Interactive 3D Visualization of Soical Network Data Using Cloud Computing

Naveed Ejaz [1], Irfan Mehmood[1], Jeong Joong Lee[1], Su Mi Ji[1], Myung Ho Lee[2], SungMmahn Anh[3], Sung Wook Baik[1*]

[1]*College of Electronics and Information Engineering, Sejong University, Seoul, Korea*
naveed@sju.ac.kr, irfanmehmood@sju.ac.kr, j2lee@sju.ac.kr, smji@sju.ac.kr, sbaik@sejong.ac.kr

[2]*Dept of Computer Science and Engineering, Myong Ji University, Korea*
myunghol@mju.ac.kr

[3]*College of Business Administration, Kookmin University*
sahn@kookmin.ac.kr

Abstract— **The social networks have revolutionized the online communication and data sharing. The researchers are now focusing on mining and analysis of large amount of social network data for a variety of purposes. However, because of the huge amount of continuously changing data, the data analysis in a daunting task. OLAP analysis is a famous data analysis method which can be used to analyze social data. This work extends our previous work in which we developed interactive 3D visual data cubes for high volume/dimension OLAP data analysis. The implementation of this scheme on traditional computing resources is much time consuming and resource intensive. The advances in cloud computing motivated us to use the cost effective cloud computing for the task of 3D visualization of social networks data. Therefore, in this paper, we propose the usage of cloud computing platforms as a possible solution for analyzing large amount of social network data.**

Keywords— **Cloud Computing, OLAP, MVC, Data Visualization.**

I. INTRODUCTION

The cloud computing provides shared computing resources to computers and other devices based on rentable infrastructure [1]. Cloud computing has reduced barriers and costs related to access the powerful computing capabilities earlier reserved for expensive enterprise systems only. It also brought new tools and techniques that surpassed enterprise system performance and scalability. Owing to the low cost, high reliability and on demand access; cloud computing is rapidly getting popular [2]. For instance, many world famous companies like Amazon, Google and Microsoft, have started their cloud services to grow their business.

We are living in the world of Web 2.0, where majority of people are using the internet and millions of those people are connected on social networking sites like Facebook, LinkedIn, MySpace, and Twitter, using blogs, and posting on YouTube and Flickr [3]. These connections between people have sparked the interest of cloud computing services and systems. IT giants are trying to use cloud computing services to develop the ways to knock into the Web 2.0 world and establish means of turning the flow of information and communication into business potential [4]. Cloud computing can also enable businesses to access information for further processing that will support their customer service capabilities.

The outstanding capabilities of cloud computing can help to solve the expensive tasks like 3D data visualization and analysis of large and complex social networking data [5]. 3D data visualization and analysis have become popular in many fields such as process control, decision support systems and scientific data analysis. For instance, in June 2011, Microsoft's High Performance Computing (HPC) team released Beta 2 of HPC Pack for Windows HPC Server 2008 clusters and LINQ to HPC R2 SP2 [6].

In this paper, we have introduced a framework in which the users can efficiently visualize and analyze large amount of social network data using cloud computing. The proposed system is capable of handling extremely large data sets. The details of the proposed framework are discussed in the subsequent sections.

II. VISUALIZATION OF SOCIAL NETWORK DATA

The proposed 3D graphical interfaces for knowledge/information navigation is geared on a framework based on MVC (Model-View-Controller) which has been implemented in WPF (Windows Presentation Foundation) [7]. This 3D space is integrated with an OLAP tool to support graphical presentation of complex business data. An example of this is a connection made between cloud computing and the social networking applications. Using our cloud computing based customer relationship management tool, businesses can search for users that are related to their company. They can then capture the information that has been posted, as well as track future conversations about their company. The information obtained from these conversations can be logged in a customized database to create customer service cases. This conversation can also help companies inform customers about products and services, as well as open doors of communication. In addition, businesses can also

