Integrating Ratings and Pairwise Preferences in Recommender Systems

Saikishore Kalloori and Francesco Ricci

Free University of Bozen - Bolzano , Italy
{ksaikishore,fricci}@unibz.it

Abstract. User preferences in the form of absolute evaluations such as user ratings or clicks are widely used in many Recommender Systems (RSs). However, such type of preferences have some disadvantages. For instance, users can not further refine the preferences for items that are scored with the same rating (e.g., are both 5 stars or both liked). In our research work, as an alternative way of modeling user preferences and compute recommendations, we have been focusing on pairwise preferences, such as, item i is preferred to item j. We aim at building RSs by combining both ratings and pairwise preferences in order to make the best use of this mixed preference data. Our results demonstrate that it is possible to effectively use pairwise preferences to generate accurate recommendations and that there are specific conditions/situations where pairwise preferences elicitation is more meaningful and useful.

1 Introduction

Most of the current research and application of Recommender Systems is based on the usage of preferences derived from absolute evaluations, such as, user ratings or clicks [10]. However, this type of preferences have few disadvantages. For instance, if a user assigns the highest rating to an item and then successively finds that she prefers another item to the first one, she has no choice but to also give to this new item the highest rating [3]. Moreover, if most of the user rated items are 5 stars, then it is difficult to understand which one the user really prefers among them. In our research work, we have been focusing on pairwise preferences as an alternative way of modeling user preferences and compute recommendations. We are considering scenarios where users compare items in pairs, indicating which one, and to what extent, is preferred [5]. Modeling user preferences in the form of pairwise preferences and identifying which recommendation situation is best suited for using these preference statements has not been explored so far.

A few existing research work [4, 2, 3] have developed recommendation techniques by using pairwise preferences. By comparing two items a numeric pair

IIR 2018, May 28-30, 2018, Rome, Italy.

score is defined; this indicates to what extent the first is preferred to the second (positive score), or if they are equivalent (null score) or if the second item is preferred to the first (negative score). Moreover, a user can also compare features of items, e.g., a user may state that "prefers a hotel with TV and kitchen to a hotel in the center but without a kitchen". In fact, in everyday life, rating items is not a natural mechanism for making decisions. For instance, we do not rate cars when we want to buy one. It is more likely that we will compare them one to one, and purchase the preferred one.

Having the goal of incorporating pairwise preferences along with ratings in RS, we have investigated the following research questions:

– RQ1: Given a collection of ratings and pairwise scores, how to effectively use this information to build RSs?
– RQ2: How to combine pairwise preferences over items and their features to effectively construct RSs and provide accurate recommendations?
– RQ3: In order to collect useful preferences, which items should a system ask a user to evaluate and which type of preferences should be asked (ratings or pair scores)?

2 Pairwise Preferences Based Recommender System

The proposed recommendation techniques combine ratings and pairwise scores to model user preferences and generate recommendations [7]. In order to address RQ1, we have proposed two pair scores prediction methods. Predicting unknown pair scores is equivalent to predicting ratings; it is a preliminary step to ranking items and generating recommendations.

MF for Pair Scores Prediction: We have proposed a matrix factorization pair score prediction and item ranking methods (MFP). It extends Matrix Factorization (MF) for ratings data sets [8] to MF for pair scores predictions. MFP replaces items (ratings) with pairs of items (pair scores) to predict a missing pair scores of a user.

NN for Pair Scores Prediction: We have designed a Nearest-Neighbor (NN) approach for predicting unknown pairwise scores that use two user-to-user similarity metrics that are better suited for user profiles formed by pairwise scores. The first metric is Goodman and Kruskal's gamma (GK), which is defined as $(P-Q)/(P+Q)$, where P is the number of item pairs that are ranked in the same order by both users (concordant pairs) and Q is the number of item pairs that are ranked in the reverse order. The second similarity metric is Expected Discounted Rank Correlation (EDRC). EDRC was proposed by [1] in order to evaluate the rank accuracy of a recommendation list. We have modified and adapted EDRC so that it can be used as a user-to-user similarity metric. EDRC is asymmetric and measures how much a user u set of pairwise preferences match a user v set of pairwise preferences. It is based on a graph representation of the pair scores where the vertices represent items and edges represent pairwise scores.

We have evaluated these techniques using pair scores collected by an item-to-item comparison GUI that was used in a movie recommender system [2].

Our experimental analysis [7] demonstrates that the proposed approaches have better ranking performance compared to other state of the art algorithms [2, 4, 9]. Hence, our results have shown that pairwise preferences can be effectively used to model user preferences and build RSs [7].

We then observed that there are many situations where preferences over features may be a natural way for the user to signal what types of items she likes. For instance, one may like a movie because of its actors. We, therefore, addressed RQ2 by extending our recommendation techniques based on item comparisons and studied how user preferences over features such as "I like Italian movies more than Indian movies" could be encoded as comparisons between items [6].

