Application of Big Data Analysis and Cloud Computing in Network Platform Building

Jichao Sun, Tao Liu*, Yong Wang, Ping Zhang, Dan Yang

School of Computer and Information, Anhui Polytechnic University, Wuhu, Anhui 241000, PR China

*Corresponding author: liutao@ahpu.edu.cn

Abstract. With the advent of the cloud era, big data has attracted more and more attention. At present, due to the expansion of the scale of chain educational training institutions and the complexity of management, the information stored in the data of management platform is mostly unstructured data model. The large amount and complexity of unstructured data become the main difficulties in big data analysis and decision-making. This paper analyzes and discusses the application of chain educational training institutions based on the background of big data, and build a platform for big data analysis and decision-making of chain educational training institutions based on cloud computing. The platform greatly improves the ability to process massive data and discovers the internal relationship and rules between big data. It aims to make better use of big data to serve chain educational training institutions, promote the continuous development and progress of chain educational training institutions, and even lay the foundation for the global chain educational training institutions.

Keywords: Chain Educational Training Institutions, Big Data Analysis, Analysis and Decision Support Platform, Cloud Computing.

1. Introduction

Today is a society with high-speed development of information, advanced science and technology, information circulation, people's communication is getting closer and closer, and life is becoming more and more convenient. Big data is the product of this high-tech era. Ma Yun, the founder of Alibaba, mentioned in his speech that the future era will not be IT era, but DT era, DT is Data Technology, which shows that big data plays an important role in business management. Therefore, with the development of the Internet and the improvement of science and technology, the application of big data in social life has been expanding, resulting in the network education management platform. The emergence and development of network education has had a positive impact on the management of education and training, making the management platforms of various chain educational and training institutions more individualized, expanded and even globalized.

2. The definition of Big Data and Big Data Analysis.

Big data is the meaning of massive data. It denotes that information is too large to be captured, managed, processed and sorted out as a more positive purpose to help enterprises make business decisions in a
reasonable time through the human brain or even mainstream software tools. Big data has the basic characteristics as 5V: Volume, Variety, Velocity, Value, Veracity [1]. Big data exists in the various walks of life, but numerous information and consultation are complex. We need to search, process, analyze, summarize its deep-seated laws.

Big data analysis is the process of studying a large number of different data types or big data to reveal hidden patterns, unknown correlations and other useful information. Such information can provide competitors with competitive advantages and bring commercial benefits, such as more effective marketing and increasing revenue. The main goal of big data analysis is to help companies make better business decisions by enabling data scientists and other users to analyze large amounts of transaction data and other data sources that may not be developed by traditional business intelligence projects [2].

With the development of Internet, cloud computing and high-speed mobile transmission technology, the world has entered the era of big data. In the era of big data, great changes have taken place in the organizational structure of network system and customer service mode of chain educational and training institutions. Because of synchronous teaching in different places and sharing the characteristics and needs of high-quality teachers and courses, the data structure of chain educational training institutions has changed from the traditional two-dimensional table structure to the massive unstructured data such as audio, image, text and video.

3. Significance of Big Data Analysis in Chain Educational Training Institutions.

With the further development of big data, various chain educational training institutions apply new teaching models and management methods. The application of big data in the management platform of educational undertakings can promote the innovation of teaching models and teaching methods, and lay a solid foundation for the implementation of the management of large-scale chain educational training institutions. At the same time, as the frontier industry of the times, private training institutions are also a trend. Therefore, under the background of big data, the application of chain educational training institutions is also an inevitable move to promote China's education industry to keep up with the trend of the times [3].

Whether the scale of training institutions expands or not will affect the effectiveness and influence of training institutions. The chain educational training institutions have emerged at the same time of improving the efficiency and expanding the scale. Because of the increasing number of trainees and other factors, the management becomes more and more complex. In recent years, with the development of Internet, cloud computing and mobile communication technology, digital education constructs an intelligent education environment through intelligent technology, flexible teaching methods and means. The application of big data in chain educational training institutions can realize the individualization of education. According to big data analysis, it can provide individualization service for the target. According to the specific learning situation and preferences of the target, it can formulate learning classes at different levels, training institutions of different subjects and solve the problems of half-way abandonment due to the transfer of incomplete training in different places. At the same time, for training institutions, while expanding scale efficiency, it facilitates financial management, staff Mobilization Management and so on. The application of big data in the education of chain educational training institutions can not only promote the continuous development of big data theory and technology, but also lay the foundation for the global development of chain educational training institutions and promote the development of training institutions and enterprises [4].