*Corresponding Author

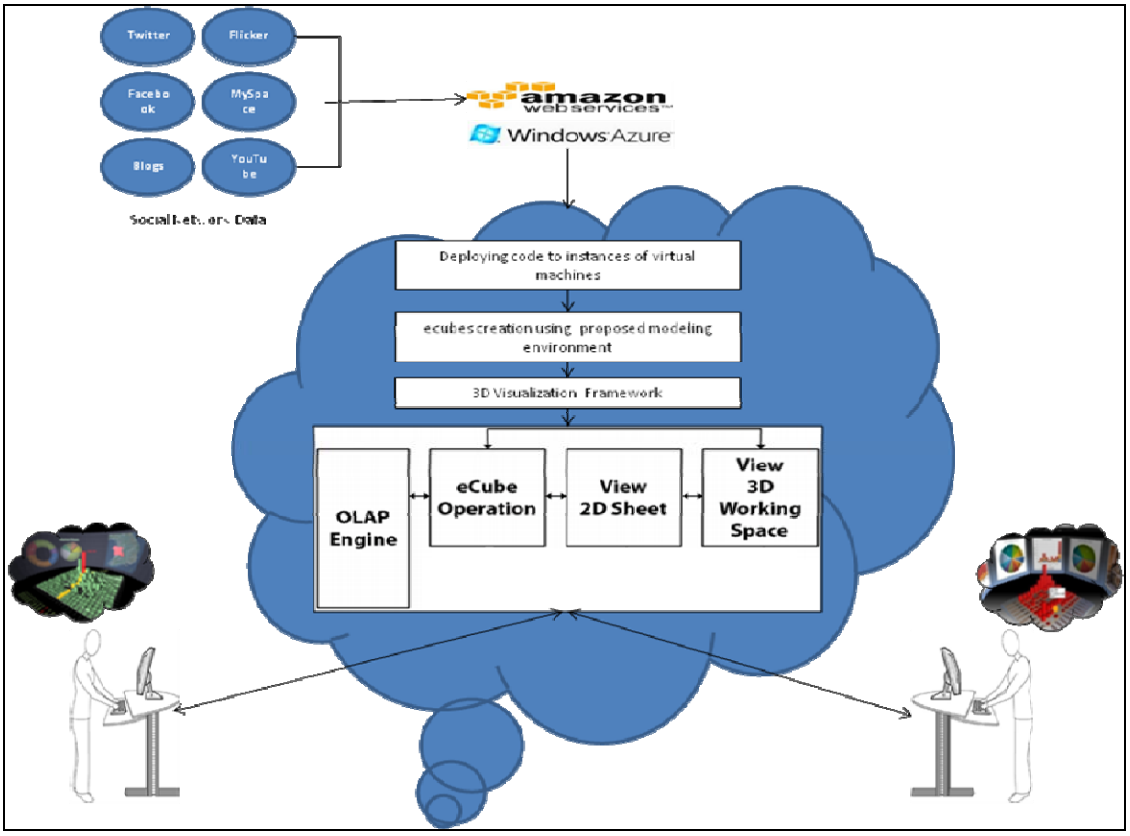

Fig. 1 Framework of the proposed system

use cloud computing to access the existing social networks, optimize search engine tools, and to connect with other businesses. It facilitates enterprises to get best-in class position in growth, and thus provide innovation and leadership by providing disciplined research and best-practice models to drive the generation, estimation, and execution of prevailing growth strategies by analyzing the large amount of data [8]. Figure 1 shows the framework of the proposed method.

III. OLAP Data Analysis and Cube Generation

OLAP analysis [9] is famous data analysis method used in many business areas. The timely analysis of huge volume of data is a daunting task. In our application, the amount of data generated by social web sites is huge and needs huge amount of processing power for analysis. An OLAP system generally uses multidimensional data called cubes for providing measurements and dimensions. Owing to large size of social networking data, the number of dimensions usually can range from hundreds to thousands.

