Multi-criteria based Item Recommendation Methods

Noor Ifada[1*], Syafrurrizal Naridho[2] and Mochammad Kautsar Sophan[3]
[1,2,3]Informatics Department, University of Trunojoyo Madura
[1] noor.ifada@trunojoyo.ac.id [2]rizal14115@gmail.com [3]kautsar@trunojoyo.ac.id

ABSTRACT

This paper comprehensively investigates and compares the performance of various multi-criteria based item recommendation methods. The development of the methods consists of three main phases: predicting rating per criterion; aggregating rating prediction of all criteria; and generating the top-N item recommendations. The multi-criteria based item recommendation methods are varied and labelled based on what approach is implemented to predict the rating per criterion, i.e., Collaborative Filtering (CF), Content-based (CB), and Hybrid. For the experiments, we generate two variations of datasets to represent the normal and cold-start conditions on the multi-criteria item recommendation system. The empirical analysis suggests that Hybrid and CF are best implemented on the normal and cold-start item conditions, respectively. On the other hand, CB should never be (solely) implemented in a multi-criteria based item recommendation system on any conditions.

Keyword: Collaborative Filtering, Content-based, Hybrid, Multi-criteria recommendation method.

INTRODUCTION

Recommendation System is a firmed research topic (G. Adomavicius & Tuzhilin, 2005; Zhang, Zhou, & Zhang, 2011). A multi-criteria recommendation system allows users to specify ratings based on various criteria (Gediminas Adomavicius, Manouselis, & Kwon, 2011; Aggarwal, 2016). For example, a user may rate tourism attraction based on attraction, accessibility, amenities, and ancillary. Such a system implements a recommendation method in which the user's preference for an item is represented as a vector of ratings corresponding to various criteria (Aggarwal, 2016).

The rating prediction algorithms can be categorized into three approaches: Collaborative Filtering (CF), Content-based (CB), and Hybrid (G. Adomavicius & Tuzhilin, 2005). The CF approach is the most popular rating prediction approach which generates the predictions based on the similarity of users or items (Su & Khoshgoftaar, 2009). The CB approach is popular for generating prediction when the information of the item's content is available (G. Adomavicius & Tuzhilin, 2005; Lops, Gemmis, & Semeraro, 2011). Meanwhile, the Hybrid approach is comprising both the Collaborative Filtering and Content-based approaches (G. Adomavicius & Tuzhilin, 2005; Burke, 2007). Researchers have proposed methods various methods based on the three rating prediction approaches (Gediminas Adomavicius et al., 2011; Fuchs & Zanker, 2012; Jannach, Karakaya, & Gedikli, 2012; Lakiotaki, Matsatsinis, & Tsoukias, 2011; Manouselis & Costopoulou, 2007). However, to the best of our knowledge, no in-depth work has been done that comprehensively compares the performance of the multi-criteria recommendation method variations.

In this paper, we study a variation of multi-criteria based item recommendation methods. The development of the methods consists of three main phases: predicting rating per criterion; aggregating rating prediction of all criteria; and generating the top-N item recommendations. The multi-criteria based item recommendation methods are varied and labelled based on what approach is implemented to predict the rating per criterion, i.e., Collaborative Filtering (CF), Content-based (CB), and Hybrid. For the experiments, we generate two variations of datasets to represent the normal and cold-start conditions on the multi-criteria item recommendation system.

The contributions of this paper are as follows: (1) Analysis of multi-criteria based recommendation methods and show the different approaches to implementing the rating prediction per criterion phase, and (2) Performance comparison of three multi-criteria based item recommendation methods.

METHOD

In this section, we explain how to develop the multi-criteria based methods for generating a top-N list of item recommendation to a target user. The input of the method is the multi-criteria rating histories of users; the item features data; and the significance score of each criterion. At this stage, we can identify and notate the set of users, items, criteria, significance score of each criterion, features, items which have been rated by each user, and features of each item. Table I lists the notations used in this paper to symbolize those sets.

