Application of Big Data Platform in Course of Java Language Programming

Dingbao Lin
Xiamen Institute of Technology, Xiamen China

Abstract—**This paper designed a big data teaching platform which adopts Hadoop technology in big data technology as its core. It provides a method to search teaching video data and integrates security technology and intelligent recognition technology. The teaching functions of this platform include teaching audio-visual presentation, intelligent data recognition, data mining sharing, data storage sharing and so on. The platform has strong advancement. To test the actual effect of this teaching platform in teaching Java language Programming, teaching contrast method was adopted. The class of computer specialty was chosen for teaching application analysis. The results show that the platform has good teaching applicability and significantly improves students' examination scores, learning interest, autonomous learning ability and innovation ability, so it owns very good application prospect.**

Index Terms—**course of Java Language Programming, big data platform, Hadoop technology, teaching application**

I. INTRODUCTION

Big data technology, also called mass data mining technology, is characterized by large capacity, multiple types and high timeliness [1]. After 2012, the word big data was mentioned more and more. Even American New York Times and The Wall Street Journal as well as Chinese Sinolink Securities and Galaxy Securities gradually introduced it in reports and started to apply big data technology. Flourishing big data technology also attracted extensive attention of educational circles. Domestic and overseas relevant researches sprang up continuously. In April 2012, the Office of Educational Technology under US Department of Education issued Teaching Promotion Through Educational Data Mining and Analysis (public comment draft) and introduced the feasibility of big data application in education field [2]. In China, researchers [3-4] indicated in succession that the introduction of big data technology in education field can facilitate educational informationization course and provide the specific direction for scientization of education decision making.

Mellody applied big data technology in teaching material specialty and proposed this technology contributed to improving teaching effect of material specialty [5]. Based on a research report about 1,093,126 English writings on the same topic written by Chinese EFL learners in 2015, Wang and Zhang explored how big data technology could promote EFL writing teaching and gain higher efficiency in teaching, stronger motivation in learning and better competence in applying [6]. Some researchers combined big data technology with chemistry course teaching, established students' learning feedback system, practice test, virtual environment related to chemistry course and net-

work course evaluation on the basis of big data technology. Through this system, the teacher could know students' learning situations. Good teaching effect was gained [7].

The above researches on big data technology are the most cutting-edge research achievements in this field. However, it is easy to find from existing researches that big data technology also has some defects in teaching application. Some researches [8] indicate that in existing data search methods, users need to determine the key words for search first, then input the key words in user interface provided by the search engine and finally screen out qualified data from mass data through the search engine. Further, users can choose data from the data screened out. In addition, watching teaching video is essential when multimedia teaching is applied. However, existing data search methods are not combined with video, and data search cannot be conducted on the basis of the video which is being played. For example, students hope to deeply understand and learn a knowledge point in the teaching video and expect to search relevant expansion of the knowledge point in video watching process, but they are not certain about the detailed name of the knowledge point. In such case, their learning progress and knowledge mastery may be affected. Meanwhile, research achievements on application of big data technology in teaching are mostly based on theoretical analysis, and there are few practice research reports. Thus, practicalness of relevant researches is insufficient, and some defects exist.

Based on existing researches, this paper designed a teaching platform by depending on Hadoop technology in big data technology. It tries to provide a method to search teaching video data, including selecting local video image in video image in video playing process; gaining the attribute information of local video image; conducting data interaction with data platform according to the attribute information so as to get data objects matched with the attribute information from the data platform. Besides, this paper studied the application effect of this platform in Java Language Programming to highlight the innovation and advantage of big data platform in teaching college computer specialty.

II. ANALYSIS OF BIG DATA PLATFORM TECHNOLOGY

A. Hadoop technology

The core of big data platform is Hadoop technology [9]. Since big data technology is still a new technology, a uniform technical framework fails to form now. Hadoop technology system is considered to be a mature technology framework system which can represent and popularize big data technology. Figure 1 shows big data structure chart based on Hadoop technology.