The tradition OLAP selection methods generally use 2D plots which place multiple dimensions as column or row factors. In such systems, a user must have some domain knowledge for the selection of appropriate dimension. In the presence of new trends or abnormality in measures, it is very difficult for users to select new dimensions and range for providing explanation of these behaviors. Moreover, most of the existing OLAP systems support 2D or partially limited 3D charts with which users have limited or no interaction. In our previous work [10], we presented an efficient and interactive 3D based OLAP cube. In this work, we provided the capibility of visualizing and manipulating 3D cubes to the users by simply interacting using key board and mouse. This method reduced the number of dimensions based on level of relationships. The user can visualize the 3D cube by using operations like slicing, drilling-down, rolling-up, and pivoting operations. The problems with this approach for the data visualization of social networking sites are that it needs high computational cost, storage resources, lack of ad hoc and ubiquiotous access. For this reason, in this paper, we used the concept of cloud computing for the 3D visualization of the data achieved from social networking sites. The usage of cloud computing renders (1) distributed architecture instead of centralized service, (2) reduces cost, (3) device and location independence, (4) increase in peak-load capacity.

IV. MVC based model for Visualization and Deployment to Cloud

The proposed system for 3D visualization of large and complex social network data is based on a MVC model (Model-View-Controller). The concept of model is used so as to represent and manage the targeted information selected by users from huge social networking data. This targeted information selected by user is a subset of raw data and statistics obtained from social networking sites. Our model is using XML (Extensible Markup Language) which helps in

efficient communication on two or more models. The XML script is composed of the definition of the 3D graphical environment, the syntax of interaction, the viewing modes and the file names of the data associated with each mode. The view is composed of an interpreter for XML and a 3D graphics renderer. The XML file is interpreted and the information is shown in a pre-defined 3D graphical environment.

For the extraction of data from social websites, we use the web data extraction services provided by Windows Azure and Amazon. These are the state-of-art services to extract web data, videos, images, files, content, etc. from social websites in to a structured form from an unstructured or semi-structured web data sources. The usage of these services renders speed, accuracy and exactness in the extraction of web data. For the selection of appropriate cloud platform, we looked for a low cost and less complex platform. There are many companies which provide PaaS (Platform as a Service) for cloud computing. Based on industry reviews and technical features [11], Google App Engine (GAE) [12] is selected. The other close alternatives are Salesforce.com [13] and Microsoft Windows Azure [14]. GAE is chosen because of its simpler API and its comparative low price. Moreover, GAE provides important services such as the Full Text Search API, MapReduce and the App Engine Pipeline API. By using GAE, tier web applications can be hosted in Google-managed data centers and then the applications are virtualized across multiple servers. The application can be monitored using a web-admin panel to see the usage details and logs etc. With the advent of App Engine Backends, introduced in version 1.5, the scalability issues of previous versions of GAE are almost resolved. For developing application using GAE, GAE Software Development Kit (SDK) [15] is needed. The SDK provides the capibility of directly deploying the application on the cloud.

The interaction of the users with 3D visualization space is monitored by the controller. The controller receives the interaction response of the users in the 3D environment of the view mode. Based on the type of interaction, the controller governs the model to perform changes based on interaction. If the state of the model is changed by the controller, the connected views are also changed. The 3D information is represented by an eCube which is composed of two selected attriburtes of the given database and their dimensions. The users can interact with the eCube in a variety of methods including selection, navigation and zoom in/out. Figure 2 presents an example of the visualized chart report. If the report includes social interaction data, the view of 3D graphical environment shows a 3D map generated from the social information of the reported data. The 3D visualization is more intutive and user friendly as compared to the 2D cubes.

The proposed system efficiently utilizes parallel computing and client side caching to accelerate responses to user requests. The result is a high-performance data visualization tool that typically generates 3D plots in under few seconds, and helps decision makers to quickly navigate through gigabyte, terabyte, and even pet byte sized data sets. It can efficiently manipulate components representing data sets, statistics, knowledge information, 3D maps, 2D images etc., using cloud computing.

V. Conclusions

This paper presents our preliminary work on 3D visualization of OLAP cube in cloud environment. The usage of cloud computing solves the problem of lack of computational resources because of the large amount of social data in social networks. The efficient 3D interactive graphical representation visual OLAP analysis yields a synergy effect among OLAP technologies, 3D graphical presentation methods, and human's cognitive capability. The usage of new Google Backend API makes it possible to execute long running background processes.

Acknowledgment

This research was supported by Seoul R & BD Program (JP090972M0214832).

