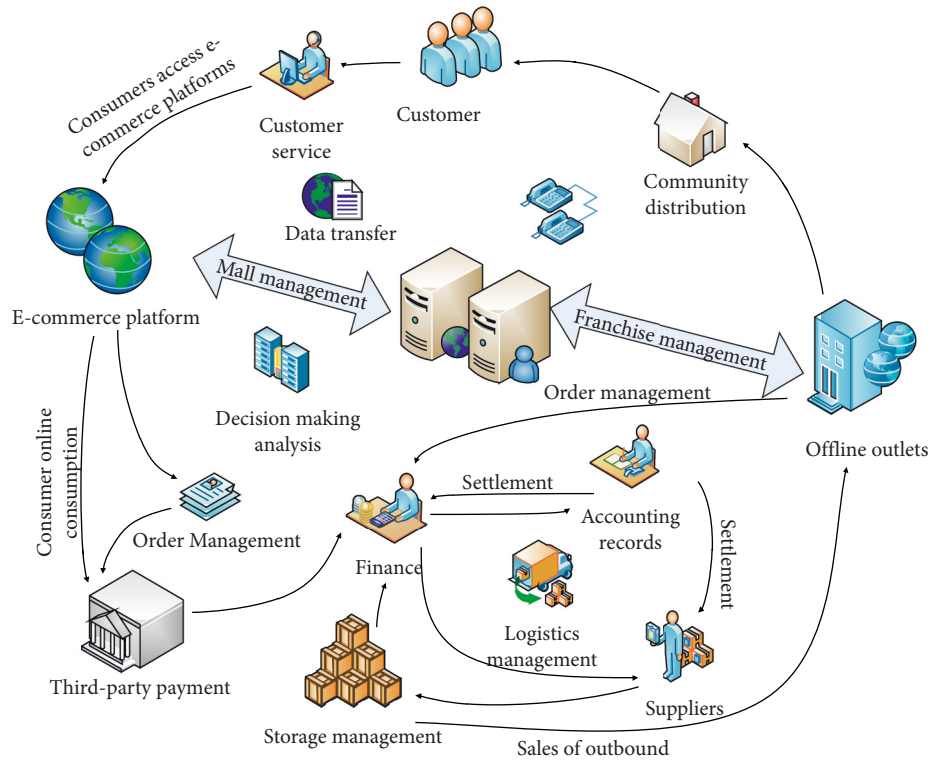


FIGURE 1: Transaction business process of e-commerce system based on big data.

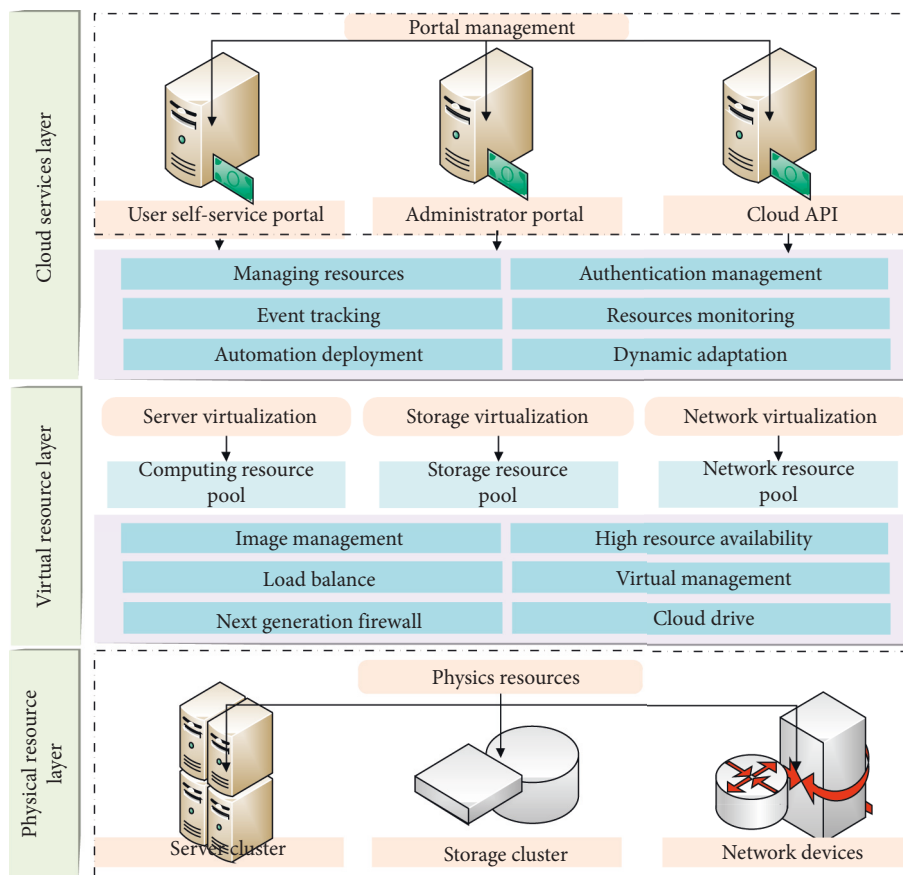


FIGURE 2: Transaction business process of e-commerce system based on big data.

and real-time reading and writing. This paper proposes a storage scheme combining the relational database MySQL with the distributed database HBase and stores different types of data in the two libraries, respectively. The data relationship between the two libraries is maintained by saving the primary key map. In the upper layer of the storage system, this paper designs the HBase-based image access API, respectively, and its organizational structure is shown in Figure 3.

