Recommender System for Fashion Design

Thiri Aung, Ei Chaw Htoon
University of Computer Studies, Yangon
padauk.april17@gmail.com, *htoon.eichaw@gmail.com*

Abstract

Recommender system applies knowledge of discovery technique to the problem of making personalized recommendations for information, products, or services. In this paper, an effective framework for combining content and collaborative filtering is used to predict new items of interest for a user. In addition, association rule mining is used to search for interesting relationships among items in a given data set and recommend suitable pairs of items. The end-use of these patterns could be for feed back into the design process and for future decisions or, in this case, to automatically adapt aspects of the site based on previous usage patterns. The proposed system is efficient to improve the user's satisfaction in the fashion design recommender system.

1. Introduction

Recommender systems help users, find and evaluate items of interest. They connect users with items to "consume" (purchase, view, listen to, etc) by associating the content of recommended items or the opinions of other individuals with the consuming user's actions or opinions. In recommender systems correlations are used to measure the extent of agreement between two users and used to identify users whose ratings will contain high predictive value for a given user [2]. Recommender systems attempt to predict items that a user may be interested in given some information about user's profile [7]. Collaborative filtering aims at predicting the user interest for a given item based on a collection of user profiles. Research started with memory-based approaches to collaborative filtering, that can be divided in user-based approaches like and item-based approaches like [4]. Collaborative filtering compares users according to their preferences. Therefore, a database of user's preferences must be available. The preferences can be collected either explicitly (explicit rating) or implicitly (implicit rating). The implicit ratings, on the other hand, are derived from monitoring the user's behavior [6]. Memory-based collaborative filtering first measures similarities between test user and other users (user-based), or, between test item and other items (item-based). Then, the unknown rating is predicted by averaging the (weighted) known ratings of the test item by similar users (user-based), or the (weighted) known ratings of similar items by the test user (item-based). Association rules have been used for many years in merchandising, both to analyze patterns of preference across products, and to recommend products to consumers based on other products they have selected. An association rule expresses the relationship that one product is often purchased along with other products. The number of possible association rules grows exponentially with the number of products in a rule, but constraints on confidence and support, combined with algorithms that build association rules with item sets of n items from rules with n-1 item item sets, reduce the effective search space. Association rules can from a very compact representation of preference data that may improve efficiency of storage as well as performance [2].

Recommendation need to employ efficient prediction algorithms so as to provide accurate recommendation to users. If a prediction is defined as a value that expresses the predicted likehood that a user will "like" an item, then a recommendation is defined as the list of n items with respect to the top-n predictions from the set of items available[1]. When buying clothes, people sometimes do not know which clothes to choose or which clothes can best fit their clothes that they have bought before. In order to help solve this problem and enhance the shopping experience of buying clothes of customers. So fashion design recommendation system will be implemented. It can also improve the customer satisfaction and increase the sale volume by the customized suggested clothes, which are suitable for the customers and they are likely to buy the products. The techniques of recommender system and its related work are described in Section 2. In Section 3, the overview design and system implementation are explained. The performance of this system and conclusion are included in Section 4 and 5.

2. Related Work and Background Theory

2.1 Related Work

The recommender system is often associated with the genesis of computer-based recommendation and collaborative filtering systems. The key attribute of the system is that it allowed recommendations to be generated based on a synthesis of the input from many other users. Machine recommendations based on the opinions of like-minded users rather than filtering items based on content, have become known as collaborative filtering. The main advantage of collaborative filtering is the ability to make serendipitous recommendations.

Collaborative filtering approaches are often classified as memory-based or model-based. In the memory-based approach, all rating examples are stored as-is into memory. In the predictions phase, similar users or items are sorted based on the memorized ratings. Based on the rating of these similar users or items, a recommendation for the test user can be generated. The advantage of the memory-based methods over their mode-based alternatives is that less parameter have to be tuned, however, the data sparsity problem is not handled in a principled manner.

In the model-based approach, training examples are used to generate a model that is able to predict the ratings for items that a test user has not rated before.

