Research on Multimedia Classification Technology in Big Data Environment

Xiaoxing Ma

Tianjin Hexi District Zhujiang Road 25# Tianjin University of Finance & Economics, 300222, Tianjin

[a]xxingl@163.com

Keywords: Big Data, Multimedia Classification, Filtration technology, Dynamic coding algorithm.

Abstract: In recent years, with the rapid development of Internet technology, construction on network has been continuously improved. In the big data environment, there are many new encapsulations for information data and coding formats, and identification of various file types. Data filtering by batch is a primary concern of big data researchers. The emergence of filtering technology on massive multimedia information in big data environment solves the problem of data identification, classification, extraction and filtering of multimedia information files in huge data resources. This paper studies the database storage and classification algorithms of multimedia in big data environment.

1. Introduction

In recent years, with the rapid development of Internet and streaming media technology, the growth of network video services has been very rapid. In 2017, the proportion of video traffic has reached 76% in Internet traffic. According to Cisco, video traffic will reach 85% by 2021, and there will be 1000000 min of video content through the network per second.[1] Such a large amount of network video traffic is undoubtedly a huge burden on the classifier, not to mention the requirement for real-time and accurate classification of network traffic, which poses a huge challenge to the classification of network video traffic.[2]

With the diversification of information and the multi-inclusiveness of the big data environment, many digital multimedia information formats have emerged, such as APE, FLAC for audio multimedia information files, mkv, HDTV, QXGA for video multimedia information files.[3] The traditional massive multimedia information filtering technology cannot identify the encoding and packaging format of above-mentioned new multimedia file, resulting in inaccurate information filtering for multimedia, and cannot cope with gushing of the huge information data. Therefore, an improved method of massive multimedia information filtering technology in big data environment is proposed.[4] Extend the underlying data of the algorithm and added dynamic multimedia learning coding to support the full format of multimedia information, and has self-upgrading learning characteristics.

Through the classification of network video service flows, Internet providers (ISPs) can better provide different levels of services according to the quality of service (QoS) requirements of different video services. [5] Due to the use of technologies such as dynamic port and address camouflage, the video stream classification method based on machine learning has become a research hotspot. Optimize the logic of the algorithm to make the algorithm more rigorous. Improve the design by adding the auxiliary logic algorithm to deal with the multi-data co-surge in the big data environment. At the same time, the multimedia information feature recognition filtering engine is added to improve the accuracy of massive multimedia information filtering in the big data environment and improve the efficiency of technical processing multimedia information. Meanwhile, the feature recognition filtering engine of multimedia information is added to improve the accuracy of massive multimedia information filtering in the big data environment, to improve the efficiency of processing multimedia information. The simulation experiment proves that the massive multimedia information filtering technology in the improved big data environment has the characteristics of accurate filtering, wide range of multimedia information recognition, easy operation and strong stability, which can meet the

  DOI: 10.25236/iccse.18.009

requirements of multimedia information filtering in big data environment.

2. Database Technology

Database technology is widely used in the following areas:

First, the application of database technology in the field of information retrieval: Through the information retrieval system, users can input relevant information according to their own needs, can find the corresponding document information in the database system, greatly improving the efficiency of information retrieval, which is mainly benefited from the wide application of database technology, it is more widely used in online book catalogue and online document management;

Second, Database technology is the most widely used in long-distance computer network systems, in the application of distributed information retrieval, the current Internet technology provides an opportunity for the expansion of database technology applications. In the current era of big data, e-commerce has been developed efficiently. Network users use computers to store data and use the network to send e-mails, transfer files and remote logins, thus realizing the real-time sharing of information data.

Third, the expert decision-making system also relies on the application of database technology. The internal business data is an important basis for decision-making at the leadership level. The database technology helps to improve the quality and efficiency of enterprise data transmission management, and thus can provide scientific decision-making reliable data support for leadership. Therefore, the most widely used application is the expert decision-making system, and the application of artificial intelligence technology is emphasized in this system, which expands the application field of database technology.

