Construction of a Big Data Intelligent Search Engine Fused with Adaptive Multi-Objective Differential Evolution Algorithm

DanHuang[1],*Dawei Zhang[2]

[1]Beihai University of Art and Design, Beihai, Guangxi, China, 536000
[2]Telecom Department Of Beihai Vocational College

Abstract

Data visualization management is a major innovation based on cloud computing big data processing technology. Under the premise of cloud computing technology, big data visualization can fully dig and collect hidden data information and use intuitive and vivid charts for performance. Big data visualization technology is a data mining technology based on cloud computing technology. It can accurately identify valuable information from a large amount of complex and changeable data and provide it to users in the form of services. It is precisely because of the strong guarantee of cloud computing that the level of data mining can be improved. And after big data digs up important information from a large amount of data, cloud computing will immediately aggregate the fragmented information, discover the laws and then analyze the overall trend of market economic development more accurately and scientifically. For visual processing technology, the method of drawing trend graphs can be used to comprehensively display the basis of visual processing, which can make the data results have intuitive and vivid characteristics and lay a good foundation for search engines to avoid decision-making errors. At this stage, many search engines in my country have their own business content and business fields. Therefore, a reasonable development of big data mining strategies under the cloud computing model will help the efficient collection of search engine data.

Keywords: Adaptive, Differential, Big Data;

1. Introduction

Using online analysis technology, in many cases only shallow data information can be obtained, but the internal connection of the data cannot be grasped. In the cloud computing mode, the use of big data mining technology can understand the essence of the data and can show the connections between different data, use patterns and concepts, etc., to fully reflect them. Currently, big data mining methods are mainly parallel, which has significant advantages in processing massive data. The data area of the original serial data processing is small, it takes a lot of time and the work efficiency is low, but it uses distributed mining technology, uses distributed systems and comprehensively uses diversified methods, such as: splitting and clustering, etc. , Reduce the time of data calculation and improve the accuracy of data calculation results. And the big data mining technology under the cloud computing mode can give full play to its parallel advantages. Compared with other serial methods, parallel mining can use computers to cluster the work of the distributed supply system. After the completion of the processing, multiple computers are used to carry out the work at the same time, which can not only greatly improve the processing level, but also

significantly reduce the cost of data processing[1].

2. Differential evolution algorithm analysis

The setting of the control parameters of the differential evolution algorithm has a great influence on the search performance and convergence speed of the algorithm and the unreasonable parameter setting may cause the algorithm to converge prematurely or stagnate. Traditional DE is very sensitive to the settings of the scaling factor F and the hybridization probability CR[2]. For different optimization problems, it is often necessary to set, test and adjust the parameters from the problem itself. Manual parameter adjustment is time-consuming, laborious and inefficient. How to set the parameters so that the algorithm can not only avoid premature but also converge quickly and keep the advantages of the algorithm simple and universal, has always been the core problem of DE research. Based on the analysis of the influence of control parameters on the performance of the algorithm, this paper proposes a white adaptive parameter selection DE. The population size NP takes a fixed value within a certain range and the individual control parameters are based on a fixed value. The sum hybridization probability CR is randomly selected within a certain range and the probability of the algorithm converging to the global optimum is improved by increasing the selection combination of parameters. The new algorithm proposed in this paper only makes improvements in parameter selection. There is no complicated adjustment strategy and hybrid technology. The algorithm flow is consistent with the standard DE and there is no need to add additional computational overhead and data storage space. It keeps the DE algorithm simple and versatile. The advantages of the algorithm are conducive to the further promotion and application of the algorithm. The intelligent search system is in the figure below.
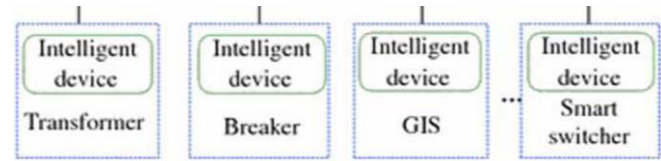

Figure1. Intelligent search system.