Figure 1 shows structural frame diagram of big data technology which adopts the form of multilayer protocol architecture. It completes data source collection, integration, management, compiling, analysis and application output from bottom to top. Different from general computer protocol management framework, the characteristics of Hadoop technology are very distinct in the above structure. For example, MAP REDUCE programming model is a very characteristic program compiling module in big data platform. The model can process biog data in a parallel manner in large-scale cluster data through reliable and high fault-tolerant mode. Original data are extracted through mapper, filtered, processed and transformed. Besides, output is processed in reducer function. Only two steps are adopted to complete data compiling and processing, which is very efficient.

B. Database technology

To achieve effective storage, downloading and transmission of platform teaching data, this system adopts SQL Server database technology to construct database management subsystem. SQL Server is a relational database management platform and can provide more secure storage function, with good flexibility and integration [10]. During construction of this system, SQL Server technology was applied to construct model system of B/S architecture. Meanwhile, Analysis Services technology provided by SQL Server was utilized. The system can fast analyze data sources collected and provide online analytical processing and data mining function for teaching information.

The above online analytical processing and data mining functions are also the core composition modules of this system. The online analytical processing function can make this system carry out information exchange with other host computers in the range of internet and LAN to gain diversified resources of big data teaching platform. Data mining function can provide matched data collection, processing, statistics and output assist for intelligent recognition after users capture the screenshot so as to provide convenience for teaching resource recognition and quick positioning.

C. Intelligent recognition technology

To make corresponding data information processed fast in teaching process after users capture video screenshot, this system adopts intelligent recognition technology as corresponding support. This technology, based on artificial intelligence system, carries out intelligent processing of video screenshot captured by users through data recognition, processing and statistics, and conducts accurate matching query in big database platform. On this basis, this technology achieves exchange of image information captured and information data in database platform, and meets users' requirement for fast search of data object. Under the support of this technology, teachers or students can directly apply big data system to search relevant information and carry out association teaching in teaching video playing process.

III. DEMAND ANALYSIS OF BIG DATA PLATFORM

A. Video demonstration function

Video demonstration function is a key function of this system. This system is mainly applied in teaching Java Language Programming. In the teaching process, the teacher needs to play the content about course teaching
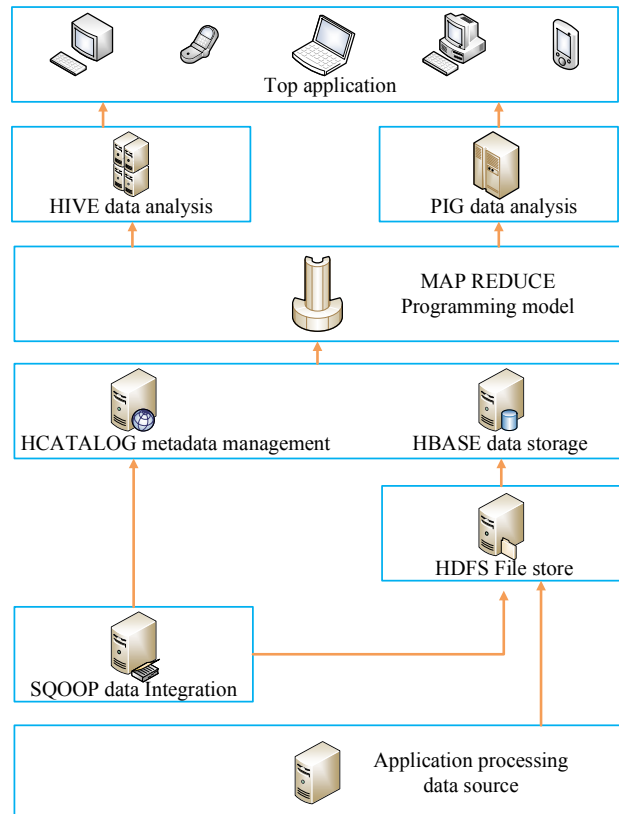

