International Journal of Multimedia and Ubiquitous Engineering
Vol. 9, No. 12 (2014), pp. 73-84
http://dx.doi.org/10.14257/ijmue.2014.9.12.07

Design and Implementation of a Dynamic Educational Content Viewer with Big Data Analytics Functionality[1]

Seong-Yeon Ju[2], Min-Ho Song[3], Ga-Ae Ryu[3], Mihye Kim[5] and Kwan-Hee Yoo[4,3,2*]

[2]Department of Business Data Convergence, Chungbuk National University
[3]Department of Digital Informatics and Convergence, Chungbuk National University
[4]Department of Software Engineering, Chungbuk National University
410 Seongbongro, Heungukgu, Cheongju, Chungbuk, South Korea
clover1007@chungbuk.ac.kr, songminho@chungbuk.ac.kr, reflexive@naver.com, khyoo@chungbuk.ac.kr
[5]Computer Science Education, Catholic University of Daegu
13-13 Hanyangro, Hayangeup, Gyeongsansi, Gyeongsangbukdo, South Korea
mihyekim@cu.ac.kr

Abstract

The use of smart mobile devices has grown considerably in recent years. A large variety of digital content has entered the market and, in particular, there has been rapid growth in sales of electronic books (e-books). Consequently, demand for more diverse digital content has increased greatly as has the demand, in particular, for smart digital content for educational purposes, including digital textbooks. This demand has been accelerated through the implementation of an educational environment termed 'Smart Education' in Korea. Digital content and digital textbooks can be displayed on smart devices using a content viewer application or an e-book reader. However, existing content viewers only support a limited range of features, most of which are platform-dependent. In addition, most content viewers do not support features that deliver analysis of big data. To address these shortcomings, here we propose a dynamic educational content viewer, which can implement big data analytics and supports multiple device platforms and a wide range of content types.

Keywords: *Content viewer, E-book viewer, Educational content viewer, Big data analytics*

1. Introduction

There has been much growth in the use of smart mobile devices in recent years, and the range of digital content available on mobile devices has also grown at a considerable pace. The electronic book (e-book) industry has also grown rapidly, spurred in part by the revival of Amazon Kindle in 2007 [1, 2]. As a consequence, demand for more diverse digital content has also grown, and demand for smart digital content for educational purposes, including digital textbooks, has grown especially quickly [3, 4]. In Korea, this has been accelerated through the implementation of an educational environment called 'Smart Education', which employs diverse and complex educational digital content.

---

[1] This paper is a revised and expanded version of a paper entitled "An Embedding Method of Dynamic Content through Big Data Analytics[13]" presented at CES-CUBE 2014, Spain, 16~20 July, 2014.

* A corresponding author

ISSN: 1975-0080 IJMUE

E-books and digital textbooks can be viewed on smart devices using a number of existing e-book viewers and content viewer applications. The most popular e-book viewers include Apple's 'iBooks 2' and Readium's 'Readium'. However, these have some limitations: iBooks2 can only run on iOS-based platform devices [5] and Readium has problems displaying interactive three dimensional (3D) objects [6], even though it supports various features of the electronic publication format 3 (EPUB3.0). In addition, neither viewers support features to process big data analytics. To address these shortcomings, we have developed a dynamic educational content viewer that can implement big data analytics and supports multiple device platforms, as well as a variety of 2D and 3D objects. This educational content viewer was implemented as a part of the cloud-based smart education system described in Refs. [7, 8]. The system enables delivery and sharing of various types of educational content through a cloud-based educational media service environment [8].

To support a wide range of educational content on multiple device platforms, we designed a content viewer based on Extensible Markup Language (XML) [2]. XML is a document format that allows information exchange among multiple devices, with applications including web services and businesses. It is straightforward to extend the XML-based content viewer so that it can interpret documents in formats including HTML and XHTML using Extensible Style sheet Language Transformations (XSLT). This enables us to display various types of XML-based educational content on multiple device platforms. That is, it is possible to develop a range of educational content that is compatible with devices running different operating systems. To broaden the range of features of our educational content viewer, we enlarged the range of 3D object control features using Open Graphics Library for Embedded Systems (OpenGL ES) rendering. In this way, various 3D objects, including interactive 3D objects can be immediately viewed and aligned with 2D-based objects, such as text, images and Scalable Vector Graphics (SVG). For big data analytics, we used Hadoop MapReduce, which is an open-source framework for processing big data. We then visualize the results of these analytics results on the content viewer using SVG. The content viewer described here is based the work reported in Refs. [2, 10-13].

