Optimal Associative Neighbor Mining Using Attributes for Ubiquitous Recommendation Systems

Kyung-Yong Jung[1], Hee-Joung Hwang[2], and Un-Gu Kang[2]

[1] School of Computer Information Engineering, Sangji University, Korea
kyjung@sangji.ac.kr
[2] Department of Information Technology Engineering,
Gachon University of Medicine and Science, Korea

Abstract. Ubiquitous recommendation systems predict new items of interest for a user, based on predictive relationship discovered between the user and other participants in Ubiquitous Commerce. In this paper, optimal associative neighbor mining, using attributes, for the purpose of improving accuracy and performance in ubiquitous recommendation systems, is proposed. This optimal associative neighbor mining selects the associative users that have similar preferences by extracting the attributes that most affect preferences. The associative user pattern comprising 3-*AU*s (groups of associative users composed of 3-users), is grouped through the ARHP algorithm. The approach is empirically evaluated, for comparison with the nearest-neighbor model and k-means clustering, using the MovieLens datasets. This method can solve the large-scale dataset problem without deteriorating accuracy quality.

1 Introduction

Nowadays, most personalized ubiquitous recommendation systems in ubiquitous commerce utilize collaborative filtering systems, in order to recommend increasingly appropriate items, which are based on the ratings of other users who have similar preference. Ubiquitous commerce that is connected to the existing e-commerce is developed. So it is interesting to provide the personalized item services suitable to the requirements and the activities of the users in ubiquitous space based on the electronic and physical space [14]. The Grouplens for Usenet news [10] was the first collaborative filtering system to deal with massive data sets. Grouplens addressed this large item set characteristic, by creating a separate item partition for each Usenet discussion group [2]. Collaborative filtering systems recommend information through building a user profile from various preferences for specific items, and comparing these preferences with other users, for the same items. The similarity of preferences between a specific user and other users is computed from the correlation coefficient. Predicting preference for a certain item is based on other users' preference for that item, and the similarity between each other. In reflecting human opinions, collaborative filtering has several advantages: filtering items that are not easily analyzed by automated processes, filtering items based on quality, and filtering serendipitous items. However, collaborative filtering does not use attributes of items in any way. In this paper, optimal associative neighbor mining using the attributes, for

H. Legind Larsen et al. (Eds.): FQAS 2006, LNAI 4027, pp. 458 – 469, 2006.

the purpose of improving accuracy and performance in ubiquitous recommendation systems, is presented. The term attribute is used for representing a primary attribute that influences the preference for the item.

The rest of this paper is organized as follows. Section 2 describes briefly the nearest-neighbor model, k-means clustering, and the threshold-neighbor model. Section 3 and Section 4 illustrate the proposed the optimal associative neighbor mining using the attributes in detail. In Section 5, the experimental results are presented. The conclusions are given in Section 6.

2 Neighbor Selections for Ubiquitous Recommendation Systems

Ubiquitous recommendation systems predict new items of interest for a user, based on the predictive relationship discovered between the user and other participants. In U-Commerce environment, to support personalization services, a system should recognize preferred items through the preference analysis such as personal profile or case of personal commerce. Based on this, ubiquitous systems recommend the items that are expected for the users to prefer according to the situation, so it can help commerce easily. Most successful research and commercial ubiquitous recommendation systems using collaborative filtering, are based on nearest-neighbor model [2], k-means clustering [1], and threshold-neighbor model [13].

2.1 Nearest-Neighbor Model

The nearest-neighbor model selects the nearest neighbors who have similar preference to the user by computing the similarities based on the preference. It only uses the neighbors who have higher correlation with the users than others. Collaborative filtering systems based on the nearest-neighbor method, works in three simple phases. First, users of the collaborative filtering system rate items that they have previously experienced. Second, the collaborative filtering system matches users with other participants of the system, having similar rating patterns. This is usually achieved through statistical correlation. The matches are selected, becoming known as neighbors of the user. Third, items that the neighbors have experienced and rated highly, but which the user has not yet experienced and rated highly, will be recommended to the user, ranked based on the neighbors to the user [2]. Collaborative filtering system should even consider some users who may give bad influences on prediction quality. It has been shown in several investigations that the ubiquitous recommendation system with the nearest neighbor model has better quality of prediction than the traditional collaborative filtering [13].

2.2 K-Means Clustering

The K-means clustering algorithm [1,12] has been shown to be effective in producing good cluster results for many practical applications and in non-hierarchical clustering methods. This algorithm initially takes the number of elements in the population equal to the final required number of clusters. The final required number of clusters is chosen, such that the points are mutually farthest apart. Next, each element in the population is examined and assigned to one of the clusters, depending on the