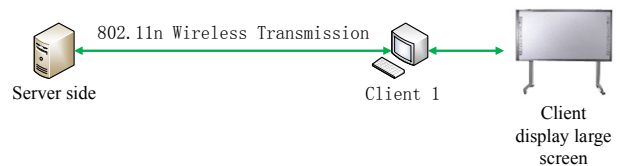Figure 1.   Big data structure chart based on Hadoop technology

Figure 2.   Video demonstration function of this system

with video demonstration function. Thus, multimedia video demonstration function of this system is very important. However, different from general multimedia teaching system, this system is based on big data technology and applies B/S structure in the design process. Video demonstration function of this system has such features as 3D structure and image processing function diversification. Fig.2 shows video demonstration function of this system.

As shown in figure 2, video demonstration function of this system is based on standard B/S architecture of computer system. The system consists of one server and 1 client. The two communicate with wireless transmission technology. The client is connected with a large display screen. Through wireless transmission and image processing, the client can download diversified teaching resources from the server in due time and display the teaching resources efficiently so that the efficiency of resource switching, selection and demonstration is higher in the teaching process.

B. Image capture function

Image capture is a major teaching function of this system. The features of this function are as follows: in the teaching process, users can choose and capture the details in video image at random, while artificial intelligent recognition technology integrated by the system can

search relevant information in big database platform according to image details captured by users. For example, the teacher captures "Java language application" in teaching. After the search, the system will skip to data search page about "Java language application". This page contains lots of relent pictures, webpages, videos and texts. The teacher can choose relevant information as teaching resources according to teaching needs.

C. Data interaction function

Data interaction function is another teaching function of this system. This function is based on Hadoop big data operation platform which includes mass data object information and is connected with internet search platform such as Baidu platform and Google platform. In the specific application, once users capture the image and text information which needs to be matched, big data platform will recognize the information and retrieve local video image information and entire image information respectively according to platform resources. On this basis, the system data interact with other data in big data platform, and the data object matched with the above information is gained from the data platform. Since system data interaction function has diversification feature, data objects retrieved are also diversified. They may be entity object information of Java language such as program statement and programming thought, or subsidiary information such as functional text description behind the specific object.

D. Resource storage sharing function

Since the system integrates SQL Server technology, it can store resource information generated in teaching process. Meanwhile, this system contains wireless transmission module. In the teaching link, big data platform resource downloading, sharing and uploading functions can be completed. With the above functions, teaching resources of Java Language Programming can be shared and stored. After multiple rounds of teaching, exclusive course teaching resource library can be created for teachers and students.

IV. BIG DATA PLATFORM DESIGN

A. Overall platform setup

Big data technology selected by this system focuses on Hadoop. Hadoop is a data processing architecture which was developed by Apache Foundation and is specialized in big data distributed system. Besides, it is also an open source software library. In overall platform setup, designers can log in relevant official websites to download software program platform and develop virtual big data human-computer interface environment. In the specific development, this platform takes Windows system as the front-end operating system, simulates big data teaching environment with the help of Cygwin software in Hadoop architecture. The setup of the whole teaching platform involves the following steps:

Firstly, install and configure Cygwin. Cygwin [11] is an interactive simulation environment and can support each version of Windows system. With this environment, front end and backstage management work of platform can be achieved. The process of installing and configuring Cygwin is also very convenient. It is just necessary to log in Cygwin official website, download Cygwin version corresponding to 64-bit WIN7 system and click to install.

Secondly, install and configure Hadoop. Hadoop is a core module of this system which is applied for big data resource analysis and processing. In the specific installation, it is necessary to log in Apache Hadoop to download. This system chooses the latest Hadoop 2.6.4 version for installation. The figure 3 shows downloading interface screenshot on the official website.