The remainder of this paper is organized as follows. Section 2 reviews the theoretical background on which this study was based, including big data, content viewers, and the cloud-based content-oriented smart education system introduced in Refs. [7, 8]. In Section 3, we describe the overall configuration and major features of the dynamic educational content viewer. In addition, we describe the process of applying big data analytics to the content viewer. In Section 4, we present our experimental results, and we summarize the results of the paper in Section 5.

2. Theoretical Background

2.1. Big Data

With the daily use of smartphones, the Internet, and the growth of user-generated content, huge amounts of data are being created and we are approaching the "dawn of the Zettabyte era". It is forecast that, by 2020, the amount of data to be managed will grow by a factor of more than 50, and that the number of servers required will increase by a factor of 10 [14]. Wikipedia defines big data as: "A collection of data sets so large and complex that it becomes difficult to process using on-hand database management tools or traditional data processing applications. The challenges include capture, curation, storage, search, sharing, transfer, analysis, and visualization." [19]. The

McKinsey Global Institute refers to big data as: “Datasets whose size is beyond the ability of typical database software tools to capture, manage, and analyze.” [15, p.1], and International Data Corporation (IDC) describes big data as: “A new generation of technologies and architectures designed to extract value economically from very large volumes of a wide variety of data by enabling high-velocity capture, discovery, and/or analysis.” [16, p.1]. According to the recently updated Garner definition, “Big Data are information assets with volumes, velocities and/or variety requiring innovative forms of information processing for enhanced insight discovery, decision-making and process automation.” [17, 18].

To efficiently handle hundreds of gigabytes of data within a reasonable timeframe, technologies in machine learning, natural language processing, signal processing, parallel processing, and relational databases are required [9]. Indeed, many new technologies are emerging to cope with big data in a more effective and economical manner. These include Schema-less databases (or NoSQL databases), MapReduce, Hadoop, Hive, PIG, WibiData, and storage technologies [19]. We analyze large sets of data using Hadoop, which is a popular open source implementation of MapReduce [14, 19]. The distributed file system of Hadoop provides support to handle hardware malfunctions, streaming data, very large datasets, simple coupled models, and comparability between different types of hardware and software platforms [14].

2.2. Content Viewers

A content viewer, or e-book viewer, is an application for reading digital content, e-books, or digital textbooks. Following the revival of the Amazon Kindle in 2007 [1], the world e-book industry has seen much growth [2]. Many content viewers (e-book readers or digital viewers) are available, including Apple’s ‘iBooks 2’ and Readium’s ‘Readium’.

Apple’s iBooks 2 is a digital textbook viewer that supports interactive digital textbooks by displaying diagrams, as well as audio and video content. It was developed from iBooks by adding learning support features, including highlighting, note taking, content searching, and glossary and dictionary searches. Through multi-touch gestures, users can also use an interactive photo gallery and pop-up animation, and can manipulate 3D objects [20]. However, iBooks2 is only available on iOS-based devices, such as iPads and iPhones.

Readium is an e-book viewer developed by Readium.org, which is a not-profit organization. It uses the EPUB 3.0 format, which is based on Web standard technologies including HTML5, JavaScript, Cascading Style Sheets (CSS), SVG, and the Document Object Model (DOM). It was launched on the Chrome Web Store in September 2012, and currently has more than 40,000 users [21]. The Readium project attempts to ensure that its open source code is readily available for EPUB 3 publications, and aims to accelerate the adoption of EPUB 3 as a universal and global digital standard publishing format [6, 21]. However, it does not support 3D objects because it is based on EPUB 3 [2, 6].

To address these issues with existing content viewers, we designed an educational content viewer which can visualize the content including text, images, SVG, video and 3D graphics. We support various 3D objects, including interactive 3D objects, and the content viewer provides features for big data analytics. It is implemented in a cloud-based educational media service environment to facilitate sharing of a wide range of educational content.