In order to encode feature preferences into comparisons between items, we have defined a rule to selectively identify the best set of item comparisons that effectively express the given feature preferences. For instance, when a user's favorite feature is *action* then we add to the data set artificial pair scores expressing that the items which have the action feature are preferred to those that are not *action*. However, since such a naive approach would generate a unnecessary large number of comparisons, we have explored alternative solutions and designed a procedure for deriving a reduced number of item comparisons: we consider only the item comparisons that refer to items that were already mentioned in other item comparisons explicitly made by users, hence focusing on a denser set of comparisons [6]. We have combined these item comparisons originated by feature comparisons along with regular item comparisons, i.e., those user preferences expressed on the items, and used them into our recommendation technology to compute a personalized ranking of items.

We have conducted an offline evaluation of our proposed methods and measured the performance of our recommendation techniques when the system uses user profiles containing an increasing amount of items preferences. Our experimental results, on the PoliMovie dataset, demonstrate that there is a benefit in using item comparisons derived from feature preferences especially in severe cold-start situation (when the users have compared a few items), and using both types of item comparisons, our recommender can improve cold start recommendations for new users, compared to a model that exploits only item comparisons [6].

Finally, in order to address RQ3, we tried to identify how and when it is appropriate to elicit pairwise preferences, i.e., when this form of user preference data is more meaningful for the user to express and more beneficial for the system. We have implemented a new technique for pairwise preferences elicitation and recommendation generation in a mobile RS, and applied it to South Tyrol Suggests (STS) app. STS recommends interesting places of interests (POIs) in South Tyrol region in Italy. STS is an Android-based RS that provides users with context-aware recommendations. For our experiments, we have implemented two RSs variants of STS and conducted a live user study. One variant is based on ratings only while the other employs our developed algorithms, gathers user pref-

http://rerex.inf.unibz.it/

erences in the form of pairwise scores and makes recommendations using them together with a possibly pre-existent rating data set.

In an A/B test comparisons of these two variants, we have demonstrated that pairwise preferences can be more effective, if compared with ratings, to model user preferences in situations and scenarios where the user has a clear objective and is looking for a specific type of items (recommendation). For instance, when a user wants to select a restaurant for a dinner with friends, is more likely to choose the best option by comparing the available ones instead of rating them. Our results show that by incorporating pairwise preferences in such scenarios, the system is able to capture user preferences effective ly and to produce better recommendations than state of the art rating-only based solutions.

3 Conclusion

We have shown that it is possible to model user preferences in the form of pairwise preferences and to effectively build RS. Pairwise preferences naturally arise and are expressed by users (directly or indirectly) in many decision making scenarios. We, therefore, believe that pairwise preferences, if they are elicited with an appropriate GUI, can become a valuable preference modeling mechanism for RSs. In our future work, we aim at further exploring active learning strategies that identify a small set of informative pairs of items for eliciting user preferences.

References

1. Ackerman, B., Chen, Y.: Evaluating rank accuracy based on incomplete pairwise preferences. In: Proc. Workshop on UCERSTI Recsys '11 (2011)
2. Blédaitė, L., Ricci, F.: Pairwise preferences elicitation and exploitation for conversational collaborative filtering. In: Proc. Hypertext & Social Media 2015
3. Brun, A., Hamad, A., Buffet, O., Boyer, A.: Towards preference relations in recommender systems. In: Preference Learning Workshop (ECML-PKDD) (2010)
4. Desarkar, M.S., Saxena, R., Sarkar, S.: Preference relation based matrix factorization for recommender systems. In: UMAP '12, pp. 63–75
5. Kalloori, S.: Pairwise preferences and recommender systems. In: Proceedings of the 22nd International Conference on Intelligent User Interfaces Companion. pp. 169–172. ACM (2017)
6. Kalloori, S., Ricci, F.: Improving cold start recommendation by mapping feature-based preferences to item comparisons. In: Proceedings of the 25th Conference on User Modeling, Adaptation and Personalization. pp. 289–293. ACM (2017)
7. Kalloori, S., Ricci, F., Tkalcic, M.: Pairwise preferences based matrix factorization and nearest neighbor recommendation techniques. In: Proc. RecSys 2016. pp. 143–146. ACM (2016)
8. Koren, Y., Bell, R., Volinsky, C.: Matrix factorization techniques for recommender systems. Computer **42**(8), 30–37 (2009)
9. Rendle, S., Freudenthaler, C., Gantner, Z., Schmidt-Thieme, L.: Bpr: Bayesian personalized ranking from implicit feedback. In: Proc. UAI '09. pp. 452–461 (2009)
10. Ricci, F., Rokach, L., Shapira, B.: Recommender systems: Introduction and challenges. In: Recommender Systems Handbook. Springer (2015)