DOI : https://doi.org/10.1007/978-3-030-27618-8_16

Context-Aware Multi-criteria Recommendation Based on Spectral Graph Partitioning

Rim Dridi[1,2(✉)], Lynda Tamine[3], and Yahya Slimani[2,4]

[1] ENSI, University of Manouba, Tunis, Tunisia
rim.dridi@ensi-uma.tn
[2] LISI Laboratory, INSAT, University of Carthage, Tunis, Tunisia
yahya.slimani@fst.rnu.tn
[3] IRIT, University of Toulouse 3, Toulouse, France
tamine@irit.fr
[4] ISAMM, University of Manouba, Tunis, Tunisia

Abstract. Both multi-criteria recommendation and context-aware recommendation are well addressed in previous research but separately in most of existing work. In this paper, we aim to contribute to the under-explored research problem which consists in tailoring the multi-criteria rating predictions to users involved in specific contexts. We investigate the application of simultaneous clustering based on the application of a spectral partitioning graph method over situational contexts in the one hand and criteria in the other hand. Besides, we conjecture that even with similar criteria-related ratings, the importance of criteria might differ among users. This idea leads us to use prioritized aggregation operators as means of multi-criteria rating aggregations. Our experimental results on a real-world dataset show the effectiveness of our approach.

Keywords: Recommender system · Multi-criteria · Context

1 Introduction

The key problem of recommendation is designing the utility function that measures the usefulness of items to target users. Traditionally, recommender systems are based on a single-criterion utility function. Some studies have begun employing multi-criteria recommender systems (MCRS) [1,10,12] that model a user's utility of an item as a vector of ratings along several criteria.

Yet, previous recommenders have highlighted the impact of context dimensions (e.g., time, location, *etc.*) on user's judgments. In this respect, several researches have been devoted to context-aware recommender systems (CARS) [2].

However, most of previous CARS still consider single item ratings while either the item criteria and their strength might evolve while context evolves.

In our work, we attempt to contribute to this under-explored research area. Specifically, we explore the idea of clustering situational recommendations

https://doi.org/10.1007/978-3-030-27618-8_16

embedding users providing similar criteria ratings to target items under similar contexts. Our assumption is that users in similar contextual situations tend to have similar interests for similar criteria. Following this assumption, we consider the joint clustering of two types of entities, where both contextual situations and criteria are simultaneously assigned to clusters. Then, users' predicted criteria ratings from the co-clusters are aggregated based on their personal preferences. We formulate the recommendation problem in terms of two sub-problems: (1) Criteria rating prediction: we transform the first sub-problem to a bipartite graph partitioning problem that we solve using the well known spectral co-clustering [5]. Then, we exploit the obtained co-clusters with a rating prediction algorithm for predicting criteria ratings. (2) Overall rating prediction: the key issue within this second sub-problem is the design of an appropriate aggregation of the criteria ratings resulting from the co-clusters. Therefore, we explore the use of two prioritized aggregation operators [3,4], where the criteria weights are computed on the basis of their priority order in accordance with the users' interests.

2 Related Work

2.1 Multi-criteria Recommender Systems

One of the popular efficient MCRS approaches is the aggregation-based one [12,14] which builds an aggregation function f (Eq. 1) that represents the relationship between the overall rating r_0 and the criteria ratings $(r_1, .., r_N)$:

$$r_0 = f(r_1, .., r_N) \tag{1}$$

In [7,10], a linear aggregation function was applied to predict the overall rating using criteria preferences. In [12], Zheng used criteria chains for multi-criteria rating predictions and conditional aggregations by viewing the criteria predictions as contexts. These criteria ratings are predicted and employed in the chain, which might lead to an accumulated loss while predicting the global rating.

2.2 Context-Aware Recommender Systems

The first category of work in this area, considers context in a single-criterion based recommendation framework. For example, in [2], a context-aware matrix factorization (CAMF) was proposed for item rating prediction.