2.3. Cloud-based Smart Education System

Cloud computing has been adopted in the education sector as a method of delivering enhanced security and more reliable services [8]. Many academic institutions are leveraging cloud-based technologies to support advanced teaching and learning processes, as well as to exploit the potential economic benefits. However, there remains a need to support an integrated cloud-based educational service environment beyond delivering and sharing learning material. Accordingly, a cloud-based smart education system has been developed by integrating a number of features for enhanced e-learning content services. These features include a platform for cloud-based smart media services, a compatible file format independent of system platform, an authoring tools for creating smart media content, a content viewer for displaying smart media content, an inference engine for providing customized learning content, and a security system to deploy a cloud-based educational media service environment [7, 8].

Figure 1 shows an overview of the configuration of the cloud-based smart education system proposed by Jeong, *et. al.*, [8].

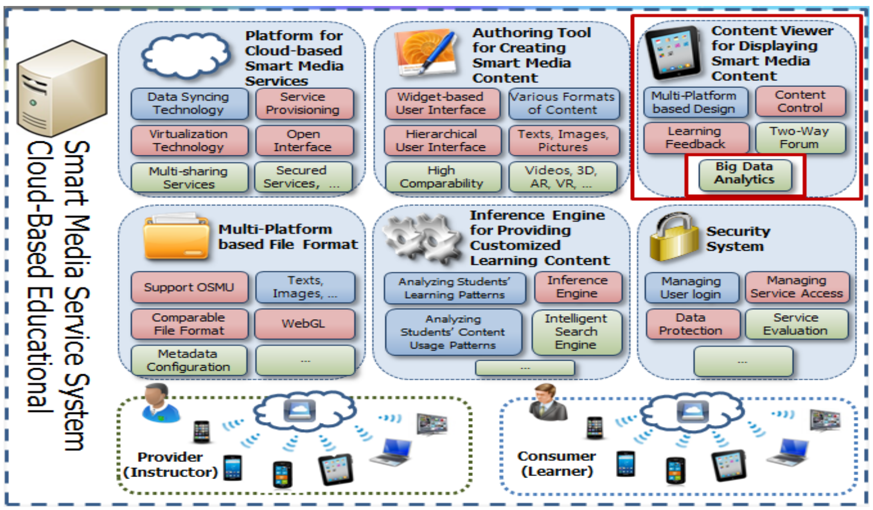


Figure 1. Configuration of the Cloud-based Educationn System [8, p.317]

Figure 2 shows the infrastructure of the six main features required to support cloud-based educational content services [8]. The content viewer is designed to run on multiple device platforms, and supports a wide range of content including text, images, videos, 2D graphics and 3D graphics object controls [2]. In this paper, we describe the content viewer for supporting more dynamic content, and the enlargement of the set of supported features by augmenting functions that deliver big data analytics, as shown in Figure 2. The big data analytics are visualized by seamlessly aligning the results thereof with text, images, other 2D and 3D objects.

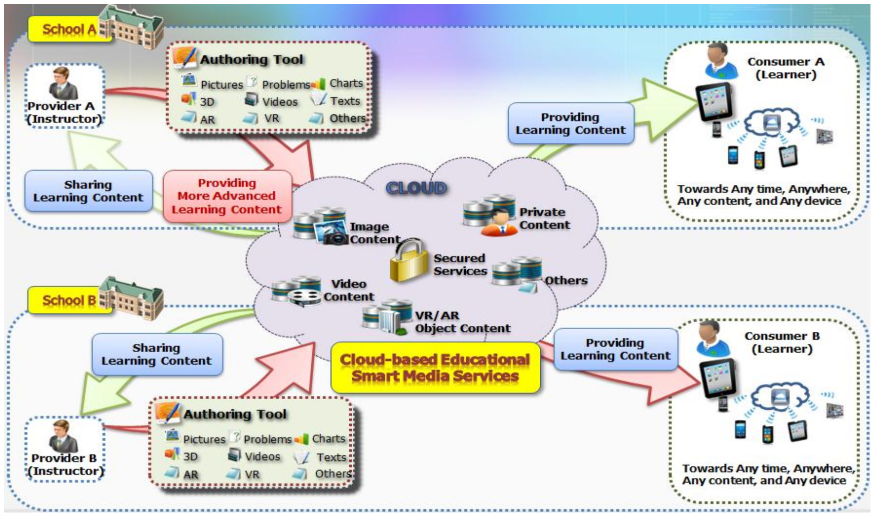


Figure 2. Infrastructure of the Cloud-based Education System [8, p.318]