The development of the multi-criteria based item recommendation method consists of three main phases: (a) predicting rating per criterion, (b) aggregating rating prediction of all criteria, and (c) generating the top-N item recommendations. Figure 1 shows the framework of the method.

Prediction Rating per Criterion

The calculation of rating prediction can be categorized into three approaches: Collaborative Filtering, Content-based, and Hybrid approach (G. Adomavicius & Tuzhilin, 2005).

<u>Collaborative Filtering Multi-criteria based Approach</u>

The Collaborative Filtering (CF) approach is the most popular rating prediction approach (Su & Khoshgoftaar, 2009). A way to implement this

Table I. Notations

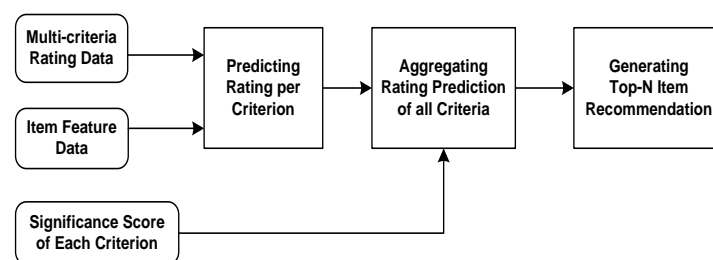| Notation	Description
$U = \{u_1, u_2, u_3, \ldots, u_p\}$	set of p users
$I = \{i_1, i_2, i_3, \ldots, i_q\}$	set of q items
$C = \{c_1, c_2, c_3, \ldots, c_l\}$	set of l criteria
$Y = \{y_1, y_2, y_3, \ldots, s_l\}$	set of the significance scores of l criteria
$\Delta = \{\delta_1, \delta_2, \delta_3, \ldots, \delta_f\}$	set of f features.
$r_{u,i}^c$	rating given by user u to item i, for criterion c
$\delta_{i,f}$	Designation of item's feature, i.e., labelled as either 1 or 0
I_u	set of items which have been rated by user u
F_i	set of features of item i

Figure 1. The Framework of Multi-criteria based Item Recommendation Method

approach is by calculating the prediction based on the similarities of users rating preferences (Gediminas Adomavicius et al., 2011; Bilge & Yargıç, 2017). Per criterion c, we calculate the similarity between user u and v using the cosine similarity function:

$$s_{u,v}^c = \frac{\sum_{i \in I_u \cap I_v} r_{u,i}^c \cdot r_{v,i}^c}{\sqrt{\sum_{i \in I_u \cap I_v} \left(r_{u,i}^c\right)^2} \cdot \sqrt{\sum_{i \in I_u \cap I_v} \left(r_{v,i}^c\right)^2}} \tag{1}$$

Once the similarities between target user u and other users are calculated, we can form the top-k nearest neighbours of user u, K_u. The CF rating prediction of user u to item i for criterion c is formulated as:

$$\alpha_{u,i}^c = \frac{\sum_{v \in K_u} s_{u,v}^c \cdot r_{v,i}^c}{\sum_{v \in K_u} |s_{u,v}^c|} \tag{2}$$

where $|K_u| \leq k$.

Content-based Multi-criteria based Approach

The Content-based (CB) approach is popular for generating prediction when the information of the item's content is available (G. Adomavicius & Tuzhilin, 2005; Lops et al., 2011). A way to implement this approach is by calculating the prediction based on the weight of each item feature (Uluyagmur, Cataltepe, & Tayfur, 2012). Per criterion, we calculate the weight of item feature f for user u using the weighting function:

$$w_{u,f}^c = \frac{\sum_{i \in I_u} \delta_{i,f} \times r_{u,i}^c}{|I_u|} \tag{3}$$

