Research on Big Data Mining and Solutions Based on Computer Cloud Computing

Jianshi Zhang1,*

1Changchun Polytechnic Information Institute, China, 130033

*Corresponding author e-mail: 23746978@qq.com

Abstract. The application of computer cloud computing has profoundly changed the way the world works and brought people an everlasting experience. In order to adapt to the development direction of information technology, we need to deeply understand the current situation of big data application and study the innovative application of data mining technology.

Keywords: Computer, Cloud Computing, Big Data Mining, Solution

1. Introduction
In the era of information explosion and convenient life, cloud computing and big data have profoundly changed our lives, changed the development mode of market economy in our country and the defects of traditional data mining [1]. It combined with cloud computing, to provide users with rapid and efficient data. Big data mainly includes a large number of data and data storage and analysis structure, massive data is mainly interactive data, it is structured or unstructured data. We can judge what happened in the past and predict what will happen in the future. Most of the data is unstructured data, which integrates the information of text, image and audio, so it is necessary to design a more reliable algorithm to read the data to extract the information contained therein. In the traditional data, there is a wide range of structured data, which is more convenient to read and easy to observe information, but in the era of big data, the reading of data is very dependent on powerful algorithms [2]. This is their main difference. But this kind of big data mining starts late in our country, the development is some imperfect, so the author first briefly introduces the big data mining for everyone, and puts forward some suggestions for its development. Hope to help our country to do this work better.

2. Several basic concepts of big data mining technology

2.1. Background conditions
With the continuous development of science and technology, the data information generated and the amount of value data needed by users have increased, data mining has been significantly affected by
the times, the traditional mining methods are single, the level is still shallow, relatively speaking, big data mining should meet the needs of the times. Nowadays, big data mining still needs to rely on advanced technology, cloud computing sustainable development [3].

2.2. Application object
The big data mining processing object is different from the traditional data mining, in which the former manages the data information of many kinds of information systems, the latter only needs to manage the data in the specific range and the specific information system [4]. The former has the characteristics of rich information, wide data sources and various types of information, but the accuracy of the data is lower. The latter is different from the former, that is, the data source is less, the acquisition method is more passive, the acquisition range is limited, and the data accuracy is higher.

2.3. Mining level
The big data mining mainly uses the network superiority to complete the data processing and the analysis, the mining scope is wider, the analysis is deeper, relatively speaking, the traditional data mining only carries on the analysis to the limited scope data information, therefore, the data type is relatively single, the data scope is more limited. In addition, the latter process, analysis of data information is vulnerable to functional constraints [5]. It can be seen that big data mining, with the help of advanced information technology such as cloud computing, can clearly analyze the goal, and at the same time, it will optimize the effect of data analysis.

3. Main architecture of big data mining system
The architecture of big data mining system shows the basic process of data mining, as shown in figure 1.

![Figure 1. Basic process of data mining](image)

3.1. Data preprocessing
The big data mining mode is behind, the data is ahead, and the deterministic model can constantly change under the influence of the data. Among them, preprocessing is mainly based on integrated technology, including data flow processing technology, remote acquisition technology, traditional preprocessing technology, multi-modal entity identification technology, etc. the integrated function of preprocessing can be effectively realized under the action of integrated technology. In the process of data preprocessing, there are light causality, heavy data phenomenon, light data integrity and heavy real-time phenomenon, which is also the main reason for the inaccuracy of big data mining results.
3.2. Data storage
When traditional data mining stores data, the system involved mainly includes file system, data warehouse and operating system. The data storage is passive and the storage mode is single. Big data mining is stored on the basis of traditional storage system and adds to distributed storage function. There are three main types of data stored, the first is unstructured data, the second is storage structure data, and the third is semi-structured data. There are two main storage methods, the first is mixed storage, second is column storage [6]. When the uncertain data is stored effectively, the uncertain data management technology and management system can be technology and management system, so as to complete the uncertain data storage activities.

3.3. Data analysis
The traditional data mining only has the processing flexibility and the analysis accuracy in the small dimension data analysis, for the multidimensional data analysis only stays in the shallow level. Big data mining for data computing, analysis, mainly with the help of fusion computing mode to process it, for a large number of dimensions, dynamic data information analysis, big data mining can be based on the integration of traditional data mining deficiencies, targeted to improve the original analysis ability [7]. At the same time, distributed file system and traditional mining algorithm are effectively combined and improved to optimize the quality of data analysis. In addition, interactive analysis can be used to analyze its content, and improve the utilization rate of mining language, fully showing the expansibility.