3. Dynamic Educational Content Viewer

3.1. Overview of the Content Viewer

The content viewer was developed to display a wide range of educational content on multiple device platforms created using an authoring tool [8, 22]. The viewer displays basic controls by transforming them in the XML format to the HTML and XHTML formats using XSLT and interactive 3D controls using OpenGL ES [2, 10, 11]. We collect big data like movie comments, social network data and web log from the related data server, and then use NoSQL analysis tools and Hadoop analytics cluster for big data analytics, where the NoSQL is used to process VOC sentiment analysis and web log mining, Hadoop to process market basket and preference analysis. Finally the analyzed summarized results can be stored into the XML format files and they will be visualized on the content viewer using SVG viewing and OpenGL ES libraries. Figure 3 shows an overview of the big data analytics and the way in which it is integrated with the viewer system.

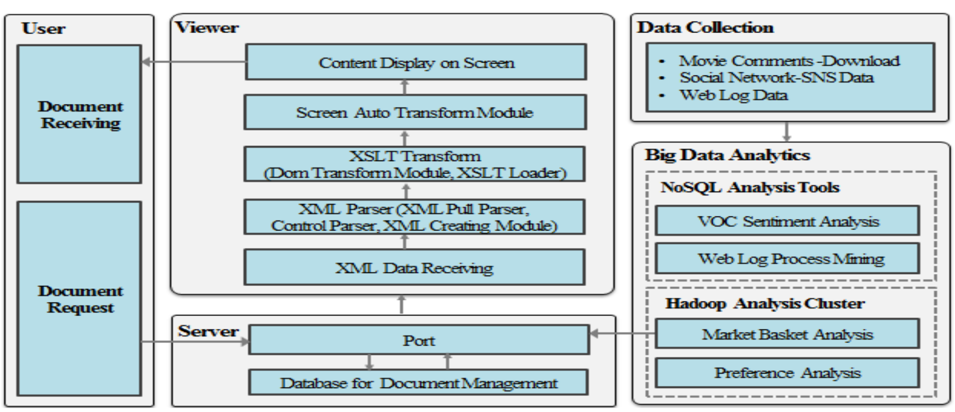


Figure 3. Overview of the Educational Content Viewer with Integrated Big Data Analysis

Figure 4 shows an overall process of the content viewer. The viewer is processed by four modules: a content download module, a content parsing module, a content control module, and a content display module.

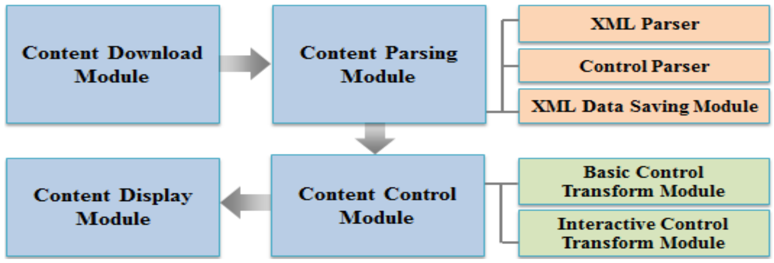


Figure 4. Overview of the Content Viewer System [2, p.415]

The content download module downloads content data that users wish to view from the system storage. The content data are transferred to a content parsing module and are parsed in page units. We employ and utilize the XML parsing techniques described in Ref. [23]. The content parsing module has three functions: an XML parser, a control parser, and an XML data saving module. The XML parser analyzes the content data of each page unit and extracts control information separately from the analyzed data, and calls the control parser. The control parser analyzes the control information and stores the parsed control data. The XML data saving module then stores the control information, such as memos, underlines, and highlights in XML format, and transmits it to the server and to the content control module. There are two categories of content control: basic controls for processing pages, texts, images, tables, formulas, and 2D objects; and interactive controls for processing 3D objects, videos, audio, annotations, quizzes, and animations. Interactive 3D controls are activated and controlled by the user via multi-touch gestures [2]. Table 1 lists the supported content controls [2, 10, 11].

