On the Application of Big Data Technology in Broadcasting: Postprint

Authors: Gao Yandong

Date: 2023-10-08T00:00:00+00:00

Abstract

The 21st century is an era of explosive growth in data and information. With the rapid development of emerging online media, self-media, and micro-media, the broadcasting and television media industry has encountered unprecedented opportunities and challenges. Leveraging technologies such as big data, cloud computing, and mobile Internet, traditional broadcasting and television media have accelerated the pace of media convergence. Specifically, through big data technology, it is possible to target interactive television terminals and mobile clients, deliver personalized program recommendations by analyzing user behavior data, provide clues on hot topics in online public opinion for converged media program production, and also address the data processing and storage challenges faced by the broadcasting and television sector. This paper begins with the concept of big data, analyzes big data technologies, and explores the current application status of big data technology in the broadcasting and television field.

Full Text

On the Application of Big Data Technology in the Broadcasting Field

Abstract: The 21st century is an era of explosive growth in data and information. With the rapid development of emerging online media, self-media, and micro-media, the broadcasting industry faces unprecedented opportunities and challenges. Leveraging big data, cloud computing, mobile Internet, and other technologies, traditional broadcasting media have accelerated the pace of media convergence. Through big data technology, personalized program recommendations can be delivered to interactive TV terminals and mobile clients by analyzing user behavior data, clues for online public opinion hotspots can be provided for converged media program production, and the data processing and storage challenges facing the broadcasting field can be addressed. This article begins

with the concept of big data, analyzes big data technologies, and explores the current state of big data applications in the broadcasting field.

Keywords: big data technology; broadcasting field; application
Classification Code: TP311.13
Document Code: A
Article ID: 1671-0134(2018)04-059-03
DOI: 10.19483/j.cnki.11-4653/n.2018.04.020
Author: Gao Yandong

In today' s increasingly competitive all-media era, traditional broadcasting media must achieve multi-dimensional convergence development with the Internet while ensuring security, providing richer resources, more diverse production models, and faster, more authoritative distribution channels. Simultaneously, facing massive amounts of data information, broadcasting media have urgent needs for information processing and storage, sampling analysis, user profiling, and visualized precision pushing. The emergence of big data technology effectively solves these problems and drives the development and transformation of broadcasting media. Researching big data technology and analyzing its applications in the broadcasting field holds significant importance for broadcasting media.

1. Big Data Technology

The Internet has created a large-scale environment for big data applications. Big data technology, a derivative of the network era, is fundamentally based on computer and information technologies to collect, analyze, and process massive datasets, thereby fully extracting data value. By utilizing local database resources and Internet data for source extraction, open-source distributed big data technology architectures are typically employed to support multiple import methods for heterogeneous data from different sources, enabling massive data storage. Through cleaning, filtering, compression, and other preprocessing steps, data indexing and query functions facilitate information retrieval within databases, while cloud computing enables massive data aggregation and analysis. In the broadcasting field, big data technology has extensive innovative applications, primarily including news public opinion analysis and intelligent topic selection, user viewing behavior and effect analysis, converged media data analysis, product precision pushing, and visualized display.

2.1 Big Data Platform Construction Technology

Big data platform construction technology encompasses multiple technologies that collectively form the big data platform, rather than existing as a single solution. These include distributed databases, distributed file systems, data processing technologies, and others.

First, distributed database technology primarily integrates network technology with database technology, taking forms such as local databases and network databases to achieve massive data storage through networking. From a technical

perspective, practical application of distributed database technology reveals that its core lies in the use of the CAP theorem. While retaining some traditional database characteristics, it differs by enabling rapid data access. Currently, commonly used distributed data platforms include Hadoop, Greenplum, Hbase, and MongoDB.

Second, distributed file system technology. An excellent distributed file system can conveniently achieve elastic horizontal scaling as data volume increases, deploying numerous servers to support the technology and thereby meeting massive data storage requirements while enhancing storage capacity. To ensure safe and reliable data usage, most systems today adopt a master/slave dual-node approach where the slave node performs real-time backup and data synchronization while the master node operates, guaranteeing data security and reliability through redundant storage. Commonly used distributed file systems include Lustre, MogileFS, Hadoop, and FreeNAS.

Finally, big data processing technology, which finds extensive application across different industries with varying manifestations. Through comprehensive analysis, big data processing technology can be categorized into three forms: batch processing and analysis, real-time stream processing, and near-real-time analysis.

2.2 Big Data Mining Technology

The core and key to big data technology lies in data mining, which encompasses data excavation, analysis, and warehousing.

