International Journal of Engineering Research-Online
A Peer Reviewed International Journal
Articles available online http://www.ijoer.in

Vol.3., Issue.2, 2015

RESEARCH ARTICLE

ISSN: 2321-7758

SERVICE RECOMMENDER SYSTEM FOR USER'S PERSONALIZED REQUIREMENTS

S.NANDHINI[1], S.PAVITHRA[2], N.POOJITHA[3], M.SANGEETHA[4]

[1,2,3]Students, [4]Assistant Professor Department of Computer Science and Engineering
Panimalar Engineering College

Article Received: 13/03/2015 | Article Revised on; 19 /03/2015 | Article Accepted on:24/03/2015

ABSTRACT

The main aim of this project is to develop a Keyword-Aware Service Recommendation method, named KASR, to address scalability and inefficiency problem in big data with traditional service recommender systems, which fails to meet user's personalized requirements.

S.NANDHINI

S.PAVITHRA

N.POOJITHA

INTRODUCTION

Big data has become the major problem in the IT Technology. Big data as the name suggests is all about handling large amount of data. It is used to describe any type of data which is structured, unstructured and grow so large which is difficult to manage with regular database and statistics tools. The rapid proliferation of data in the last decade leads to the scalability and inefficiency problems in the existing service recommender systems. In the current hotel service recommender systems are providing only basic information like overall ratings about the hotels. This information will not be sufficient for current user to get recommendations as per each user personalized requirements, because different user has diverse preferences. Thus the traditional service recommender systems fails to meet user personalized requirements.

Current service recommendation methods are usually used. Content-based approaches recommend services similar to those the user preferred in the past. Collaborative filtering (CF) approaches recommend services to the user that users with similar tastes preferred in the past. Hybrid approaches combine content-based and CF methods in several different ways.

In CF based systems, users will receive recommendations based on people who have similar tastes and preferences, which can be further classified into item-based CF and user-based CF. In item-based systems; the predicted rating depends on the ratings of other similar items by the same user. While in user-based systems, the prediction of the rating of an item for a user depends upon the ratings of the same item rated by similar users.

As the time flies so fast in busy daily life, every user wishes to get recommendation service quickly as per his/her preferences. And in this work, we will take advantage of a user-based CF algorithm to deal with our problem. User based collaborative filtering algorithm is used to get appropriate recommendations to each user as per his/her requirements. For that, our proposed system is implemented in hadoop and on map reduce framework, a cloud computing tool [14].

EXISTING WORK

In most existing service recommender systems, such as hotel reservation systems and restaurant guides, the ratings of services and the service recommendation lists presented to users are the same. They have not considered users' different preferences, without meeting users' personalized requirements. Most existing service recommender systems are only based on a single numerical rating to represent a service's utility as a whole. In fact, evaluating a service through multiple criteria and taking into account of user feedback can help to make more effective recommendations for the users.Existing Approaches solve the scalability problem by dividing dataset. But their method doesn't have favorable scalability and efficiency if the amount of data grows.

PROPOSED WORK

Service recommendation method, for user's personalized requirements, is proposed in this paper, which is based on a user-based Collaborative Filtering algorithm [6].In KASR, keywords extracted from reviews of previous users are used to indicate their preferences. Moreover, we implement it on HADOOP Map Reduce as its computing framework. In KASR, keywords are used to indicate both of users' preferences and the quality of candidate services. A user-based CF algorithm is adopted to generate appropriate recommendations. KASR aims at calculating a personalized rating of each candidate service for a user, and then presenting a personalized service recommendation list and recommending the most appropriate services to him/her. Moreover, to improve the scalability and efficiency of our recommendation method in, we implement it by splitting the proposed algorithm into multiple Map Reduce phases.

MODULE DESCRIPTION

1 .BIG DATA AND ENVIRONMENT

Huge Collection of data is retrieved from open source datasets that are publicly available from major Travel Recommendation Applications. Big Data Schemas were analyzed and a Working Rule of the Schema is determined. The CSV(Comma separated values) files were read and manipulated using Java API that itself developed by us which is developer friendly ,light weighted and easily modifiable.