In the era of big data, the analysis needs, sharing needs, and decision-making needs of various data require the support of database technology. This requires close integration of the application status of database technology in the era of big data, pertinence application of database technology, and continuous efforts and innovation for future development, expanding the application field of database technology to better serve people.[6]

In the era of big data, multimedia technology and database are combined to achieve the advantages of multimedia technology and database technology, to offset their own deficiencies, to build a new style database, enrich the multimedia data interface, and also do personalized design according to specific requirements, effectively solve the technical problems faced by multimedia in the process of database introduction, effectively avoiding database security problems, and ensuring that the security of the database is comprehensively improved.

Specifically, in the era of big data, from the perspective of database technology, in the application process of multimedia database, we are still working to solve the following specific problems: First, personalize according to the needs of multimedia data information; Second, how to ensure the compatibility of multimedia data in the integration process, Third, the problem of multimedia data and user interaction is outstanding. Multimedia databases must be able to support structured data and unstructured data in order to better perform the value of combining multimedia technology with database technology.

3. Algorithm optimization

Aiming at the problems of obsolete algorithms and lack of rigor of logic in traditional multimedia information filtering technology, the corresponding optimization and improvement have been made. Algorithmic expansion and design are added to assist the logic algorithm to solve the problems in the traditional filtering technology. [7] The multimedia information feature recognition filtering engine is designed to enhance the accuracy of the massive multimedia information filtering technology in the improved big data environment. The simulation experiment proves that the proposed mass multimedia information filtering technology has high accuracy and recognition of multimedia information filtering, low filtering time, high efficiency, strong stability, etc., can meet the multimedia information filtering requirements in big data environment.

3.1 Optimization expansion algorithm.

In the big data environment, the underlying algorithm support library in the multimedia information filtering technology is too old to distinguish the emerging digital high-definition encoding and packaging format, resulting in many new multimedia information resources not being recognized and filtered. To this end, the dynamic coding algorithm is used to replace and update the underlying support library of the original algorithm. The dynamic coding algorithm refines and summarizes multimedia information according to the common feature arrangement and has the characteristics of self-upgrading and self-learning. The dynamic coding algorithm expression is:

$$fd'' = \frac{(f_{(e \in link)}\ e^2)^{2e} \sum \prod d}{2f_s} \Rightarrow \frac{1}{2_{link}} s. \tag{1}$$

In Eq. 1: d is the big data space; s is the large data space data volume; $f_{(e \in link)}\ e^2$ is the big data space feature function. The above dynamic coding algorithm expression is a steady state dynamic coding algorithm expression. With the change of d and s values, the dynamic coding algorithm expression performs self-derived conversion to realize self-upgrading and self-learning functions.

3.2 Improve logical rigor.

In the big data environment, the traditional multimedia information filtering technology algorithm is not rigorous enough, and the logic bugs appear dynamically. When the data volume suddenly increases in the big data environment, the logical retrieval are abnormal, which causes the traditional algorithm to crash and the multimedia information data bursts to gushing. In response to this problem, an auxiliary logic algorithm is added to the above dynamic coding algorithm to enhance the stability and logical rigor of the algorithm. The auxiliary logic algorithm retrieves, analyzes, identifies, confirms and extracts a series of processes according to the internal packaging of the multimedia information resources in the big data environment. The result data is automatically returned to the total algorithm, the dynamic coding algorithm is used to identify and confirm. The auxiliary logic algorithm expression is as follows:

$$\sin kmv^n = \varphi_i^m \Rightarrow fw'_{i-1} \frac{c^{2 \wedge m}}{link \rightarrow \prod d_i}. \tag{2}$$

In Eq. 2: The value range of n , m , and i is determined by the big data resource coefficient in the network space and meets the constraint conditions ($n < m \in$ large data space resource quantity, $i \neq 0$), $link \rightarrow \Pi d_i$ represents the mapping process of dynamic data. fw'_{i-1} represents the second constraint of dynamic data. When a new encapsulation format for multimedia information data appears in the big data environment, the auxiliary logic algorithm performs feature processing according to the newly-encoded data encapsulation format of the multimedia encapsulation format, and returns the new processed encapsulation feature tag back to the underlying coding support library to achieve self-upgrading function. Add active execution to auxiliary logic algorithm execution code to ensure that the scans the dynamics of multimedia information real-time in the big data environment. It provides guarantee for the accurate extraction of subsequent filtering engines.

