ABSTRACT
This thesis investigates application of clustering to multi-criteria ratings as a method of improving the precision of top-N recommendations. With the advent of ecommerce sites that allow multi-criteria rating of items, there is an opportunity for recommender systems to use the additional information to gain a better understanding of user preference. This thesis proposes the use of the relevant set correlation model for a clustering-based collaborative filtering system. It is anticipated this novel system will handle large numbers of users and items without sacrificing the relevance of recommended items.

Categories and Subject Descriptors
H.3.3 [Information Storage and Retrieval]: Information Search and retrieval – information filtering. H.3.3 [Information Storage and Retrieval]: Systems and Software – User profiles and alert services.

General Terms

Keywords
Multi-Criteria Recommender System, Relevant Set Correlation, Clustering.

1. INTRODUCTION
The key goal of a recommender system is to find items of interest to a given user. In order to reach this goal, a key task for the system is to accurately understand the user’s preference for items that he/she has experienced. User preference is typically expressed in the form of a rating on an ordinal scale. Collaborative filtering systems utilize this rating information to learn about a user’s interests, and to find a community of like-minded users from which they can infer a user’s preference for an unseen item.

While very successful, collaborative filtering system based on single criterion rating ignore the reason behind the user’s preference for an item. Two users can like (or dislike) an item for different reasons, and these reasons play an important role in determining groups of users that truly share the same preferences.

It has been suggested in [5] that the inclusion of multi-criteria ratings can potentially extend the capabilities of recommender systems. Many sites now allow users to provide ratings on multiple aspects of an item. For example, Yahoo! Movies [1] allows users to provide an overall grade for a movie (on an A+ to F scale) while also grading four aspects of the movies: acting, directing, visuals and story. The popular restaurant guide Zagat.com [2] allows users to rate a restaurant based on four criteria: food, service, décor and cost. Each of these criteria provide more insight into the nature of the user’s preferences. The challenge is to properly utilize this information to generate recommendations.

One way that multi-criteria rating can be used to improve recommendations is by employing them to cluster groups of similar users or items. An advantage of this approach is that the clusters would represent groups of users with not only similar ratings but similar reasons for these ratings. These preferences of the group can then be employed to recommend interesting items to members of the group. Clustering has been applied to collaborative filtering recommender systems [14] to improve the scalability of these systems. However, this improvement often comes at the expense of predictive accuracy. Recent work by [7][17] has pointed to the possibility that clustering can improve scalability without losses in accuracy providing that the clustering algorithm meets certain conditions.

The goal of this research effort is to employ a novel clustering algorithm, the relevant-set correlation (RSC) model to multi-criteria recommender systems. It is anticipated that the application of this algorithm to multi-criteria ratings will address many shortcomings of existing multi-criteria systems. These shortcomings include the dependence on a particular similarity metric to find user neighborhoods, the assumptions that users employ consistent preference structure on all occasions, that the individual criteria are independent, and that the overall rating criterion (if present) is dependent on all criteria instead of a subset of available criteria.

It is hoped that the research proposed in this work will contribute to field of recommender systems by extending existing techniques of multi-criteria recommendation. Furthermore, it should provide insights on the effects dataset characteristics on multi-criteria recommendation.
2. RELATED WORK

Even though there has not been much research has been done in applying clustering to multi-criteria recommender systems, preliminary results indicate that this method is promising. One approach to clustering multi-criteria ratings is based on extensions of the best performing clustering algorithms in the single rating problem [7]. [19] utilizes probabilistic latent semantic analysis to cluster Yahoo! Movie ratings and they discover that their approach performs better in selecting Top-N recommendations than the equivalent single rating approach. In a different vein, [12] applied multi-linear singular value decomposition (MSVD) to cluster the multi-criteria ratings of a restaurant recommender system. The recommendation space, which consists of <User X Item X Criteria>, is represented as 3-order tensor and MSVD is applied to reduce the dimensionality of the space. The nearest neighbors of the target user is then found in the reduced space and used to generate top-N recommendations. This method was found to outperform a similar method [18] that was applied to single criterion ratings.

Other approaches to utilizing multi-criteria ratings in recommender systems do not employ clustering. In [3] traditional nearest neighbor collaborative filtering is extended by utilizing multi-criteria rating information in the similarity calculation. They also examine the use of aggregation functions based on linear regression for top-N recommendation. Both of these approaches outperformed traditional Pearson-correlation based recommender system, with the aggregation functions outperforming the methods based on the extension of similarity calculation. This would indicate that not only are multi-criteria ratings important but there is some consistency in the order of preference of the criteria.