3.4. Data display
Big data mining is superior to traditional data mining in data display. The presentation of mining results is mainly image and animation. At the same time, users can analyze the validity of data. Traditional data mining for dynamic data, multidimensional data display, the visual effect is poor, and causality cannot be directly presented. Therefore, in order to improve the visual analysis effect, we should also study the interactive mining, memory in situ, system visualization and other technologies in depth to ensure that node communication, visualization technology aging and other problems are effectively solved.

4. Examples of big data mining solutions based on cloud computing
A library managed by cloud computing big data mining technology is shown in figure 2.
4.1. Analysis of library network data flow

Library is a place where data flow is very large. In addition to the control of book information, classification, location, storage and quantity of data, it is necessary to make accurate statistics and analysis who have borrowed books, the frequency of books lending, the length of books lending, the flow of library people, etc. In order to provide readers with humanized service and targeted information push, so as to make the library service more considerate. It is very important to analyze the video network data produced by users scientifically and effectively. The premise of effective analysis of network data is to save and mine its data stream. Network data stream refers to the research on the theme and communication of the Internet through Internet crawler. For emergencies, people often analyze data through historical data, but sometimes this analysis is not scientific, because the changing trend of a lot of information is not transmitted according to the traditional information flow. Therefore, it needs to be optimized by the powerful computing power of the computer to capture it better. Figure 3 show the big data mining book management information statistics of cloud computing.
4.2. Collect the reader's short text information through the network and solve it through data mining technology

In the age of the Internet, the management of libraries is constantly introducing new technologies to improve and optimize, and people are willing to share lessons, tips and other experiences on the Internet while trying to find an excellent life experience. The short text is a widely concerned form of information, and a lot of information can be obtained effectively through simple search, but many times the text also contains a lot of garbage information, so some of the information generated by users is not necessary to be collected. We need to analyze and excavate the text. The information which is of low importance to the management and operation of the library is excluded, and the information which can enhance the management level of the library and help the readers and the library to understand each other is retained, so that the information can be collected and pushed efficiently.

4.3. Exploring the visual data technology of library related information

The application of visual data in library is still in its infancy, but information visualization data technology is very many in other industries. Many media annotate keywords on images, extract image features, convert images into text, and transform some text into image description. This is very important for data mining and analysis.

5. Conclusion

The emergence and existence of big data has brought us a variety of conveniences, its value and future prospects are worth looking forward to, and at the same time, when big data is developed and perfected, Traditional data management methods or some ways of thinking are changing to varying degrees, both opportunities and challenges. Even though there are still some problems that need to be solved in the development of big data mining, it is a very effective application compared with traditional data mining, which can meet the higher needs of our modern life and provide us with more convenience. But technology is like a double-edged sword.
References

[1] Deng Zhonghua, Liu Weiwei, Lu Yingjun. Big Data Mining Connotation and Solution Research [J]. Based on Cloud Computing Information Theory and Practice, 2015(07):103-108.

[2] Shang Ting. Big Data Mining and Solutions [J]. Based on Cloud Computing China New Communications, 2018.

[3] Li Kaimin, Zeng Zili. Big Data Mining Connotation and Solution Research [J]. Based on Cloud Computing Computer Fans, 2018, No.91(04): 39.

[4] Zhang Zhigang. A Study on the Mining Algorithm of Frequent Items Set in Cloud Computing Environment [D]. Nanjing normal University, 2014.

[5] Liu Zhen. Big Data Mining Connotation and Solution Research [J]. Based on Cloud Computing Science and Technology Wind, 2017000(019):39-39.

[6] Wang Shuliang, Ding Gangyi, Zhong Ming. Considerations on Spatial Data Mining under Big Data [J]. Journal of the Chinese Academy of Electronic Sciences, 2013, 008(001):8-17.

[7] Rao Zheng Chan, Pu Tianyin. Big Data Mining Connotation and Solutions [J]. Under Cloud Computing Electronic Technology and Software Engineering, 2018, 000(013):154-155.