Unlikely, the second category of work which is closest to ours, explores the exploitation of context information in addition to multi-criteria ratings to provide more accurate predictions [9,14]. Li et al. [9] defined a 4-order tensor recommendation space, where the contextual information and the multi-criteria ratings are considered besides the users and items. This tensor was then reduced by using the relevant context to find the closest neighbors based on the multi-linear singular value decomposition. Recently, Zheng [14] integrated context information into four MCRS baselines. The independent and dependent methods were used for the multi-criteria rating predictions step, and the linear and conditional aggregation methods for the rating aggregations step.

Beside the differences in the used prediction methods, what basically differentiates our proposal is considering that criteria are both item and user-dependent.

3 Context-Aware Multi-criteria Recommendation Framework

3.1 Basic Notation

User's Situational Context. A user's situational context refers to the situation characterized by a user involved in a specific surrounding context. We represent distinct pairs (user, context) as distinct contextual situations. Let users set, noted Us is represented by $Us = \{u_1, .., u_k\}$, where k is the total number of users, and contexts Co are represented by $Co = \{co_1, .., co_l\}$, where l is the total number of contexts. A contextual situation is built up as an entity noted s_{ij}, represented by a contextual situation that implicitly refers to the pair user u_i in context co_j. For care of the simplicity of the notations, s_{ij} is noted as s_i where i is in the range $1..m$ leading the whole set of situations noted as $S = \{s_1, .., s_m\}$.

Criteria. The criteria set contains rated item aspects involving in situational contexts. The set of entities referring to rated item criteria is noted $C = \{c_1, .., c_n\}$, where n is the number of criteria considered for rating an item.

Situational Bipartite Graph. A situational bipartite graph is a triple G = (S, C, E) where S, C are the two vertex sets and E is the set of edges that connect nodes from vertex S to vertex C such as (E = $<s_i, c_j>$ | $s_i \in$ S, $c_j \in$ C).

3.2 Situational Bipartite Graph Co-clustering

We focus on extending the conventional rating prediction process using a co-clustering method to find sub-groups of contextually similar users and criteria that these users are interested in. Our driving hypothesis is the following:

H: "Users in similar contextual situations tend to have similar interests for similar criteria".

To solve the partitioning problem, we employ the popular spectral co-clustering algorithm [5] which approximates the normalized cut of the bipartite graph to find co-clusters. An approximate solution to the optimal normalized cut may be found via the decomposition of the normalized $m \times n$ rating matrix R as follows: $R_n = D_1^{-1/2}$ R $D_2^{-1/2}$, where D_1 is the diagonal matrix with entry i equal to $\sum_j R_{ij}$ and D_2 is the diagonal matrix with entry j equal to $\sum_i R_{ij}$. Then, the singular value decomposition of the resulting matrix $R_n = U \Sigma V^\top$ provides the desired partitions of the rows and columns of R. U is an $m \times m$

matrix, Σ is an $m \times n$ diagonal matrix, and V^T is the transpose of an $n \times n$ matrix. The columns of U and V are called the left and right singular vectors respectively. A subset of the left singular vectors will give the users' situational contexts partitions, and a subset of the right singular vectors will give the criteria partitions. Later, the singular vectors are used to build the matrix Z.

$$Z = D_1^{-1/2} U D_2^{-1/2} V$$

Finally, the resulting matrix Z is decomposed using k-means++ to obtain the desired co-clusters to be used as input to the prediction process detailed below.