Table 1. The Supported Content Controls [2, p. 416, 10, 11]

Category	Control Type	Definition
Basic Control	Page & Chapter	Each page/chapter layout control
	Text	Content information of a page/chapter control
	Image	Image file control (*.png, *.jpg, and *.gif)
	Table	Table control
	Math	Formula control
	2D Object	2D object control (line, circle, polyline, polygon, …)
Interactive Control	3D Object	3D object control (*.3ds)
	AR	AR (Augmented Reality)/VR (Virtual Reality) control
	Video	Video data control (*.mp4)
	Audio	Audio data control (*.mp3)
	Annotation	Annotation data control (memo, underline, highlight, …)
	Quiz	Quiz control for testing learning achievement
	Animation	Animation control that responds to user actions

Basic and interactive controls are processed differently. Basic controls include page and text manipulation, and are implemented by the content transform module, which converts basic control elements to HTML format using the XSLT loader, JavaScript execution, and auto transform modules. It then provides the content information to users by adjusting the content resolution corresponding to an individual smart device. Interactive content controls are managed by the interactive control collector, and interactive controls are viewed on the user’s screen by creating classes for each control action [2]. The content viewer module displays the transmitted content on the smart device, and this module handles device dependence of particular smartphones.

3.2. Big Data Analytics

Figure 5 shows the process of big data analytics in the educational content viewer [12]. First, we save the data to be analyzed to storage using the Hadoop distributed file system (HDFS), and filter the data and perform a statistical summary operation using MapReduce (*i.e.*, using the Map() and Reduce() methods). We then save the statistical results to a database and visualize them on the user's screen using SVG.

In the filtering process, the movies to be counted are imported to Map()s using the mapper, the outputs of which are partitioned per Reducer using the values with the same key. We can then obtain the number of comments on each movie by counting the occurrences of each keyword in the Reducers. As many output files are created as the number of Reducers. Figure 6 shows the number of comments on each movie produced via MapReduce.

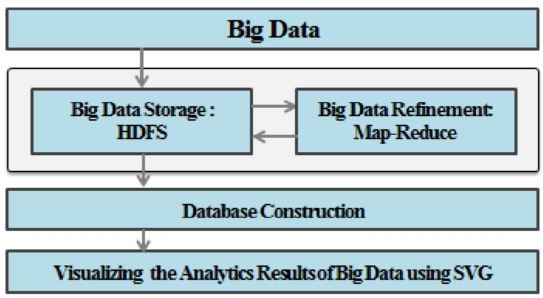


Figure 5. The Process of Big Data Analytics in the Content Viewer

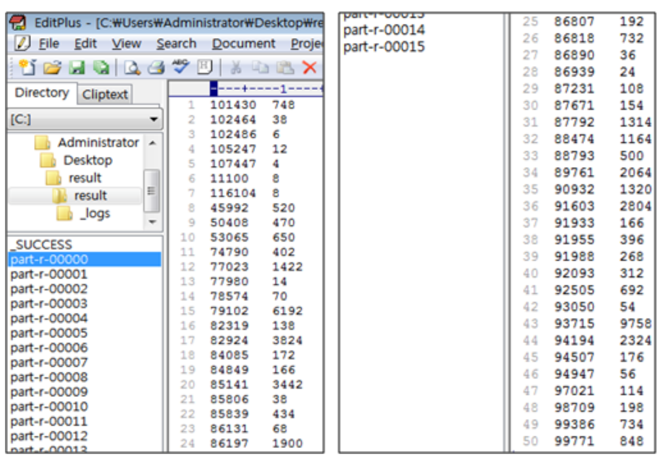


Figure 6. Examples of the Results of Big Data Analysis using MapReduce 4

Experiments

The educational content viewer system was implemented on iOS and Android SDK 4.0.4. Here, however, we only describe examples of the content viewer implemented using Android. Examples of iOS-based devices can be found in Ref. [2]. Figure 7 shows examples of basic controls, including images and text, and Figure 8 shows an example of interactive controls for a 3D object image of the Earth. These 3D images are provided to users as a class using OpenGL ES, and can be controlled using user multi-touch gesture actions; for example, if users hold the 3D earth sphere with their fingers and drag it, the user can see the other side of the Earth. Users can also change the size of a 3D object by zooming in or out.

