Journal of Applied Sciences

ISSN 1812-5654

science alert

ANSInet
an open access publisher
http://ansinet.com

Journal of Applied Sciences 14 (14): 1653-1657, 2014
ISSN 1812-5654 / DOI: 10.3923/jas.2014.1653.1657
© 2014 Asian Network for Scientific Information

Integration of FPGA, Cloud and Opinion Mining

[1]N. Bharathi, [2]G.R. Brindha, [2]B. Santhi and [3]P. Neelamegam
[1]Department of Computer Science, School of Computing,
[2]Department of Information Communication and Technology, School of Computing,
[3]Department of Electronics and Instrumentation, School of EEE,
SASTRA University, Thanjavur, Tamil Nadu, India

Abstract: Inter-domain analysis and studies are growing rapidly, which is fruitful in improving the efficient usage of individual field. This analysis embeds Opinion Mining in Cloud Computing and Field programmable Gate Array (FPGA). The explosion of social networks and blogs provide a new set of issues and openings by means of information searching and retrieval. These opinions play major role in influencing everything, such as purchase of product, availing service also even the presidential candidate that the people support. One of the proposed modules is having opinion mining tool to process the opinion corpus. Organizations can use this to determine user's perception about their products and competitors and consumers can get decision about products. Though many researchers have developed these types of tools, their accessibility is limited and corpus size is small. Keeping in mind, the objective of this study is to combine domains and thereby increasing the effective utilization of tool, the proposed framework is implemented globally in a cloud network. Besides this, cloud provides services such as storage, communication and computation. Since cloud has the capability to process large volume of data at a time, it need to increase the speed in storage, retrieval and process. Here comes the domain FPGA which solves this issue and fasten the service usage and provision.

Key words: Opinion mining, cloud computing, FPGA, service management, reconfiguration

INTRODUCTION

The development of Web 2.0 provides sharing of experiences, emotions and many facts among the Internet users and this increases the demand of storage and services. Internet users are involving themselves more aggressively and producing corpus, which include reviews and suggestions about products or services. This makes the Internet data volume rate to increase second by second. These data can be processed effectively through Opinion Mining, which is an emerging research area and used as a competitive intelligence of a business or a particular user. So the data set of Internet about a product or services or a person can be collected and mined for decision making by opinion mining analysis, which gives qualitative and quantitative information. The research on opinion mining started by 2006 and only during 2010 to 2012, the necessity drag the researchers to do more work in this field. All research articles concentrated on different types of classification algorithm, retrieval and normalizing the text. To prove their classification accuracy and novelty many studies were come out such as movies (Dhoshi *et al*., 2010), bikes (Brindha and Santhi, 2012a), games, electronic goods etc., (Ramkumar *et al*., 2010). Moreover though studies for different data set such as news reviews (Tan and Zhang, 2008), teaching and learning (Chen and Weng, 2009), performance analysis of team leader (Brindha and Santhi, 2012b), the availability of the opinion mining tool in cloud is missing. Though this is an emerging area, many researchers studied mostly about classification of reviews and none of them focus on embedding an opinion mining application in a cloud service to offer a categorized large amount of data in a single storage that ease the access.

Cloud computing provides software, platform and infrastructure as a service. The cloud services can be accessed from wherever and whenever people need (Dingguo *et al*., 2011). It realized the novel idea of anything as a service. At this scenario, managing resources privately is waste of time and man power. Cloud is working based on usage and billing concept. It is well suited for storage services besides others which are the basic need for opinion mining application. The issue is normally cloud service takes more response time, since data storage is more. Currently many researchers have considered the effective usage of cloud computing for

Corresponding Author: N. Bharathi, Department of Computer Science, School of Computing, SASTRA University, Thanjavur, Tamil Nadu, India

scientific applications (Hoffa *et al*., 2008; Deelman, 2010). Gao and Kang (2012) gave cloud simulation scheduling algorithm based on Quality of Service, which showed the algorithm reduced simulation time and increased resource utility rate. Ding *et al*. (2012) built a collaborative manufacturing resource sharing platform based on cloud computing that implement description of information, resources and knowledge during manufacturing process. They have combined cloud computing and ontology technique.

This proposed methodology provides the above scheme and deciding on the issue of cloud usage by integrating FPGA (Field Programmable Gate Array) in its architecture. Microprocessors are generally suitable for any kind of common applications, in which I/O usage is more than CPU usage. Specific chip like Application Specific Integrated Circuits (ASIC) is better preference for implementing unusual computation intensive applications than all-purpose microprocessors. The concern here is, since the fabrication of ASIC is done at the production unit itself, reconfiguration cannot be performed by the customer. FPGA overcomes this issue by providing the facility of configuration at the customer site. So, the user can morph the FPGA configuration as and when he is needed on the fly (Zhang *et al*., 2012). Researchers have proved the execution speed of FPGA in compared with general purpose microprocessor by nearly 100 times for computation intensive applications (Zerigui *et al*., 2008; Bharathi and Neelamegam, 2011; Rajagopalan *et al*., 2012).

