Analysis on Computer Information Processing Technology Integrating Big Data Technology

Yuanyi Chen1,*

1College of Internet of Things, Jiangxi Teachers College, China, 335000

*Corresponding author e-mail: cyy321123@sohu.com

Abstract. In the era of big data, massive data information is produced every moment. Therefore, the traditional computer information processing technology (hereinafter referred to as IPT) has been unable to meet the actual needs of data information processing, which requires us to integrate big data technology (hereinafter referred to as BDT), which will continue to promote the development of computer information technology. China has gradually entered the era of big data. With the application of computers by users, the information data gradually increases, and various problems arise. The computer IPT is facing more and more challenges. In the face of various challenges, computer technology needs to constantly improve the ability of information and data processing. With the application of computer, the information data is increasing gradually, and the computer IPT is facing more and more challenges. In the face of various challenges, computer technology needs to constantly improve the ability of information and data processing. First of all, this paper analyzes the common computer IPT. For the application of BDT, this paper puts forward the future development of computer IPT.

Keywords: Big Data Technology, Information Processing Technology, Future Development

1. Introduction

With the growth of the number of network users, trillions of GB of data are generated every day. Therefore, the era of "big data" has come, which brings new challenges to computer IPT. We must update the traditional ideas, which will better handle computer data [1]. In the era of big data, all sectors of society have gradually formed a close system through information data, which will enhance the communication and communication between people through Internet data transmission. Through IPT, we can help people improve their quality of life. In the era of big data, we must search for useful information in massive data, which requires us to integrate BDT. Through the integration of BDT, we can effectively enhance the quality and level of computer information processing, which can help people better process information data. However, there are still many problems in the computer information
processing, which requires continuous innovation of computer IPT and meet the needs of people. Therefore, we must strengthen the integration of computer IPT and BDT [2].

2. Integration of IPT and BDT

2.1. The process of data mining

Data mining provides a way to get value from data. There are many kinds of data mining algorithms, but they all have similar processes. At the same time, we need to use various algorithms together, which will solve a variety of practical problems. The general process of data mining consists of translation data, preprocessing data and data modeling [3]. Translation data can transform the original data into recognizable data format. Then through the data input, we can complete the data selection and sampling, which realizes the value selection of the original data. Preprocessing data is the initialization of data, which will improve data quality. Translation data contains many interference items and incomplete data, which must be preprocessed by certain methods. There are many common data preprocessing methods, such as data filtering, data variable conversion, missing value processing, bad data processing, data standardization, main component analysis, attribute selection, data specification, etc. Data modeling is a multi-dimensional analysis of preprocessed data. By establishing mathematical model, we can mine information data [4]. The general process of data mining can be shown in Figure 1.

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

2.2. Principle of K-means clustering algorithm

K-means clustering algorithm is a clustering method based on similarity classification proposed by J.B. Macqueen. It is an algorithm based on distance. K-means clustering algorithm is one of the most commonly used information classification methods in data mining. First, determine the types of data to be processed. For example, it can be set to K value, which is to take k points as the first clustering center. Secondly, the distance between all data and K initial cluster centers is calculated. Thirdly, according to the algorithm, the data is divided according to the nearest distance to K initial clustering centers. Fourth, a new cluster center is calculated after partition. Fifth, repeat the second, third and fourth steps until the two clustering centers tend to overlap. Finally, we can determine the final cluster center [5]. The flow chart of K-means clustering algorithm is shown in Figure 2.
2.3. Computer information storage technology

Information storage technology is an effective measure taken in the first time after obtaining relevant data through information acquisition technology. By saving the data information, we can store the data in the database system according to the specific classification rules. Through search technology, users can get the demand information from the database at any time. Under the background of big data, the preservation of massive data information has long-term and integrity, which is a major problem faced by computer storage technology. Therefore, to maximize the storage time of data information has become the main task of IPT at this stage. The most common deep web technology of computer information storage technology is deep network space technology. GFS distributed storage technology developed by Google is widely used, which can compress huge data and information resources in a short time. Through GFS distributed storage technology, we can reduce the overall storage of data and query time, which can fully improve the utilization of disk space [6]. The GFS configuration topology is shown in Figure 3.