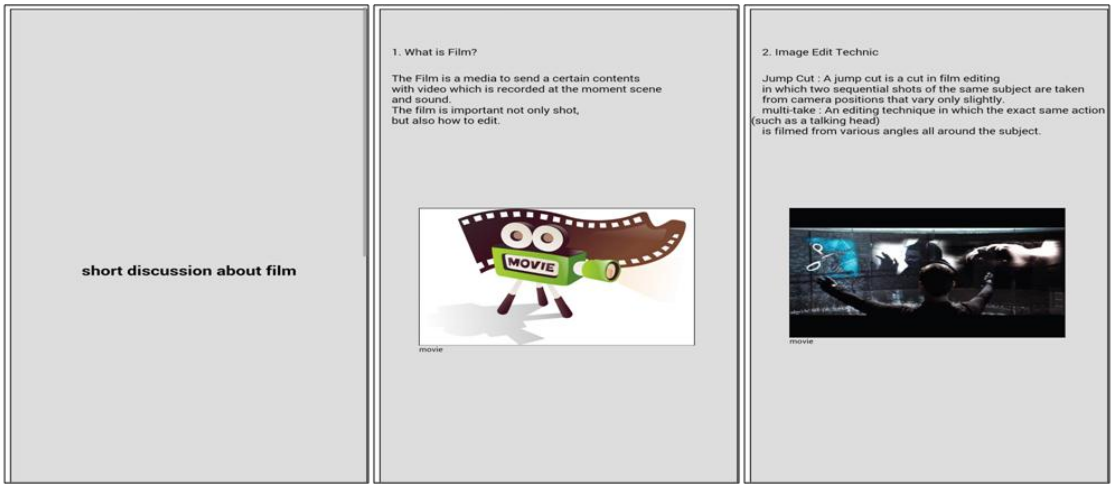


Figure 7. Examples of basic Controls (Images and Text)

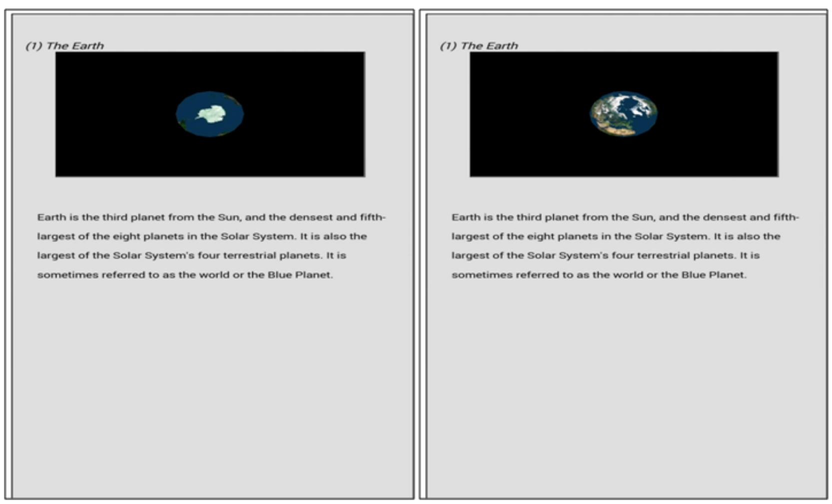
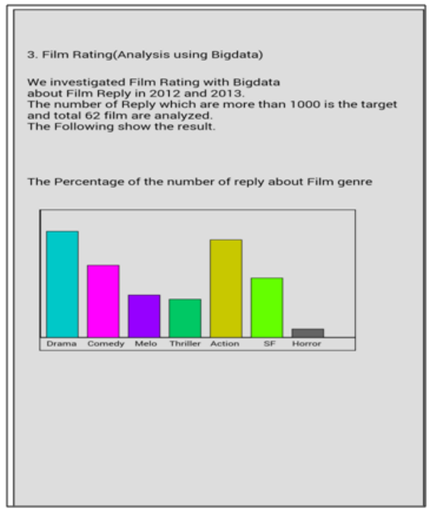


Figure 8. Examples of Interactive Control Over a 3D Object

Figure 9(a) shows the result of big data analytics used on data hosted by the Ministry of Science, ICT and Future Panning (MSIP), and National Information Society Agency (NIA) in 2013 [24]. Figure 9(b) shows the visualization of the analysis result using the educational content viewer. We analyze the user preferences of movies, which contain over 10,000 comments. We selected 62 movies out of 870, which had more than 10,000 comments. We categorized the 62 movies into 7 genres, as shown in Figure 9(a) [12].