4. The Application Category of Big Data Analysis in Chain Educational Training Institutions

The application of big data analysis in chain educational training institutions mainly includes the following aspects:
4.1. Financial management
The tuition fees are different for different subjects, different levels and different learning time. Due to the difference of training institutions, each training institution should have its own management data and then classify and summarize them.

4.2. Registration management
(1) Recognition of trainees' characteristics. This is the starting point for educational and training institutions. Feature recognition mainly includes learner's basic identity recognition, learning habits recognition, learning preferences recognition, learning difficulties recognition and so on.

(2) Selection of training institutions. Learning starting point and learning path should also be different. For different needs of training projects, such as IELTS training, TOEFL training, computer training, adult education training courses, postgraduate entrance examination tutorial classes and so on. For different levels, choose basic classes, strengthen classes, sprint class, etc. For different subjects, each kind of tutorial class can also be refined. Different parallel classes can be chosen according to different locations and choices of teachers.

4.3. Learning management
(1) Selection, matching and remote sharing of learning resources. There is no learning resource suitable for all learners, and the learning resource suitable for oneself is the high-quality learning resource. Therefore, an important task of educational and training institutions is the screening of learning resources. The core of screening is to evaluate and classify the quality, form and difficulty of learning resources. By matching the selected resources with the learner's learning path, learners can begin to learn. The better the match, the better the learning effect. Distance sharing can be used for many purposes. One is to review and consolidate after class and do homework. Second, it can be used as an online distance education classroom. Third, we can exchange and learn from each other, upload information and so on. The resource sharing system based on big data of chain educational and training institutions is shown in Figure 1.

(2) Intelligent monitoring and intelligent service of learning process. Once learners have problems in learning path, learning difficulty, learning progress and learning effect, the intelligent platform can provide timely intelligent reminder, intelligent change of learning path and learning resources, online answering and other services.

(3) Intelligent evaluation of learning effect. Learning path and learning resources are individualized and intellectualized, so the evaluation of learning effect also needs to be intellectualized. For every point of learners' progress in the learning process and the adjustment of each learning path, the evaluation mechanism should follow up in real time and respond in time.
5. Key Problems to be Solved in Big Data of Chain Educational Training Institutions

Firstly, we need to filter out false, wrong and worthless data by big data filtering and big data mining technology, and improve the management security, efficiency and economy of unstructured and semi-structured data through cloud storage system.

Secondly, the trainee characteristic data, trainee behavior data, trainee relationship data, server operation and log data collected by sensor networks and monitoring devices have the characteristics of low value, fragmentation, heterogeneity and high redundancy. Thus, improving the comprehensiveness, scientificity, accuracy and timeliness of big data mining is the key to timely and quickly discovering the correlation between big data, the development trend of data relations, data sudden anomalies and the hidden value.

Thirdly, in the process of storing, adding, intercepting, deleting, retrieving and updating big data, the simple script language pretreatment of traditional data system cannot resolve too complex big data structure, nor can it meet the needs of multi-user high concurrent reading and writing, efficient storage and access of massive data, high availability and high scalability of the system in the era of big data. In addition, the optimization space and data operation ability of traditional database are limited, and the quality life cycle of big data cannot be effectively monitored. The Internet of Things, the Mobile Internet, Sensor Networks and Traditional Internet Networks have become managers.

6. The Relation between Big Data of Chain Educational Training Institutions and Cloud Computing

The application of big data in chain educational training institutions focuses on the collection, filtering, analysis, mining and knowledge discovery of user service-related data. It focuses on providing reliable big data support for the management of chain educational training institutions in terms of scientific decision-making, service guarantee of business departments. The application of cloud computing focuses on providing chain educational training institutions with supercomputing, mass storage, resource virtualization management and on-demand payment platform for big data processing, storage and mass data high-speed transmission by building private clouds or signing service agreements with cloud service providers. Therefore, big data and cloud computing are complementary and inseparable [5]. How to filter noise signals and improve the value density of big data based on cloud service platform
is the key to effectively reduce the cost of big data application and improve the efficiency of big data
decision-making in chain educational training institutions [6]. In addition, the integration and sharing of
big data in the cloud of chain educational training institutions, seamless connection of individual trainee
data, and the process of data analysts' supervision, analysis and value mining of big data are important
factors to ensure the availability of big data and scientific decision-making of chain educational training
institutions. Therefore, only relying on the advantages of cloud computing, such as super-large scale,
virtualization, high scalability, resource allocation on demand and low application cost, can chain
educational training institutions scientifically build a data processing system architecture with
distributed computing, storage and dynamic scalability, and realize distributed computing and storage
of big data in multiple nodes, thus ensuring the analysis and decision-making of big data. Efficient, fast,
real-time and economical.