3. Overview of big data processing technology

3.1. Big data collection technology

According to different collection methods, big data collection can be divided into multiple types, such as centralized collection and distributed collection. Each method has its own advantages and disadvantages. Below we can use the advantages of these methods to analyze. First of all, centralized collection can master and control various data, but distributed collection has certain flexibility. When collecting big data, if you want to collect both internal data and information between the enterprise and the enterprise, we can set up different servers within the enterprise to store the data shared by the enterprise. The use of distributed computing methods and the joint use of various collection methods will help improve the level of data collection[3]. Therefore, a centralized collection method can be used within an enterprise, a distributed collection method can be used between enterprises and an enterprise and a distributed collection method can be selected for the organization between central servers. In addition, combining different structure types, big data can be divided into multiple types of data, such as structured, unstructured and semi-structured data. When collecting data, you must first systematically analyze the types of data. According to different types, give full play to the advantages of cloud computing, such as fault tolerance and expansion, so as to achieve the purpose of information isomorphism and effectively complete data docking[4]. The intelligent search storage system is in the figure below.
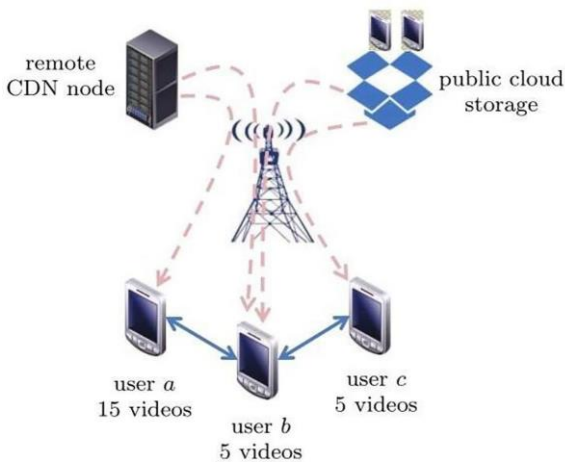
Figure2. Intelligent search storage system.

3.2. Big data storage technology

The original data storage is often a single-node warehouse and its capacity space is quite small. It has lost all carrying capacity for a large amount of data in the new era. Although it has view and indexing capabilities, it is still unable to truly be due to space constraints. Meet the actual needs of the development of modern society. Especially in the new era, the slow running of data in the past can no longer meet the requirements of modern society for data analysis and processing speed. As far as cloud computing is concerned, most of its storage methods are columnar, which is conducive to accurately and scientifically distinguishing the attributes of data and can fundamentally realize the storage and classification of data according to attributes. At the same time, the required data can be extracted in the first time by using the search attribute, which significantly improves the data processing level and makes the system function more and more prominent. And according to different attribute classifications, there are other features, that is, the ability to scientifically arrange data attributes according to the degree of similarity and when querying some uncertain attributes, you can master the similar attributes so as to lay a solid foundation for subsequent query data. Significant data compression effect, reducing problems caused by incorrect queries[5]. The intelligent search application system is in the figure below.
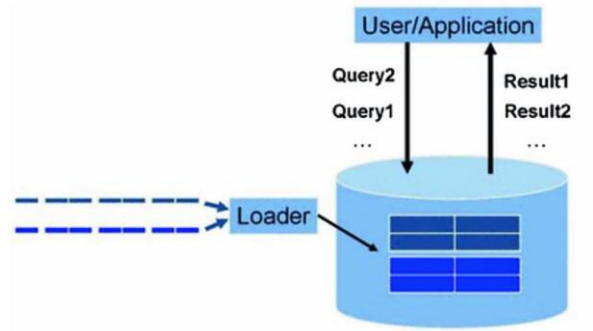

Figure3. Intelligent search application system.

3.3. Big data online analysis technology

For big data systems, online technology is the core content, cumbersome and complex data analysis links, the focus is on decision-making analysis and the actual results are provided to users. Generally, for the use of online analysis methods, the entry point should be to analyze comprehensive data, build a multi-dimensional model and obtain the overall final analysis result, which provides a powerful reference for relevant personnel to make correct decisions. And the salient feature of online analysis and processing is to analyze data and use warehouse and online analysis technology together, so that in addition to calculating massive data, it can also analyze data comprehensively. The intelligent search user system is in the figure below.
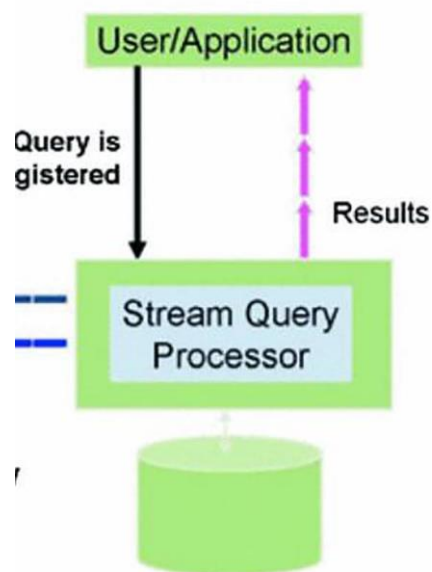

Figure4. Intelligent search user system.

4. Big data search engine construction model

4.1. Set the main keyword according to the product name

The main keywords are mainly the name of the product and the category to which the product belongs. Some companies call the main keywords the category keywords and category keywords. The main keywords are the most likely to be searched by users. They are generally called high-frequency words, which refer to words that users frequently search. These keywords reflect the user's search habits and search characteristics. For example, if a user wants to buy a cotton garment, some people search for cotton clothes, some people search for overcoats, some people search for cotton jackets and some people search for cotton jackets. Then these are the name of the product or the category or category the product belongs to, so they are all main keywords[6]. The setting of the main keywords also changes constantly according to the business situation of the enterprise. For example, a company mainly promotes short-sleeved shirts in the summer and cotton-padded clothes in the winter. If the main keywords set are different, they attract different customers. Companies set main keywords based on product names and product categories, which can allow more users to search for corporate products, attracting a lot of traffic. With high traffic and large crowds, the probability of purchase will increase and the sales of the company will increase. . Therefore, combining these frequently searched words, designing a keyword suitable for the enterprise and performing bidding ranking is an important method for setting the main keyword. The intelligent search task system is in the figure below.
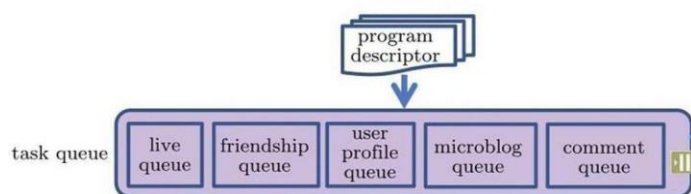