3.3 Rating Prediction Algorithm

Criteria Rating Predictions. The Algorithm 1 aims to provide, as an output, the criteria predicted ratings for each co-cluster of situational contexts and criteria. As stated in the algorithm, for each co-$cluster_k$, we can extract a rating sub-matrix $R_k \in \mathbb{R}^{m_k \times n_k}$ from the original rating matrix $R \in \mathbb{R}^{m \times n}$, m_k and n_k denote respectively the number of users' situational contexts and criteria in co-$cluster_k$. Then, we use the Matrix Factorization (MF) [8] as the rating prediction algorithm on each obtained sub-matrix R_k due to its efficiency and scalability. In line 2, the algorithm calls the $MatrixFactorization$ function. This routine applies the MF algorithm where we assume there are F hidden factors, which capture users' situational contexts features and criteria features to model users' preferences. Matrix factorization algorithm works by decomposing the $m_k \times n_k$ rating sub-matrix R_k into the product of two lower dimensionality matrices. Users' situational contexts are represented by a $m_k \times F$ matrix called P, where each row of P would represent the strength of the associations between a user's situational context and the features. In order to relate users' situational context with criteria, the latter are also represented by a matrix called Q, where each row of Q would represent the strength of the associations between a criterion and the features. P and Q are learned using stochastic gradient descent method by minimizing the rating prediction errors. The predicted preference $\hat{r}_{ij}$ of a user's situational context s_i for a criterion c_j can be computed as follows:

$$\hat{r}_{ij} = p_i {q_j}^T \tag{2}$$

Overall Rating Prediction. We make the first attempt to apply "Scoring" and "And" prioritized aggregation operators [3,4] for overall rating prediction. The criteria weights depend on users' preference order of criteria extracted on the basis of their expressed criteria ratings. Besides, regarding the problem of contextual recommendation at hand, we conjecture that the criteria strength also varies in accordance with users' contexts. Hence, the prioritized operators allow flexible personalization of the overall rating prediction by considering the criteria weights based on users' criteria preferences under different contexts.

Algorithm 1: Criteria Rating Prediction for each Co-cluster

Input: Rating matrix with multicriteria: $R \in \mathbb{R}^{m \times n}$, the number of co-clusters: L, and the number of factors: F.

begin

  for *each co-cluster* $k \in \{1, .., L\}$ **do**

    R_k=ExtractSubmatrix (R,co-$cluster_k$)

    P_k,Q_k=MatrixFactorization(R_k,F)

    for *each* $i \in P_k$ **do**

      for *each* $j \in Q_k$ **do**

        for *each* $t \in \{1, .., F\}$ **do**

          $\hat{r}_{ij} = p_{i,t} \times q_{j,t}$

  Output: Criteria predicted ratings

The importance weight computation of a criterion c_i, with i $\neq$ 1, depends on users' preference order of criteria, and depends also on both the weight associated to criterion c_{i-1}, and the preference of c_{i-1}. The user preference ordering of the considered criteria is based on computing an average score for each criterion in accordance with the users expressed criteria ratings. More formally, let $C = \{c_1, ..., c_N\}$ be a set of ordered criteria, where c_1 presents the most preferred criterion and c_N is the least one. We indicate by w_p the importance weight of the criterion $c_p \in C$ for a given item and user's context. The weights associated with the ordered criteria are computed as follows:

– The weight associated with the most important criterion c_1 is set to be 1.
– The weights of the other criteria c_p for $p \in$ [2, N], are computed as follows:

$$w_p = w_{p-1}.r_{p-1} \tag{3}$$

r_{p-1} denotes the preference rating given by a user on criterion c_{p-1} of an item. We define in the following a new way in which the function f (See Eq. 1) is defined according to the mentioned prioritized aggregation operators.

– **Prioritized "Scoring" operator (F_s):** This operator calculates the overall item rating r_0 from several criteria evaluations, where the weight associated with each criterion depends both on the weights and on the preferences of the most important criteria. The higher the satisfaction degree of a more important criterion, the more the satisfaction degree of a less important criterion impacts the overall rating. F_s is defined as: $F_s : [0, 1]^N \longrightarrow$ [0, N]

$$r_0 = F_s(r_1, .., r_N) = \sum_{p=1}^{N} w_p.r_p \tag{4}$$

For example, let us consider that a user is looking for an hotel. His choice depends on two criteria c_1 = "comfort" and c_2 = "inexpensiveness" with $c_1 > c_2$. An hotel with a "comfort" degree of 1 and an "inexpensiveness" degree of 0 would have an overall rating of 1.