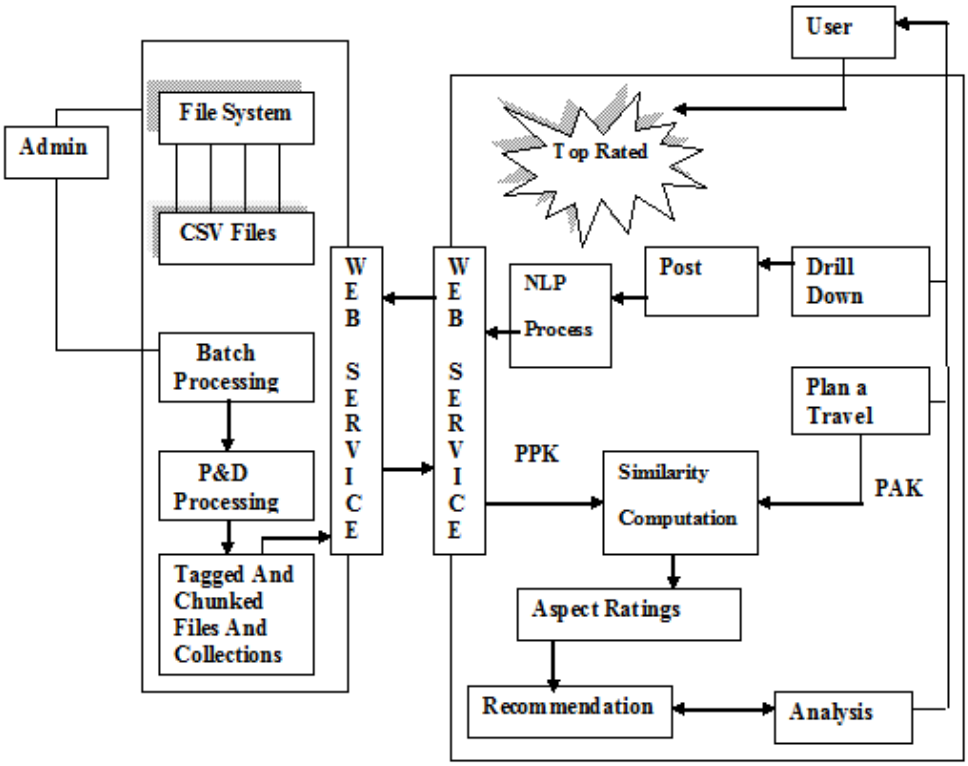
2. BATCHING AND PREPROCESS

The Traditional View of Service Recommender Systems that shows Top-K Results are displayed with Paginations with which a user can navigate Back and Forth of the Result sets. All Services Ratings and Reviews of Each Hotels are listed. POS (Parts of Speech) Tagger and Chucker Process are done on each and every review of all hotels for all countries in a Parallel and Distributed Manner as Batch jobs. The Master Job is Split up into 'n' no of small Batch jobs based on the slave machines Connected with the Master. POS Tagger tags each words of a review with its tags and the Clunker Process will take POS tagged output as input for Groping the Words based on meaning of the Review.

