ISSN 2395-1621

A Keyword Aware Service Recommendation Method For Hotel Recommendation using Big data

#1Anjali Yadav, #2R.B Thakur

1anjaliyadav2807@gmail.com

#1IOKCOE,Savitribai Phule University,Pune

ABSTRACT

Service recommender systems have been shown as valuable tools for providing appropriate recommendations to users. In the last decade, the amount of customers, services and online information has grown rapidly, yielding the big data analysis problem for service recommender systems. Consequently, traditional service recommender systems often suffer from scalability and inefficiency problems when processing or analysing such large-scale data. Moreover, most of existing service recommender systems present the same ratings and rankings of services to different users without considering diverse users' preferences, and therefore fails to meet users' personalized requirements. In this paper, we propose a Keyword-Aware Service Recommendation method, named KASR, to address the above challenges. It aims at presenting a personalized service recommendation list and recommending the most appropriate services to the users effectively. Specifically, keywords are used to indicate users' preferences, and a user-based Collaborative Filtering algorithm is adopted to generate appropriate recommendations. To improve its scalability and efficiency in big data environment, KASR is implemented on Hadoop, a widely-adopted distributed computing platform using the MapReduce parallel processing paradigm. Finally, extensive experiments are conducted on real-world data sets, and results demonstrate that KASR significantly improves the accuracy and scalability of service recommender systems over existing approaches.

Keywords— **Recommender system, preference, keyword, big data, MapReduce, Hadoop,Navie bayes, data mining**

ARTICLE INFO

Article History

Received :12th January 2016

Received in revised form :

13th January 2016

Accepted :19th January , 2016

Published online :

21st January 2016

I. INTRODUCTION

1.1 Background

IN recent years, the amount of data in our world has been increasing explosively, and analyzing large data sets—so-called "Big Data"—becomes a key basis of competition underpinning new waves of productivity growth, innovation, and consumer surplus [1]. Then, what is "Big Data"?, Big data refers to data sets whose size is beyond the ability of current technology, method and theory to capture, manage, and process the data within a tolerable elapsed time. Today, Big Data management stands out as a challenge for IT companies. The solution to such a challenge is shifting increasingly from providing hardware to provisioning more manageable software solutions [2]. Big data also brings new opportunities and critical challenges to industry and academia [3], [4].

Similar to most big data applications, the big data tendancy also poses heavy impacts on service recommender systems. With the growing number of alternative services, effectively recommending services that users preferred has become an important research issue. Service recommender

0S. Meng and W. Dou are with the State Key Laboratory for Novel Software Technology, the Department of Computer Science and Technology, Nanj-ing University, Nanjing 210023, China. E-mail: shunmei89@gmail.com, douwc@nju.ed-u.cn.

0X. Zhang and J. Chen are with the Faculty of Engineering and IT, Univer-sity of Technology, Sydney, NSW

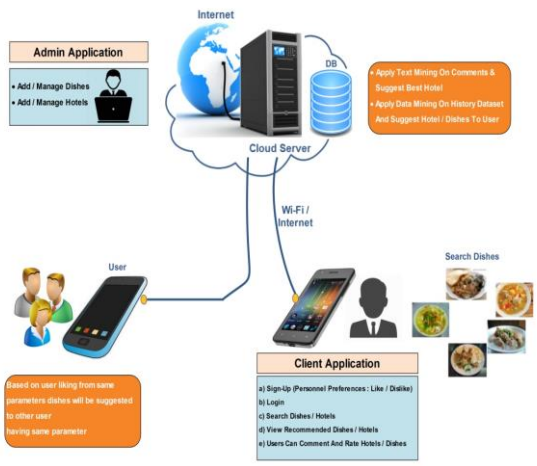
2007,Australia. E-mail: {xyzhanggz, jinjun}@gmail.com.

Manuscript received 1 Dec. 2013; revised 15 Dec. 2013; accepted 16 Dec. 2013. Date of publication 15 Jan. 2014; date of current version 14 Nov. 2014. Recommended for acceptance by V. Misic.
For information on obtaining reprints of this article, please send e-mail to: reprints@ieee.org, and reference the Digital Object Identifier below.
Digital Object Identifier no. 10.1109/TPDS.2013.2297117
systems have been shown as valuable tools to help users deal with services overload and provide appropriate rec-ommendations to them. Examples of such practical appli-cations include CDs, books, webpages and various other products now use recommender systems [5], [6], [7]. Over the last decade, there has been much research done both in industry and academia on developing new approaches for service recommender systems [8], [9].