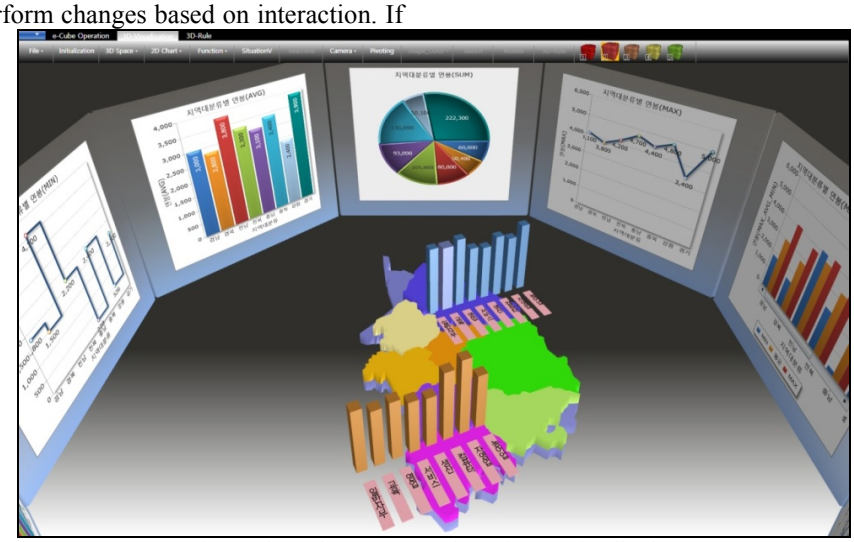

Fig. 2 Information Representation using 3D visualization

REFERENCES

[1] B.P.Rimal, Eunmi Choi, I.Lumb, A Taxonomy and Survey of Cloud Computing Systems INC, IMS and IDC, 2009. NCM '09. Fifth International Joint Conference on , 2009 , Page(s): 44 - 51

[2] Jie Song, Yao Junfeng, Wu Chengpeng, Cloud computing and its key techniques, Electronic and Mechanical Engineering and Information Technology (EMEIT), 2011 , Page(s): 320 – 324

[3] Puteri N.E. Nohuddin, Frans Coenen, Rob Christley, Christian Setzkorn, Yogesh Patel, Shane Williams , Finding "interesting" trends in social networks using frequent pattern mining and self organizing maps ,Original Research Article *Knowledge-Based Systems*, *Volume 29*, *May 2012*, *Pages 104-113*

[4] Chen, Li Yizeng , Chen Xingui, Fangning, Overview and analysis of cloud computing research and application , 2011 , Page(s): 1 – 4

[5] Y. Tanahashi, Chen Cheng-Kai, Marchesin, S. Kwan-Liu Ma, An Interface Design for Future Cloud-Based Visualization Services , 2010 , Page(s): 609 – 613

[6] R.Pathak, S.Joshi, , S.Ahmed, D. Mishr, Optimizing HPC and parallelization for computation Nanotechnology in MCCS environment, 2009 , Page(s): 712 – 715

[7] M.J. Mahemoff, L.J. Johnston, Handling multiple domain objects withModel-View-Controller, 1999 , Page(s): 28 – 39

[8] V. Tosic, H. Wada, A. Guabtni, K. Lee, A. Liu, Management towards reducing cloud usage costs,Network Operations and Management Symposium (LANOMS),2011 .

[9] S. Chaudhuri , U.Dayal, An Overview of Data Warehousing and OLAP Technology, ACM SIGMOD Record, 26(1), (1997), pp.65-74

[10] Su Mi Ji, Beom Seok Lee, Kyoung Il Kang, Sung Gook Kim, Cheolwhan Lee, Oh-young Song, Joon Yeon Choeh, Ran Baik, Sung Wook Baik "A Study on the Generation of OLAP Data Cube Based on 3D Visualization Interaction", 2011 International Conference on Computational Science and Its Applications.

[11] Noordhuis,P.,Heijkoop,M,Lazovik,A.; Mining Twitter in theCloud: ACase Study Cloud Computing (CLOUD), 2010 , pp. 107 - 114

[12] [online] Google App Engine - http://code.google.com/appengine/

[13] Salesforce Force.com - http://www.salesforce.com/platform/

[14] MS Windows Azure - http://www.microsoft.com/windowsazure/

[15] [online] http://code.google.com/