3. DIGGING IN BIG DATA & SERVICE RECOMMENDER APPLICATION

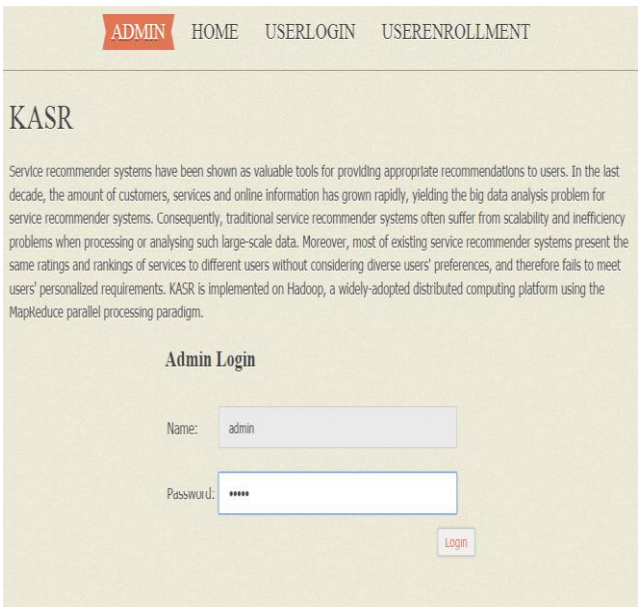
The CSV Files in distributed Systems are invoked through Web Service Running in the Server Machine of the Host Process through a Web Service Client Process in the Recommendation System [11].The data that Retrieved to the Recommendation Systems are provided with a clean GUI and can be queried on Demand. Each and Every process on the Recommendation Application invokes Web Service which uses light weighted traversal of data using XML. The Users can Review each hotel and can post comments also. The Reviews gets updated to the CSV Files as it get retrieved. A User can scan or schedule a Travel highlighting his requirements in a detailed way that shows the Preference Keywords Set of the Active User. A Domain Thesaurus is built depending on the Keyword Candidate List and Candidate Services List. The Domain Thesaurus can be Updated Regularly to get accurate Results of the Recommendation System.

ARCHITECTURE DIAGRAM:

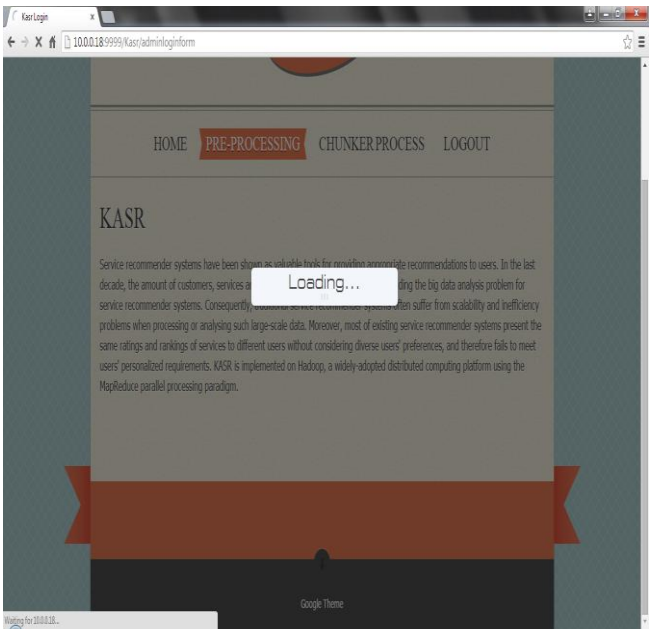


SCREEN SHOTS :

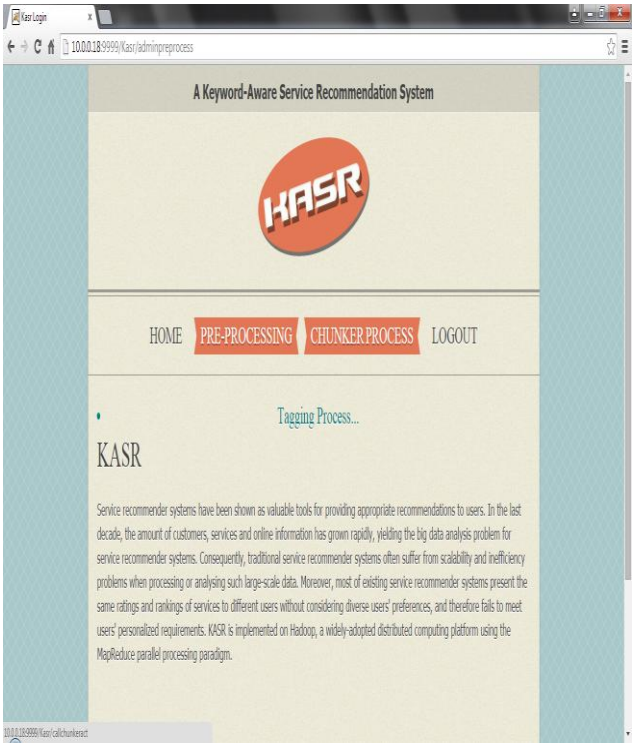
Admin login:

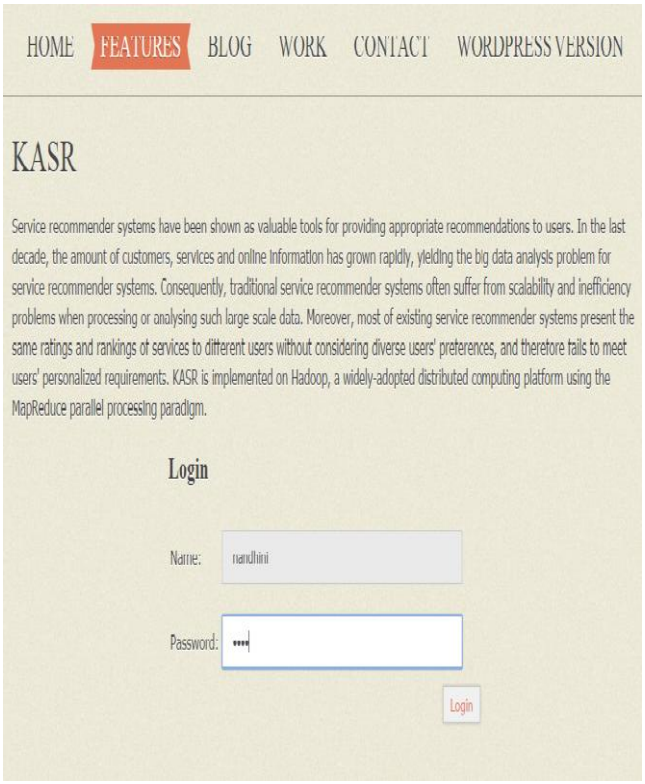


Preprocessing-Tagging:

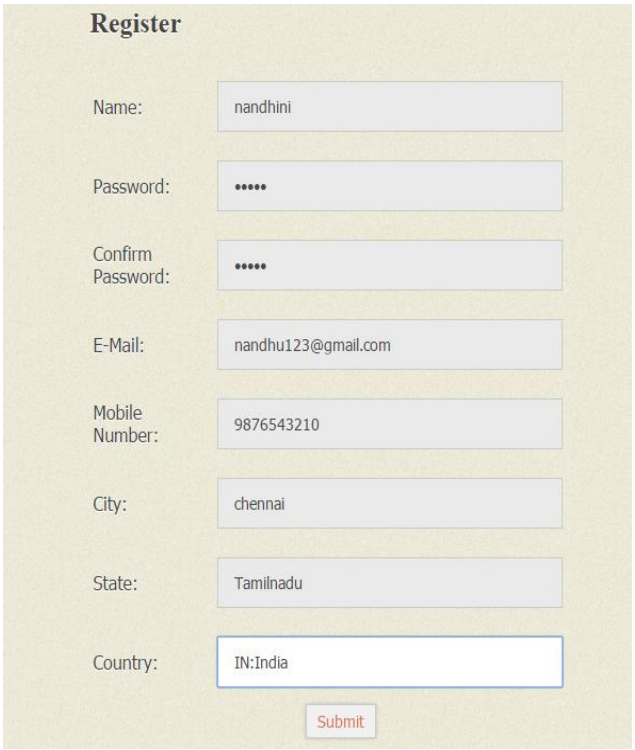


Chunking process:

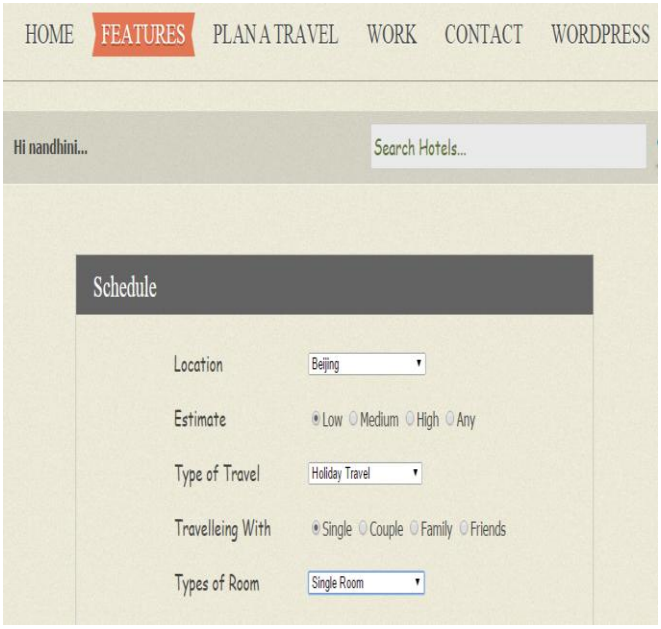


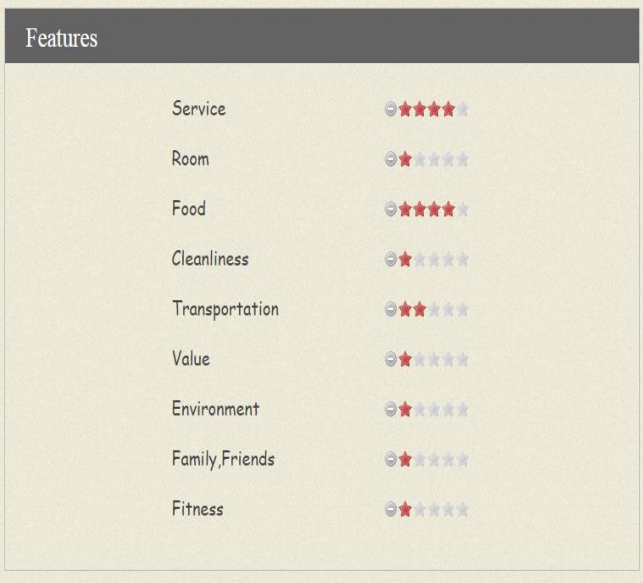


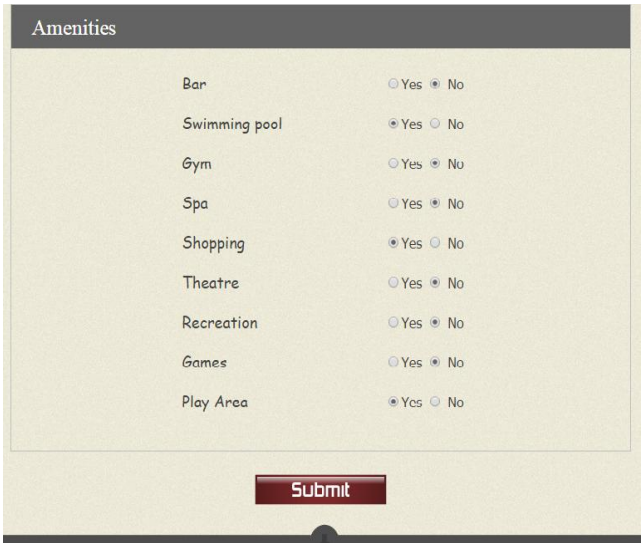
User's Enrollment and login:



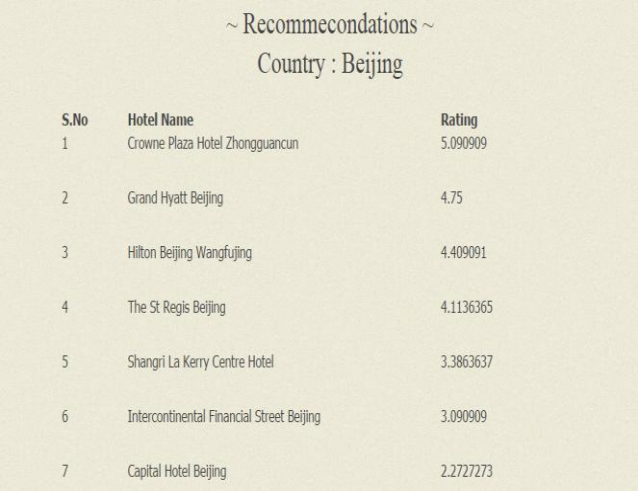
User's requirements are given:







OUTPUT:



~ Recommecondations ~

Country : Beijing

S.No	Hotel Name	Rating
1	Crowne Plaza Hotel Zhongguancun	5.090909
2	Grand Hyatt Beijing	4.75
3	Hilton Beijing Wangfujing	4.409091
4	The St Regis Beijing	4.1136365
5	Shangri La Kerry Centre Hotel	3.3863637
6	Intercontinental Financial Street Beijing	3.090909
7	Capital Hotel Beijing	2.2727273

4. MAP REDUCE AND HADOOP

To improve the scalability and efficiency of service recommendation methods in big data environment we use Hadoop. HADOOP is an open source cloud computing platform inspired by map reduce framework of Google to overcome the problems of processing large volume of data. The preferences of active users and previous users are transformed into their corresponding preference keyword sets PAK and PPK respectively. PAK represents Preferences of active user Keyword set and PPK represents Preferences of Previous user Keyword set.

Active user*:*

An active user refers to a current user needs recommendation .Active user can give his/her preferences by selecting keywords from a keyword-candidate list. The active user should also select the importance degree of the keywords as "1" represents the general, "3" represents important and "5" represents very important.

Previous user:

Previous User refers to people who has visited the particular hotel and have given his/her reviews that will be updated in the database. Keywords will be extracted from those reviews. In the review snippet collection we tag each review based on pos (parts of speech). We remove meaningless tags and stop words in the reviews snippet collection. To remove the commoner morphological and in flexional endings from words in English we use the Porter Stemmer Algorithm.

Each review will be converted into a keyword according to the keyword-candidate list and domain thesaurus. If the review contains a keyword as in the domain thesaurus, then the corresponding keyword should be extracted into the Preferences of Previous user Keyword set (PPK).

Similarity computation:

In similarity computation, reviews of previous users who have similar tastes to an active user should be identified; thereby we may be able to find the neighborhoods of the active user based on the similarity of their preferences. If the intersection of PPK and PAK is null set, then it implies no similarity keyword between the previous user and active user. Thereby reviews that are not similar between the active and previous user will be removed.

5. KASR and Analysis:

The Chunked Reviews of the Similar User List is retrieved and the Keywords corresponding to the User is analyzed for its Valence and Arousal. Valence Means Weather the Keywords Means a positive or Negative thing and Arousal answers, how much it is? Ratings are given for each Domain based on the Valence and Arousal for each User of each hotel. The Overall Hotel Rating is now manipulated by taking average values of each Rating of several users of a particular hotel. Now ranking is done for all hotels based on Ratings and will be sorted based on Bubble Sort Algorithm to have the Most appropriate personalized Recommendation for the User. The Results will be analyzed with Graphical Views so as to understand easier.

CONCLUSION

To improve the scalability and inefficiency problem in existing service recommender systems, we go for a method named KASR-keyword aware service recommendation method. In KASR Active user gives his/her requirements by selecting keywords from keyword candidate list. Besides the active user should also select the importance degree of the keywords. User-based collaborative filtering algorithm is used to generate appropriate recommendations. We do POS tagging and chunking process to get a meaningful keyword while extraction. We distinguish the positive and negative preferences from the previous user reviews via NLP, valence and arousal so that we get more accurate predictions. Improved version of KASR is implemented on a Mapreduce framework, parallel computing and in Hadoop platform to address the big data problems so that it improves the scalability and efficiency over existing approaches.

REFERENCES

[1]. J. Manyika et al., "Big Data: The Next Frontier for Innovation,Competition, and Productivity," 2011.

[2]. C. Lynch, "Big Data: How Do Your Data Grow?" Nature, vol. 455,no. 7209, pp. 28-29, 2008.

[3]. F. Chang, J. Dean, S. Ghemawat, and W.C.Hsieh, "Big table: ADistributed Storage System for Structured Data," ACM Trans.Computer Systems, vol. 26, no. 2, article 4, 2008.