– **Prioritized "And" operator (F_a):** This operator models a situation where the overall rating r_0 strongly depends on the importance of the least satisfied criterion. If it is the most important criterion, the value of the least satisfied criterion is considered as the overall rating merely. F_a is defined as follows: $F_a : [0,1]^N \longrightarrow [0, 1]$

$$r_0 = F_a(r_1, .., r_N) = \min_{p \in [1,N]} (\{r_p\}^{w_p}) \tag{5}$$

Let us come back again to the previous example. c_1 = "comfort" and c_2 = "inexpensiveness" with $c_1 > c_2$. Here, an hotel with a "comfort" degree of 1 and a "inexpensiveness" degree of 0 would have an overall rating of 0. So in this case, the under-satisfaction of the inexpensiveness criterion cannot be compensated by the satisfaction of the "comfort" criterion.

4 Experimental Evaluation

4.1 Experimental Settings

The only suitable dataset with respect to our evaluation purpose is TripAdvisor data [6] since: (1) user's context is available based on a contextual dimension which refers to the *season*. This contextual dimension is derived from the trip date expressed in months in the dataset (e.g., March, April and May are the spring season months). (2) Users' ratings of seven individual criteria, plus one overall rating are provided. The used criteria are: *value for the money*, *quality of rooms*, *the hotel location*, *cleanliness of the hotel*, *experience of check-in*, *overall quality of service and business services*. There are a total of 22.130 ratings given by 1502 users on 14.300 hotels. The bipartite graph modeling is built upon $m = 3916$ users situational contexts connected to $n = 7$ criteria.

We measure the performance by mean absolute error (MAE) on this dataset by adopting a training-testing methodology for both parameter tuning and evaluation. For this purpose, we fixed a splitting ratio of training/test of 80/20. For comparison, we used a single rating approach (BiasMF [8]), multi-criteria rating approaches (Agg [1], CluAllCrit [10], CIC [12], CCA [12], CCC [12]) and a context-aware rating approach (CAMF [2]).

4.2 Research Hypothesis Validation

To validate our research hypothesis **H** (See Sect. 3.2), we perform a statistical analysis to determine the strength of the relationships between contextually similar users according to their criteria importance. More precisely, we run a correlation analysis on all the users providing criteria preferences of similar items in similar context situations from the real-world TripAdvisor dataset. First, we compute the importance of each criterion for each user to identify users preferred criteria according to their contexts [11]. Having computed the users criteria importance, we examine the strength of the relationship between these users with respect to their criteria importance through the computation of the Spearman's rank correlation coefficient. To interpret the strength of the obtained correlation coefficient values, we use the rule of thumb (See Fig. 1). We can clearly see from Fig. 1, the high percentage of the very strongly correlated users in similar situations. This result shows that the majority of contextually similar users achieve a fairly strong positive correlation coefficient with respect to their interests for similar criteria which represents a good agreement between contextually similar users on criteria importance order. Hence, we could conjecture that the more similar the users contexts, the more these users tend to have similar criteria importance which provides a strong support for our research hypothesis **H**.

