Using Images in Context-Aware Recommender Systems

Sabri Boutemedjet and Djemel Ziou

Département d'informatique
Université de Sherbrooke, QC, Canada J1K 2R1
{sabri.boutemedjet,djemel.ziou}@usherbrooke.ca

Abstract. In this paper, we propose a unified probabilistic framework for product recommendation which uses both images and user's contextual situation to predict accurately the ratings. In addition, this framework suggests highly rated and diversified products to reach better user satisfactions in conformance with researches in consumer psychology. Experimental results show that images improve the usefulness of recommendation comparatively with state-of-art methods.

Keywords: Recommender systems, context-awareness, content-based image suggestion, information filtering, ranking by diversity, clustering

1 Introduction

The widespread of Internet has promoted many e-commerce services over the world wide Web. For instance, on eBay.com or Amazon.com, it is possible to sell almost everything such as books, DVDs, clothes, etc. Generally, consumers purchase products to satisfy their long-term needs which are relatively stable, regular and refer to periodic preferences. For example, a user interested by fashion, would like to receive periodically items (product highlights, news) related to new fashion clothes, shoes or accessories. Recommender systems are software tools which predict the buyers long-term needs in order to suggest relevant products satisfying these needs. They help users to save a valuable search time spent before purchasing products by reducing the number of choice alternatives. For instance, Amazon.com suggests products to its users based on their historical data of ratings. A rating is a numerical value defined on an ordered scale and quantifies the users interest in the rated item (explicit preference indicator). From online retailers' point of view, recommender systems constitute an efficient advertisement strategy which personalizes highlighted products in order to acquire new potential consumers and retain existing ones.

Recommender systems predict the buyers' needs based on the collected historical data. The historical data can be seen as a user-product matrix where each entry (u, p) is the rating provided by user u to the product p. Due the availability of a huge amount of products, the proportion of empty entries in the user-product matrix is extremely high. Then, recommender systems first start by predicting the missing ratings (empty entries) corresponding to unseen products using information filtering techniques [2]. After that relevant items are identified as those having the highest predicted rating. It has been noticed in literature that the accuracy of the rating prediction is increased by exploiting the both information about the user's context and products.

Context-awareness in recommenders has been motivated from researches in consumer psychology which recognize the dependence of user long-term needs on the time, location, and any information about the physical environment surrounding the user [3]. It introduces an additional level of personalization by considering the influence of the external environment of the user on his/her appreciation of the products [21, 6]. For instance, location-based recommender systems exploit the contextual information defined by the user's geographical location (captured from user's mobile device) to suggest personalized advertisements of products in neighboring commerces.

Images constitute an other important factor influencing the usefulness of the recommendations. Note that many products such as jewelery or clothes are adopted by users because of their visual appearance which defines their look-and-feel in terms of the color, shape, and texture [10]. In some cases, the semantic information extracted from images may lead to better discrimination among image categories [13]. In the domain of marketing, images have been used as efficient means in advertisements since they can convey meanings that cannot be expressed using words [16]. For instance, images have been used successfully to present "highlighted products" in the Web site of many online retailers such as Amazon.com or eBay.com. This presentation style is motivated mainly by high persuasion power of images. In fact, a qualitative study published in 2005, shows how users are influenced by the product's visual appearance which carries information about aesthetics (emotional pleasure), functionality (number of offered options), or quality [8]. So far, the persuasive power of images on consumers has not been taken into account by rating prediction algorithms but only to present products. Thus, existing recommender systems do not model explicitly user long-term needs that related to the visual appearance of products expressed as "like product X of look-and-feel Y".

Once we have collected the data about users, products, contexts, and ratings, we need to design algorithms which model these data in a feature space to predict the missing ratings of unseen products. The majority of existing recommendation algorithms rank products by the predicted rating only. However, consumer psychology researches have shown that the variety-seeking behavior of the consumer pushes him/her to reduce the redundancy of some product attributes in consecutive purchase occasions. In fact, by diversifying choices, users reduce the risk of uncertainty caused by lack of expertise on some products, complement other already purchased products, or simply to avoid boredom [15]. In other words, the predicted rating is not a sufficient criterion to better satisfy consumers. Let us consider the example illustrated in Fig.1 which shows the suggestion lists of three products obtained using two ranking strategies where "laptop" is the product category having the highest predicted rating. If the rat-

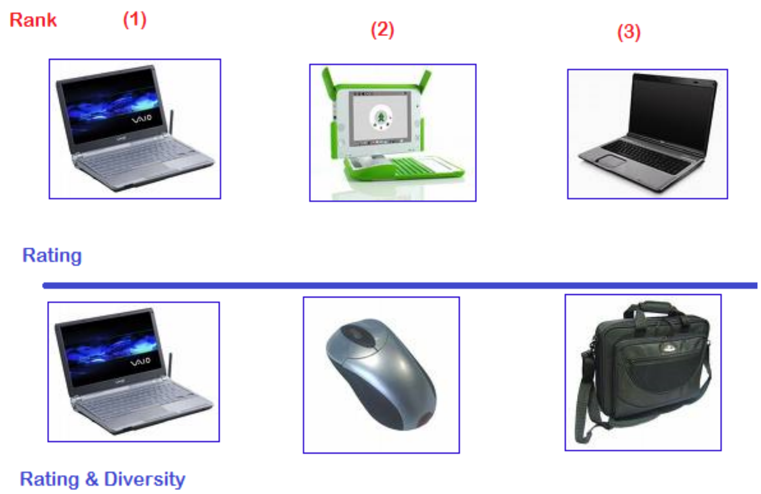


Fig. 1. Example of two suggestion strategies for given user interested by a "laptop".

ing is the only criterion for ranking products, it is natural that the first three suggestions are laptops since similar items receive similar ratings. Note that the second suggestion obtained by "rating & diversity" is more useful for both the user and online retailers since it provides diversified and complementary suggestions. Therefore, to reach highest user satisfaction, it is important to consider also the consumption history of each user to rank products by both predicted rating and diversity.