[4]. W. Dou, X. Zhang, J. Liu, and J. Chen, "Hire Some-II: Towards Privacy-Aware Cross-Cloud Service Composition for Big DataApplications," IEEE Trans. Parallel and Distributed Systems, 2013.

[5]. G. Linden, B. Smith, and J. York, "Amazon.com Recommendations:Item-to-Item Collaborative Filtering," IEEE Internet Computing,vol. 7, no. 1, pp. 76-80, Jan. 2003.

[6]. M. Bjelica, "Towards TV Recommender System Experiments withUser Modeling," IEEE Trans. Consumer Electronics, vol. 56, no. 3,pp. 1763-1769, Aug. 2010.

[7]. M. Alduan, F. Alvarez, J. Menendez, and O.Baez, "RecommenderSystem for Sport Videos Based on User Audiovisual Consumption,"IEEE Trans. Multimedia, vol. 14, no. 6, pp. 1546-1557,Dec. 2012.

[8]. Y. Chen, A. Cheng, and W. Hsu, "Travel Recommendation by Mining People Attributes and Travel Group Types from Community-Contributed Photos," IEEE Trans. Multimedia, vol. 25, no. 6,pp. 1283-1295, Oct. 2013.

[9]. Z. Zheng, X. Wu, Y. Zhang, M.Lyu, and J. Wang,"QoS RankingPrediction for Cloud Services," IEEE Trans. Parallel and DistributedSystems, vol. 24, no. 6, pp. 1213-1222, June 2013.

[10]. W. Hill, L. Stead, M. Rosenstein, and G. Furnas,"Recommendingand Evaluating Choices in a Virtual Community of Use," Proc.SIGCHI Conf. Human Factors in Computing System (CHI '95),pp. 194-201, 1995.

[11]. P. Resnick, N. Iakovou, M. Sushak, P. Bergstrom, and J. Riedl,"GroupLens: An Open Architecture for Collaborative Filtering ofNetnews," Proc. ACM Conf. Computer Supported Cooperative Work(CSCW '94), pp. 175-186, 1994.

[12]. R. Burke, "Hybrid Recommender Systems:Survey andExperiments," User Modeling and User-Adapted Interaction, vol. 12,no. 4, pp. 331-370, 2002.

[13]. G. Adomavicius and A. Tuzhilin, "Toward the Next Generation of Recommender Systems: A Survey of the State of- the-Art and Possible Extensions," IEEE Trans. Knowledge and Data Eng., vol. 17,no. 6, pp. 734-749, June 2005.

[14]. D. Agrawal, S. Das, and A. El Abbadi, "Big Data and Cloud Computing:New Wine or Just New Bottles?" Proc. VLDB Endowment,vol. 3, no. 1, pp. 1647-1648, 2010.

[15]. J. Dean and S. Ghemawat, "MapReduce Simplified Data Processing on Large Clusters," Comm. ACM, vol. 51, no. 1,pp. 107-113,2005.

[16]. G. DeCandia, D. Hastorun, M. Jampani, G. Kakulapati, A.Lakshman, A. Pilchin, S. Sivasubramanian, P. Vosshall, and W.Vogels, "Dynamo: Amazons Highly Available Key-ValueStore," In Proc. 21st ACM SympOperating Systems Principles,pp. 205-220, 2007.

[17]. M. Isard, M. Budiu, Y. Yu, A. Birrell, and D. Fetterly, "Dryad: Distributed Data-Parallel Programs from Sequential Building Blocks," Proc. European Conf. Computer Systems, pp. 59-72, 2007.

[18]. S. Ghemawat, H. Gobioff, and S. T. Leung, "The Google File System," Proc. 19th ACM Symp. Operating Systems Principles, pp. 29- 43, 2003.

[19]. L. Zhang, "Editorial: Big Services Era: Global Trends of Cloud Computing and Big Data," IEEE Trans. Services Computing, vol. 5, no. 4, pp. 467-468, Fourth Quarter, 2012.

[20]. Z. Luo, Y. Li, and J. Yin, "Location: A Feature for Service Selection. in the Era of Big Data," Proc. IEEE 20th Int'l Conf .WebService,pp.515-522,2013