Genre	Percent(%)
Drama	25%
Comedy	17%
Melo	10%
Thriller	9%
Action	23%
SF	14%
Horror	2%
Total	100



(a) (b)

Figure 9. (a) The Results of Big Data Analysis and (b) Visualization of these Data

5. Conclusion

We have described the development of a dynamic educational content viewer that supports multiple smart device platforms and a wide range of content types. In addition, the content viewer enables big data analytics and visualization of the results. The content viewer was implemented as a part of a cloud-based smart education system, which integrates a number of features required to deploy a cloud-based educational media service environment, delivering a variety of smart educational content on a small-scale cloud infrastructure [7, 8]. We developed the content viewer for iOS and Android-based smartphones, and demonstrated features including big data analytics. Our results indicate that the content viewer should be able to not only display various types of content on multiple device platforms by automatically adjusting the content resolution to fit the corresponding device screen but also handle interactive 3D objects easily using multi-touch gestures, support big data analytics, and respond in a reasonable speed.

The content viewer has a number of important advantages over existing content viewers; however, there remain a number of areas with significant scope for further improvements. The viewer described here is a prototype, and further implementations of 3D interactive controls are required, including different types of math formulae, quizzes and questions. In addition, a study into the design and implementation of a database system for content synchronization should be performed because synchronization between smart devices is essential in a cloud-based service. Furthermore, more in-depth research into big data analytics should be carried out via the establishment of a Hadoop ecosystem to support enhanced security, resulting in a more robust educational content viewer.

Acknowledgements

This research was financially supported by Ministry of Education (MOE) and National Research Foundation of Korea (NRF) through the Human Resource Training Project for Regional Innovation, and supported by the MSIP(Ministry of Science, ICT and Future Planning), Korea, under the ITRC(Information Technology Research Center) support program(NIPA-2014-H0301-14-1022) supervised by the NIPA(National IT Industry Promotion Agency).