1.2 Motivation

With the success of the Web 2.0, more and more compa-nies capture large-scale information about their custom-ers, providers, and operations. The rapid growth of the number of customers, services and other online informa-tion yields service recommender systems in “Big Data” environment, which poses critical challenges for service recommender systems. Moreover, in most existing service recommender systems, such as hotel reservation systems and restaurant guides, the ratings of services and the ser-vice recommendation lists presented to users are the same. They have not considered users’ different preferen-ces, without meeting users’ personalized requirements. Following is an example in hotel reservation system illus-trating such a case.
Example 1. Alice and Tom are respectively browsing a hotel reservation website to reserve a hotel in Kowloon, Hong Kong. But the ratings and recommendation list of the hotels provided by the website to them are the same. Assuming there are three hotels in Kowloon: A, B and C. Comparing the three hotels, A is convenient to the air-port and has a shopping mall nearby; B has convenient transportation with an underground station nearby and owns comfortable accommodation equipment.

II. SYSTEM ARECHITECTURE



III. MATHEMATICAL MODELLING

A mathematical model is a description of a system using mathematical concepts and language. The process of developing a mathematical model is termed mathematical modeling. Mathematical models are used in the natural sciences (such as physics, biology, earth science, meteorology) and engineering disciplines (such as computer science, artificial intelligence), as well as in the social sciences (such as economics, psychology, sociology, political science). Physicists, engineers, statisticians, operations research analysts, and economists use mathematical models most extensively. A model may help to explain a system and to study the effects of different components, and to make predictions about behavior**.**

Mathematical Model.

Set Theory :

Let s (be a main set of) ≡ {SDB, LDB, C, A, S, MR, AO}

where,

SDB is the copy of the server database. This database is responsible for storing user information related to cloud interactions.

LDB is a set of local database that a user owns. It consists of data tables having data items related to the products and their sales transactions.

C is a set of all clients using the server database and mining services from the server. And (c_1 , c_2 , c_3, ...c_n) Є C.

A is a set of algorithms applied on the input data to get mining results.

S is the server component of the system. The server is responsible for registering, authenticating and providing associations to the end user.

MR is a set of mining rules that are applied on the input dataset provided by the client from his LDB. And (mr_1 , mr_2 , mr_3, ...mr_n) Є MR

AO is a set of associations that are extracted from the input and a form the output of the system.

Functionalities :

SDB' = RegisterUser(uid, password, fullname, address, country, contact, email);
password = SHA1(input_password);
U = AuthenticateUser(uid, password, SDB');
LDB1 = ManageProducts(pid, product name, cost);
LDB2 = ManageBilling(transactions, items);
LDB = LDB1 + LDB2
ED(Encoded data) = EncodeTransactions(LDB2, EncodingAlgorithm(EA));
UPLOAD(ED);
AO = Apply Mining(ED);
Results = Decode(Download(AO));

IV. FUTURE SCOPE

Advanced data mining Algorithm can be used for more accurate results using sufficient dataset. In our future work, we will do further research in how to deal with the case where term appears in different categories of a domain thesaurus from context and how to distinguish the positive and negative preferences of the users from their reviews to make the predictions more accurate.

V. CONCLUSION

In this paper, we have proposed a keyword-aware service recommendation method, named KASR. In KASR, key-words are used to indicate users' preferences, and a user-based Collaborative Filtering algorithm is adopted to generate appropriate recommendations. More specifically, a key-word-candidate list and domain thesaurus are provided to help obtain users' preferences. The active user gives his/her preferences by selecting the keywords from the keyword-candidate list, and the preferences of the previous users can be extracted from their reviews for services according to the keyword-candidate list and domain thesaurus. Our method aims at presenting a personalized service recommendation list and recommending the most appropriate service(s) to the users.

ACKNOWLEDGEMENT

W. Dou and J. Chen are the corresponding co-authors of this paper. This paper was partly supported by project National Science Foundation of China under Grant 91318301, and 61321491; National Key Technology R&D Program of the Ministry of Science and Technology under Grant 2011BAK21B06.

REFRENCES

[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, "Bigtable: A Distributed 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, "HireSome-II: Towards Privacy-Aware Cross-Cloud Service Composition for Big Data Applications," IEEE Trans. Parallel and Distributed Systems, 2013.

[4]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.

[5]M. Bjelica, "Towards TV Recommender System Experiments with User Modeling," IEEE Trans. Consumer Electronics, vol. 56, no. 3,

42. 1763-1769, Aug. 2010.
43.

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

