Chapter 35
Personalized Recommendation Mechanism Based on Collaborative Filtering in Cloud Computing Environment

Xinling Tang
School of Electronics and Information Engineering, Hunan University of Science and Engineering, Yongzhou, China

Yonghong Tan
School of Electronics and Information Engineering, Hunan University of Science and Engineering, Yongzhou, China

Hongyan Xu
Department of Public Science Education Committee, Zhengzhou Shuqing Medical College, Henan Zhengzhou, China

Yanjun Gong
School of Electronics and Information Engineering, Hunan University of Science and Engineering, Yongzhou, China

ABSTRACT

With the advent of cloud computing era and the dramatic increase in the amount of data applications, personalized recommendation technology is increasingly important. However, due to large scale and distributed processing architecture and other characteristics of cloud computing, the traditional recommendation techniques which are applied directly to the cloud computing environment will be faced with low recommendation precision, recommended delay, network overhead and other issues, leading to a sharp decline in performance recommendation. To solve these problems, the authors propose a personalized recommendation collaborative filtering mechanism RAC in the cloud computing environment. The first mechanism is to develop distributed score management strategy, by defining the candidate neighbors (CN) concept screening recommended greater impact on the results of the project set. And the authors build two stage index score based on distributed storage system, in order to ensure the recommended mechanism to locate the candidate neighbor. They propose collaborative filtering recommendation algorithm based on the candidate neighbor on this basis (CN-DCF). The target users are searched in candidate neighbors by the nearest neighbor k project score. And the target user's top-N recommendation sets are predicted. The results show that in the cloud computing environment RAC has a good recommendation accuracy and efficiency recommended.

DOI: 10.4018/978-1-7998-5339-8.ch035

1. INTRODUCTION

The rapid growth of the scale of the Internet brings the problem of information overload. It is difficult for users to obtain valuable information from massive data. A sharp decline is in the efficiency of the use of information. Personalized recommendation is as an important means of information filtering. By the analysis of the user's interests and historical behavior, we can recommend potential items of interest to the user, in order to effectively solve the problem of information overload (Tunkelang, 2011). Collaborative filtering is one of the most successful personalized recommendation technologies. The basic idea is to search for the target user affinity neighbors' users, and the neighbor who is interested in the project is recommended to target users (Breese, Heckerman, & Kadie,1998). At present, collaborative filtering has become the core technology of electronic commerce, social networking and video sharing Internet applications (Adomavicius & Tuzhilin,2005; Manzato & Goularte, 2012).

Due to the characteristics of cloud computing (Mell & Grance, 2011) and large scale distributed processing architecture, the internet applications deployed to the cloud computing environment in the data storage location, data storage capacity, and scale applications, etc. are different from the traditional model. Therefore, the traditional collaborative filtering algorithm applied directly to the cloud computing environment will produce the following problems

1. **The Accuracy of Recommendation Problem:** Traditional collaborative filtering recommendation algorithm based on the user's rating data is stored on the server side of the database. The recommendation algorithm obtains the k nearest neighbor of the target user by computing the similarity between users. And according to the k nearest neighbors, the recommended results are generated. However, in cloud computing environment, in order to improve the efficiency of data parallel processing and I/O performance by the cloud service providers, we construct data center based on nothing shared (SN) architecture. Under the SN architecture, each node in the data center has independent local storage, and the data is distributed to each node (Ghemawat, Gobioff, & Leung, 2003). Under the framework of SN system, according to the stored data block strategy of Internet application, projects may be distributed to each node. User access to the project, the evaluation of the data will be distributed to the corresponding data stored in each node. Traditional centralized collaborative filtering recommendation algorithm may lose some of the recommended values in local nodes to calculate the similarity of users, which leads to the reduction of the accuracy of the algorithm.
2. **Recommended Latency:** Cloud computing platform as its business operations environment of Internet companies generally have large-scale applications. Business operation process will produce a large amount of rating data. The traditional collaborative filtering recommendation algorithm is used to deal with this kind of applications. Frequent traversal of the scoring matrix will lead to a sharp increase in the recommended delay. At the same time, ratings data transmission delay between nodes will further exacerbate the problem of high latency recommendation algorithm. This is clearly unable to meet the needs of real-time applications recommended.
3. **Network Overhead Problem:** Because the user interest will change with time, we need to regularly update the similarity matrix in order to improve the accuracy of the traditional collaborative filtering recommendation algorithm. Especially the collaborative filtering recommendation algorithm based on user is update more frequently. In cloud computing environment, the score data are distributed to store. Each storage node and the recommended node will frequently exchange score