After downloading, Hadoop configuration file needs to be decompressed. Finally, users can start Hadoop in Cygwin and operate ./start-all.sh to start Hadoop. Later, configuration of Ant, TortoiseSVN, Eclipse and other sub-environments is completed respectively according to installation prompts. In this way, big data resource program compiling can be achieved under Cygwin environment. After external system of platform is set up,

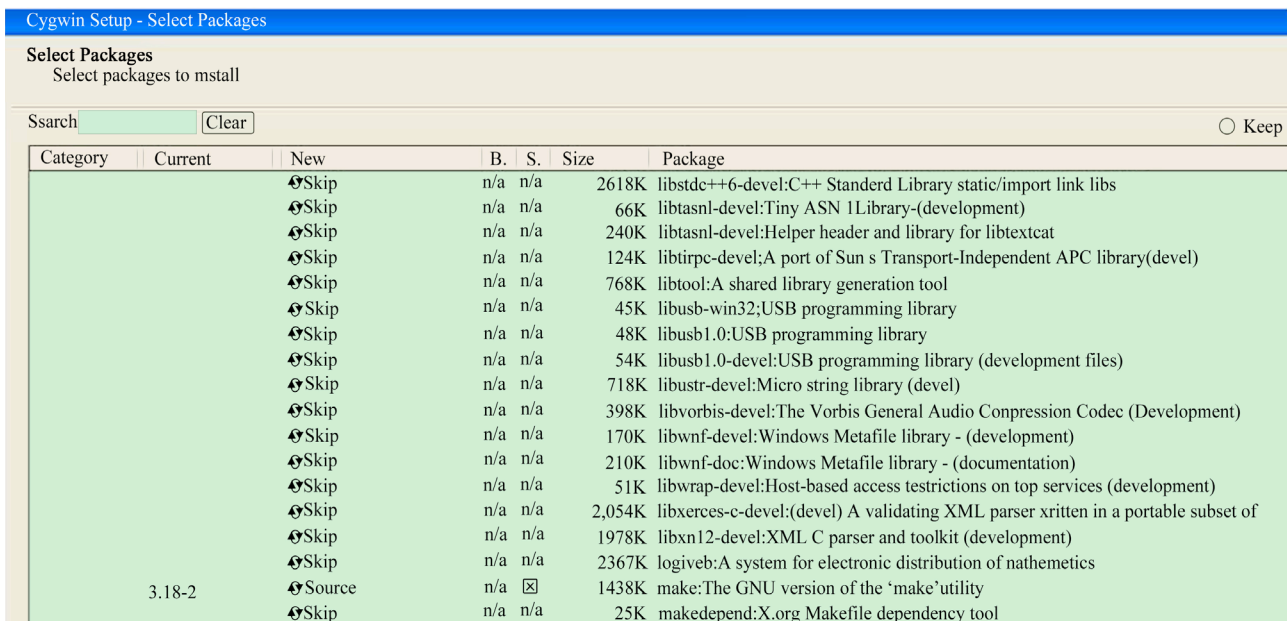| Category	Current	New	B.	S.	Size	Package
		↻Skip	n/a	n/a	2618K	libstdc++6-devel:C++ Standerd Library static/import link libs
		↻Skip	n/a	n/a	66K	libtasnl-devel:Tiny ASN 1Library-(development)
		↻Skip	n/a	n/a	240K	libtasnl-devel:Helper header and library for libtextcat
		↻Skip	n/a	n/a	124K	libtirpc-devel;A port of Sun s Transport-Independent APC library(devel)
		↻Skip	n/a	n/a	768K	libtool:A shared library generation tool
		↻Skip	n/a	n/a	45K	libusb-win32;USB programming library
		↻Skip	n/a	n/a	48K	libusb1.0:USB programming library
		↻Skip	n/a	n/a	54K	libusb1.0-devel:USB programming library (development files)
		↻Skip	n/a	n/a	718K	libustr-devel:Micro string library (devel)
		↻Skip	n/a	n/a	398K	libvorbis-devel:The Vorbis General Audio Conpression Codec (Development)
		↻Skip	n/a	n/a	170K	libwnf-devel:Windows Metafile library - (development)
		↻Skip	n/a	n/a	210K	libwnf-doc:Windows Metafile library - (documentation)
		↻Skip	n/a	n/a	51K	libwrap-devel:Host-based access testrictions on top services (development)
		↻Skip	n/a	n/a	2,054K	libxerces-c-devel:(devel) A validating XML parser xritten in a portable subset of
		↻Skip	n/a	n/a	1978K	libxn12-devel:XML C parser and toolkit (development)
		↻Skip	n/a	n/a	2367K	logiveb:A system for electronic distribution of nathemetics
	3.18-2	↻Source	n/a	☒	1438K	make:The GNU version of the 'make'utility
		↻Skip	n/a	n/a	25K	makedepend:X.org Makefile dependency tool