HBase adopts the column data storage model and stores and processes data according to column family [17]. For each column family, HBase stores the data in each cell as a key-value pair, a combination of the Row Key, the column family name, the column name, and Timestamp. The value is the data in the cell. The underlying storage structure of HBase is distinct. It divides the table into multiple regions in the direction of rows, and each region will be split after the data growth reaches a certain threshold. Region consists of one or more stores. Each store stores the data of a column family within the scope of the current Region row, and each store is composed of a MemStore and zero or more StoreFile. StoreFile is used to store persisted data in HDFS in the format of HFILE. MemStore is a read-write cache sorted by primary key. When the data in the cache reaches the threshold value, it will be flushed to StoreFile. In summary, the data in each cell of an HBase table is ultimately stored in HDFS as an HFile data block, so the process of packaging small files is implicit when storing files. HBase-based mass image storage technology solves the storage problem of mass unstructured data in electronic commerce system. However, the main underpinning of the specific business processes of the system is structured data that not only requires real-time reading and writing, but also often requires various complex operations on it. Although HBase has the real-time reading and writing ability of massive data, its support for multiple complex queries such as multicondition query, multitable joint query, and fuzzy query is not good enough. Therefore, the relational database MySQL is adopted in this paper to solve the storage problem of business data of e-commerce system. The main data in the Web mall system can be divided into modules such as commodity management, order management, user information management, and authority management according to the business functions of the system. Because there is a close connection between goods, orders, and users, they are collectively referred to as user shopping module and described uniformly.

The commodity list is the core, in which a commodity can only belong to a certain category and also to a brand, so there is a many-to-one relationship between commodity, commodity category, and brand. Commodity attribute table is the summary of the characteristic of a certain classification of goods, so commodities attribute table and classification table formed a many-to-one relationship, and attribute value table describes a variety of possible values of commodity attribute. A product can have multiple attributes, and each attribute has a specific value, so the table and the attribute value of goods have a many-to-many relationship. Order table describes the user's shopping information, a user's order corresponds to one of the records, and each order can

package multiple goods, so the order table and the user table have many-to-one relationship, and the product table has many-to-many relationship. User browsing history table and user favorite table, respectively, describe the historical information of browsing commodities and the information of user favorite commodities. Shopping cart data table records the commodities added to the shopping cart by the user. After the user logs in the system again, the status of shopping cart can be restored by querying this table.

Permission management is an important module of the background management system to control the access of administrators at all levels. The design of this module will directly affect the security of the system. Permission refers to the operation that a module can do. This system describes the permission information through the way of URL prefix. If the current administrator access to a URL is not blocked, it means that he has the corresponding operation authority of the URL. Role table records all role information in the system. Role is also called permission group. Each role is composed of multiple permissions, and one permission can belong to multiple roles, so the relationship between permission table and role table is many-to-many. The administrator table records all the administrator information in the system, and the administrator table has many-to-many relationship with the role table; that is, an administrator can be assigned multiple roles at the same time, and a role can correspond to multiple administrators.

*2.3. Product Recommendation Subsystem Design.* The commodity recommendation subsystem is based on the user transaction data in the existing e-commerce system, which is analyzed by using data mining algorithm, and the results are presented to the users. Firstly, this paper analyzes the user's purchasing behavior through *K*-means algorithm and realizes the function of purchasing recommendation for specific users. Then, the order information of all users is analyzed through the Apriori algorithm, and the correlation relationship among commodities is mined, so as to realize the function of recommending related commodities for specific commodities. In order to meet the requirements of efficient processing of massive data, this paper implemented the parallelization of *k*-means algorithm and Apriori algorithm on the Hadoop platform [18]. The following introduces the basic idea of parallelization of data mining algorithm and the process of analyzing historical data and realizing commodity recommendation through data mining algorithm.

The specific implementation method of user purchase recommendation function based on *k*-means is as follows: *K*-means clustering analysis is carried out on the historical behavior data of all users, and the clustering algorithm is executed once for each commodity category. When all clustering in the *K*-means algorithm converges, or the algorithm reaches the specified number of iterations, the algorithm ends, and a set of centroid data obtained is the user level corresponding to the current commodity classification. The basic process of *k*-means algorithm is shown in Figure 4. Its basic idea is as follows: first, traverse the original data set

stored in HDFS, and select  $K$  points randomly as cluster centers from them. After a round of traversing, save the initial  $K$  cluster centers in HDFS. And then enter the iterative update cluster center stage; each iteration started a process graph, and each record category of the original is set to the nearest cluster center; then calculate all the points in the cluster mean as a new cluster center; each iteration is to choose the last calculated cluster center, after the number of iterations set like this. Finally, a Map task is launched to categorize each node in the original data set into the final cluster center.

The basic idea of Apriori algorithm's recommendation function is to scan the order table, combine multiple commodity IDs associated with each order in the order table into a record, and dump these commodity combination records into HDFS. After the dump is successful, the Apriori algorithm is used to mine the correlation between commodities, and the frequent item sets that meet the minimum support requirements are calculated, as well as the corresponding support degrees of these frequent item sets. Finally, these are dumped from HDFS to MySQL to save the data table for the frequent item set. The parallelization of Apriori is also realized by using the MapReduce model, and its basic process is shown in Figure 4. The design idea is as follows: first, traverse the original data set, and select the candidate 1-item set by removing duplicate items. Then, the candidate 1-item set is compared with the original data set to calculate the support degree of each item, and then the candidate 1-item set whose support degree is not less than the minimum support degree is screened out to obtain the frequent 1-item set. Next, candidate 2-item sets are generated from frequent 1-item sets, and frequent 2-item sets are obtained by comparing them with the original data sets. Iterate in this way until the candidate  $k$ -item set is generated. The candidate  $k$ -item set is compared with the original data set. If there is a frequent  $k$ -item set, the execution of iteration is continued. If not, we end up with the frequent  $(k - 1)$ -item set.