2.2 Personalization Systems

Personalization is to provide user with what they want or need without requiring them to ask for it explicitly. Personalization means knowing who the user is what the user wants and can recognize a specific user based on a user profile. Personalization includes how to find and filter the information that fits the user's preferences and needs, how to represent it and how to give the user tools to reconfiguration the systems. The interest in personalization has increased as a way to filter information and reduce information overload.

2.3 Content-Based and Collaborative Filtering

The simplest form of a personalization system is the content-based system. It merely provides products with similar features to the user once he entered a key word. There are no personal ratings of other users which could have an influence on the outcome. [5]

Collaborative filtering approaches are often classified as memory-based on model-based. In the prediction phase, similar user or items are sorted based on the memorize ratings. Based on the ratings of these similar users or items, a recommendation for the test user can be generated. Nowadays, the fact that the tourism as an enterprise is dealing with many activities. If the better services are promoted to users, more efficiency and success will be achieved. It is necessary to use powerful methods for retrieving information and making good correlation between parts of the tour, and scheduling the activities [4].

Recommender systems apply data analysis techniques to the problem of helping users find the items they would like to purchase at Ecommerce Sites by producing a predicted likeliness score or a list of top-N recommended items for a given user.

The goal of a collaborative filtering algorithm is to suggest new items or to predict the utility of a certain item for a particular user based on the user's previous linking and the opinions of other like-minded users. Collaborative Filtering works by collecting user feedback in the form of ratings for items in a given domain and exploit similarities and differences among profiles of several users in determining how to recommend an item or how to give the prediction for the active user's interest. A subset of users is chosen based on their similarity to the active user, and a weighted combination of their ratings is used to produce predictions for the active user.

2.4 Association Rule Mining

Association rule mining is a common technique for performing market basket analysis. The intent is to gain insight into customers buying habits and discover groups of products that are commonly purchased together. Association rules capture relationships among items based on patterns of co-occurrence across transactions. Considering each user profile as a transaction, it is possible to use the Apriori algorithm and generate association rules for groups of commonly liked items.

Association rule mining is a two-step process:

1. Find all frequent itemsets: By definition, each of these itemsets will occur at least as frequently as a pre-determined minimum support count.
2. Generate strong association rules from the frequent itemsets: By definition, these rules must satisfy minimum support and minimum confidence.

Minimum support means that the minimum number of the items that is bought on all transacitions. The set of frequent item set is determinied by consisting of those candidate item set. And then, prune or remove the set of frequent item set is not subset of those candidates. The remaining frequent item set become the pair of items which are greater than or equal with minimum support.

2.5 Similarity Measure

Similarity is a powerful way to retrieve interesting information from large repositories. There are many similarity measure algorithms. They are-

- Pearson Correlation Coefficient
- Uncentered Pearson Correlation Coefficient
- Squared Pearson Correlation Coefficient
- Averaged dot product
- Cosine Correlation Coefficient
- Covariance
- Euclidian distance
- Manhattan distance
- Mutual Information
- Spearman Rank-Order Correlation
- Kendall's Tau

The first 9 are linear correlation coefficient whereas the last 2 are nonparametric or rank correlation coefficients. In this system, Pearson Correlation is used.[3]

The threshold of Correlation Coefficient ranges from -1.00 to +1.00. The value of -1.00 represents a perfect negative correlation while a value of +1.00 represents a perfect positive correlation. A value of 0.00 represents a lack of correlation. A correlation coefficient of +1 means perfect positive correlation. A correlation coefficient close to 0 means no correlation. A correlation coefficient of -1 means perfect negative correlation

The following Pearson Correlation Coefficient function is used to compute the similarity measure between the user's preference functions.

$$S(a,b) = \frac{\sum_{i=1}^{N}(x_{a,i} - \bar{x}_a)(x_{b,i} - \bar{x}_b)}{\sqrt{\sum_{i=1}^{N}(x_{a,i} - \bar{x}_a)^2 * \sum_{i=1}^{N}(x_{b,i} - \bar{x}_b)^2}}$$

Where, S (a, b) is the similarity of user a and b, N is the number of items, $x_{a,i}$ and $x_{b,i}$ are the ratings given to the item i by user a and b. $\bar{x}_a$ and $\bar{x}_b$ are the average ratings (mean) of user a and b.