7. Construction of Big Data Analysis and Decision Platform Based on Cloud Computing.
Analyzing the process of producing big data, the big data of chain educational training institutions are
mainly produced by students, sensor equipment, servers, monitoring equipment and so on. The platform
for analyzing big data of chain educational training institutions is mainly composed of five parts: data
source, collection layer, organizational component layer, processing layer and business layer. With the
increase of the total amount of big data, the cost of storing and calculating big data takes up a fast growth
trend in the total cost of big data application of chain educational training institutions. Therefore, cloud
computing architecture should be adopted in the design of big data distributed computing architecture
of chain educational training institutions to ensure that the big data analysis and decision-making system
has the characteristics of safety, reliability, high integration, reasonable configuration, low energy
consumption and easy expansion. Combining the design requirements of big data distributed computing
architecture of chain educational training institutions with the characteristics of cloud computing
technology, the target architecture of big data computing platform of chain educational training
institutions is shown in Figure 2.

Data acquisition (ETL), responsible for source data acquisition, cleaning, conversion and loading,
including:
1. Load the original data into Hadoop platform.
2. Loading the processed data into distributed database and master data warehouse
Hadoop cloud platform:
   It is responsible for storing massive traffic data, providing parallel computing and unstructured data
   processing capabilities, and realizing low-cost storage, low latency and high concurrent query
   capabilities [7].
Distributed Database (MPP):
Storage, processing, Association and aggregation of business data, and provide distributed
computing to support data depth analysis and data mining capabilities, export KPI and highly aggregated
data to the main data warehouse.
Master Data Warehouse (in conjunction with MPP):
Store index data, KPI data and highly aggregated data.
Data Open Interface:
Provide big data platform capabilities to big data applications.
Figure 2. Target Architecture of Big Data Platform for chain educational training institutions

Considering the characteristics of Big data environment and the decision-making needs of Big data in the target structure of Big data platform of chain educational training institutions, this paper adheres to the design principles of constructing a unified data operation platform, scheduling and allocating system resources according to business priorities, flexible expansion and allocation of system resources, low system construction and operation costs in the design of Big data analysis and decision-making system platform. The organization of big data analysis and decision-making platform system of chain educational training institutions based on cloud computing is shown in Figure 3.

Figure 3. chain educational training institutions Based on Cloud Computing Big Data Decision Analysis Platform
The system mainly consists of five parts: distributed infrastructure layer, infrastructure virtual layer, cloud basic service layer, big data analysis layer and big data decision layer. Virtualization infrastructure layer provides cloud data center infrastructure services for chain educational training institutions. It mainly consists of cloud virtual computing, cloud virtual storage, security system, basic services, platform services, software services and so on [8]. The upper virtualized resource management and scheduling module can scientifically allocate and manage the cloud resources for the operation of the lower system according to the security requirements of big data calculation, storage and data transmission of chain educational training institutions, so as to ensure the security, efficiency, economy and dynamic control of the operation of virtualized infrastructure. Based on the support of cloud computing technologies and services such as flexible cloud computing, virtual private cloud, flexible cloud storage, security and balance, cloud basic service layer provides three cloud service modes: infrastructure as a service, platform as a service and software as a service for chain educational training institutions [9]. Big data processing and analysis platform is based on the support of cloud basic service layer. Firstly, it filters, integrates and extracts the original big data to improve the value density and data availability of big data. Then, data processing, calculation, analysis and decision-making are completed through distributed system computing, real-time stream computing, big data storage and management, context search, etc [10]. Based on the support of the bottom of the platform, managers of chain educational training institutions can improve the efficiency, availability, controllability and economy of big data processing and application by visual analysis and development of big data, development of application programs and management of big data platform. Based on the support of the lower platform layer, the application layer of big data decision-making completes the selection decision-making of training institutions, the functional decision-making of training institutions, the analysis and prediction of trainees' reading needs, the analysis and prediction of students' needs, the evaluation of teachers, and the related big data decision-making of big data support for chain educational training institutions [11, 12]. It provides strategic decision-making, system management and operation, and the quality assurance of teachers' service for chain educational training institutions. Providing reliable big data decision support for sustainable development of proof and service productivity.