Figure 3. Hadoop 2.6.4 downloading interface screenshot on the official website

platform connection certification is required to make sure the platform can interconnect with external network system during the application. The detailed access configuration codes are as follows:

(1) Add or modify configuration section <configSections>

<section

name="log4net"

type-"l094net. Config. L094 Net Configuration Section Handler"/>

<section name="casClientConfig"

type="DotNetCasClient.Configuration.CasClientConfig uration,DotNetCasClient"/>

(2) Add configuration section<log4net>

<log4net>

<appendername= "RollingFile" type="l094net. Appender. RollingFileAppender">

<file

value="logskDotNetCasClient.1094net.109"/>

<appendToFile value="true"/>

<rollingStyle value="Date"/>

<datePattem value="yyyyMMdd">

<lockingModel type=" l094net. Appender. FileAppender + MinimalLoek" >

<layout type="l094net. Layout. PattemLayout">

<eonversionPattem Value='% date[%thread]%-5level%logger{5}.

%message%newline¨/>

</layout>

</appender>

<root>

<! - Levels(from lowest to highest): ALL|DEBUG |INFO|WARN|ERROR

FATAL|OFF|-->

<level value: "DEBUG¨/>

<appender-ref ref=-"RollingFile"/>

</root>

</log4net>

Under the function of the above codes, the interconnection function of the whole platform and external network is constructed so that the system owns external connection function.

B. Video demonstration design

Video demonstration as front-end display system of this system can play course teaching video and animation resource, and provide students with visual teaching materials. In the specific design, this part regards B/S structure as the organization center to achieve remote downloading, uploading, sharing and playing of video resources. Moreover, multimedia high-definition projection screen is used as the demonstration hardware, and wireless device, artificial intelligence equipment and projection equipment are equipped to complete display and sharing of teaching video resources. Figure 4 shows the design diagram of video demonstration design module.

As shown in Figure 4, the wireless module in this part is mainly used to transmit video resources between the client and the server, while artificial intelligence module is
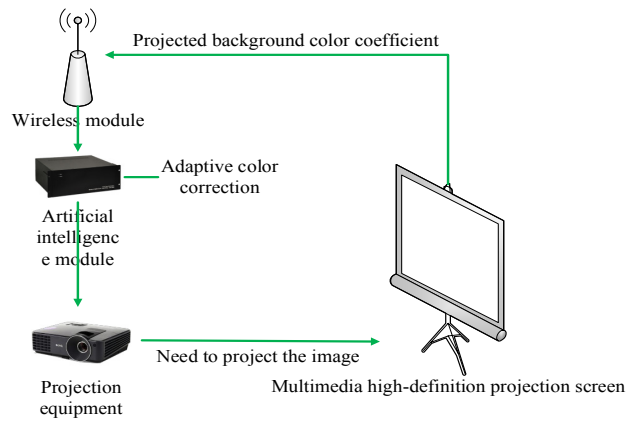

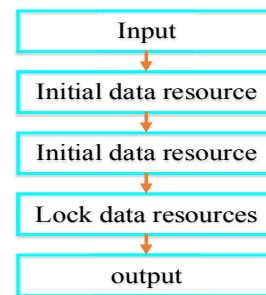Figure 4. Design diagram of video demonstration system

Figure 5. Big data collection analysis process design with MapReduce calculation model

applied for adaptive color correction to get the highest-definition video demonstration effect. In this way, video demonstration quality of this system improves greatly.