**2.4. Online Shopping Mall Subsystem Design.** Web mall system is the entire electronic commerce platform and the outside interaction window, mainly including friendly interactive pages, humanized shopping process, and perfect function of the background management system, in order to provide users and system administrators with convenient and quick shopping and management experience. The Web mall system adopts Struts2, Spring, Hibernate, and other frameworks for building and adopts layered mode to organize the code. From top to bottom, there are interaction layer, business logic layer, and data access layer, respectively. Each layer relies on the next layer and provides services to the next layer. According to the different functions of the system, the Web system can be divided into the foreground shopping system, and the background management system of two parts, in which the foreground shopping system provides users with commodity browsing, order shopping, and other functions. Background management system for the administrator provides commodity management, store

management, member management, transaction management, system settings, and other functions.

System function design is mainly on the system of each function module design. This paper designed the system by the front desk shopping mall and the background management system of two parts; the front desk shopping system provides a quick and convenient shopping platform for the public users, while mall background management system provides the functions for system administrators and operational personnel perfect and steady operation management platform. The function modules of the foreground shopping system and the background management system are shown in Figure 5.

The business process of the front desk shopping system is as follows. After entering the front page, the user can view the product display, product classification, promotion activities, and the products recommended by the system in different fields in the front page. Users select a specific commodity category item on the navigation bar to enter the commodity list page. In addition to displaying all commodities under a certain category in pages, the page also provides the function of screening commodities through attribute combination. In order to make up for the shortage of fuzzy matching in commodity classification and attribute screening, a keyword search box is set on the side of the navigation bar. The system realizes fuzzy query function according to the keywords entered by the user, and the query results are also presented to the user in the way of list. After the user clicks on a commodity goods list links, they can enter the goods details page, and the page shows the commodity pictures and detailed description, parameter information, commodity information consultation and evaluation, and after-sale service, such as information, and the system also provides a special here field-associated recommendation, recommended for the user to buy the goods of the highest frequency combination. The homepage displays the detailed information of the user's purchase recommendation and related product recommendation. The page displays the information calculation resulting from the recommendation subsystem. The Web system accesses the recommendation subsystem through remote invocation technology and allocates computing tasks to it. After the recommendation subsystem analysis is completed, the analysis results are notified and saved in the database. The Web system extracts the analysis results and displays them in the foreground after processing. In addition, the background management system mainly includes administrator login, commodity management, member management, transaction management, website information management, and system setting functions.

### 3. Analysis of the Impact of Big Data Processing on E-Commerce

The development of big data technology has brought new development space to e-commerce, and new sales model has become the mainstream direction of e-commerce development. This paper mainly analyzes and tests the impact of big data processing on e-commerce. In the experiment, the

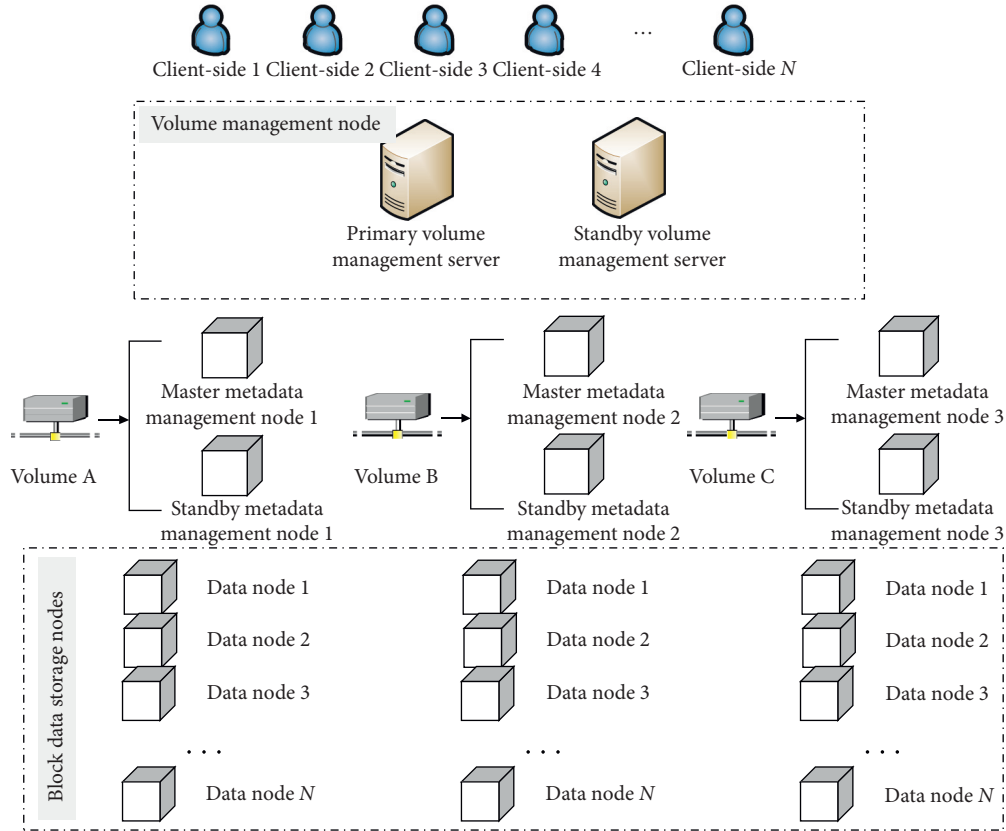


FIGURE 3: The organizational structure of the data storage system.

system operating environment is divided into three parts, which are Hadoop cloud platform, HBase distributed database, and Web system, respectively. After the successful installation of the operating environment, the Web system is deployed to Tomcat server. And the precision marketing, information security assessment, logistics path optimization system is deployed to the Hadoop cluster. After successful deployment, test the system functionality. A total of 8 PCs are used in the experiment. The main work assignment is that one of them acting as a Name Node, acting as a task manager, and the other 7 PCs act as a Data Node, acting as a worker. The software and hardware configuration of the eight PCs are all the same.