This study provides an innovative combination of FPGA and Opinion mining, which are embedded in cloud computing to get the maximum benefit of application. In this way the opinion mining application of car review is embedded in a cloud framework which is supported by FPGA to improve the response time.

MATERIALS AND METHODS

Architecture framework: The proposed framework shown in Fig. 1 includes Opinion Mining, Cloud and FPGA as major components. The opinion mining application is accessing the storage service, preference and enhancement service and decision making service.

In order to provide these services, cloud service developer, service provider and service management processes works together. Large volume of data is assured when using cloud and total response time increases. To reduce the response time FPGA is used in which the services are configured as hardware blocks (Bharathi and Neelamegam, 2012). The hardware includes any one of the blocks such as storage, communication (Zerigui *et al*., 2008; Berube *et al*., 2004) and computation resources. Opinion mining system without FPGA is implemented in MATLAB 10 under Core 2 Duo processor with frequency 2.10 GHz. The system with FPGA is simulated in Xilinx 9.2 with same platform.

OPINION MINING APPLICATION PHASE

In this module a corpus with 200 reviews about a car had been taken as data set. Data may be in any form, such as raw opinion or normalized data set. Raw reviews are taken directly from the websites that contains opinions about products or persons. Then the reviews are being feed to the pre process where the following normalization

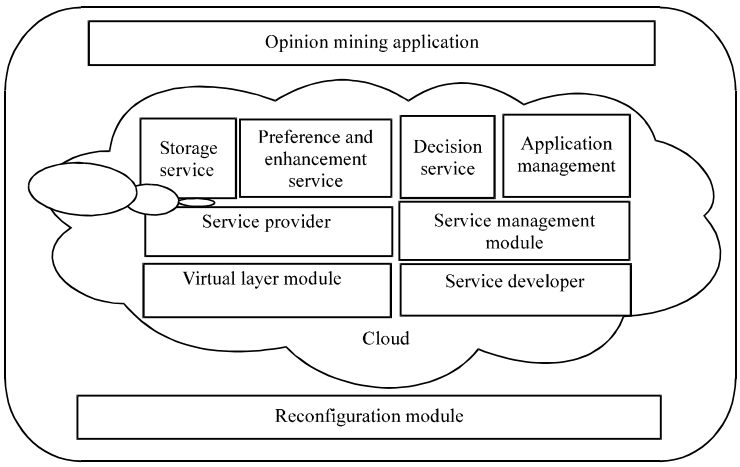

Fig. 1: Opinion mining, cloud and FPGA integrated architecture

was done. Tokenization process cut the text as consequential elements called token such as symbols, phrases and words. The majority of reviews contain stop words i.e., articles and prepositions. Removals of stop words reduce the corpus size and give way for effective searching and process. Next step is case folding to form proper noun, which make upper case letters into lower case.

The unnecessary afflictions ('mines') are chopped ('mine') in stemming. Lemmatization provides terminology and morphological analysis that gives the base dictionary word. Now the remaining word vector is given as input to train classifier. After additional filters like tagger and shallow parser, testing makes the entire matrix robust to be an input for classifier such as NB-Naïve Baise or SVM-Support Vector Machine etc., in sentiment lexicon process, ranking, negation detection and sentiment orientation are being done. After this, based on review words for each preference ranking is posted. This normalized data is taken as input for the proposed algorithm.

For any product, parameters are ordered based on preferences.The review is classified into 3 types, 'Poor', 'Normal' and 'Good' which takes the values 1, 2, 3, respectively for each above said feature. The features are ranked (matrix) based on review word extraction. The proposed algorithm used this matrix as input, which is stored as corpus storage in cloud. Then preference and enhancement service takes preference from the user to provide suggestions and product position for the manufacturer. Based on the input preference, weight value is allotted and given as input for the next module, decision service. In this module mean value for the rows are calculated and based on mean value, range is set for 'Poor', 'Normal' and 'Good' ranking count. Then these three counts provide the purchase idea to the user.

The proposed algorithm used in these modules is given below.