C. Design of data collection analysis

As stated above, the design of data collection analysis part is mainly implemented with artificial intelligence recognition technology and big database search technology. In this process, once users capture detail module in the video, the system will call MapReduce calculation model in Hadoop technology to recognize can compute the detail module. The specific data collection analysis mode is shown in Figure 5.

As shown in Fig.5, MapReduce model is called for data analysis during data collection analysis, which needs multiple processes. The above processed can be subdivided into map stage and reduce stage. Initial data collection is mainly accomplished in map stage, while data matching and locating are finished in reduce stage. The data are finally outputted in the form of <key, value>. In addition, the system needs to acquire data information of exchange platform through the module in design of data collection analysis module. The detailed operation thought is as follows: establish a database especially used for data exchange; when the system needs to connect the data on another big data platform, it requests to gain data through calling data request interface on external hardware of database.

Under the function of the above program, communication platform is mainly responsible for data exchange and transmission services in the internal part of data exchange platform. In this way, information exchange and transmission between this system and other big data platforms of internet can be achieved. The application interface of big data platform teaching system is shown in Figure 6.

Figure 6.   application interface of big data platform teaching system

V.   ANALYSIS OF BIG DATA PLATFORM TEACHING APPLICATION

To test the teaching effect of this system in Java Language Programming, Class (3) (experimental class) and Class (4) (control group) of 2014 computer specialty in XX University (XX Province in China) were chosen for teaching application analysis. In the specific practice, traditional teaching method was applied for Class (4), while big data platform was adopted for Class (3). The detailed teaching process includes the following:

Firstly, case display. The teacher applied video demonstration module of the system to display teaching case to students. Secondly, data recognition. When the teacher explained some key knowledge points, or students had doubts about a program in video resource, the teacher chose corresponding teaching content through data collection recognition function of this system, and the system would automatically skip to big data search page for automatic search. Thirdly, result display and sharing. After the search, the teacher showed search result to students on video demonstration interface. Besides, the teacher might utilize research result to carry out diversified associative and disperse teaching. Fourthly, resource sharing and reconstruction. In teaching process or at the end of teaching, the teacher could apply the platform to store students' learning experience, upload it to big data platform and share it as user-defined teaching resource.

After the teaching for 48 class hours was finished, the teacher conducted statistics of learning indexes of students in both classes, and summarized the results as shown in Tables I and II.

According to Table I, the average score of final examination of experimental class is significantly superior to that of control class, and the statistical results have differences. The data in Table II truly reflect big data platform learning of experimental class. In other words, the introduction of this platform can improve students' learning interest in Java Language Programming, and contributes to enhancing their autonomous learning ability and innovation ability. Thus, this platform plays an important role in improving teaching quality of Java Language Programming.

The above research results fully embody the teaching effect of big data platform in Java Language Programming. As informatization and datamation develop and popularize

TABLE I.
STATISTICAL TABLE OF AVERAGE SCORES OF FINAL EXAMINATION OF EXPERIMENTAL CLASS AND CONTROL CLASS (SCORE, ---$,X+S$)

Index	Average score		t	P
	Experimental class	Control class		
Theory score	44.5	21.7	3.35	0.001
Practice score	46.3	30.3	3.51	0.002
Total	90.8	52.0	4.12	0.003

TABLE II.
LEARNING EFFECT SELF-EVALUATION IN EXPERIMENTAL CLASS (SCORE, ---$,X+S$)

Evaluation item	Experimental class
Improve learning interest	9.45+2.35
Beneficial to improving autonomous learning ability	9.21+1.38
Beneficial to improving innovation ability	8.92+3.21

continuously, the introduction of big data technology in course teaching of computer specialty can not only enrich teaching means and improve teaching advancement, but also collect mass teaching information in a short time by intelligence and large capacity of big data, provide more basis for teaching design, enhance dynamic sense and int4ractive sense of classroom teaching, create more diversified teaching situation and facilitate great leap of course teaching effect. Meanwhile, this paper offered a method to search the video to be played form big data platform. Students can gain data object matched with the attribute information from the data platform in the process of watching teaching video. Such technology can help students to master new knowledge fast and enhance their autonomous learning ability and learning interest. Of course, since successful cases about application of big data technology in teaching are few in practice, teachers need to continuously sum up the experience, and rationally choose teaching methods in order to reach the purpose of optimizing teaching quality and cultivating practical and innovative talents for computer specialty.

