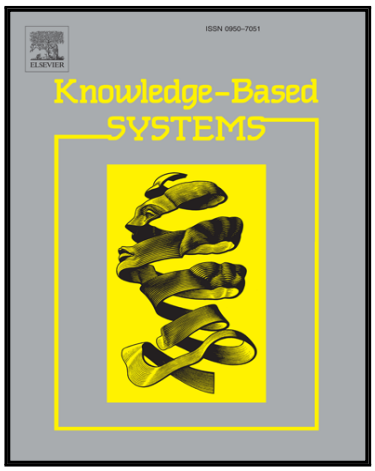

Service Recommendation based on Quotient Space Granularity Analysis and Covering Algorithm on Spark

Yi-wen Zhang , Yuan-yuan Zhou , Fu-tian Wang , Zheng Sun , Qiang He

PII: S0950-7051(18)30066-2
DOI: 10.1016/j.knosys.2018.02.014
Reference: KNOSYS 4222

To appear in: *Knowledge-Based Systems*

Received date: 22 July 2017
Revised date: 5 February 2018
Accepted date: 6 February 2018

Please cite this article as: Yi-wen Zhang , Yuan-yuan Zhou , Fu-tian Wang , Zheng Sun , Qiang He , Service Recommendation based on Quotient Space Granularity Analysis and Covering Algorithm on Spark, *Knowledge-Based Systems* (2018), doi: 10.1016/j.knosys.2018.02.014

Service Recommendation based on Quotient Space Granularity Analysis and Covering Algorithm on Spark

Yi-wen Zhang[1], Yuan-yuan Zhou[1*], Fu-tian Wang[1], Zheng Sun[2], Qiang He[3]

1. Key Laboratory of Intelligent Computing & Signal Processing; Ministry of Education, Anhui University, Hefei, Anhui, China;

2. School of Science, Arcadia University, Glenside, America;

3. School of Software and Electrical Engineering, Swinburne University of Technology, Melbourne, Australia.

Email addresses: zhangyiwen@ahu.edu.cn; yyzhouahu@qq.com; wft@ahu.edu.cn; zsun@arcadia.edu; qhe@swin.edu.au

Abstract− The rapid growth of Web services has made it a challenge for users to find appropriate Web services because it is very difficult for traditional Web service recommendation approaches to process the large amount of service-relevant data. To address this issue, this paper proposes CA-QGS (**C**overing **A**lgorithm based on **Q**uotient space **G**ranularity analysis on **S**park), a scalable approach for accurate Web service recommendation in large-scale scenarios. CA-QGS first clusters users and Web services based on users' past quality experiences on co-invoked Web services. It then performs granularity analysis on the clustering results to identify users and Web services that are similar to the target user and Web service, and employs the collaborate filtering technique to predict the quality of the target Web service for the target user. This way, appropriate Web services can finally be recommended to the target user. To increase the efficiency of CA-QGS, we parallelize CA-QGS on Spark. Extensive experiments show that CA-QGS outperforms existing approaches in both recommendation accuracy and efficiency.

Keywords: Covering algorithm, Quotient space, Recommender systems, Association matrices

1. Introduction

Big data analytics is being employed in many areas to improve the quality and value of a variety of services [1]. In the big data environment, more and more services are deployed in the cloud to provide different functionalities [2]. According to Programmable Web, an online Web service repository, the number of published Web services has increased by four times since 2009. The statistics published webservices.seekda.com, a Web service search engine, also indicate an exponential growth in the number of published Web services in the past several years [38]. The rapid increase in the number of Web services, as well as users, has generated a large amount data on users, Web services and users' experiences on Web services. It is essential for researchers to explore and analyze this big data to extract useful information for predicting the quality of Web services and recommending Web services to users. In recent years, the collaborative filtering (CF) technique, such as item-based [21, 23, 24, 25] and user-based [17, 18, 22] methods, have been widely employed in the implementation of recommender systems for Web services. Its fundamental theory is to analyze the historical quality values of Web services