Figure5. Intelligent search task system.

4.2. Set attribute keywords according to product characteristics and attributes

Users' search habits are different. Some users are accustomed to searching for product functions, features and attributes. Therefore, in addition to setting primary keywords, companies must also set attribute keywords, or use a combination of primary keywords and attribute keywords, so that the company's products can be accurate Searched by users to achieve the purpose of corporate sales. Attribute keywords are the brand, functions, features, effects and attributes of a product. For example, you just mentioned the cotton clothing purchased by the user. The user may want to buy cotton clothing with black color, short style and Nike brand. These are all attributes of the product and belong to attribute keywords. If the company does not set suitable attribute keywords, it may miss many target customers. Therefore, the combination of main keywords and attribute keywords can accurately locate the needs of users. The smaller the product, the more accurate the attribute keywords and the more users will be trust. Therefore, companies can set two types of hot-searched attribute keywords and precise attribute keywords and use different combinations to optimize the setting of keywords to meet the needs of different users. The intelligent search scheduler system is in the figure below.
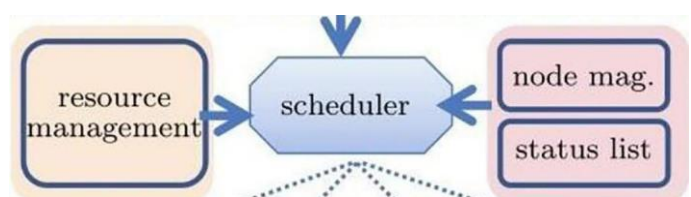

Figure6. Intelligent search scheduler system.

4.3. Set marketing keywords and intention keywords according to the combination of user psychology and product characteristics

Marketing keywords refer to methods of promotion or marketing. For example, users will search for keywords such as special prices, free shipping and buy one get one free. Such promotional and marketing words belong to marketing keywords. Different types of companies have different

marketing keywords. For example, for consumables such as facial masks, users may be attracted to buy one get one free, because they consume a lot of money every month and can buy more products, which are repetitive purchases; for example; For the cotton clothing just exemplified, users may only buy one piece of cotton clothing in a winter and will not buy cotton clothing with repeated patterns, so marketing keywords such as buy one get one free will not attract users. Therefore, companies should set reasonable marketing keywords based on the characteristics of their products to attract users to buy.

5. Conclusion

The latest generation of search engine platforms can analyze big data and large amounts of smaller files at the same time and convert them in real time. Compared with the previous data mining technology, in the process of processing distributed parallel data, big data mining technology mainly adopts the method of combining different types of calculation modes of computing mobile data. When analyzing data groups with a large data cube and more complex dimensional attributes , Its data management system can play the role of delayed query. Generally speaking, the big data mining technology under the cloud computing mode can effectively process a large amount of data with different structures in a short time.

References

[1] Jing Chen, Zhu Quanmin, Hu Manfeng, Guo Liuxiao, Pritesh narayan. improved gradient descent algorithms for time-delay rational state-space systems: intelligent search method and momentum method[J]. Nonlinear Dynamics,2020(prepublish).

[2] Nonlinear Research; Studies in the area of nonlinear research reported from Jiangnan University (Improved Gradient Descent Algorithms for Time-delay Rational State-space Systems: Intelligent Search Method and Momentum Method)[J]. Journal of Mathematics,2020.

[3] Chun Yan,Zhaochuang Ou,Wei Liu,Qiang Xu. Research on UBI Auto insurance pricing model based on adaptive SAPSO to optimize the fuzzy controller[J]. International Journal of Fuzzy Systems,2020,22(2).

[4] Searchspring; Searchspring and nextopia announce merger to deliver intelligent search and merchandising platform for ecommerce teams[J]. Journal of Technology & Science,2020.

[5] Soumesh Chatterjee,Pronob K. Ghosh,Biman K. Saha Roy. PMU-based power system component monitoring scheme satisfying complete observability with multicriteria decision support[J]. International Transactions on Electrical Energy Systems,2020,30(2).

[6] Qian Fan,Xiaolin Meng,Chengquan Xu,Jiayong Yu. Solution method for ill-conditioned problems based on a new improved fruit fly optimization algorithm[J]. Journal of Applied Geodesy,2020,14(1).