VI.   CONCLUSIONS

Big data teaching system is a new intelligent and information-based teaching system which is different from previous teaching platforms. Teaching experiment contrast results in this paper also verify that compared with traditional multimedia equipment teaching, this system has remarkable value for improving teaching quality of Java Language Programming. Therefore, this system is worth further popularizing in other courses of college computer specialty.

REFERENCES

[1]   Abuín, J.M., Pichel, J.C., Pena, T.F., & Amigo, J., "BigBWA: approaching the Burrows–Wheeler aligner to Big Data technologies," *Bioinformatics*, vol. 31, no. 24, pp. 4003-4005., December 2015. http://dx.doi.org/10.1093/bioinformatics/btv506

[2]   Pietraszek, T., & Tanner, A., "Data mining and machine learning—towards reducing false positives in intrusion detection," *Information security technical report*, vol. 10, no. 3, pp. 169-183, September 2005. http://dx.doi.org/10.1016/j.istr.2005.07.001

[3]   He, Z.L., Xiao, X.H., He, Y.H., "Research on Information Technology with the Consideration of Computer Professional Education in Big Data Era," *Advanced Materials Research*, vol. 1014,

pp. 429-433, July 2014. http://dx.doi.org/10.4028/www.scientific.net/amr.1014.429

[4] Chen, S.Y., Meng, L.J., Zhang, H.Y., "Educational policy evidence in big data era: enlightenment of evidence-oriented concept on China's education management modernization and decision-making scientification," *Global Education*, vol. 43, no. 2, pp. 121-128, April 2014.

[5] Mellody, M., Council, N., "Big Data in Materials Research and Development: Summary of a Workshop," *Teaching Statistics*, vol. 17, no. s1, pp. 8-12, October 2014.

[6] Wang, Z., Zhang, Y., "Big Data Methodology and EFL Teaching Innovation— A Case Study of One Million-odd EFL Writings on the Same Topic," *Technology Enhanced Foreign Language Education*, no. 5, pp. 3-8, September 2015.

[7] Cooper, M.M., Underwood, S.M., Bryfczynski, S.P., & Klymkowsky, M.W., "A short history of the use of technology to model and analyze student data for teaching and research," *Tools of chemistry education research*, vol. 1166, pp. 219-239, April 2014.

[8] Zhang, Y., Cai, X., Li, A.M., Liu, Y., Yao, X.H., "Application of Distributed File System MPP Database Mashup Architecture in Telecom Big Data Platform," *Telecommunications Science*, vol. 29, no. 11, pp. 12-16, December 2013.

[9] Leverich, J., & Kozyrakis, C., "On the energy (in) efficiency of hadoop clusters," *ACM SIGOPS Operating Systems Review*, vol. 44, no. 1, pp. 61-65, January 2010. http://dx.doi.org/10.1145/1740390.1740405

[10] Qi, X.A., Hou, Q.J., "The Use of SQL Server Database," *Manufacturing Automation*, vol. 32, no. A7, pp. 30-32, March 2010.

[11] Wang, F.G., Yao, L., & Ma, Z.C., "Simulation Platform for Intelligent Electricity Consumption Terminal," *Jisuanji Xitong Yingyong- Computer Systems and Applications*, vol. 21, no. 7, pp. 127-130, July 2012.

AUTHOR

Dingbao Lin is a Lecturer of Xiamen Institute of Technology, Xiamen 361021, China His research interests include Big Data and Java Language. (e-mail: abcdlindb@tom.com)