Multi-criteria decision analysis (MCDA) methods have also been employed in this area since they are ideally suited to the nature of the data [13]. These systems typically employ multi-attribute utility theory (MAUT) for recommendation. The user’s overall preference is represented as a linear additive value function of the individual criteria and the weights for each criteria are learnt from past user ratings by employing linear programming [11].

While the use of clustering for multi-criteria recommender systems may be new, many different clustering algorithms have been applied to recommender systems. Most of these algorithms have sacrificed predictive accuracy for scalability. However, some clustering algorithms [17][7] are able to outperform traditional nearest neighbor collaborative filtering because they possess the following properties:

- They are able to cluster users and items inter-dependently. In generating clusters it is important to remember that what needs to be represented is the preference patterns of a subgroup of users on a particular subset of items. This is different from representing the preferences of a subgroup of users on all items. It seems that representing groups of users that like or dislike specific subsets of items is an important part of improving the predictive accuracy of the recommender system.
- Successful algorithms allow users and items to belong to multiple clusters. Since different subsets of preferences account for the membership of a user in a given cluster, it makes sense then to allow users to belong to many different clusters in order to adequately represent the true variety of their preferences. Similarly, items can belong to many different clusters depending on the subset of criteria that cause users to like or dislike them.
- To achieve superior performance these algorithms are able to decouple preference from numerical ratings. An individual’s preference is essentially the reason for liking or disliking an item. As noted in [14] the rating only an indirect representation of this preference and users employ different “personal scoring function” to convert their preferences into ratings. Normalization and other techniques for decoupling preferences from ratings are thus a necessary part of any successful clustering approach to recommender systems.

As demonstrated in [19] the combination of a clustering algorithm that has these properties along with informative multi-criteria rating can greatly improve the ability of a recommender system to suggest relevant items. As is discussed in the next section, the Relevant Set Correlation (RSC) model meets the criteria for a successful clustering approach to recommender systems while avoiding the disadvantages of Flexible Mixture Models (FMM) and Probabilistic Latent Semantic Indexing (PLSA).

3. RELEVANT SET CORRELATION (RSC) CLUSTERING

3.1 RSC Model

The RSC model has been used to cluster many different types of data including categorical, image data and protein sequences. However, it has not yet been applied to the recommender system problem. The RSC model uses a shared-neighbor clustering strategy [9]. In shared-neighbor methods the similarity of two objects is not directly assessed by a metric, but instead is dependent on the number of neighbors they have in common. Therefore, two objects A and B are similar if their respective neighborhoods (relevant sets) contain many common neighbors. Thus, the RSC model is able to form clusters containing objects that are strongly correlated.

Let $S$ be a dataset drawn from some domain $D$. For every item $q \in S$, we assume the existence of a unique ordering $\pi(q) = (q_1,q_2,\ldots,q_\|S\|)$ of the items of $S$, where $i < j$ implies that $q_i \in S$ is deemed more relevant or similar to $q_j$ than $q_q$. The ranking $\pi(q)$ induces a collection of relevant sets $Q(q,k) = \{\pi(q,i) \mid 1 \leq i \leq k\}$ for each choice of set size $1 \leq k \leq \|S\|$. The ranking function is called the oracle. The RSC model uses this ranking information and only this ranking information to determine significant aggregations (clusters) of items. Details of the method are available in [8]. The goal is to partition the dataset into a set of clusters $P = \{P_1, P_2,\ldots, P_n\}$ such that if an item $a \in P_i$, then the relevant set returned by the oracle for a would rank any item of the partition set $P_i$ higher than any item not in $P_i$.

In the single rating case of recommender systems, one oracle $\pi(q)$, is required to produce relevant sets for each object (users or items)
to be clustered. For example, to cluster users based on a single rating, the RSC model employs the oracle (which is a ranking function) to provide a ranked list of similar users. The ranked lists of all users are compared and two users are deemed similar if they share a large number of common neighbors. The final clusters, therefore, consists of highly correlated users whose share common neighbors, and presumably, common preferences.