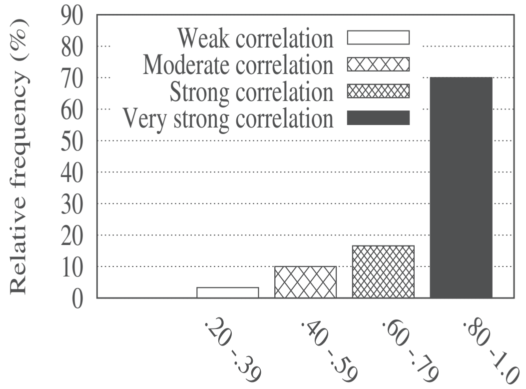
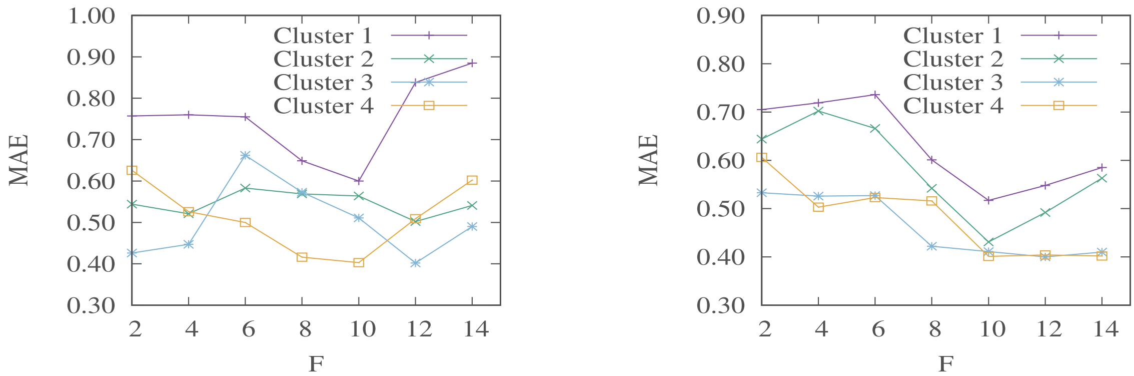


Fig. 1. Distribution of the correlation measures between users' criteria importance in similar contexts

4.3 Evaluation of the Prioritized Aggregation Operators

We begin by tuning the latent factor number F which is one of the important parameters for matrix factorization. As shown in Fig. 2(a), we can observe, when F is equal to 12 the MAE of our proposal using "And" operator declines to the lowest in cluster 2 and cluster 3. So, we come to a conclusion that F = 12 is a better choice for both cluster 2 and cluster 3. For cluster 1 and 4, the MAE of "And" operator model shows a good prediction accuracy when F = 10. While the prediction accuracy of "Scoring" operator model in all clusters improves as the number of latent factors reaches 10 (Fig. 2(b)). Then, we assess in this experimental scenario, the effectiveness of the "Scoring" and "And" prioritized operators for improving the overall rating prediction in comparison with the standard "Average" operator. Particularly, to evaluate the joint effect of the aggregation operators and the number of co-clusters on rating prediction accuracy, we experiment different numbers of co-clusters ranging from 2 to 10. From Fig. 3, we can observe that the "Scoring" operator (resp. the "And" operator) achieves an average improvement of 19.9% (resp. 14.6%) over the "Average" aggregation operator for a number of co-clusters ranging from 5–8. This result confirms the effectiveness of the prioritized combination of the considered criteria in the co-clusters, which allows flexible personalization of the overall prediction results according to users' preferences. The "Scoring" operator is the best performing operator in these comparisons due to the appropriateness of the importance order of relevant criteria in accordance with users' contexts. Fig. 3 also reveals that the prediction accuracy is affected by the number of co-clusters. We can observe that the accuracy slightly increases as the co-clusters number increases from 2 to 4 since the information within each co-cluster is more tied to users. However, when the co-clusters number continues to increase, the prediction accuracy tends to be steady. This observation could be explained by the fact that increasing the number of co-clusters would lead to divide the rating matrix into several more small sub-matrices. Yet, the criteria rating prediction using the MF algorithm requires a sufficient volume data to provide accurate predictions. Thus, under a reasonable threshold of data provided by the co-clusters, the criteria aggregation process can not achieve good results, which have a downside effect on the prediction quality. Therefore, we fix the number of co-clusters to 4 for the prioritized operators and 3 for the "Average" operator.



