

Analysis of Computer Information Processing Technology Under the Background of Big Data

Yixiang Jiang^a, Qingsong Zhang, Chengting Zhang^b, Xuefeng Gao

China Tobacco Zhejiang Industrial CO.,LTD. Zhejiang. China

^ajiangyx@zjtobacco.com. ^btitanbyron@126.com,

Abstract. Large data, big data, refers to a large set of data sets. This kind of data collection has massive data information, which needs to be processed with more intelligent, rapid and efficient means to obtain value information. In the background of large data, the traditional computer information processing technology has been difficult to adapt to the processing of mass information. Therefore, continuous technological innovation is a significant trend in the development of computer information processing technology. This paper, based on large data and computer information processing technology, combines the characteristics of the information age in the new period, and explores the application and development of computer information processing technology under the background of large data.

Keywords: big data, computer information processing technology, application and development.

1. Introduction

With the popularization and development of computer Internet, all kinds of information data show the trend of "geometric" growth. The explosion of information and data has directly opened the prelude of the era of big data. Computer information processing technology is the basis for computer to complete information processing, instruction release and task execution. Advanced and efficient computer information processing technology can help people to solve problems and realize the needs better. It should be pointed out that the development of computer information processing technology at the present stage can not fully adapt to the demand of information processing in the era of big data. Therefore, the information processing technology in the era of big data still faces great opportunities and challenges, and it is very necessary to promote the development and perfection of the computer information technology [1].

2. Definition and Characteristics of Large Data

2.1 The Definition of Large Data

The essence of large data is a type of information asset. For large data, more scholars and institutions think that it is a kind of valuable information assets. Through new processing mode, precise insight and powerful optimization ability, the intrinsic value can be explored to the greatest extent. In addition, different scholars have different definitions of large data, some scholars believe that large data is a class of data sets, the size of large data is far larger than the traditional database, it has a large number of information data, and the types of various kinds of data are also different. To what extent the data of large data are large, there is no clear regulation in the industry at present. It is generally believed that the data over 10TB-1PB is large data (1TB=1024GB, 1PB=1024TB).

2.2 Characteristics of Large Data

Large data is a concept proposed in recent years. In the process of practice, it is found that it has the characteristics of large data capacity, diversified data form, and fast data processing. The large data capacity mentioned here refers to the huge data information needed by computers to process information in the era of large data. At the same time, the growth rate of these information is very fast, and the speed of information growth is even increased from the original TB to ZB (1ZB=1024PB). The diversity of data forms is another significant feature of large data. The data types of large data are very complex and present a variety of characteristics. For example, the common data

types in large data include web pages, text, video, audio, pictures and so on. These data are in a variety of formats, which all give computer information. Processing has brought some challenges. The rapid data processing is another significant feature of large data.[2] Compared with the traditional data processing, the speed and quality of information processing in the era of large data has been improved. Because the large data is dynamic, time increasing and changing, it requires more efficient and quick information processing technology to deal with the data information, thus extracting the real-time value of the information.

Table 1.Characteristics of large data

capacity	Huge data capacity
type	Diversity of data types
speed	The speed of data acquisition and change
complex	Data sources, formats, and content are complex
quality	The data is true and reliable

3. Computer Information Processing Technology Under Large Data Background

3.1 Information Acquisition and Storage Technology

The first step of computer information processing is to obtain information. Information acquisition needs to monitor the source of information in real time.[3] At the same time, the required information is stored in various pre designed databases by some technical means, and then the corresponding data information is provided to the software platform in need. In the process of information acquisition, DEEP WEB data perception and acquisition technology is widely used at this stage. DEEP WEB belongs to a network deep space technology, which has the characteristics of large information scale, fast information change speed, wide distribution of information and special way of information access. DEEP WEB technology can carry out high-quality data integration, and integrate and analyze integrated data. At present, the widely used data storage technology is distributed data storage technology developed by Google Corporation, that is, GFS technology. Distributed data storage technology has been widely applied in Baidu, IBM and other giant companies. The concept of column storage is used in the development of distributed storage technology. The so-called column storage is data storage in columns, compared with the traditional row storage mode, its data compression and circulation speed are faster. At this stage, the way of row and column hybrid storage is also widely used. This storage structure improves the efficiency of data loading and disk space utilization, and is a promising technology.

Table 2.DEEP WEB's directory coverage

	The number of DEEP WEB databases	Coverage rate (%)
Completeplanet.com	70000	15.6
Lii.org	14000	3.1
Turbo10.com	2300	0.5
Invisible.web.net	1000	0.2

3.2 Classification Analysis and Cluster Analysis

The main purpose of computer information processing in the era of big data is to obtain valuable information and serve the management and strategic decisions of enterprises. The classification analysis and cluster analysis technology implement two different functions [4]. The classification analysis technology is mainly used to predict the possible behavior of the next stage of the customer, and the cluster analysis is mainly applied to the mining of value data. Classification analysis technology first classifies the data points, and then determines the new data points. On the basis of formulating a series of reasonable assumptions and objective conditions, we can reasonably predict the next stage behavior of customers. There are some differences between cluster analysis technology

and classification analysis technology. It is a method to analyze data in the case of not knowing the relevant limiting factors. It divides data into several object groups, then analyzes the data in each object group according to a set of standards, and then extracts the value information [5].



Fig.1 classification analysis technical steps

3.3 Data Analysis Technology and Visualization Technology

In the background of large data, the computer information processing technology can be attributed to the following types, which are affective analysis, network analysis and spatial analysis. Among them, the emotion analysis technology is the subjective analysis of the natural language; the network analysis technology analyzes the data according to the difference of the network characteristics, and the spatial analysis technology uses the technology of polar topology, geometry and geo coding to make a comprehensive statistical analysis of the data. In order to help people to explore the value of data more comprehensively and accurately, and to understand the results of data analysis objectively, it is necessary to use visualization technology to further adjust and optimize the results of data analysis. At present, visualization technology is mainly applied to pictures, charts creation and animation simulation. For example, Clustergram is a widely used visualization technology, which is mainly used in the data allocation process.

4. Summary

To sum up, this paper first introduces the basic concepts and characteristics of large data. It holds that large data has three basic features: large data capacity, complex data type and fast data processing. Secondly, three computer information processing technologies are introduced, which are information acquisition and storage technology, classification analysis and clustering analysis technology. Data analysis and visualization technology. The author thinks that under the background of big data, computer information processing technology will develop in three main directions. The first is the development of computer networks towards the cloud computing network. The computer network structure at the present stage has a strong dependence on the hardware, which restricts the development of large data, but the cloud computing technology is only dependent on the hardware architecture, and it has a broader space for development. Second, computer technology will develop

towards a more open direction. By defining the network architecture autonomously, we can separate network information from hardware, and then achieve more open information processing.

References

- [1]. Chen Shuai. Analysis of computer information processing technology under the background of big data [J]. China new communications, 2018 (6).
- [2]. Marshall. Computer information processing technology in the background of big data analysis [J]. industrial b:00187-00187.
- [3]. Li Junfang, Xu Xiao Ya. Analysis of computer information processing technology under the background of big data era [J]. information and computer (theoretical version), 2017 (2): 45-47.
- [4]. Zhang Wenjuan, Sun Dan. Analysis of computer information processing technology in the era of big data [J]. technology information, 2017 (36).
- [5]. Wenjin Yu ; Yixiang Jiang, Mobile RFID Mutual Authentication Protocol Based on Hash Function[C]. IEEEExplore, pp. 358-361, 2018.