Analysis of High Performance Computing System for Agricultural Big Data

Mingxing Liu

School of Information Engineering, Zhengzhou University of Industrial Technology, Xinzheng 451150, Henan, China

Abstract. With the development of the information age, people pay more and more attention to the big data industry. In many industries, big data will be used for corresponding processing and analysis. However, at present, in the agricultural industry, there are still few researches on agricultural data processing by using big data. Therefore, in the agricultural data processing, the use of high-performance computing system to process its massive data can play a certain role in promoting the transformation and upgrading of agriculture, and it can play a better practical significance and use value in promoting agricultural development.

Keywords: agriculture; Big data; High performance; Computing system.

1. Introduction

China is a populous country. In order to solve the food security problem of the people, China's investment in agriculture is increasing continuously. With the rapid development of information technology, information management has become an important part of China's agricultural economy. It is also a necessary link to transform traditional agriculture into modern agriculture, and information management has also played an important role in China's agricultural development. In the development of agriculture in China, there are some problems, such as the shortage of resources, the deepening of pollution, the difficulty of docking in production, and the imbalance between the domestic market and the international market. Information technology has played an important role in solving these problems and promoted the agricultural construction in China. Agricultural system has highly complex characteristics, which will not only produce a large amount of original data, but also need to use these data for in-depth analysis, and find potential value in the analysis. Only by deeply integrating big data technology with agriculture can information technology play a huge role in agriculture and enhance agricultural value.

2. Brief introduction of agricultural big data

Agricultural big data is the data in the agricultural industry and agricultural economy, which is characterized by high accuracy, wide range of data, and large amount of available value. By using big data technology, agriculture can get scientific data management in its later development, and at the same time, its data can be deeply analyzed and reasonably utilized. The realization of agricultural networking can make modern agriculture more intelligent and accurate, and at the same time promote the rapid development of agricultural informatization.

2.1 Current situation of agricultural big data

With the continuous rapid development of the Internet, the speed of information dissemination has been greatly improved, which also brings many possibilities. China's relevant departments continue to support agriculture, so that the speed of China's economic growth continues to increase. With the help of the power generated by capital, the share of big data in the agricultural market can be quickly occupied, making sufficient preparations for the subsequent development. However, according to the current development of China's agriculture, the efficiency of its production needs to be kept low, and there is still much room for improvement, and agricultural big data technology can also play a key role in this respect.

In foreign countries, among its agricultural industries, the United States focuses on the development of its large-scale agriculture, the overall layout of crops is relatively stable, and there are large-scale farms, and the division of labor in each link is also very clear, with a relatively complete industrial chain; The Netherlands, on the other hand, focuses on the fine management of its agriculture and builds a greenhouse factory suitable for plant growth; In the development of agricultural industry in Israel, water-saving agriculture is emphasized. With the continuous development of Internet of Things technology, many foreign agricultural planting technologies adopt mechanical production, and at the same time, it has a relatively high degree of automation, which saves time and effort. In today's era of big data, the visualization of various data in agricultural industry has gradually become a new development trend.

2.2 Prospects of Agricultural Big Data

The data is not embodied in the form of physical objects, and it is difficult to feel intuitively. Agricultural big data is a collection of many data collected in the agricultural industry, and it needs to be deeply analyzed, followed by induction and integration, thus forming the final agricultural big database, which plays an obvious role in the development of the agricultural industry and contains a lot of information. Firstly, it includes a lot of data about the natural environment needed by seeds, the amount of chemical fertilizers and pesticides used in the process of plant growth, variety research and development, and planting. Secondly, it contains a large amount of real-time data in the processing and marketing of agricultural products. In the whole industrial chain of agriculture, big data is involved, and its scale is relatively large and its content is complex. Agriculture has its own characteristics in its position and time, so it is necessary to analyze all aspects of agriculture in different time. Judging from the big data in the current agricultural industry, its development direction can focus on the direction of refinement. At the same time, the big data generated in agricultural management can be analyzed and modeled in real time, so as to realize intelligent management. At the same time, agricultural resources can be shared and professionals can be trained.

3. High performance computing system

With the continuous increase of computing demand, all countries have made development plans for high-performance computers, and continuously optimized the performance of high-performance computers, and the energy consumption of their computer systems has also been continuously improved. By using high-performance computing system, the experiments that are difficult to complete at present can be carried out, which can not only save the huge cost of real experiments, but also avoid the negative impact on the environment. With the development of society, the competition brought by globalization is becoming more and more fierce. High-performance computing system has been applied to all aspects of life, and its performance is outstanding. Only by fully exerting the overall performance of high-performance computing system can it better contribute to the improvement of social economy and provide corresponding services for the development of science and technology. High-performance computing system was first applied in the area of computing technology, including scientific computing, simulation engineering and simulation engineering, signal processing, image processing, data visualization and business volume calculation.

