Establishment of a M-Learning Model Supported by Big Data

Hanwei Dong, Zhuying Lin, Zhangqin Huang
Beijing University of Technology. Software Engineering. Beijing, China

ABSTRACT: Development of technology including big data and mobile internet has brought convenience to people for the requirement of mobile learning (m-learning) via portable devices. However, issues regarding storage and sharing of mass data has turned out to be a hindering factor on the development of m-learning application. Targeting at storage and sharing of educational resource data, the essay provides a solution to m-learning on a big-data platform, hence set up a m-learning model supported by big data technology.

1. INTRODUCTION

It is becoming a trend to study via the Internet in the educational territory. Types of online educational resources not only include text documents, but also cover multimedia files such as videos, audios, and pictures. Huge quantities of data which expands in an exponential scale has brought huge challenge to its storage requirements. Traditional storage solution does not meet such big-data storage sharing requirements and obstacles emerge on educational resource sharing and categorizing. To solve such problems, cloud computing—a new type of computing method based on the Internet came to life. It adopts multiple computing and network technology, e.g. Grid computing, distributed computing, parallel process, online storage and load balance, henceforth brings possibility to mass data storage and processing.

As an fundamental distributed architecture for big data, Hadoop composes a powerful, distributed cluster system upon inexpensive general hardware resources. Users are able to develop distributed program without knowledge to the details of lower levels. The platform, which is mainly used to storage and computing on mass data, is capable to process not only on structured data, but also on non-structural data such as video, audio and text, and further provides reliable service to the requirements of online learning or m-learning in an big-data environment. It is this essay’s intention to adopt such a big-data platform and figure out one solution to data storage and sharing on educational resources for m-learning, and establish a m-learning model which is big-data supported.

2. ANALYSIS AND DESIGN ON ARCHITECTURE

2.1 Background and Method

As a most innovated method, m-learning has attracted most attention in this big-data supported environment. However, m-learning platforms available nowadays provide only a few basic functions and cannot meet actual requirements that users raise during their learning process. At the same time, newly-emerged educational patterns, smart mobile applications, social network, online storage systems, together with massive open online courses (MOOCs) which just appeared recently, have provided large quantities of data and application, and has brought great difference to education. It is statistically noted that, by June 2012, global smart phone user population exceeded 1,000 million and registrants on Facebook over 2,200 million. By the end of 2014, there had been more than 10 million users who registered on the MOOC website Coursera. It is obvious that human being has stepped into a big-data educational era.

It is therefore necessary to set up a novel and personalized m-learning platform based upon the concept of a feasible model, to satisfy the practical requirements of m-learning users. To provide such a solution, it is correct to adopt big data, m-learning and other advanced technology to educational field and provides service to mass educational data.

2.2 Design of Educational Resource Model on Hadoop

The M-learning platform is set up upon Hadoop, and is composed of a master unit and multiple slave units. Mass educational resource is stored inside HDFS which supports batch data processing and m-learning in a Hadoop computing model. Both
Map and Reduce functions operate in parallel to execute query and computing tasks. Not only does this architecture solve the problem of mass data storage, but also improves data throughput capacity, as well as reduces network load pressure, and thus provides better experience for teachers and students to get access to educational resources. The model of Hadoop-based educational resource architecture is listed in Figure 1.

<table>
<thead>
<tr>
<th>user level</th>
<th>teachers</th>
<th>students</th>
</tr>
</thead>
<tbody>
<tr>
<td>cloud terminals</td>
<td>smart phones</td>
<td>pads servers laptops</td>
</tr>
<tr>
<td>cloud service layer</td>
<td>Internet</td>
<td>mobile applications</td>
</tr>
<tr>
<td>application layer</td>
<td>Internet explorers</td>
<td>Logical abstraction layer</td>
</tr>
<tr>
<td>resource abstraction layer</td>
<td>Hadoop</td>
<td>Physical resource management</td>
</tr>
<tr>
<td>education resources</td>
<td>Audio Video PPT Text</td>
<td></td>
</tr>
</tbody>
</table>

Figure 1 The model of Hadoop-base educational resource architecture

Teachers and students at the user level may get access onto cloud service layer via cloud terminals such as smart phones, pads, servers or laptops, and submit requests on data via Internet explorers and mobile applications to resource abstraction layer, big data cloud storage system. The lower-level system responds to clients’ requests, calls related application and provides proper services such as index, management, storage and backup. The system and Hadoop platform combine to fulfill user requests and feedback results to the terminal. This well extensible architecture can solve storage problems and found a solid basis for the development of m-learning.

2.3 Design of an m-learning model

Hadoop can solve the problem of mass data storage and thus sets up a good basis for a m-learning platform. Figure 2 illustrates a graph of m-learning architecture platform.

It is explicated that teachers can view, add or delete his/her own courses, and manage course contents. Resources besides course scope can also be uploaded for all viewers to get access. Meanwhile, teachers are granted to review homework materials submitted by students and make comments. Messages will be sent and conversations will be made when necessary. Besides teachers may choose to join discussion groups and answer questions raised from students.

Students, on the other hand, may add, view or delete his/her courses, download shared materials and view notifications. Students are also welcomed to share documents with other users. Students will submit homework, review remarks and interact with other users such as form a discussion group.