**3.1. Achieving Precision Marketing in E-Commerce.** Cloud computing can be used for unified resource management and optimization of data resources connected to the Internet, and massive high-quality data can be used to optimize marketing services. Precision marketing based on cloud computing is to integrate various network resources obtained from multiple channels so as to achieve a wider range of cloud platforms. Combined with the cloud computing, the network marketing resources on the cloud platform are integrated. By constantly improving its comprehensive coverage and corresponding calculation effect, accurate insight into consumer behavior can be realized, and more ideal precision marketing can be achieved. For example, a collaborative filtering recommendation algorithm

can be applied to the recommendation of a cloud platform system, the collaborative filtering algorithm is the essence of the calculation of similarity, and similarity calculation is the core of the target users nearest neighbors selection of data sets, through the combined filtering algorithm and association rules, and can filter out low similarity of users, keep high correlation between the user, and select the optimal nearest neighbor, so as to improve the accuracy of recommendation [19]. The selection of the nearest neighbor is determined by the similarity value between users or projects, and the calculation of the similarity value directly affects the performance of collaborative filtering algorithm. Cosine similarity is the basic method of similarity calculation, and its calculation is as follows:

$$\text{sim}(i, j) = \frac{\sqrt{\sum_{c \in I} R_{i,c}^2 \cdot R_{j,c}^2}}{\sum_{c \in I} R_{i,c} \cdot R_{j,c}}, \quad (1)$$

where  $\text{Sim}(i, j)$  is the similarity between user  $i$  and user  $j$ ,  $i$  is the common rating item set of user  $i$  and user  $j$ , and  $R_{i,c}$ ,  $R_{j,c}$  and  $c$  are, respectively, the rating of user  $i$  and user  $j$  to item  $c$ . In addition, it is necessary to eliminate the difference between the average user scores. Here, users are introduced to correct the average scores of all projects, as shown in

$$\text{sim}(i, j) = \frac{\sqrt{\sum_{c \in I} (R_{i,c}^2 - R_i^2) \cdot \sum_{c \in I} (R_{j,c}^2 - R_j^2)}}{\sum_{c \in I} R_{i,c} \cdot R_{j,c}}, \quad (2)$$

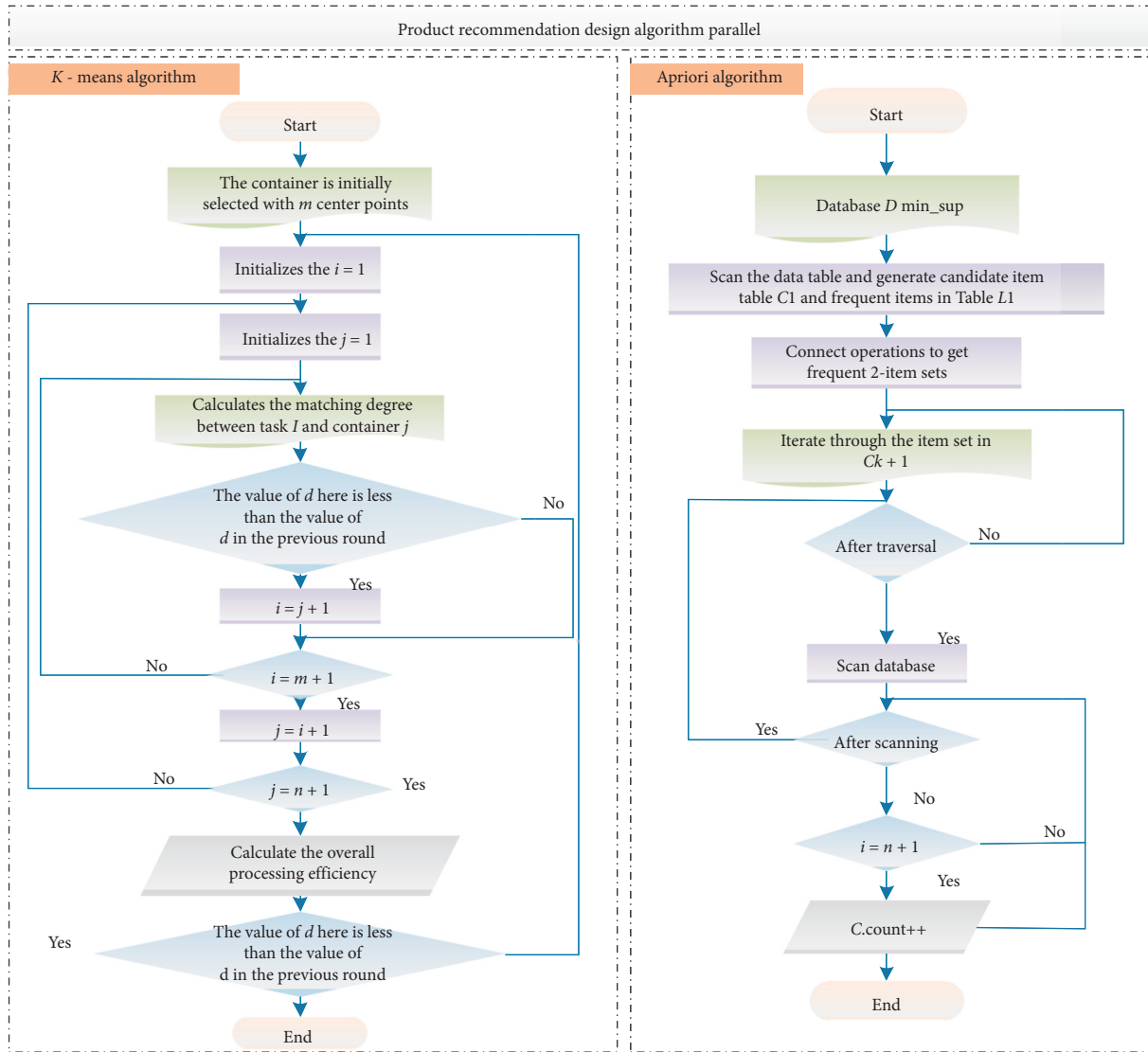


FIGURE 4: K-means and Apriori's parallel algorithm for commodity recommendation.