2.6 Recommender System

Recommender systems are a specific type of information filtering (IF) technique. Information filtering refers to selection of data objects from a continuously changing dynamic stream of information. The goal is to select only those objects relevant to user's interest [7]. Recommendation algorithms are content-based filtering, collaborative filtering, knowledge engineering or rule-based filtering, demographic approach and hybrids approach. Recommended systems attempt to present the user information items (movies, music, books, news, and web pages) and the user is interested in. To do this the user's profile is compared to some reference characteristics. These characteristics may be from the information item (the content-based approach) or the user's social environment the collaborative filtering approaches). Recommender systems typically used techniques from collaborative filtering, in which proximity measures between users who are formulated to generate recommendations, or content-based filtering, in which users are compared directly to user's interests. The proposed system uses similarity measures between users, but also directly measures the user's interests that make them appealing to specific users.

3. System Implementation and Design

3.1. System Design

In this paper, for the creation of user profile, the user creates profile by filling the form and save the user information which provide an easier way to gather accurate information about the users. It is shown in Figure 1.

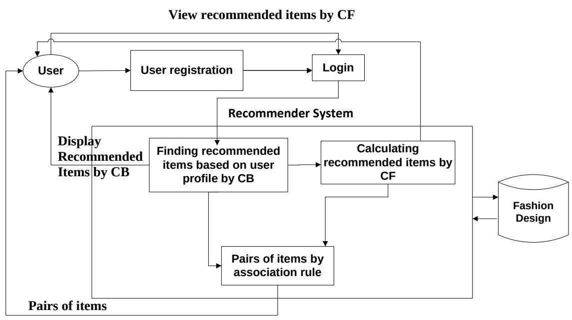

Figure 1. Overview System

First, if user is a new member, user must register first. User must fill detail information and preference data and then to enter login. If you were member to enter directly login. In this phase, check valid or invalid login username and password will be displayed. After checking, history of interested items is displayed from user profile by using content-based filtering. And then, user can request recommend items that you want to browse. After requesting, recommended items can be browsed by calculating similarity items using Pearson Correlation Coefficient collaborative filtering from rating database. And then, user can choose to interest the recommended items. In this phase, it can be updated item rating. Finally, suitable pairs of items are displayed by association rules with Apriori algorithm from selection-pair database.

3.2 System Implementation

There are three main steps for the user in the proposed system. First step**,** user checks member or not member. If member, the user will come system logs in by name and password, and then check valid or invalid. User gives rating, see recommend items, selected item information. If user is not member, go to registration, fill information. After registration, the recommended items can be seen by Content-based filtering and the user can log in recommender system for requesting similar recommend items. In home page, the recommended items can be selected. After selecting, pairs of items can be shown by association rule with frequent itemset. When user enters the next time in this system, selected items will be seen from user profile by using content-based filtering.

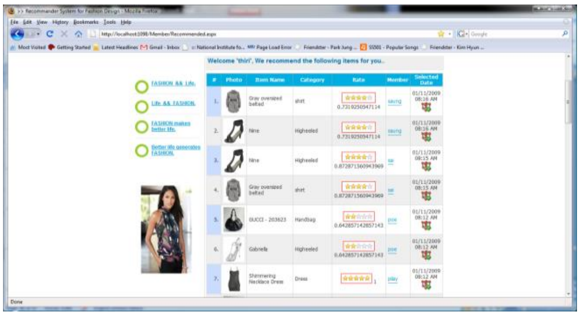

Figure3. View Recommend items

Second step, by clicking view recommended, recommended items are displayed for user by calculating user preferences in Figure 3. Third step, if user select desire item from similar recommended items, the system will recommend the pairs of items which are greater than or equal minimum support and confidence value defined by the administrator with association rule as shown in Figure 4.