Figure 2. Flow chart of K-means clustering algorithm

Figure 3. GFS configuration topology

2.4. Data aware acquisition processing
The traditional IPT can’t effectively complete the acquisition and perception of data information. But perception acquisition is an essential part of data processing, we must pay attention to it. Through the deep web technology, we can effectively integrate all kinds of data information together, which can lay the foundation for the subsequent data extraction work. Deep level network technology itself covers a large number of dynamic data information, which has the characteristics of strong distribution and particularity. Data mining and data analysis are two kinds of technology. By setting off the two, we can have a significant effect on business information collection and processing. E-commerce platform can filter sales data through technology, which can be targeted for product adjustment. In the era of big data, we can mine and analyze data through data perception acquisition and processing technology, which will effectively improve the value of information.

2.5. Data information security processing

Information security is the core of computer IPT, which can effectively avoid information leakage. Through the data information security processing, we can make the information get the security protection. Information security processing mainly from two aspects. On the one hand, we need to establish and improve the computer information security system. On the other hand, we need to develop computer information security protection products. Through the establishment of school enterprise practice training mechanism, colleges and universities can cultivate more talents to invest in the construction of information security protection system. In addition, military information has strict requirements on the security of classified information, which needs to be paid enough attention. Therefore, we should strictly follow the standard system for scientific management, which will avoid the leakage of classified information. Taking software as an example, the website group formed by software creation can adopt hybrid management, which will carry out security design from many aspects, such as identity authentication, communication confidentiality, access control, residual information protection, security audit, communication integrity, non-repudiation, software fault tolerance, resource control, etc.

3. Development direction of IPT

3.1. Information acquisition, processing and transmission technology

The primary task of computer information processing is to retrieve data. At the same time, any information processing means is based on the data acquisition. Data acquisition is to monitor the target information source in real time. By collecting and storing the required data into the prefabricated database, we can provide information input to various software system platforms. The second step of information processing is information processing, which is to classify and process the data stored in the database. Finally, we can transfer the processed data to users through data transmission technology, which realizes the data transmission.

3.2. Information storage technology

Computer information storage technology is to store the retrieved information in the structural database. When the user needs some information, the user can directly call out the corresponding information from the database by searching and other operation means. In the era of "big data", data has the characteristics of huge data information and rapid information change, which requires us to reasonably
apply computer information storage technology, which will realize the long-term and stable storage of a large number of data.

3.3. **Information security technology**

We must adapt to the characteristics of information security in the era of "big data", which will speed up the development of information security technology. We can proceed from the following three aspects. First, the construction of computer information security system. We must increase the training of computer information security personnel, which is a necessary condition for the construction of computer information security system. Second, accelerate the research and development of "big data" information security technology products. Third, improve the detection of key data. "Big data" has a large amount of information, which makes it difficult to detect each data. Therefore, we need to focus on strengthening the detection of important data.

3.4. **Development direction of IPT**

Due to the limitation of computer hardware, computer network has many limitations. Therefore, we need to transform the computer network into cloud computing network, which is the development trend of computer IPT in the "big data" era. In the future, the concept of computer network development is to separate hardware and network data, which will transform cloud computing into cloud computer network. In the future, the computer information network will form a "big data" network system, which can’t be separated. Therefore, computer IPT does not rely on the development of a single system, which will become the joint research and development of many companies.

4. **Conclusion**

The arrival of big data era brings important opportunities and challenges to computer information processing, which requires us to strengthen the research and development of computing IPT. Through IPT, we can improve the efficiency of computer information processing, which will ensure the security of data information. Therefore, the information processing technical personnel must master the existing computer IPT. Through information storage, security processing and other content, China can continuously improve the computer IPT.

**References**

[1] Li Junfang, Xu Xiaoya. Analysis of computer IPT under the background of big data era [J]. Information and computer (theoretical Edition), 2017 (2): 45-47.

[2] Zheng Xinxiao. Analysis of computer IPT under the background of "big data" era [J]. Science and technology innovation guide, 2016 (17): 72-73.

[3] Liang Dingyi. Analysis of computer IPT under the background of "big data" era [J]. Electronic testing, 2017 (17): 117-118.

[4] Wang Mingyu, Dai Weibo, Zhang Shiqi. Analysis of computer IPT in the era of big data [J]. Electronic components and information technology, 2018 (02): 1-3.

[5] Wang Chunju. Research on computer IPT in the era of big data [J]. Communication world, 2016
(06): 92-94.

[6] Chen Haiyu. Discussion on computer IPT under the background of "big data" era [J]. Computer products and circulation, 2020 (05): 6.