PowKMeans:
A Hybrid Approach for Gray Sheep Users Detection and Their Recommendations

Honey Jindal, Department of Computer Science and Engineering, Jaypee Institute of Information Technology, Noida, India
Shalini Agarwal, Department of Computer Science and Engineering, Jaypee Institute of Information Technology, Noida, India
Neetu Sardana, Department of Computer Science and Engineering, Jaypee Institute of Information Technology, Noida, India

ABSTRACT
This article describes how recommender systems are software applications or web portals that generate personalized preferences using information filtering techniques, with a goal to support decision-making of the users. Collaborative-based techniques are often used to predict the unknown preferences of the user based upon his past preferences or the preferences of the similar users that have already been identified. A user which has a high correlation with any group of users is known as white user whereas the users which have less correlation with any group of users are known as gray-sheep users. The presence of gray-sheep users affects the accuracy of the model, and generates inaccurate predictions. To improve the prediction accuracy, it is important to differentiate gray-sheep users from white users. Experimental results show that PowKMeans is effective in improving the prediction accuracy by 4.62%. It has also shown reduction in Mean Absolute Error by 0.7757.

KEYWORDS
Collaborative Filtering, Gray-Sheep Users, K-Means++, PowKMeans, Recommender System

1. INTRODUCTION
Recommendation systems are effective software used to solve the problem of information overload over the internet. They can automatically recommend items to the target users’ according to observed user preference information, purchase behaviour, evaluation behaviour and so on. Recommendation systems are extensively used in various applications ranging from e-commerce to digital library and knowledge management. Due to the broad usage of recommendation system, it becomes tedious to recommend someone appropriate without filtering the data for relevant choices. It depends upon several factors like user rating given to a collection of items based upon their satisfaction level, their likes and dislikes, age, gender, occupation, region or locality, community, etc. Some of the popular websites that are using a recommendation engine to filter choices are Amazon, Hulu, Netflix, FlipKart, Last.fm, Pandora, etc. Recommendation systems automatically recommend items to the target user based on their past purchased behaviour. For example, a customer usually expresses his interest in an item- either by viewing a product description or by placing the item in his “Shopping cart.” Thus. the

DOI: 10.4018/IJITWE.2018040106

Copyright © 2018, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.
customer will likely receive recommendations for related and/or additional products. These products can be recommended based on the top overall sellers on a site or on an analysis of the past buying behaviour of the customer.

Collaborative filtering is one of the well-suited techniques applied on recommendation systems. Collaborative filtering techniques recommend related items to the user by analyzing similar user’s preferences. Using collaborative filtering, description of items is not required and this technique recommends items regardless of the type of items. Despite the advantages of collaborative filtering, some problems are introduced by its own, namely, cold-start, data sparsity and gray sheep problem. Collaborative filtering failed to recommend items when the purchase history of the user is not available. This problem is known as user cold-start problem (Jindal & Singh, 2014). Due to the large size of user-item matrix, rating details are extremely sparse, which gives rise to the data sparsity problem. Another problem occurs in collaborative filtering, when the opinions of users do not consistently agree or disagree with any group of people will receive inaccurate recommendations. These types of users are recognized as Gray sheep users (Ghazanfar & Prugel-Bennett, 2011).

In this paper, we focused on gray sheep users; a popular problem in a recommendation system. A user which has a high correlation with a group of users is known as white user and a user which has less correlation and/or no correlation with other users is known as a gray sheep user. Therefore, when we use collaborative filtering for predicting user’s preferences, precise recommendations cannot be generated. The existence of gray sheep users have two negative effects on recommender systems: 1) these users cannot receive precise recommendations 2) they have negative recommendations on other users’ recommendations (Ghazanfar & Prugel-Bennett, 2014). An appropriate solution for reducing the effect of the gray-sheep users in the performance of recommendation system is separating these users from other users. Therefore, we proposed a novel method, PowKMeans which separates the gray sheep users from the white users. Then, we further use content based user profiling method to generate the recommendations for these identified users.

1.1. Major Contributions

In this paper, we make the following contributions:

1. We propose, PowKMeans, a novel clustering algorithm to detect gray sheep users;
2. We propose a content based user profiling method for recommendations to the gray sheep users.

1.2. Organization

The rest of the paper is organized as follows: Section 2 provides the broad overview of gray-sheep users. Section 3 summarizes the related work. Section 4 presents a proposed method in detail. Section 5, shows the result and simulation of work done. Section 6 includes conclusion and future work.

2. GRAY-SHEEP USERS

Due to the advancement in online-recommendation sites, lots of users with diverse taste have utilized the recommendation system broadly. These recommendation systems help the user to identify their desired product within a short-span of time while saving their time to reach the preferred stores and purchase the product. The major challenge with the recommendation systems is that to provide precise recommendations which are compatible with users’ needs.

One of the widely used techniques in recommendation systems is collaborative filtering, which is used for suggesting or recommending items to users whose preferences match with other users in the recommendation system. The main assumption of collaborative filtering is that a given user will prefer items which are used/purchased/liked by people who have similar preferences as the given user. In such recommender systems, the profiles of the users are expressed in terms of their preferences for
Related Content

Dimensions of Business-to-Consumer (B2C) Systems Success in Kuwait: Testing a Modified DeLone and McLean IS Success Model in an E-Commerce Context
[www.igi-global.com/chapter/dimensions-of-business-to-consumer-b2c-systems-success-in-kuwait/140849?camid=4v1a](www.igi-global.com/chapter/dimensions-of-business-to-consumer-b2c-systems-success-in-kuwait/140849?camid=4v1a)

A Novel Architecture for Deep Web Crawler
[www.igi-global.com/article/novel-architecture-deep-web-crawler/52804?camid=4v1a](www.igi-global.com/article/novel-architecture-deep-web-crawler/52804?camid=4v1a)
Requirement Estimation and Design of Tag software in Web Application
[www.igi-global.com/article/requirement-estimation-and-design-of-tag-software-in-web-application/115932?camid=4v1a](www.igi-global.com/article/requirement-estimation-and-design-of-tag-software-in-web-application/115932?camid=4v1a)

Experiences with Software Architecture Analysis of Usability
[www.igi-global.com/article/experiences-software-architecture-analysis-usability/2648?camid=4v1a](www.igi-global.com/article/experiences-software-architecture-analysis-usability/2648?camid=4v1a)