The CB rating prediction of user u to item i for criterion c is formulated as:

$$\beta_{u,i}^c = \frac{\sum_{f \in F_i} w_{u,f}^c}{|F_i|} \tag{4}$$

Hybrid Multi-criteria based Approach

The Hybrid approach is comprising both the Collaborative Filtering and Content-based approaches (G. Adomavicius & Tuzhilin, 2005; Burke, 2007). A way to implement this approach is by calculating the prediction based on the average of CF and CB rating predictions (Aggarwal, 2016). The Hybrid rating prediction of user u to item i for criterion c is formulated as:

$$\gamma_{u,i}^c = \frac{\alpha_{u,i}^c + \beta_{u,i}^c}{2} \tag{5}$$

Aggregating Rating Prediction of All Criteria

Given the rating prediction of user u to item i for each criterion c, we now need to aggregate the prediction of all criteria. The aggregated prediction of all criteria is formulated as (Aggarwal, 2016):

$$\theta_{u,i} = \frac{\sum_{c \in C} y_c \times \hat{r}_{u,i}^c}{\sum_{c \in C} y_c} \tag{6}$$

		Item											
		i_1			i_2			i_3			i_4		
		c_1	c_2	c_3	c_1	c_2	c_3	c_1	c_2	c_3	c_1	c_2	c_3
User	u_1	3	3	4	2	4	4	4	1	3	4	3	4
	u_2	?	?	?	3	3	1	2	3	4	?	?	?
	u_3	5	4	4	2	3	3	3	1	3	3	3	4
	u_4	4	4	3	2	2	3	2	3	4	1	4	3

Figure 2. A toy example of multi-criteria rating data

		Feature				
		δ_1	δ_2	δ_3	δ_4	δ_5
Item	i_1	1	1	1	0	1
	i_2	1	1	1	1	0
	i_3	1	0	0	0	1
	i_4	1	0	0	1	1

Figure 3. A toy example of item feature data

	Criteria		
	c_1	c_1	c_3
Score (y)	2	3	1

Figure 4. A toy example of the criteria significance scores

where

$$\hat{r}_{u,i}^c := \begin{cases} \alpha_{u,i}^c & \text{case CF approach} \\ \beta_{u,i}^c & \text{case CB approach} \\ \gamma_{u,i}^c & \text{case Hybrid approach} \end{cases} \quad (7)$$

Generating Top-N Item Recommendation

The list of top-N list item recommendation for target user u, $Top_u(N)$, is generated based on the aggregated rating predictions. In this case, item i is listed in the $TopN$ if $\theta_{u,i}$ is within N highest values in $\theta_{u,*}$

Toy Example

This section shows examples of how the multi-criteria based methods for generating a top-N list of item recommendation to a target user are developed based on the CF, CB, and Hybrid approaches. Figure 2, Figure 3, and Figure 4 respectively show the toy examples of multi-criteria rating data, item features data and criteria significance scores. The complete calculation of the three main phases in the multi-criteria based item recommendation method is presented in Figure 5.

Figure 2 shows that there are four users $U = \{u_1, u_2, u_3, u_4\}$, four items $I = \{i_1, i_2, i_3, i_4\}$, and three criteria $C = \{c_1, c_2, c_3\}$. The set of users' items are formed as $I_1 = \{1,2,3,4\}$, $I_2 = \{2,3\}$, $I_3 = \{1,2,3,4\}$, and $I_4 = \{1,2,3,4\}$. On the other hand, Figure 3 shows that there are five item features data $\Delta = \{\delta_1, \delta_2, \delta_3, \delta_4, \delta_5\}$, in which the set of items' features are formed as $F_1 = \{1,2,3,5\}$, $F_1 = \{1,3,4,5\}$, $F_1 = \{1,5\}$, and $F_1 = \{1,4,5\}$. Meanwhile, Figure 4 shows that the significance score of each criterion is $y_1 = 2$, $y_2 = 3$, and $y_3 = 1$.