Administrators are obliged to categorize different users using different labels, and provide registration\logon service. The privilege of this role is to review all shared resources, publicize or delete general-available notifications and other management tasks.

3. MODELING OF BIG-DATA CLOUD STORAGE

3.1 A cloud storage model based on Hadoop

In accordance with analysis result to distributed file system together with computing model of Hadoop, the essay proposes a cloud storage model based on Hadoop. The model is set up with different logical nodes, one as master node and the others as slave nodes. Master node is consisted of Namenode and Jobtracker. Namenode, which is composed with high-performance servers, is mainly responsible to administrate metadata of the distributed file system of Hadoop. Jobtrack, as the master to MapReduce, is mainly used to initiate, monitor and dispatch different TaskTracker processes, and assign designated Map & Reduce tasks. The system also adopts a secondary master which is meant to prevent possible service suspension. Figure 3 illustrates the functional model of the cloud storage system.
During the application of this model, original educational resources including PPTs, videos and audios are stored into distributed file systems. Namenodes from the Master set up indexes of the file system to manage data stored inside different Datanodes. Submitted requests raised from teachers and students are self-defined APIs, which are user-personalized Map and Reduce functions. JobTracker forward Map and Reduce functions to slaves to receive final results after MapReduce computing.

3.2 Design of cloud storage platform via Haodop.

In an laboratory experiment, it is the essay’s intention to set up a secure and reliable mass education resource storage system based upon Hadoop. The platform need to be highly reliable to fully integrate different types of resources and provide storage and access solutions, and provides unified access APIs to terminal users. The storage platform needs to be scalable to adapt to dynamic transformation of physical resources, and to be fault-tolerant to effectively handle single-node malfunction, data loss or other problems. The cloud storage platform based upon Hadoop is functionally illustrated in Figure 4.

Key technology of Hadoop, HDFS, is a distributed storage system, which is able to store large quantities of educational data in different blocks, and store all metadata of the file system for a convenient administration. Backup to the stored resource is backup to prevent possible data loss and make sure all data is securely stored.

4. M-LEARNING BASED ON EDM

To set up a big-data cloud storage platform based on Hadoop does not only solves the problem of resource management, but also founds a technical basis to the development of m-learning.

Thanks to its advantageous physical traits such as touch ability, portability, Internet-access and functional sensors, as well as abundant applications, students enjoy brand new study experience. Acknowledged research result have explained that to utilize portable devices during the study process can attract students’ interest and focus on study, while can also help them better understand the knowledge required to grasp. This is why research on m-learning turns out so important in educational experience.

Time has witnessed the rapid development and adoption of data mining in different fields in the past few years. Driven by informatization and Web2.0, educational data mining (EDM) is paid increasing attention. Therefore, to adopt EDM into construction of a m-learning system, will be able to provide full-scale information gathering and analysis, moreover better predict user habits and personalize configurations.

4.1 EDM technology

EDM mainly process data. Different educational environment let out diversified resources, e.g., documents, microblog messages, audios or videos. From a perspective of data mining, EDM should be scheduled as three different passes, pre-procession, data mining, and evaluation. Pre-procession is a process that alters original data into a data-mining available format. This process is usually labeled as four different phases, data clearance, data integration, data reduction and data alternation. Because data mining heavily relies upon the quality of data to be analyzed, pre-procession is considered to be a vital part of the entire process.

Intention of data mining is to set up models from analyzed data, which are mainly used for prediction or description. A prediction model is used to speculate unknown data based upon known ones, while a description model is used to find out new pattern from known data sets. There are different data mining methods adopted by the two models, such as classification, regression, clustering, and correlation.

The experimental data is categorized into three different sets, i.e., training set for model training, verification set for model optimization, and test set for evaluation.

Nowadays, EDM is adopted into below-listed fields:
1. Visualization: to demonstrate information or knowledge in a figurative way.
2. Students Modeling: to locate a student’s study...
pattern via retrieved data through automated modeling work, in which Bayes Network, Sequence Pattern Mining and Logic regression are mainly adopted to better form a personalized model.

3. Students prediction: to make predictions upon known data, e.g., to make predictions on grade or academic performance according to recorded information.

4. System recommendation: to recommend courses, learning documents or methods

5. Self-adaptive system: to automatically adapt the system itself to better facilitate students’ usage

4.2 EDM in m-learning

Upon research to EDM technology, it is discovered that to adopt EDM in m-learning does not only benefit teachers by allowing them to better evaluate students’ study performance, but also provides better assistance to students for higher study efficiency by prediction, evaluation and recommendation.

The recommendation system is composed of presentation layer, recommendation engine, business logic layer, learning resource layer and monitoring layer. Figure 5 illustrates the thorough design of this system:

![Figure 5 EDM-based recommender systems design](image)

With the development of mobile technology, m-learning as a brand new study method, has been accepted by more and more educational participants. The pattern combined with recommendation cooperative filtering algorithms, forms a sub-module on a big data supported, m-learning platform, which is directed by the idea of adopting EDM to form a personalized study platform.

5. CONCLUSION

A well-established m-learning platform should be set up by well-designed architecture, interaction pattern and educational resources. This platform does not only provide study resources, but also form a new, harmonious and long-lasting environment to better facilitate educational participants to communicate in a more convenient way, and eventually achieves the goal for any person to be able to study whenever, wherever, whatever possible.

REFERENCES