First, big data mining technology requires distributed computing platforms to excavate massive database contents, providing data support for analysis. This technology combines network technology with traditional data mining techniques, using computing platforms for data excavation and processing. Special circumstances may warrant simplified operations to ensure effective underlying platform operation.

Second, big data analysis technology processes data purposefully based on big data storage, management, and processing technologies, converting data into reports. Common analysis techniques include log analysis, file attribute analysis, text sentiment analysis, and natural language processing.

Finally, big data warehousing technology refers to the integration and storage of collected massive datasets.

3.1 Application Requirements of Broadcasting Big Data Platform

Broadcasting systems have specific requirements for big data platform utilization.

First, business requirements. Broadcasting systems must collect and integrate behavioral data and public opinion data from Internet, television, and new me-

dia users through big data platforms. After collection, the platform must mine and analyze this data to investigate user behavior patterns and preferences, thereby enabling new operational models for broadcasting systems. Collected user data can create video models for visual display and analysis of viewing conditions. For instance, real-time viewership numbers can be tracked according to different programs and channels to enable intelligent video recommendations and real-time queries. Additionally, broadcasting big data platforms should meet these requirements: (1) Provide practical case capabilities and establish commercial big data platforms with structured data processing capabilities, supporting B/S mode for visualized management and development; (2) Implement high-speed access and rapid computation for massive data based on the MapReduce distributed computing framework and HDFS distributed storage mechanism; (3) Establish MPP architecture and columnar storage databases for PB-level data analysis, and build HDFS technology extension and encapsulation platforms for PB-level unstructured data processing and analysis.

Second, performance requirements. Broadcasting big data platform construction must fulfill data collection, organization, analysis, and query requirements, necessitating flexible, controllable, and scalable mechanisms. Performance requirements include: (1) Data loading speed must not be less than 1TB/h; (2) Data processing speed must not be less than 1TB/h; (3) Complex data statistical calculations must complete within 15 seconds.

3.2 Construction Content of Broadcasting Big Data Platform

Broadcasting big data platform construction comprises two main components: the big data platform itself and personalized portal development. Personalized portal construction is particularly crucial, as differentiation strategy represents an important competitive approach in the broadcasting industry. Therefore, broadcasting big data portal construction must emphasize personalization and differentiation to meet users' individual needs.

First, big data platform construction. When constructing a broadcasting big data platform, industry development and potential future business transformation requirements must be considered. Consequently, the platform must support structured data and leverage computer architecture to demonstrate comprehensive data, advanced technology, and convenient openness. Construction content includes research platform setup, operation and maintenance platform construction, operating system installation, platform deployment, dynamic node addition and removal, MapReduce job management, platform monitoring, parameter tuning, platform configuration, mining tools, and system algorithms.

Second, personalized portal construction for broadcasting big data. Based on the big data platform' s capabilities for data collection, integration, and analysis, broadcasting enterprises can establish personalized portals. This process is vital for user data collection and analysis, encompassing user access records, chat data, forum posts, and other information. This data is integrated and

stored through de-formatting processes, and processed using behavioral, relational, semantic, and visual algorithmic tools to build the personalized portal. Leveraging powerful cloud computing capabilities, the big data platform can classify broadcasting users according to different criteria to achieve precise content recommendation and personalized choices.

Typically, personalized portal construction should focus on: (1) User behavior data analysis, which primarily involves using the big data platform to collect user data and analyze preferences and behaviors, including browsing data, playback data, on-demand data, advertising data, live channel data, and time-shifted data; (2) Associated recommendations during on-demand and live streaming, which analyze user data to establish preference models and recommend related content to enhance user satisfaction and advertising precision; (3) Internet user analysis. In the Internet era, with widespread mobile and online media usage, user data collection must incorporate Internet user data. Deep mining of user preferences through integration of mobile, broadcasting, and Internet networks provides data support for content acquisition, advertising placement, content production, and ratings evaluation. Personalized recommendation represents the best manifestation of big data application. In the big data era, analyzing users' real-time data and click behavior enables understanding of user preferences for pushing relevant content. In recent years, major media portals and mobile platforms have leveraged big data collection and analysis capabilities to discover user preferences and push relevant programs, thereby enhancing user satisfaction.