PROCEDURE

Corpus storage:

Step 1: Set survey_matrix as primary input

Preference and enhancement service:

Step 2: Get preference for price, mileage, modish, facilities and performance
Step 3: Assign weights to preference
Step 4: Compute preference

Decision service:

Step 5: Find the rank (1, 2, 3) positions
Step 6: Set range for 'poor', 'normal' and 'good'
Step 7: Use the range to get count of poor, normal and good ranks:

- Consumers get suggestion from the output and they can also read particular reviews
- Manufacturers get ideas from feature values in such way that the lacking point and which feature is to be improved

CLOUD PHASE

This module provides Corpus Service, Preference and Enhancement service and Decision making service. These services acquire storage, computation and communication resources of the cloud. Application management maintains log about the applications and controls service access. Service provider maintains a registry of services available. Service developer releases new services based on the requirements of the customers. Service management keeps track of the service usage and billing. Virtual layer module hides the underlying software and hardware details of the cloud.

FPGA PHASE

The service developers generate configuration bits of the services and are termed as service instance. The service instance can be of any kind such as, storage, computing or communication service as specified in service management. The service instances are configured into the FPGA for execution whenever it is needed. The service instances are configured dynamically when the application demands for the service. The reconfiguration of dynamic part is performed when the requested service is not currently configured in the FPGA. It takes place at run time and only dynamic part of the FPGA is reconfigured as shown in Fig. 2.

RESULTS AND DISCUSSION

To certify the proposed model with respect to its effectiveness based on response time, a corpus contains car review of 200 data is taken. Using opinion mining with the effective basement of FPGA and cloud architecture a full-fledged decision support structure for car purchase is implemented, which gives the suggestion that the review of particular car is 'Good'. Cloud is better in terms of cost effective by maintaining data privately in any industry or organization. Cost for installation and maintenance are more in private corpus when compared to cloud. When

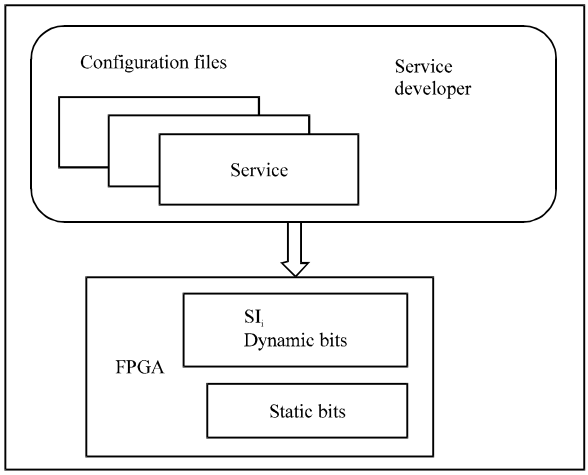

Fig. 2: FPGA reconfiguration structure

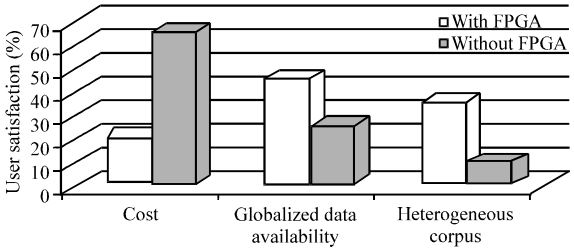

Fig. 3: User satisfaction with and without cloud

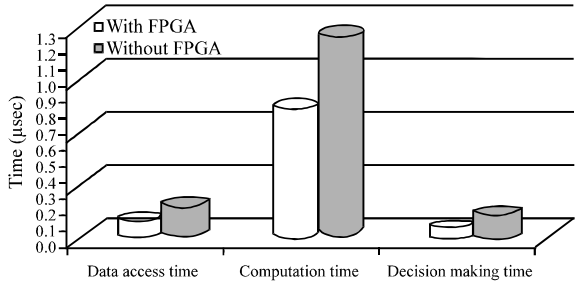

Fig. 4: Time comparison with and without FPGA

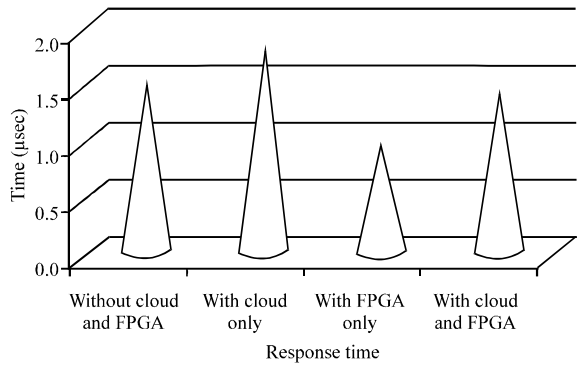

Fig. 5: Response time comparison

accessing data outside cloud such as locally stored data or data collected from Internet cannot provide accurate information for decision making. But in cloud with its ubiquitous nature vast and variety of information is available and hence decision making is accurate and easier. The analysis shows that the cloud integration improves the application in the following ways.

The observation is made for system with cloud and without cloud as shown in Fig. 3. Three parameters are analyzed for this case viz., cost, globalization of data availability and heterogeneous data obtained.