(a) "And" operator model (b) "Scoring" operator model

Fig. 2. F variation on the prioritized operators

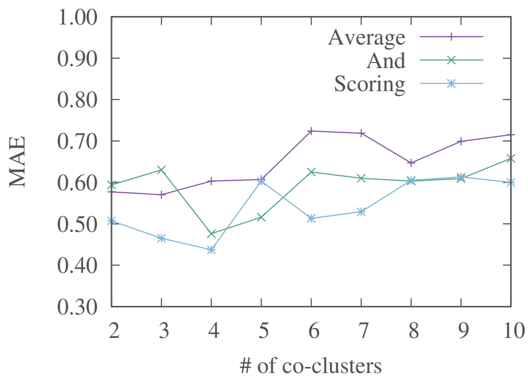


Fig. 3. Effectiveness of prioritized operators

4.4 Comparison Effectiveness Evaluation with Baselines

The multi-criteria baselines results are reported from the published corresponding research papers referenced in Table 1 using their optimal parameters and using the same dataset we used in our experiments. While the results of the other categories of baselines are obtained from the toolkit CARSKit [13].

In Table 1, *IR Scoring* and *IR And* indicate the improving rate achieved using the "Scoring" and the "And" operators respectively. According to Table 1, our proposed approach is able to outperform the baselines by achieving higher prediction accuracy. More precisely, our model based "Scoring" operator allows achieving a considerable improvement of +72.1%, +72.9% and +62.4% over Agg, CIC and CCA models respectively. The same trend of improvement holds for the model based on the "And" operator. These results could be explained by the fact that the multi-criteria Agg, CIC and CCA models use either a traditional way for predicting multi-criteria ratings, a linear aggregation, or both which may decrease prediction accuracy. The multi-criteria algorithm based on clustering (CluAllCrit) which uses a linear aggregation degrades the prediction results compared with other multi-criteria algorithms. Therefore, our model allows a huge improvement over CluAllCrit (+482.4% by the "Scoring" operator and +434.7% by the "And" operator), this may be because the problem with the automatic criteria coefficients obtained by the linear aggregation function. Even when employing a clustering technique to enhance prediction results, using such coefficients in the aggregation process may generate many rating prediction results with negative values or outside of the [1..5] scale. Comparing with the CCC model, which considers criteria dependency to predict the criteria ratings and uses conditional aggregations, there is a little difference in the accuracy results between this latter model and ours. These results reveal that there might exist complementary criteria affecting the user's choice for choosing an item. Meanwhile, using a conditional aggregation may not always be a good choice, since CIC model which uses a conditional aggregation performs worse than CCA model which uses a linear function.

For the contextual baseline, CAMF works better than the majority of baselines but still outperformed by our model (+46.2% using the "Scoring" operator and +34.2% using the "And" operator); this may be because it does not take extra information such as multi-criteria ratings.

Overall, our results indicate that particularly in situations where different criteria ratings are available, it can be advantageous to consider the criteria strength with respect to user's context. This explanation is corroborated by cross-comparing the results obtained using the prioritized operators in the one hand versus the average aggregation and the CAMF approach on the other hand. We can see that the MAE decreased from 0.639 to 0.570 when leveraging context and decreased more to less 0.480 when additionally applying the prioritized operators.

Table 1. Comparison results for the rating prediction task

Category	Algorithms	MAE	IR Scoring	IR And
Traditional single	BiasMF [8]	0.894	+104.5%	+87.8%
Multi-criteria rating approaches	Agg [1]	0.752	+72.1%	+57.9%
	CIC [12]	0.756	+72.9%	+58.8%
	CCA [12]	0.710	+62.4%	+49.2%
	CCC [12]	0.460	+5.3%	−3.5%
	CluAllCrit [10]	2.545	+482.4%	+434.7%
Context-aware rating approach	CAMF [2]	0.639	+46.2%	+34.2%
Our model	Average	**0.570**	-	-
	Scoring	**0.437**	-	-
	And	**0.476**	-	-

5 Conclusion

In this paper, we have proposed a context-aware recommendation approach that relies on multi-criteria rating predictions. The key characteristics of the proposed approach consist in jointly clustering users involved in contextual situations while rating items with respect to multiple facets. For this purpose, we used the spectral graph partitioning method. The obtained co-clusters provide partial user's item ratings that are aggregated using prioritized aggregation operators which allow tailoring the criteria strengths to the user's preferences.