4. Application of Big Data Technology in Broadcasting Field

4.1 Application in Broadcasting System Operation and Maintenance Management

Operation and maintenance management represents one application of big data technology in broadcasting systems. Broadcasting involves numerous unstructured data types that increase operational and maintenance difficulty. With continuous technological development and progress, the broadcasting industry faces increasingly massive data volumes. To better manage and maintain broadcasting systems, enterprises must transform and strengthen big data technology application to achieve digital management models. Specifically, big data platforms should be built for program production and data storage based on digital models. Broadcasting systems comprise multiple platforms including batch deployment platforms, batch installation operations, operation and maintenance platforms, and job management platforms. Big data technology application enables timely fault location when system failures occur, improving operation and maintenance efficiency. Furthermore, real-time analysis of broadcasting system operation conditions allows timely understanding of system status and load conditions, enabling scientific prediction of system issues and effective trou-

bleshooting through early detection and treatment, thereby providing reliable guarantees for normal system operation.

4.2 Application in Broadcasting Business Operation Management

Broadcasting industry business operation requirements primarily include collecting, integrating, and analyzing data from television, Internet, and new media to predict user preferences and behaviors for business development. Big data technology application enables integration and analysis of this information on big data platforms to reasonably predict user preferences and establish video models. For example, user browsing data can simulate user profiles for intelligent classification by age, preference, and content, enabling intelligent recommendations for different categories. Additionally, user quantity and viewing duration information can serve as decision-making data to scientifically guide broadcasting business operations. In practice, the MapReduce distributed computing framework enables rapid access and high-speed computation of user data. MPP architecture databases process PB-level structured data, while HDFS technology processes PB-level unstructured data. Evidently, big data technology enables broadcasting systems to more accurately analyze user preferences for targeted content recommendation, meeting user needs and improving satisfaction rates.

4.3 Application in Broadcasting User Data Management

Building personalized portals requires effective user data management, strengthening data utilization through integration and analysis to scientifically predict user preferences and behaviors. Big data technology application enhances broadcasting user data management. First, networks enable communication with users to obtain user data. Currently, broadcasting enterprises collect user data on a per-network-user basis through sampling modes for data collection and analysis, using limited collected data to reflect user viewing conditions. In recent years, the proliferation of mobile phones, tablets, and widespread mobile device usage has inevitably created a large mobile user base. In response, broadcasting enterprises should strengthen interaction and convergence among Internet, mobile, and broadcasting networks to achieve data sharing, using big data technology to mine user commonalities and personalities to guide content acquisition, ratings evaluation, and content production. Big data technology application enables reasonable collection and utilization of viewing data, strengthening user data analysis and exploitation while enhancing broadcasting enterprise decision-making levels.

4.4 Application in News Public Opinion Analysis and Intelligent Topic Selection

Through probes and crawler technologies, various information data from mainstream news websites, forums, blogs, Weibo, WeChat, and mobile clients can be actively captured. Big data analysis of these massive datasets extracts enormous implicit information to provide critical market data, enabling public opinion

guidance and control, supporting news discovery and dissemination effect analysis, and achieving analysis and mining of Internet media communication effects. This also facilitates hot news recommendations, hot Weibo clues, similar event correlations, hot topic trend analysis, assistance in pre-reporting topic selection, support for user participation analysis (reading/sharing/commenting), and mining of user attention and tendency analysis. In the tide of media convergence, the broadcasting industry faces an epoch-making innovative revolution, with disruptive transformations in content form, communication channels, and distribution methods. Broadcasting media should fully utilize big data technology to research and build big data technology platforms suitable for broadcasting applications, achieving the convergence transformation from the IT era to the DT era.

References

[1] Song Wenjuan. A Brief Discussion on Big Data Technology and Its Application in Broadcasting Systems[J]. Liaoning Broadcasting Television Technology, 2014(4): 13-15.

[2] Wang Jusen. Exploration of Big Data Technology Application in the Broadcasting Field[J]. Network Security Technology and Application, 2016(2): 9-10.

[3] Lin Jian. Analysis of Big Data Technology Application in the Broadcasting Field[J]. Radio & TV Journal, 2017(6): 215-216.

[4] Tan Shan. Exploration of Big Data Technology Application in the Broadcasting Field[J]. Modern Industrial Economy and Informatization, 2016, 6(23): 79-80.

[5] Liu Xinrong, Chen Kai. Exploration of Big Data Technology Application in the Broadcasting Field[J]. Engineering Technology: Citation Edition, 2016(6): 00291-00292.

[6] Li Lei. Research on Broadcasting Industry Application Under "Big Data" [J]. Communication World, 2017(13): 291.

(Author' s Institution: Jilin Television Station)

Note: Figure translations are in progress. See original paper for figures.