where  $R_i$  and  $R_j$  are the average ratings of all items by user  $i$  and user  $j$ , respectively. At present, the collaborative filtering algorithm cannot meet the high demand of service consumers and merchants, so formula (3) is needed to solve the problem of data sparse:

$$S = \frac{N_{i,j}}{R_{i,c} \cdot R_{j,c}} \quad (3)$$

The concrete process of the precision marketing refers to Figure 6. Enterprises should first establish the consumer database, and a precise marketing strategy is closely tied to the implementation of the data analysis, especially against the background of e-commerce, consumer data by measurement, calculation, and analysis, because consumer database can help companies develop more accurate marketing strategy, according to the database consumers who can provide analysis of more accurate prices, products, routes, and promotional activities, so as to achieve more accurate communication, advertising of products and

services, one-to-one communication, and other value-added services.

This experiment will test the recommended efficiency of the algorithm, that is, the running speed of the algorithm. Taking the algorithm running time as the evaluation index, the shorter the algorithm running time is, the stronger the algorithm's computing power is, and the higher the recommendation efficiency is. Conversely, the longer the algorithm running time is, the weaker the algorithm's computing power is, and the weaker the recommendation efficiency is. The data set input for the experiment contains data information sets of different sizes, namely, 100,000, 200,000, 300,000, 400,000, and 500,000. The algorithm is run under different Data Node nodes (2 nodes, 3 nodes, 5 nodes, and 7 nodes, respectively), and the running time of the algorithm under different nodes is compared. The experimental results are shown in the figure above. Compared with the traditional mode, the improved algorithm proposed in this paper



runs faster, and the recommendation system has higher recommendation efficiency.

**3.2. Searching for the Optimal Path of E-Commerce Logistics.** The problem of how to choose the route in logistics distribution is a typical NP problem. For the problem of logistics distribution route, many scholars have carried out in-depth design and exploration. In the system simulation method, the logistic trend in reality cannot provide sufficient guarantee for the feasibility of the simulation system. In the precise solution method, the solution will cost too much time, and the efficiency is too low because of the large amount of data. In man-machine interaction method, the executor must have the professional knowledge of logistics distribution, but the distribution personnel are more subjective, in the process of distribution of the choice of the path that has high arbitrariness. All of these limitations limit the use of these methods. The heuristic algorithm is a method derived from inductive reasoning and analysis of past experience, which is used to solve practical problems. The goal is to obtain a solution to the problem under acceptable conditions, which can not only save calculation time, but also meet the needs of problem solving. Because of the advantages of simple implementation and high efficiency, heuristic algorithm has attracted great attention from experts in the industry and has been developed rapidly in recent years.

Real ant colony algorithm is a new population heuristic algorithm, which solves some difficult problems by simulating the process of searching for food in the shortest path of ant colony from the nest to the food source in nature [20–24]. It has positive feedback, parallel computation, and strong robustness. At the beginning of the algorithm, there are  $m$  logisticians and  $n$  customer points, and the first element of  $m$  logisticians is set to the customer point where it is currently located. At this point, the amounts of pheromones on each path are equal. Suppose that  $C$  is a relatively small constant, below, for each logisticians person  $k$ , the path

memory vector records the serial numbers of all customer points, and  $k$  has visited in the order of access. Assume that the current position of logisticians  $k$  is  $i$ , and then the probability that he chooses customer  $j$  as the next object to visit is

$$P_{ij}^k(t) = \begin{cases} \frac{(\alpha + \beta)e^{\tau_{ij}(t) \cdot \eta_{ij}(t)}}{\sum_{s \in jk(i)} [\tau_{ij}(t) \cdot \eta_{ij}(t)]^{\alpha + \beta}}, & \text{if } j \in jk(t), \\ 0, & j \notin jk(i). \end{cases} \quad (4)$$

For example, the ant colony algorithm is used to calculate the logistics distribution routes of some orders of a logistics company. The company has 20 delivery vehicles. The fuel consumption of the trucks is 25 L/100 km, and the driving speed of the trucks is 50 km/h. The trucks need to deliver to 9 customers. The coordinate of the logistics center is (450, 350), and the coordinate of the 9 peripheral distribution points and the cargo demand, as well as the optimized distribution path using the ant colony algorithm model, are shown in Figure 7.

In this distribution scheme, the total length of the distribution path  $T$  is 13.42 km. The distribution length of line 1 is  $1T = 7.34$  km, and the distribution length of line 2 is  $2T = 6.08$  km. Therefore, using ant colony algorithm to solve the problem of logistics distribution can quickly and effectively find the best path of logistics distribution.