3.3 Multimedia information feature recognition filtering engine.

In the improved in big data environment, the design of multimedia information feature recognition filtering engine is added method of massive multimedia information filtering technology to enhance the improved technology filtering effect. The feature recognition filtering engine consist multimedia information feature and matching module. The multimedia information feature comparison module transmits the information data according to the dynamic coding algorithm, and performs multimedia information feature tag identity DNA matching binding processing on the information data. The

multimedia information that meets the binding condition is sent to the feature filtering classification module perform filtering classification process, and the information that fails to be authenticated by the multimedia information feature matching module is re-identified by the initial module. The multimedia information feature comparison module adopts the multimedia information kernel construction algorithm, which has the characteristics of high recognition rate and high accuracy compared with the traditional filtering algorithm. At the same time, the algorithm writes a string of dynamic identity codes on the bottom of the identified multimedia information data. The code itself does not affect the data content of the original multimedia information, only used for identification, and only this technology can recognize this code.

4. Experiment and analysis

The simulation experiment tests the accuracy and time-consuming of the improved multimedia information filtering technology. The experimental test platform selects seven ultra clear web video streams with representative for analysis: Video streaming like Tecent Video and iQIYI, instant video communication like Weixin video, webcast video like CCTV sport and BBC, P2P client video like Douyin and Http download video.

The experiment takes the traffic in the real network, extracts the network flow in different time periods, adopts the traditional multimedia information filtering technology and the improved multimedia information filtering technology respectively, and compares the filtering quantity in the test time. Obtain the amount of information filtering, and calculate the filtering accuracy. The test time is 60 minutes, there are 6 groups, each group is 10 minutes, and the time span is from February 2017 to July 2017. The sample consists of a quintuple (time, source IP address, destination IP address, protocol, size of video stream). Each video stream lasts for 30 minutes, and a total of 840 streams are counted. Compare the time and filtering effect of improved multimedia information filtering technology for big data environment and traditional multimedia information filtering technology.

The classification accuracy results are average of the above various feature selections: the dynamic coding algorithm has higher classification accuracy rate. The network multimedia file classification is the key to realize QoS stystem, and for the network flow classification task requiring high concurrency and low delay. In other words, the dynamic coding algorithm is indispensable, which can improve the classification efficiency and the accuracy of classification.

5. Summary

This paper introduces the key technologies of multimedia classification in big data environment, and the development of classification filtering algorithm and summarize its advantages, perform simulation experiment to compare accuracy and operation time in the classification, the experiment data shows that the improved method of dynamic coding algorithm for massive multimedia information filtering technology in big data environment. The design of dynamic coding algorithm is added to the improved method of massive multimedia information filtering technology in big data environment to enhance the improved filtering effect. It has the characteristics of high accuracy of multimedia information filtering, good recognition effect, less filtering time, high efficiency and strong stability. It can meet the claim of filtering of multimedia information in big data environment.

References

[1] Zhou Yan, Zeng Fanzhi, Zhang Zhifei, Research on efficient retrieval algorithm of mass digital images for manufacturing process [J].Journal of Natural Science, Xiangtan University, 2014, (4):433-437.

[2] Niu Y, Fang L, Sun S, et al. The Design of Book Sorter Base on Radio Frequency Identification[J]. Journal of Applied Science and Engineering Innovation, 2018, 5(1): 18-21.

[3] Yang Yan, YANG Yan,The improvement of mass multimedia information filtering technology

in large data environment Improvement of mass multimedia information filtering technology in big data environment[J]. Journal of Xi'an Polytechnic University, 2017 issue 4.

[4] Li R, Yu R, Wang X. Information Resources Sharing Security in Cloud Computing[J]. Journal of Applied Science and Engineering Innovation, 2018, 5(3): 65-68.

[5] Wu Zheng, Dong Yuning. Research on feature selection method of network video traffic classification [J]. computer engineering and applications, 2018, (6): 7-13.

[6] Zhao Mingyan, Li Zeping. A popularity prediction algorithm based on video features and historical data [J].Computer and Modernization, 2018, (2): 49-53.

[7] Feng Yaya, Yuedong. Design and Implementation of Distributed Retrieval System for Power Video Large Data [J]. Computer Technology and Development, 2016, (12): 186-189.