References

[1]. W. J. Kim, "Status and Challenges of the e-Book Market", KOTA, 2010 SPRING, vol. 52, **(2010)**, pp. 72-22.
[2]. S.-Y. Ju, J.-S. Jeong, S.-O. Kwon, M. Kim and K.-H. Yoo, "Design and Implementation of a Smart Educational Content Viewer", Journal of Next Generation Information Technology, vol. 4, no. 8, **(2013)**, pp. 412-422.
[3]. J. S. Roh and Y. J. Lee, "Analysis on Current Status and Issue of the World e-Book Market", KOCCA (Korea Creative Content Agency) Focus, vol. 60, **(2012).**
[4]. "Price Waterhouse Coopers (PWC)", Global entertainment and media outlook, **(2012),** pp. 2012-2016.
[5]. C. S. Park, "Analysis of the Apple's Digital Textbook Ecosystem", KERIS Issue Report, RM 2012-3, **(2012),** Readium.org, Readium, http://readium.org (accessed on **May 2014).**
[6]. J.-S. Jeong, J.-O. Kim, D.-H. Kim, S.-O. Kwon, S.-Y. Ju, J.-W. Lee, C. Park, M.-S. Lim, D.-Y. Kim, R.-H. Jang, J. Fiaidhi and K.-H. Yoo, "Contents Oriented Cloud Educational Systems, International Conference on Convergence Content, vol. 10, no. 2, **(2012)**, pp. 457-458.
[7]. J.-S. Jeong, M. Kim and K.-H. Yoo, "A Content Oriented Smart Education System based on Cloud Computing", International Journal of Multimedia and Ubiquitous Engineering, vol. 8, no. 6, **(2013)**, pp. 313-328.
[8]. "Wikipedia", Big Data, http://en.wikipedia.org/wiki/Big_data (accessed on **May 2014).**
[9]. J.-S. Jeong, J.-O. Kim, D.-H. Kim, S.-O. Kwon, S.-Y. Ju, J.-W. Lee, C. Park, M.-S. Lim, J.-J. Jeong, J.-S. Jang and K.-H. Yoo, "XML-based Educational Content Viewer on Cloud Services", Proceedings of HCI 2013, **(2013)** January 31-Februray 01, pp. 1250-1253, Gangwon, Korea.
[10]. J.-S. Jeong, J.-O. Kim, D.-H. Kim, S.-O. Kwon, S.-Y. Ju, J.-K. Kim, C. Park, M.-S. Lim and K.-H. Yoo, "Design and Implementation of a 3D Integrated Smart Education Content Viewer", International Conference on Convergence Technology, vol. 2, no. 1, **(2013)**, pp. 254.
[11]. S.-Y. Ju, J.-S. Jeong, J.-K. Kim, S.-O. Kwon and K.-H. Yoo, "Design and Implementation for interworking system of a Smart Educational Content Viewer and Big data", Proceedings of the Korean Big Data Service Society, Autumn Conference 2013, **(2013)** December 06, pp. 24-26, Chungbuk, Korea.
[12]. S.-Y. Ju, M. H. Song, G. A. Rye, M. Kim and K.-H. Yoo, "An Embedding Method of Dynamic Content through Big Data Analytics, Proceedings of the CES-CUBE 2014, **(2014).**
[13]. S. T. Kim, "Era of Big Data Opening up a New Future", National Information Society Agency, **(2013).**
[14]. J. Manyika, M. Chui, B. Brown, J. Bughin, R. Dobbs, C. Roxburgh and A. H. Byers, "Big data: The next frontier for innovation, competition, and productivity, McKinsey Global Institute, **(2011)** May.
[15]. E. Semenovskaia, L. Popescu and T. Sloniewski, "Industry Developments and Models, Big Data in CEE, **(2012).**
[16]. D. Lancy, http://www.opentracker.net/article/definitions-big-data, Opentracker, **(2013),** (accessed on **May 2014).**
[17]. "Stack Exchange", http://programmers.stackexchange.com/questions/151955/what-is-the-definition-of-big-data, Answered by Doug Lancy, **(2012)** (accessed on **May 2014).**
[18]. T. Rodrigues, "10 emerging technologies for Big Data", Tech Republic, http://www.techrepublic.com/blog/big-data-analytics/10-emerging-technologies-for-big-data/, 2012 (accessed on **May 2014).**
[19]. C. S. Park, "Apple's Digital Textbook Ecosystem Analysis", KERIS, RM2012-3, **(2012).**
[20]. "Chrome Web Store", http://chrome.google.com/webstore/detail/empty-title/fepbnnnkkadjhjahcafoaglimekefifl, (accessed on **May 2014).**
[21]. S.-O. Kwon, J.-O. Kim, J.-S. Jeong, S.-Y. Ju and K.-H. Yoo, "Design and Implementation of Educational Content Authoring Tool for Smart Devices", Journal of the Korea Contents Association, vol. 13, no, 12, **(2013)**, pp. 1-8.
[22]. J.-W. Lee and H.-S. Yueon, "Design and Implementation of Android based Total Weather Information Application using XML Parsing Techniques", Journal of Digital Contents Society, vol. 12, no. 4, **(2011)**, pp. 611-618.

[23]. "The First Competition on Big Data (hosted by MSIP and NIA)", http://contest.kbig.kr/, (accessed on **May 2014).**

Authors

Seong-Yeon Ju

She received the B.S degree in the Information Statistics from Chungbuk National University in 2010. She is currently working towards M.S degree on Business Data Convergence from Chungbuk National University, Korea. Her main research interests include computer graphics, e-learning, and data analysis.

Min-Ho Song

He received the B.S degree in the Department of Computer Education from Chungbuk National University in 2014. He is currently working towards M.S degree on Digital Informatics and Convergence from Chungbuk National University, Korea. His main research interests include computer graphics, e-learning, and integral imaging.

Ga-Ae Ryu

She received the B.S degree in the Department of Cyber Security and Police from Kwangju University in 2014. She is currently working towards M.S degree on Digital Informatics and Convergence from Chungbuk National University, Korea. Her main research interests include computer graphics, e-learning, and security of smart content.

Mihye Kim

She received her Ph.D. degree in Computer Science and Engineering from New South Wales University, Sydney, Australia in 2003. She is currently an Associate Professor in the Department of Computer Science Education at Catholic University of Daegu, South Korea. Her research interests include knowledge management and retrieval, e-learning, digital textbooks, computer graphics, and cloud computing.

Kwan-Hee Yoo

He is a professor in the Department of Computer Education and the Department of Information Industrial Engineering at Chungbuk National University, Korea. He received his B.S. in Computer Science from Chonbuk National University in 1985, and his M.S. and Ph.D. in Computer Science from KAIST (Korea Advanced Institute of Science and Technology) in 1988 and 1995, respectively. His research interests include computer graphics, 3D character animation, and dental/medical applications.