When the application is used without cloud, the cost incurred is more since the software should be purchased individually as depicted in Fig. 3. But in cloud the software cost is less. Maintenance cost can be reduced since it is accessed globally by more number of users. The next parameter considered is globalized and volume of data. Cloud is used to assure the coverage of more geographical locations which is not available/accessible through Internet. In cloud, users can have more volume of segregated data in domain basis. So, the advantage is assurance of accuracy when the data size is larger. But when cloud is included, response time is more because of larger volume of data. Without cloud as illustrated in Fig. 3, the cost is high and also the globalized data availability is less.

Figure 4 shows the response time of the opinion mining system under four cases viz., without cloud and FPGA, with FPGA only, with cloud only and with cloud and FPGA. When the system is with FPGA only, then the data set won't be large as compared to with cloud. So the time taken to process and produce decision will be less. Hence FPGA is added to improve the response time of the system as demonstrated in Fig. 5 When the users are using the proposed system overcoming the negative of cloud and using the advantage of cloud with FPGA base, they can have full-fledged perfect system.

CONCLUSION

The assembled architecture comprises three different fields namely opinion mining, cloud and FPGA. The inter-domain applications normally provide efficient usage and effective performance which is proved here. The architecture describes opinion mining analysis and cloud application clearly with system configuration. The comparative analyses are pasteurized in results and discussions part for individual and integrated domains.

Many studies concentrated on cloud storage and services, mining sentimental reviews and FPGA incorporation, individually but none of studies concentrated on integrating these three areas. In future the researchers can develop any such model which can be used for decision support information.

REFERENCES

Berube, P., M. MacGregor and J.N. Amaral, 2004. FPGA implementation and experimental evaluation of a multizone network cache. Microprocessors Microsyst., 28: 237-252.

Bharathi, N. and P. Neelamegam, 2011. FPGA based linear heating system for measurement of thermoluminescence. Meas. Sci. Rev., 11: 207-209.

Bharathi, N. and P. Neelamegam, 2012. A reconfigurable framework for cloud computing architecture. Proceedings of the International Conference on Embedded and Cloud Computing, June 2-3, 2012, India.

Brindha, G.R. and B. Santhi, 2012a. A novel opinion mining technique for product review based on preferences. Res. J. Applied Sci. Eng. Technol., 4: 5084-5088.

Brindha, G.R. and B. Santhi, 2012b. Application of opinion mining technique in talent management. Proceedings of International Conference on Management Issues in Emerging Economies, August 17-18, 2012, Thanjavur, India, pp: 127-132.

Chen, Y.L. and C.H. Weng, 2009. Mining fuzzy association rules from questionnaire data. Knowl. Based Syst., 22: 46-56.

Deelman, E., 2010. Grids and Clouds: Making workflow applications work in heterogeneous distributed environments. Int. J. High Perform. Comput. Appl., 24: 284-298.

Dhoshi, L., J. Krauss, S. Nann and P. Gloor, 2010. Predicting movie prices through dynamic social network analysis. Procedia Social Behav. Sci., 2: 6423-6433.

Ding, B., X.Y. Yu and L.J. Sun, 2012. A cloud-based collaborative manufacturing resource sharing services. Inform. Technol. J., 11: 1258-1264.

Dingguo, Y., C. Nan and T. Chengxiang, 2011. Research on trust cloud-based subjective trust management model under open network environment. Inform. Technol. J., 10: 759-768.

Gao, W. and F. Kang, 2012. Cloud simulation resource scheduling algorithm based on multi-dimension quality of service. Inform. Technol. J., 11: 94-101.

Hoffa, C., G. Mehta, T. Freeman, E. Deelman, K. Keahey, B. Berriman and J. Good, 2008. On the use of cloud computing for scientific workflows. Proceedings of the IEEE Fourth International Conference on eScience, December 7-12, 2008, Indianapolis, IN., USA., pp: 640-645.

Rajagopalan, S., R. Amirtharajan, H.N. Upadhyay and J.B.B. Rayappan, 2012. Survey and analysis of hardware cryptographic and steganographic systems on FPGA. J. Applied Sci., 12: 201-210.

Ramkumar, V., S. Rajasekar and S. Swamynathan, 2010. Scoring products from reviews through application of fuzzy techniques. Expert Syst. Appl., 37: 6862-6867.

Tan, S. and J. Zhang, 2008. An empirical study of sentiment analysis for Chinese documents. Expert Syst. Appl., 34: 2622-2629.

Zerigui, A., X. Wu and Z.Q. Deng, 2008. Communication of mobile rover based on FPGA, DSP and wireless communication. Inform. Technol. J., 7: 374-377.

Zhang, J., Q. Wu and J. Chen, 2012. Research on design method of dynamic partial reconfigurable system. J. Software Eng., 6: 21-30.