Publication year: 2011
Source: Applied Soft Computing, Available online 20 November 2011
Chih-Fong Tsai, Chihli Hung
Recommender systems, which recommend items of information that are likely to be of interest to the users, and filter out less favored data items, have been developed. Collaborative filtering is a widely used recommendation technique. It is based on the assumption that people who share the same preferences on some items tend to share the same preferences on other items. Clustering techniques are commonly used for collaborative filtering recommendation. While cluster ensembles have been shown to outperform many single clustering techniques in the literature, the performance of cluster ensembles for recommendation has not been fully examined. Thus, the aim of this paper is to assess the applicability of cluster ensembles to collaborative filtering recommendation. In particular, two well-known clustering techniques (self-organizing maps (SOM) andk-means), and three ensemble methods (the cluster-based similarity partitioning algorithm (CSPA), hypergraph partitioning algorithm (HGPA), and majority voting) are used. The experimental results based on the Movielens dataset show that cluster ensembles can provide better recommendation performance than single clustering techniques in terms of recommendation accuracy and precision. In addition, there are no statistically significant differences between either the three SOM ensembles or the threek-means ensembles. Either the SOM ork-means ensembles could be considered in the future as the baseline collaborative filtering technique.

Highlights

 This paper is the first study to examine the performance of clustering ensembles for collaborative filtering. In this paper, clustering ensembles by cluster-based similarity partitioning algorithm (CSPA), hypergraph partitioning algorithm (HGPA), and majority voting are compared with two single clustering techniques (i.e.k-means and SOM) in terms of accuracy, precision, and recall. The experimental results show that clustering ensembles outperform single clustering techniques. This allows future studies proposing novel clustering ensemble techniques to not only consider single clustering techniques as the baselines, but also compare either SOM or k-means ensembles for further validation in order to reach a more reliable conclusion.

Read Full Article