**3.3. Providing Information Security Assessment in Cloud Computing Environment.** In the cloud computing environment, the threat and vulnerability of a single asset can be investigated, respectively, so as to obtain the risk level of the entire e-commerce system by integrating the risks of all assets [25–27]. In the case of existing security measures, the probability of occurrence of a threat event and the influence degree of a threat event are expressed as follows:

$$P(T_i) = \frac{1}{P(A_i)} \cdot \left[ 1 - \prod P(V_{ij}) \right] \prod_{k=1}^l Sa_{ik}, E(T_i) = \frac{1}{P(A_i)} \cdot \left[ 1 - \prod P(V_{ij}) \right] \prod_{k=1}^l Sa_{ik} \cdot \sum_{i=1}^s (w_t d_{it}) \prod_{k=1}^l Sb_{ik}. \quad (5)$$

The risk of a single asset is determined by the value of the asset and the impact of the threat to the asset. According to

the above calculation, we can get the risk value of this asset as follows:

$$R(A_i) = A_i \sum_{i=1}^n \left\{ \frac{1}{P(A_i)} \cdot \left[ 1 - \prod P(V_{ij}) \right] \prod_{k=1}^l Sa_{ik} \cdot \sum_{i=1}^s (w_t d_{it}) \prod_{k=1}^l Sb_{ik} \right\}. \quad (6)$$

In this paper, this part is described as risk assessment, and the corresponding risk assessment theoretical model is designed. This model first identifies the threat and vulnerability of a single information asset. Threat is the external

cause of risk, which may come from many factors such as system, natural environment, or human. In the e-commerce system based on cloud computing, it is necessary to fully consider the evaluation of the cloud service mode of the

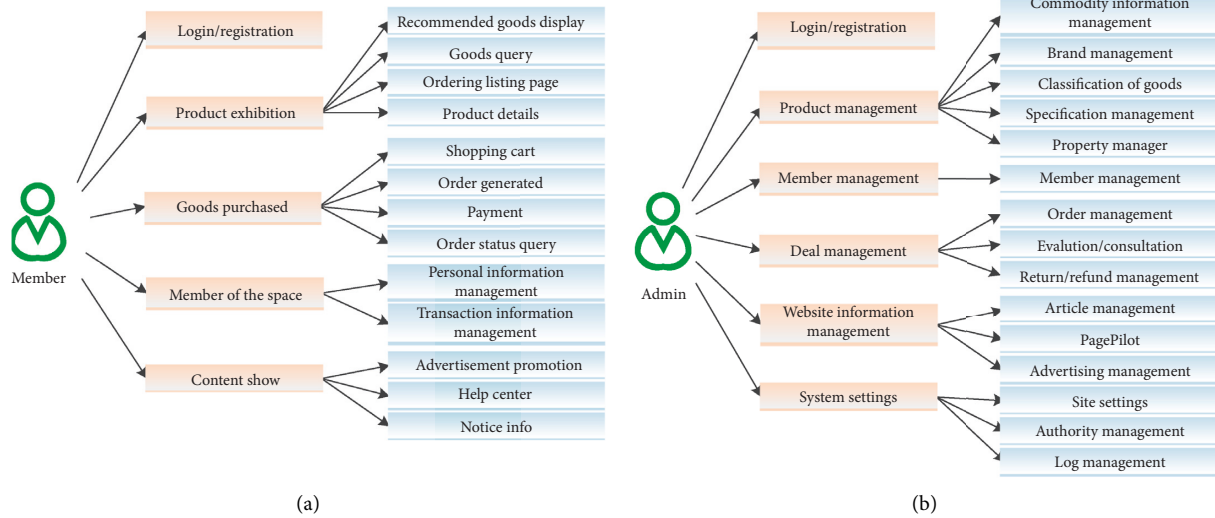


FIGURE 5: Front-stage shopping system and background management system function modules. (a) Front desk shopping system. (b) Backstage management system.

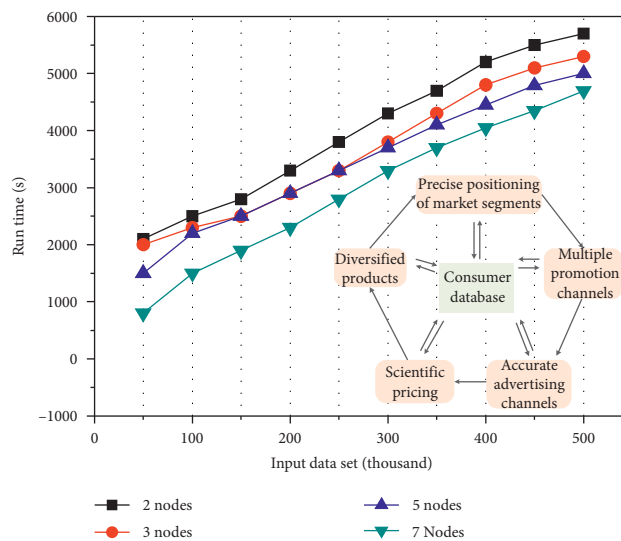


FIGURE 6: The specific process of precision marketing and the test results of recommendation efficiency.

system and the security scale of the service provider. Vulnerability is the internal cause of risk, cloud platform corresponding infrastructure is different, and from the scale, whether to allow a privileged user data access, data storage location, isolation, and restoration measures, e-commerce platform can review sex, and long-term survivability of vulnerability identification can lead to security incidents. Its model is shown in Figure 8.

According to the collected information on assets, threats, and vulnerability, the questionnaire survey results were combined with expert ratings to assign the corresponding value and numerical calculation of each risk element, and the risk information of a single asset was obtained in Figure 8. According to the risk grade classification standard, the risk value is divided into grades: low risk level, medium risk level,

high-risk level, and extra high-risk level. It seems to be that the system risks are in the medium level; among them, the hardware and the risk of e-commerce platform are the worst, the electronic commerce system of physical security, system access, and data backup control exist, and the fragility of the problems to be settled urgently in addition to cloud data security and security management configuration reflects the risk that cannot be ignored. In view of this series of problems, the corresponding improvement strategies are proposed, such as improving the physical environment of the computer room, configuring access control rules with a higher security level, negotiating with cloud service providers for a more optimized cloud data storage strategy, strengthening the professional skills of technical personnel, and improving the assignment of responsibilities.