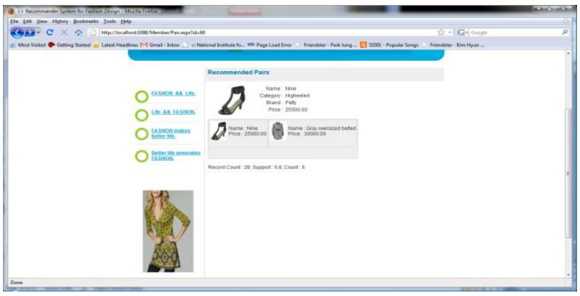

Figure4. Recommended Pairs with Minimum Support

In the backend site, only an analyst log in to this side with the administrator account and password. In this side, there are three menus to analyse the required detail information of the recommender system such as Members, Item Category , and Item. In member page, all members of the system can be viewed. And then, the item category can be seen by clicking "Item Catagory" link. By clicking "New" button, the analyst can add,delete and edit the item category name and the image of its item. The analyst can see the item list by clicking "Item" link. By clicking "New" button, the analyst can add,delete and edit the item category name and the image of its item. For the detail information of every sub-item, the analyst can see by clicking the Detail button.

4. System Performance

In this system, Fashion design database contains 1000 items and 100 users. All calculation rating is done based on user profile. For the computation of similarity, we assume the weight for age (child, between 15 to 25,between 26 to 35 such as 10,7.5), season (hot, cold such as 1,5), color (Black, Blue, White, Pink, Grey, Red such as 6,5,4,3,2,1) and made by for items (Singapore, Thailand, UK, US such as 4,3,2,1) respectively. Firstly, for items 1000, user 100, the similarity is 0.6. Secondly, the similarity is 0.8 for user 200 and 0.95 for user 500.

Table 1. Comparison of the differences of age weight

Child	10	3
Age15-25	7	2
Age26-35	5	1
rating	efficient	Lack of data

Table 2. Comparison of the differences of season weight

For season weight		
Cold	5	2
Hot	1	1
Rating	efficient	Lack of data

In the proposed system, age and season weight are assumed as priority. The range is considered from 0.5 to 1 for calculating similarity. The system will recommend the items which are greater than or equal threshold 0.5. In that case, some of the data will be lacked for assuming child weight (15, 10, and 5) and season weight (9, 6). So, child weight (10, 7, 5) and season weight (5, 1) is efficient and effective for the proposed system. The accuracy is approximately perfect for the similarity.

5. Conclusion

Recommender Systems have emerged as powerful tools for helping users find and evaluate items of interest. The presented scheme uses a content-based predictor to enhance existing user data and then provide personalized suggestions through collaborative filtering. Association rule mining is to discover interesting patterns from the data set. The items are recommended by using Content-based and Collaborative filtering. In addition, suitable pairs of items are shown by association rule. By using these techniques, this system can help manager design different type of store layout items that are frequently purchased together can be placed in close proximity in order to further encourage the sale of such items together. This system provides the users with most effective and efficient information that is relevant to the user's request.

6. References

[1] B. Sarwar, G. Karypis, J. Konstan, J.Redl, "Item-based Collaborative Filtering Recommendation Algorithms", Proceedings of the Tenth International Conference on World Wide Web, Pp 285-295, 2001.

[2] B Schafer, "Application of Data-Mining to Recommender Systems", Encyclopedia of data warehousing and mining, 2002.

[3] JL. Herlocker, JA Konstan, LG Jerveen, "Evaluating Collaborative Filtering Recommender Systems", ACM Transactions on Information System, 2004.

[4] JWang, AP de Vries, MJT Reinders, "Unifying User-based and Item-based Collaborative Filtering Approaches by Similarity Fusion", Proceedings of the 29th annual, 2006.

[5] M. Garden, G. Dudes, "Mixed Collaborative and Content-Based Filtering with User-Contributed Semantic Features", Of The National Conference on Artificial, 2006.

[6] M. Grcar, "User Profiling Collaborative Filtering", Proceedings of the conference on data mining, 2004.

[7] T. Denis, "Recommender system based on user-generated content", 2003.