For our target user u_2, we are generating the top-N item recommendation $TopN_2$ based on the aggregated rating prediction $\theta_{2,1}$ and $\theta_{2,4}$. The rating prediction per criterion based on CF, CB, and Hybrid approaches are respectively calculated using Equation (2), (4), and (5). Whereas the aggregated rating prediction is calculated using Equation (6). Figure 5 shows that the $Top_2(N)$ generated based on CF, CB, and Hybrid based methods are respectively {1,4}, {4,1}, and {1,4}.

RESULTS AND DISCUSSION

In this section, we present the empirical analysis by conducting experiments to evaluate the performance of the multi-criteria based item recommendation methods built in this paper.

Dataset and Experiment Procedure

This paper uses a tourism multi-criteria rating dataset that consists of 77 users, 50 touristic attractions or items, and 900 rating data. Table II lists the details of the dataset. Our experimentations implement the 5-fold cross-validation evaluation approach, in which each fold has a training set D_{train} and a test set D_{test}.

For the experiments, we generate two variations based on the tourism dataset to represent the normal and cold-start conditions:

- **TN**: The dataset is refined such all items and users in D_{test} occurred several times in D_{train}. This dataset represents the normal condition.

- **TCS**: The dataset is refined such that there occurred items in D_{test} that have not been rated by users in D_{train}. This dataset represents the condition in which the cold-start item problem occurs.

Evaluation Criteria

Our recommendation methods build the model using D_{train} and utilize it afterwards to generate a top-N item recommendations for target users in D_{test}. In this case, the evaluation is conducted by comparing the top-N list of item recommendations for a user u, $Top_u(N)$, to the ground-truth items listed in D_{test}, GT_u. We use the AP (Average Precision) evaluation metric to measure the performance of recommendations. The AP score of the first N list of item recommendations for a target user u is formulated as:

$$AP_u(N) := \sum_{n=1}^{N} Precision_u(N) \cdot \mathbb{I}(Top_u(n) \in GT_u) \quad (8)$$