4. Application of High Performance Computing System in Agricultural Big Data

4.1 Characteristics of Big Data

In modern agriculture, big data plays an important role. In big data, not only a large amount of data information is stored, but also its application value is high. At present, with the rapid development of Internet technology and cloud computing, the effective application of big data has become an important link in China's agricultural modernization. Big data has the following

characteristics: First, the volume of its data is very large, and all major terminal devices are continuously generating real-time data, while meteorological and environmental changes are also continuously increasing; Secondly, the speed is very fast. With the continuous optimization and update of technology, the speed of data information generation has also been improved. The third is diversity. For crops, there are many factors that can affect them, so it is necessary to collect all kinds of information continuously. The fourth is authenticity. In order to let decision makers know the actual situation, agricultural related data must be true and reliable, so as to provide accurate basis for decision makers to make decisions. Fifth, the value is very high, and the application value of big data is very high. Therefore, it is necessary to continuously mine the information in big data so that it can continuously provide corresponding support for decision-making.

4.2 Types of Big Data

First of all, it is the monitoring data in the agricultural Internet of Things, which refers to the Internet of Things technology that can be applied in agriculture. It can comprehensively understand agricultural information from multiple angles, and at the same time, it can intelligently process these information to make agricultural production more intelligent. Secondly, meteorological data, which is very important for agriculture, has a great impact on the overall growth and development of crops, and will have a certain impact on the quality and yield of crops. The third is resource and environment data, which provides a lot of spatial information data for agricultural resources with the support of earth observation data. With the continuous upgrading and optimization of remote sensing technology, it plays an increasingly important role in agricultural production. Fourthly, the data of agricultural economic monitoring. In the rapid development of agricultural economy, it is necessary to use the data of agricultural economic monitoring to obtain the overall development of agricultural economy and guess the related development direction.

4.3 Problems of high-performance computers in agricultural big data processing

With the continuous expansion of the system, high-performance computers encountered a series of problems when processing agricultural data. Therefore, it is difficult to adapt to its rapid development only by relying on traditional technology. High performance computing needs to solve many problems such as energy consumption, programming, communication and storage. At present, R&D personnel have come up with many methods to calculate and process a large amount of data, but their ability is still insufficient to cope with the increase of data. In the characteristics of agricultural big data, there are various types, and the processing time is short, and the data is uncertain. Many Internet companies at home and abroad have successively launched their own big data processing systems, applying many technologies to the processing of big data. The establishment of data system needs continuous optimization and improvement in energy efficiency, physical space occupation and other fields.

4.4 Application of High Performance Computing System in Agricultural Big Data

3.4.1 Application of Agricultural Big Data Platform

Among agricultural information resources, agricultural big data platform plays a very important role. If we want to analyze the massive data in agricultural information, we need a high-tech application platform. Only with the support of platform technology, agricultural technology can give full play to the advantages of its data function, and integrating platform big data technology into the scope of agricultural field can play a real-time monitoring and early warning role in China's agricultural market, so that under the relevant guidance of smart agriculture, corresponding production management can be carried out to guide decision makers to make correct decisions.

In the analysis of agricultural data, data clustering analysis is a very important method, which can analyze the data gathered in a cluster and analyze their common characteristics, so as to find out the potential laws, and give corresponding guidance for farmers' subsequent production and management, and the opinions put forward are scientific and reasonable. Among data mining, clustering mining

algorithm is the most commonly used, which can deeply analyze and mine the correlation between events. According to the differences between data, or the similarities between similar data, data can be divided into many categories. Among the same species, their data have high commonalities, while among different species, there are some differences.

For the practical application of agricultural data, spectral clustering algorithm is effective and widely used when mining its data. The similarity of the algorithm is generally calculated by Gaussian kernel function, and the elements in the similarity matrix S are Sij:

$$S_{ij} = \exp(-\frac{||x_i - x_j||}{2\sigma^2}) \quad (1)$$

i,j=1,2,⋯， n

Sparse the similar matrix, and the processed matrix is defined as the connection matrix W.