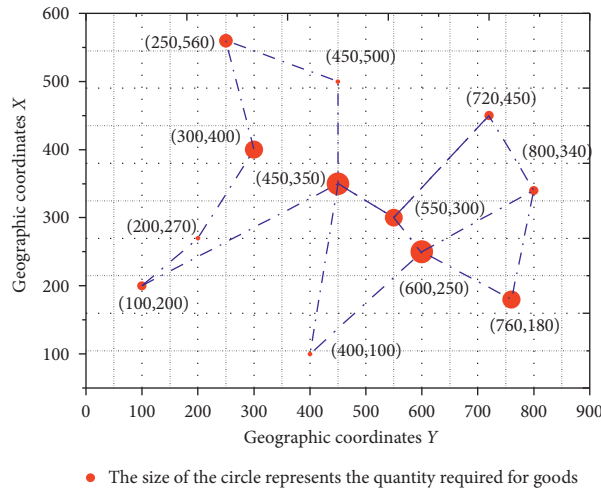


FIGURE 7: Distribution route and demand for goods at each distribution point.

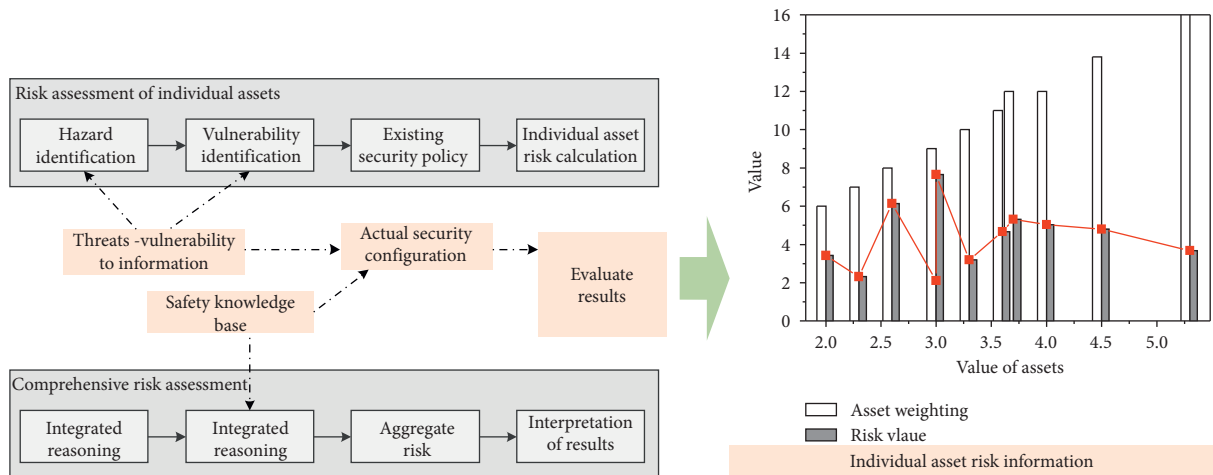


FIGURE 8: Distribution route and demand for goods at each distribution point.

To conduct an effective investigation on all aspects of information assets of the e-commerce system and evaluate the current security risk level of the system through numerical method, on the one hand, it provides a suitable control strategy for the system’s own security performance, and, on the other hand, it also has a certain reference significance for how to choose cloud service providers.

#### 4. Conclusion

This article starts with the analysis of the existing electronic commerce system, summarizes its characteristics, and analyzes and solves its existing problems. Firstly, the relational database My SQL and distributed database HBase are used to store unstructured data such as pictures. Secondly, considering the characteristics of large amount of data in e-commerce system and complex calculation of data mining algorithm, this paper uses MapReduce to realize the parallelization of data mining algorithm and builds a Hadoop based commodity recommendation subsystem on this basis. Finally, JavaEE technology is used to design a set of Web

mall system with complete functions, which is built on Hadoop and HBase. This paper realizes the communication among Web system, HBase image storage system, and commodity recommendation system through Webservice telecommunication technology and organizes them into an organic whole. Finally, based on the impact of cloud computing, mobile e-commerce is analyzed, including relevant theories, service mode, architecture, core technology, and the application in e-commerce and other contents, to comprehensively grasp the performance advantages of cloud computing. Combined with the advantages of cloud computing for data storage and processing, it is used for data analysis and processing in the mobile e-commerce system to avoid the disadvantages of the traditional e-commerce where large-scale data cannot be processed in time. E-commerce enterprises need to seize the development opportunity of the era of big data and collect and analyze massive information with the help of big data processing technology. The value behind the data is mined, the data information is efficiently utilized, the development direction of the enterprise is adjusted through the analysis of the results, the precision

marketing work is done well, the service is provided with better quality for consumers, the core competitiveness of e-commerce enterprises is improved, and the development of e-commerce is promoted.

## Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

## Conflicts of Interest

The authors declare that they have no conflicts of interest.