where

$$Precision_u(N) := 100 \cdot \frac{|Top_u(N) \cap GT_u|}{N} \quad (9)$$

Predicting Rating per Criteria where $I_1 = \{1,2,3,4\}$, $I_2 = \{2,3\}$, $I_3 = \{1,2,3,4\}$, $I_4 = \{1,2,3,4\}$, $F_1 = \{1,2,3,5\}$, and $F_4 = \{1,4,5\}$				Aggregating Rating Prediction of all Criteria where $S = \{2,3,1\}$		Top-N
Collaborative Filtering (CF) Approach where $k = 2$						
Criteria / 1	$s^1_{2,1} = \frac{(3\times2)+(2\times4)}{\sqrt{3^2+2^2}\times\sqrt{2^2+4^2}} = 0.87$ $s^1_{2,3} = \frac{(3\times2)+(2\times3)}{\sqrt{3^2+2^2}\times\sqrt{2^2+3^2}} = 0.92$ $s^1_{2,4} = \frac{(3\times2)+(2\times2)}{\sqrt{3^2+2^2}\times\sqrt{2^2+2^2}} = 0.98$	$\alpha^1_{2,1} = \frac{(0.92\times5)+(0.98\times4)}{\lvert0.92\rvert+\lvert0.98\rvert} = 4.48$	$\alpha^1_{2,4} = \frac{(0.92\times3)+(0.98\times1)}{\lvert0.92\rvert+\lvert0.98\rvert} = 1.97$	$\theta_{2,1} = \frac{(4.48\times2)+(4.00\times3)+(3.48\times1)}{2+3+1}$ $= 4.08$	$\theta_{2,4} = \frac{(1.97\times2)+(3.52\times3)+(3.48\times1)}{2+3+1}$ $= 3.00$	$Top_2(N) = \{1,4\}$
2	$s^2_{2,1} = 0.86$ $s^2_{2,3} = 0.89$ $s^2_{2,4} = 0.98$	$\alpha^2_{2,1} = 4.00$	$\alpha^2_{2,4} = 3.52$			
3	$s^3_{2,1} = 0.78$ $s^3_{2,3} = 0.86$ $s^3_{2,4} = 0.92$	$\alpha^3_{2,1} = 3.48$	$\alpha^3_{2,4} = 3.48$			
Content-based (CB) Approach						
Criteria / 1	$w^1_{2,1} = \frac{(1\times3)+(1\times2)}{2} = 2.50$ $w^1_{2,2} = \frac{(1\times3)+(0\times2)}{2} = 1.50$ $w^1_{2,3} = \frac{(1\times3)+(0\times2)}{2} = 1.50$ $w^1_{2,4} = \frac{(1\times3)+(0\times2)}{2} = 1.50$ $s^1_{2,5} = \frac{(0\times3)+(1\times2)}{2} = 1.00$	$\beta^1_{2,1} = \frac{2.5+1.5+1.5+1.0}{4} = 1.63$	$\beta^1_{2,4} = \frac{2.5+1.5+1.0}{3} = 1.67$	$\theta_{2,1} = \frac{(1.63\times2)+(1.88\times3)+(1.38\times1)}{2+3+1}$ $= 1.71$	$\theta_{2,4} = \frac{(1.67\times2)+(2.00\times3)+(1.67\times1)}{2+3+1}$ $= 1.83$	$Top_2(N) = \{4,1\}$
2	$w^2_{2,1} = 3.00$ $w^2_{2,2} = 1.50$ $w^2_{2,3} = 1.50$ $w^2_{2,4} = 1.50$ $w^2_{2,5} = 1.50$	$\beta^2_{2,1} = \frac{3.0+1.5+1.5+1.5}{4} = 1.88$	$\beta^2_{2,4} = \frac{3.0+1.5+1.5}{3} = 2.00$			
3	$w^3_{2,1} = 2.50$ $w^3_{2,2} = 0.50$ $w^3_{2,3} = 0.50$ $w^3_{2,4} = 0.50$ $w^3_{2,5} = 2.00$	$\beta^3_{2,1} = \frac{2.5+0.5+0.5+2.0}{4} = 1.38$	$\beta^3_{2,4} = \frac{2.5+0.5+2.0}{3} = 1.67$			
Hybrid Approach						
Criteria / 1	CF + CB	$\gamma^1_{2,1} = \frac{4.48+1.63}{2} = 3.05$	$\gamma^1_{2,4} = \frac{1.97+1.67}{2} = 1.82$	$\theta_{2,1} = \frac{(3.05\times2)+(2.94\times3)+(2.43\times1)}{2+3+1}$ $= 2.89$	$\theta_{2,4} = \frac{(1.82\times2)+(2.76\times3)+(2.57\times1)}{2+3+1}$ $= 2.42$	$Top_2(N) = \{1,4\}$
2	CF + CB	$\gamma^2_{2,1} = \frac{4.00+1.88}{2} = 2.94$	$\gamma^2_{2,4} = \frac{3.52+2.00}{2} = 2.76$			
3	CF + CB	$\gamma^3_{2,1} = \frac{3.48+1.38}{2} = 2.43$	$\gamma^3_{2,4} = \frac{3.48+1.67}{2} = 2.57$			

Figure 5. Examples of generating the top-N item recommendation for the target user u_2 ($Top_2(N)$)