8. Conclusion
At present, the era of big data has come. As an important component of means of production, big data has become an important decision-making basis for chain educational training institutions to enhance service productivity, transform management and service modes, gain competitive advantages and build wisdom chain educational training institutions. With the growth of customer service demand and the change of service mode of chain educational training institutions, the big data environment of chain educational training institutions presents 5 "V" characteristics: Volume, Variety, Velocity, Value and Veracity. The rapid increase of the total amount of unstructured data and the proportion of big data will lead to the rapid increase of chain educational training institutions. The Big data environment of educational and training institutions is more complex and changeable. The difficulty of collecting, transmitting, storing, managing and processing big data is increasing rapidly. The scientific, controllability of cycle and return on investment of unstructured big data analysis are also affected. In order to ensure the scientific and efficient analysis and decision-making process of unstructured big data in chain educational training institutions, the chain educational training institutions should focus on the improvement of customer demand and service productivity, insist on "data-driven" as the basis of unstructured big data analysis and decision-making, and constantly strengthen the usability of unstructured big data in collection, noise filtering, value extraction and storage. Only by improving the real-time analysis, human-computer interaction, scientific evaluation and feedback optimization level of unstructured data with controllability management, can we effectively improve the scientific and usability level of big data decision-making in chain educational training institutions, and provide reliable big data decision-making support for personalized customer activities. Moreover, the technologies of super computing, mass storage, virtualization management and hot backup redundancy of cloud computing platform enable chain training institutions to scientifically and rapidly configure and
schedule cloud computing resource sharing pools (resource pools include network, server, storage, application software and cloud services) according to the needs of big data decision-making. Provide sufficient IT infrastructure and system resource support for big data analysis of chain educational training institutions.

Acknowledgements
This work was supported by the projects of National Natural Science Foundation of China (Grant No. 61300170), Anhui province higher education promotion plan (No. TSKJ2017B30), the Nature Science fund of Anhui province (Grant No. 1908085ME128), the Nature Science Foundation of Education Department of Anhui province (Grant No. KJ2018A0123), the Nature Science fund major project of Anhui province (Grant No. KJ2017ZD50), Shanghai Key Laboratory of Materials Laser Processing and Modification (Grant No. MLPM2017-3), Anhui Natural Science Foundation (Grant No. 1908085MF183), the Safety-Critical Software Key Laboratory Research Program (Grant No. NJ2018014), Anhui Polytechnic University Program (Grant No.2015YQ18).

References
[1] Jagadish H V, Gehrke J, Labrinidis A, et al. big data and its technical challenges [J]. Communications of the ACM, 2014, 57 (7): 86 - 94.
[2] Lynch, Clifford. Big data: How do your data grow [J]. Nature, 2008, 455 (7209): 28 - 29.
[3] Sin K, Muthu L. APPLICATION OF BIG DATA IN EDUCATION DATA MINING AND LEARNING ANALYTICS - A LITERATURE REVIEW [J]. ICTACT Journal on Soft Computing, 2015.
[4] Davenport H, Thomas. How strategists use “big data” to support internal business decisions, discovery and production [J]. Strategy & Leadership, 2014, 42 (4): 45 - 50.
[5] Silva Y N, Reed J M, Tsosie L M. Map Reduce-based similarity join for metric spaces [C]//Proceedings of the 1st International Workshop on Cloud Intelligence. ACM, 2012: 3.
[6] The rise of “big data”on cloud computing: Review and open research issues [J]. Information Systems, 2015, 47: 98 - 115.
[7] Chikhale M R. Data Mining: Exploring Big Data Analytics, Hadoop and Mapreduce[J]. International Journal of Engineering Sciences & Research Technology, 2014, 3 (8).
[8] Huerta-Canepa G, Lee D. A virtual cloud computing provider for mobile devices[C]// Acm Workshop on Mobile Cloud Computing & Services: Social Networks & Beyond. ACM, 2010.
[9] Tian W, Su S, Lu G. A Framework for Implementing and Managing Platform as a Service in a Virtual Cloud Computing Lab[C]// International Workshop on Education Technology & Computer Science. IEEE, 2010.
[10] Janssen M, Voort H V D, Wahyudi A. Factors influencing big data decision-making quality [J]. Journal of Business Research, 2017, 70: 338 - 345.
[11] Monino, Jean-Louis. Data Value, Big Data Analytics, and Decision-Making [J]. Journal of the Knowledge Economy, 2016.
[12] Lee S, Kim N, Kim J. A Multi-dimensional Analysis and Data Cube for Unstructured Text and social media[C]// IEEE Fourth International Conference on Big Data & Cloud Computing. IEEE, 2015.