The experiments shows that: (1) the prioritized operators outperform basic average aggregation but that improvement is achieved only with a limited number of co-clusters and that (2) the co-clusters of contextual situations and criteria provide relevant signals about the users' perceptions about item aspects.

In the future, we plan to evaluate our recommendation framework on other datasets allowing a multi-dimensional-based context evaluation. Within this line of work, we will support our model with an in-depth analysis of the users' ratings on item aspects in various contexts and study the correlation between them. This analysis would give insight into the relevance of extending the bipartite graph to deal with different context nodes and the usefulness of filtering relevant interactions between contexts and item criteria before applying the aggregation.

References

1. Adomavicius, G., Kwon, Y.: New recommendation techniques for multicriteria rating systems. IEEE Intell. Syst. **22**(3), 48–55 (2007)
2. Baltrunas, L., Ludwig, B., Ricci, F.: Matrix factorization techniques for context aware recommendation. In: RecSys 2011, New York, USA, pp. 301–304 (2011)
3. da Costa Pereira, C., Dragoni, M., Pasi, G.: Multidimensional relevance: a new aggregation criterion. In: Boughanem, M., Berrut, C., Mothe, J., Soule-Dupuy, C. (eds.) ECIR 2009. LNCS, vol. 5478, pp. 264–275. Springer, Heidelberg (2009). https://doi.org/10.1007/978-3-642-00958-7_25

4. da Costa Pereira, C., Dragoni, M., Pasi, G.: A prioritized “and” aggregation operator for multidimensional relevance assessment. In: Serra, R., Cucchiara, R. (eds.) AI*IA 2009. LNCS (LNAI), vol. 5883, pp. 72–81. Springer, Heidelberg (2009). https://doi.org/10.1007/978-3-642-10291-2_8
5. Dhillon, I.S.: Co-clustering documents and words using bipartite spectral graph partitioning. In: Proceedings of the 7th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp. 269–274. KDD, New York (2001)
6. Jannach, D., Zanker, M., Fuchs, M.: Leveraging multi-criteria customer feedback for satisfaction analysis and improved recommendations. J. IT Tourism **14**(2), 119–149 (2014)
7. Jhalani, T., Kant, V., Dwivedi, P.: A linear regression approach to multi-criteria recommender system. In: Tan, Y., Shi, Y. (eds.) Data Mining and Big Data, pp. 235–243. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-40973-3_23
8. Koren, Y., Bell, R., Volinsky, C.: Matrix factorization techniques for recommender systems. Computer **42**(8), 30–37 (2009)
9. Li, Q., Wang, C., Geng, G.: Improving personalized services in mobile commerce by a novel multicriteria rating approach. In: WWW, NY, USA, pp. 1235–1236 (2008)
10. Liu, L., Mehandjiev, N., Xu, D.L.: Multi-criteria service recommendation based on user criteria preferences. In: RecSys 2011, pp. 77–84. ACM, New York (2011)
11. Sreepada, R.S., Patra, B.K., Hernando, A.: Multi-criteria recommendations through preference learning. In: CODS, pp. 1:1–1:11. ACM, New York (2017)
12. Zheng, Y.: Criteria chains: A novel multi-criteria recommendation approach. In: IUI 2017, New York, NY, USA, pp. 29–33 (2017)
13. Zheng, Y., Mobasher, B., Burke, R.D.: CARSKIT: A java-based context-aware recommendation engine. In: ICDMW, Atlantic City, USA, 14–17, pp. 1668–1671 (2015)
14. Zheng, Y., Shekhar, S., Anna Jose, A., Kumar, S.: Integrating context-awareness and multi-criteria decision making in educational learning. In: SAC, ACM (2019)