Table II. Dataset

Total number of users (p)	Total number of items (q)	Total number of items (l)	Significance of Criteria (Y)	Total number of features (f)	Total number of rating data
77	50	4	{2,2,3,1}	26	900
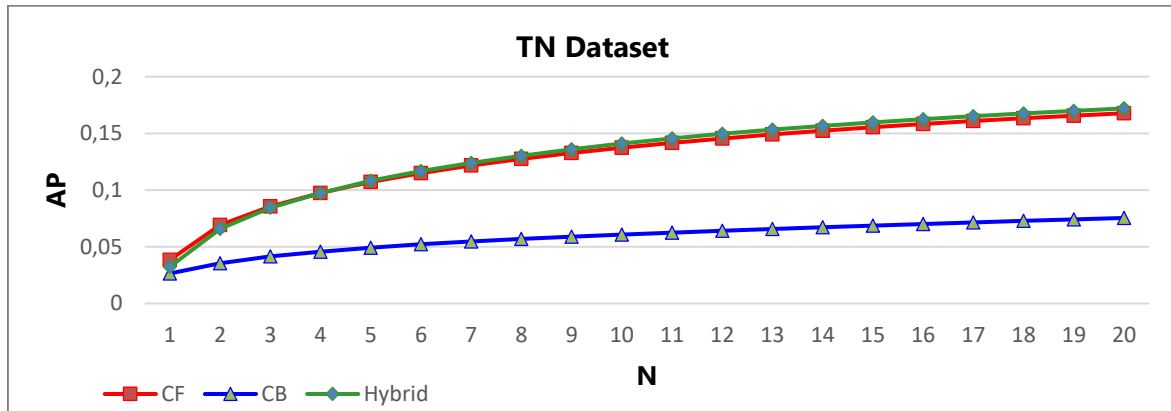
Figure 6. The performance comparison on *TN* dataset
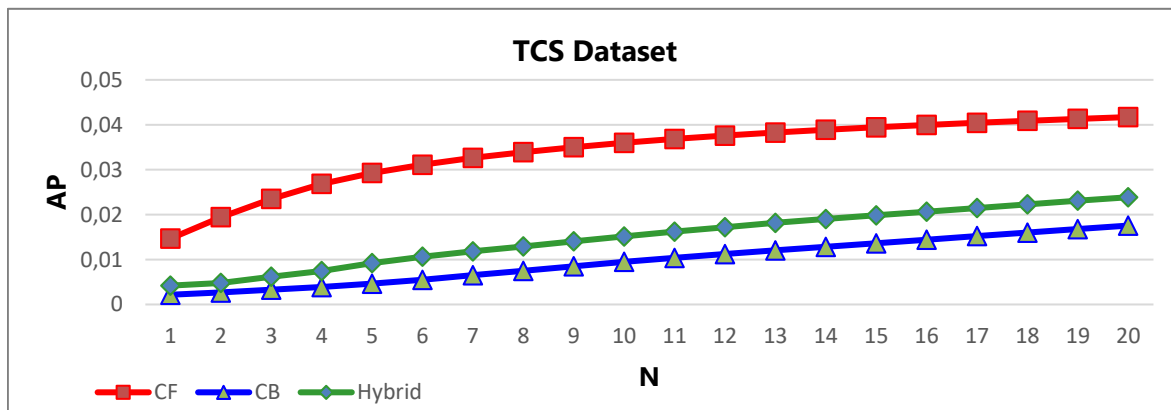
Figure 7. The performance comparison on *TCS* dataset

It is to be noted that $\mathbb{I}(\cdot)$ results 1 when the condition within the bracket is fulfilled, or 0 otherwise.

Performance Comparison

In this sub-section, we compare the performance at $N = \{1 \cdots 20\}$ of the three developed three multi-criteria based item recommendation methods. For ease of explanation, we label the methods based on the rating prediction per criterion approach implemented, i.e., CF, CB, and Hybrid.

Figure 6 shows the performance comparison of the three methods on the LN dataset. We can observe that CF and Hybrid have comparable performance, while CB performs the worst. Note that Hybrid slightly outperforms CF when $N \geq 5$. These results suggest that, on a normal

condition, Hybrid is the best method to be implemented in a multi-criteria based item recommendation system.

Figure 7 shows the performance comparison on the LCS dataset. We can notice that CF achieve the best performance, followed by Hybrid and CB. These results advise that a multi-criteria based item recommendation system must implement a CF method on a cold-start item condition.