To accommodate multi-criteria ratings, we assume that $S$ is a set of objects drawn from some domain $D$, and $F$ is a set of oracles for $S$, defined as follows. For every object $q \in D$ and every oracle $f \in F$, we assume the existence of a unique ordering $(s_1, s_2, \ldots, s_k)$ of the objects in $S$, where $i < j$ implies that $s_i$ is deemed by oracle $f$ to be more relevant or similar to $q$ than $s_j$. The relevancy ranking for $q$ with respect to oracle $f$ indicates the collection of sets $Q(q,k) = \{s_{i_1}, \ldots, s_{i_k}\}$ for each choice of set size $1 \leq k \leq |S|$. This means that for each criterion, an oracle, $f$, is employed to produce a ranking of objects similar to the target object.

We propose that for multi-criteria recommender systems the oracle’s ranking is based on the SimRank [10] between vertices of the bipartite graph that represents the relationships between users and items on a single criterion. The edge of each oracle’s bipartite graph is the normalized rating of a particular user for a given item on a single criterion. This is the multi-oracle extension of the relevant-set correlation model as applied to multi-criteria ratings. This multi-oracle extension can be applied to both user-based and item-based recommendations. To generate user-based recommendations, the users are clustered based on their multi-criteria recommendations using the extended RSC model with a separate oracle for each criterion. Once the clusters are generated, recommendations are generated for a target user based on the nearest neighbors in the cluster(s) to which the target user belongs.

This RSC model has many advantages that make it highly applicable to multi-criteria recommender systems. It avoids the problems encountered by using similarity metric which may have undesirable effects in high-dimensional spaces [6]. Furthermore, it has the advantages of allowing clusters to overlap so that users (or items) can belong to more than one cluster. By utilizing bipartite graphs as the basis of oracle ranking the relationships between users and items (connected by ratings) is maintained. This allows items to influence the formation of user groups and vice versa. Also, by using the shared neighbor approach the preferences of users are decoupled from the ratings. The RSC-based approach would also avoid the problem that pLSA-based approaches have, which is the tendency to converge on a local optima in EM-based parameter estimation. Furthermore, the shared neighbor approach is less sensitive to violation of the assumption of independence between criteria, unlike MAUT-based systems.

4. Proposed Plan of Study

The research effort is still in its earliest stages. However, the goal of this effort is two-fold. The first goal, is to evaluate the performance of the RSC-based multi-criteria recommender system with respect to other clustering-based [19] and non-clustering based [3] multi-criteria systems. The goal is to compare the ability of the RSC-based system to recommend highly relevant items. We will, therefore, examine the precision of the ranked list produced by the RSC, pLSA, Extended Similarity and Aggregation Function based multi-criteria recommender systems at different levels of precision (1% to 30% in increments of 5). In addition, an examination of the temporal effects of multi-criteria ratings on system performance will be addressed, by evaluating the precision of systems when ratings in the training set are specifically chosen to be older than those in the test set. This should allow us to detect temporal trends in ratings that may affect system performance.

Our second goal is to investigate the effects of dataset characteristics and preference structure on multi-criteria recommender system performance. This will be achieved by conducting a series of experimental studies on a simulated dataset. The need for simulation is evident since many aspects of the dataset must be manipulated in order to study the effects of these factors.

Prior research in [16] has noted that multi-criteria recommender systems perform well when data is sparse. The goal is to understand why by examining the performance of systems under various conditions of sparsity Also under consideration, is the effects of various distributions of objects among clusters, the effect of the sizes of the clusters, and the manner in which a skewed distribution of ratings affects the ability of a multi-criteria system to generate highly relevant ratings.

The simulation study will also allow the examination of effects of the preference structure on the recommender systems performance. Some techniques for multi-criteria recommender systems [19][3] assume that the overall rating is dependent on all the criteria. However, research into contextual effects on recommendation [4] demonstrates that the context has an effect on how the user evaluates recommender system results. We hope to examine the effect of users employing various subsets of criteria, at different times on the performance of multi-criteria recommender systems.

5. Future Work and Conclusions

While it is too early to make definitive conclusions, the application of the RSC-model to recommender systems seems like a promising approach. The goal is to determine how robust this approach is in the face of sparse rating, changing preference structures, missing information and skewed distributions of item rating and membership in clusters. It is hoped that this work will also provide greater insight into the effects of context on recommender systems.

6. REFERENCES