## References

- [1] P. Feng, "Big data analysis of e-commerce based on the internet of things," in *Proceedings of the 2019 International Conference on Intelligent Transportation, Big Data & Smart City (ICITBS)*, pp. 345–347, Changsha, China, January 2019.
- [2] I. A. T. Hashem, I. Yaqoob, N. B. Anuar, S. Mokhtar, A. Gani, and S. Ullah Khan, "The rise of "big data" on cloud computing: review and open research issues," *Information Systems*, vol. 47, pp. 98–115, 2015.
- [3] C. Stergiou and K. E. Psannis, "Efficient and secure big data delivery in cloud computing," *Multimedia Tools and Applications*, vol. 76, no. 21, pp. 22803–22822, 2017.
- [4] H. Chen, R. H. L. Chiang, and V. C. Storey, "Business intelligence and analytics: from big data to big impact," *MIS Quarterly*, vol. 36, no. 4, pp. 1165–1188, 2012.
- [5] S. Akter and S. F. Wamba, "Big data analytics in e-commerce: a systematic review and agenda for future research," *Electronic Markets*, vol. 26, no. 2, pp. 173–194, 2016.
- [6] O. Sohaib, M. Naderpour, W. Hussain, and L. Martinez, "Cloud computing model selection for e-commerce enterprises using a new 2-tuple fuzzy linguistic decision-making method," *Computers & Industrial Engineering*, vol. 132, pp. 47–58, 2019.
- [7] H. El-Sayed, S. Sankar, M. Prasad et al., "Edge of things: the big picture on the integration of edge, IoT and the cloud in a distributed computing environment," *IEEE Access*, vol. 6, pp. 1706–1717, 2017.
- [8] P.-J. Wu and K.-C. Lin, "Unstructured big data analytics for retrieving e-commerce logistics knowledge," *Telematics and Informatics*, vol. 35, no. 1, pp. 237–244, 2018.
- [9] A. M. Talib and F. O. Alomary, "Cloud computing based e-commerce as a service model: impacts and recommendations," in *Proceedings of the International Conference on Internet of Things and Cloud Computing*, pp. 1–7, Cambridge, UK, March 2016.
- [10] B. M. Balachandran and S. Prasad, "Challenges and benefits of deploying big data analytics in the cloud for business intelligence," *Procedia Computer Science*, vol. 112, pp. 1112–1122, 2017.
- [11] Z. Song, Y. Sun, J. Wan, L. Huang, and J. Zhu, "Smart e-commerce systems: current status and research challenges," *Electronic Markets*, vol. 29, no. 2, pp. 221–238, 2019.
- [12] A. Behl, P. Dutta, S. Lessmann, Y. K. Dwivedi, and S. Kar, "A conceptual framework for the adoption of big data analytics by e-commerce startups: a case-based approach," *Information Systems and E-Business Management*, vol. 17, no. 2–4, pp. 285–318, 2019.
- [13] S. Juncai and Q. Shao, "Based on cloud computing e-commerce models and its security," *International Journal of E-Education, E-Business, E-Management and E-Learning*, vol. 1, no. 2, p. 175, 2011.
- [14] T. Oliveira, M. Thomas, and M. Espadanal, "Assessing the determinants of cloud computing adoption: an analysis of the manufacturing and services sectors," *Information & Management*, vol. 51, no. 5, pp. 497–510, 2014.
- [15] A. Galletta, L. Carnevale, A. Celesti, M. Fazio, and M. Villari, "A cloud-based system for improving retention marketing loyalty programs in industry 4.0: a study on big data storage implications," *IEEE Access*, vol. 6, pp. 5485–5492, 2017.
- [16] T.-P. Liang and Y.-H. Liu, "Research landscape of business intelligence and big data analytics: a bibliometrics study," *Expert Systems with Applications*, vol. 111, pp. 2–10, 2018.
- [17] E. Yadegaridehkordi, M. Nilashi, L. Shuib et al., "The impact of big data on firm performance in hotel industry," *Electronic Commerce Research and Applications*, vol. 40, Article ID 100921, 2020.
- [18] I. Lee, "Big data: dimensions, evolution, impacts, and challenges," *Business Horizons*, vol. 60, no. 3, pp. 293–303, 2017.
- [19] L. T. Wright, R. Robin, M. Stone et al., "Adoption of big data technology for innovation in B2B marketing," *Journal of Business-to-Business Marketing*, vol. 26, no. 3-4, pp. 281–293, 2019.
- [20] A. Bhimani and L. Willcocks, "Digitisation, "big data" and the transformation of accounting information," *Accounting and Business Research*, vol. 44, no. 4, pp. 469–490, 2014.
- [21] S. Shakya, "An efficient security framework for data migration in a cloud computing environment," *Journal of Artificial Intelligence and Capsule Networks*, vol. 1, no. 1, pp. 45–53, 2019.
- [22] J. Zhang and G. Qu, "Physical unclonable function-based key sharing via machine learning for IoT security," *IEEE Transactions on Industrial Electronics*, vol. 67, no. 8, pp. 7025–7033, 2019.
- [23] X. Ma, K. Zhang, L. Zhang et al., "Data-driven niching differential evolution with adaptive parameters control for history matching and uncertainty quantification," *SPE Journal*, vol. 26, no. 2, pp. 993–1010, 2021.
- [24] W. Wang, Z. Gong, J. Ren, F. Xia, Z. Lv, and W. Wei, "Venue topic model-enhanced joint graph modelling for citation recommendation in scholarly big data," *ACM Transactions on Asian and Low-Resource Language Information Processing (TALLIP)*, vol. 20, no. 1, pp. 1–15, 2020.
- [25] J. Yang, M. Xi, B. Jiang, and H. Song, "Robust six degrees of freedom estimation for IIoT based on multibranch network," *IEEE Transactions on Industrial Informatics*, vol. 17, no. 4, pp. 2767–2775, 2021.
- [26] S. Qi, Y. Lu, W. Wei, and X. Chen, "Efficient data access control with fine-grained data protection in cloud-assisted IIoT," *IEEE Internet of Things Journal*, vol. 8, no. 4, pp. 2886–2899, 2021.
- [27] J. Qian, X. Cheng, B. Yang et al., "Vision-based contactless pose estimation for human thermal discomfort," *Atmosphere*, vol. 11, no. 4, p. 376, 2020.

Copyright © 2021 Rongrui Yu et al. This is an open access article distributed under the Creative Commons Attribution License (the “License”), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. Notwithstanding the ProQuest Terms and Conditions, you may use this content in accordance with the terms of the License. <https://creativecommons.org/licenses/by/4.0/>