Additionally, the poor performance of CB confirms that this method should not be implemented solely in a multi-criteria based item recommendation system on any conditions. It is also worth it to note that CF can be more effective than CB at providing recommendations for cold-start items condition since CF performance is better than that of CB on the LCS

dataset, contradicting the results of previous researches (G. Adomavicius & Tuzhilin, 2005; Aggarwal, 2016).

CONCLUSION

In this paper, we develop three multi-criteria based item recommendation methods. The methods are varied and, therefore, labelled based on what approach is implemented to predict the rating per criterion, i.e., Collaborative Filtering (CF), Content-based (CB), and Hybrid. For a deeper analysis, we generate two variations of datasets to represent the normal and cold-start conditions of multi-criteria item recommendation system. The experiment results suggest that Hybrid and CF are best implemented on the normal and cold-start item conditions, respectively. On the other hand, CB should never be (solely) implemented in a multi-criteria based item recommendation system on any conditions.

Nevertheless, we believe that this study needs to be further expanded. Our recommendations are to implement: (1) other functions of CF, CB and, Hybrid approaches for predicting the rating per criterion; or/and (2) other functions for aggregating rating prediction of all criteria. Alternatively, we may build an application to test the user experience on the multi-criteria item recommendation system based on the CF and Hybrid methods developed in this paper.

REFERENCES

Adomavicius, G., Manouselis, N., & Kwon, Y. (2011). Multi-criteria recommender systems. In *Recommender systems handbook* (pp. 769-803): Springer, Boston.

Adomavicius, G., & Tuzhilin, A. (2005). Toward the Next Generation of Recommender Systems: A Survey of the State-of-the-art and Possible Extensions. *IEEE Transactions on Knowledge and Data Engineering, 17*(6), 734-749. doi:10.1109/tkde.2005.99

Aggarwal, C. C. (2016). *Recommender Systems: The Textbook*. Switzerland: Springer International Publishing.

Bilge, A., & Yargıç, A. (2017). Improving accuracy of multi-criteria collaborative filtering by normalizing user ratings. *Anadolu Üniversitesi Bilim Ve Teknoloji Dergisi A-Uygulamalı Bilimler ve Mühendislik, 18*(1), 225-237.

Burke, R. (2007). Hybrid Web Recommender Systems. In *The adaptive web* (pp. 377-408): Springer Berlin Heidelberg.

Fuchs, M., & Zanker, M. (2012). *Multi-criteria ratings for recommender systems: an empirical analysis in the tourism domain*. Paper presented at the International Conference on Electronic Commerce and Web Technologies.

Jannach, D., Karakaya, Z., & Gedikli, F. (2012). *Accuracy improvements for multi-criteria recommender systems*. Paper presented at the 13th ACM Conference on Electronic Commerce.

Lakiotaki, K., Matsatsinis, N. F., & Tsoukias, A. (2011). Multicriteria user modeling in recommender systems. *IEEE Intelligent Systems, 26*(2), 64-76.

Lops, P., Gemmis, M., & Semeraro, G. (2011). Content-based Recommender Systems: State of the Art and Trends. In *Recommender Systems Handbook* (pp. 73-105): Springer US.

Manouselis, N., & Costopoulou, C. (2007). Analysis and classification of multi-criteria recommender systems. *World Wide Web, 10*(4), 415-441.

Su, X., & Khoshgoftaar, T. M. (2009). A Survey of Collaborative Filtering Techniques. *Advances in Artificial Intelligence, 2009*(January), 4:2-4:2.

Uluyagmur, M., Cataltepe, Z., & Tayfur, E. (2012). *Content-Based Movie Recommendation Using Different Feature Sets*. Paper presented at the World Congress on Engineering and Computer Science Vol I, San Francisco, USA.

Zhang, Z.-K., Zhou, T., & Zhang, Y.-C. (2011). Tag-Aware Recommender Systems: A State-of-the-Art Survey. *Journal of Computer Science and Technology, 26*(5), 767-777. doi:10.1007/s11390-011-0176-1