Then, the processed W matrix is transformed to obtain the degree matrix D:

$$D = (d_i) = \sum_{j=1}^{n} w_{ij} \quad (2)$$

i,j=1,2,⋯， n

The difference between the degree matrix D and the weight connection matrix W constitutes the Laplace matrix:

$$L = D - W \quad (3)$$

Properties of normalized Laplacian matrix:

$$L_{sym} = D^{-1/2}LD^{-1/2} = I - D^{-1/2}WD^{-1/2} \quad (4)$$

$$L_{rw} = D^{-1}L = I - D^{-1}w \quad (5)$$

3.4.2 Application of Agricultural E-commerce Big Data Development System

In the agricultural e-commerce big data system, various functions can be realized, such as detailed statistics of user activity, specific detection of malicious purchases, visual analysis of purchases, intuitive mastery of e-commerce data and corresponding support for subsequent measures.

3.4.3 Application of data visualization technology

Among the visualization technologies, the area covered and the fields covered are relatively wide. Visualization of data integrates intuitive and interactive performance. Most visual retrieval methods are based on the original data, producing visual graphics, and processing and rendering them accordingly. Among the final requirements of visualization, the embodiment should be simple and easy to understand, so that non-professionals can have an intuitive understanding of the specific situation. However, the complexity of data visualization itself is not completely equal to that of data itself. The ultimate goal of visualization is to enable customers to get effective information in the data quickly. With the update of information technology, massive data analysis and deep learning in agricultural production can use information technology to effectively solve the problems. Generally, data visualization is integrated with the function of data analysis, and the development of data visualization needs continuous optimization with the development of society. Compared with data processing, data display needs to emphasize its functional practicality, so that more people can have a deep understanding of data visualization. At present, data visualization is more and more difficult to meet the needs of the current information situation, so it is necessary to make a breakthrough in the direction of evidence-based analysis and decision-making.

3.4.4 Application of High Performance Computing System

In a high-performance computing center, the core is the ability of computing, while the core of agricultural big data is the storage of data. In the process of data storage and data calculation, one of them is needed to dominate, manage all hardware and software information, and share its data information at the same time. It is also necessary to coordinate the resource allocation among agricultural data, schedule the data, store it later and finally back it up. Generally, the management party stores data. From the overall hardware configuration, there is no high requirement for the speed of calculation when storing data. At the same time, the data has a great demand for caching, and for the storage and backup of big data, it can be processed in the background in a time-sharing manner.

In this way, a complete coordination mechanism can be built. Secondly, by integrating computing power and storage function, the accuracy of agricultural data processing can be effectively improved. For agricultural data centers, the amount of data is relatively large, so high-performance computing is needed to play an important role.

5. Summary

The proportion of agriculture in China is relatively large, so a detailed and in-depth analysis of agricultural big data can greatly promote the development of agriculture in China. Big data is based on informatization, so it is promoted upwards. In the future agricultural big data, its development direction is to use the data opening and data sharing mode to finally promote the comprehensive construction of transaction resources, so that transactions and products can form a common development mode, thus promoting the all-round development of big data. In this process, in order to accelerate the development of agricultural big data, it is necessary to promote the sharing of resources and integrate its transactions, so as to create large-scale data transactions, so that data can be actively followed, thus promoting the development of China's overall industrial chain.

6. References

[1] Sun Zhongfu, Du Keming. Prospect of research and application of big data in smart agriculture [J]. China Agricultural Science and Technology Herald, 2013(06): 63-71.

[2] Li Youfeng, Wang Zhigang. Research on Smart Agriculture Architecture Based on Dynamic Cloud [J]. Computer Technology and Development, 2014(03): 190-193.

[3] Shi Lianmin, Chen Zhifeng. Application of Internet of Things in Smart Agriculture [J]. Research on Agricultural Mechanization, 2013(06) :250-252.

[4] Hu Yiwen. Research and application of data mining in agricultural production [J]. Times Agricultural Machinery, 2019,0(3):27-28.

[5] Zhou Xiao, Tan Junshan. Research on the application of big data in smart agriculture [J]. Southern Agricultural Machinery, 2019,50(14):74.

[6] Xu Xiaobing, Chen Yimin, Shi Hong, et al. The application of big data in modern agriculture [J]. China Seed Industry, 2017 (7): 35-36.

[7] Gu Yongquan. Analysis on the development and application of big data in modern agriculture [J]. Shanghai Agricultural Science and Technology, 2017 (4): 11-12.

[8] Han Nan. Analysis of ways to promote the application of agricultural big data in modern agriculture [J]. Agricultural Economy, 2017 (8): 25-27.