In this paper, we present content-based image suggestion (CBIS) which investigates the added-value of images and user contextual situations in making useful and diversified recommendations. We present our unified probabilistic approach which models seamlessly the uncertainty of the long-term needs of consumers, image collection, and the diversity of suggestions. This paper is organized as follows. In the next Section, we present recent advances in recommender systems. Then, we detail two ways of using images in improving the usefulness of recommendation algorithms. Experimental results are presented in Section 4. Finally, we conclude the paper with a summary of the work.

2 Related research work

During the last two decades, many relevant issues have been addressed in literature to increase the usefulness of recommender systems. In the following two subsections, we categorize recent advances in recommender systems at the levels of predicting accurately the ratings and ranking products by diversity.

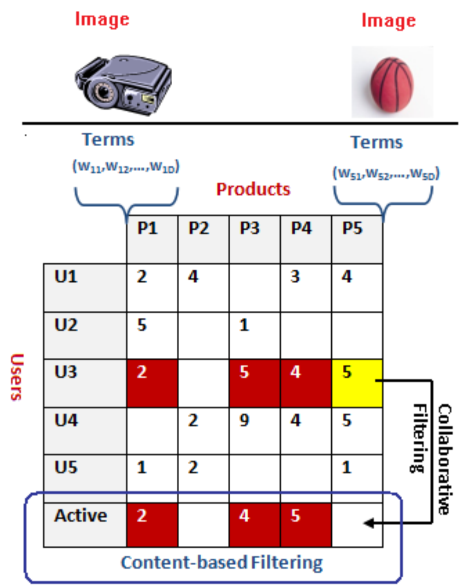


Fig. 2. Three major techniques of rating prediction. CBIS exploits the power of product images during the prediction of ratings.

2.1 Rating prediction

Recommender systems employ Information filtering (IF) technologies to predict the missing ratings for unseen products (e.g. empty entries in the user product matrix) [2]. In IF literature, there exist three families namely content-based filtering (CBF), collaborative filtering and hybrid methods. CBF employs information retrieval (IR) techniques in representing user profiles using content descriptors that are mainly defined by the textual information extracted from product captions, surrounding text in Web pages, etc (see "Terms" in Fig. 2). The principle of CBF methods is that items similar to those preferred by the user in the past, will be preferred in the future. For a given user, CBF classifies the ith product represented by word features $(w_{i1}, \dots, w_{iD})$ in the category of relevant or irrelevant products [17]. However, the major shortcoming of CBF is its inability to recommend to the user "unexpected" items different from what he/she has already rated in the past.

Collaborative filtering (CF) methods identify the neighbors of the user (other users with similar needs) based on ratings they provided on the same products. The neighbors are identified by analyzing the correlation among the rows of the user-product matrix. For example, in Fig. 2, $U3$ is the neighbor the active user since both of them liked the products P_3 and P_4 and disliked P_1. In CF, we find either distance-like methods [18] or model-based clustering-like techniques such as [12, 11, 19]. CF methods consider items as a categorical variable (i.e. unique index for each item) and are unable to suggest unseen items (novel products). Hybrid methods take advantages of both CF and CBF and identify both the neighbors of the user and products categories in making rating predictions [20]. For context-awareness, the authors in [1] define the contextual information as location, time, and compagnon. Then, many reduced user-product matrices specific to each context, are derived from the historical data. Then, a collaborative filtering technique is employed on the reduced matrix to predict the empty entries for a given context.

2.2 Ranking items by diversity

A natural way to promote the diversity is to eliminate the redundancy by considering the dissimilarity of each image with respect to previous consumptions according to some distance metric. Some normalization of term weights of the query can be resolved before accurately computing the distance metric [9]. Information retrieval has addressed the issue of ranking documents by diversity by a general two stage procedure as follows [22, 7]. The first document is selected as the one being the most similar to the query (topic). Then, documents are inserted successively into the result set according to both their similarity to the query and the redundancy they provide with respect to the already retrieved documents. The authors in [22] penalize the results with lower number of covered subtopics. The authors in [7], employ a probabilistic model for retrieval where the prior distribution (over word features) is updated successively each time a document is selected within the result set. Diversity-ranking methods have shown to reach better user satisfactions in retrieval tasks.

3 Content-based image suggestion

In this section, we investigate the use of product images instead of textual features, as shown in Fig. 2 in modeling the entries of the user-product matrix. Therefore, we consider images as a contextual information at the level of products defining their look-and-feel. For users, we investigate the added-value of the context defined by the external environment (location and time) in refining product recommendations.

3.1 Notations

We consider the following representation of the user-product matrix extended with both visual and contextual information. We have a set of users $\mathcal{U} = \{1, 2, \ldots, N_u\}$, a set of images $\mathcal{V} = \{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_{N_v}\}$, and a set of possible contexts $\mathcal{E} = \{1, 2, \ldots, N_e\}$. Each $\boldsymbol{v}_k$ is a visual descriptor used to represent the content (color, shape, texture) of products. For instance, it may carry information about the shape, color, or the texture present in images. We define the context as a combination of two attributes: location $\mathcal{L} = \{in - campus, out - campus\}$ inferred from the Internet Protocol (IP) address of the subject, and time as $\mathcal{T} = (weekday, weekend)$ i.e $N_e = 4$. The rating is expressed explicitly on an ordered scale defined as $\mathcal{R} = \{1, 2, \ldots, N_r\}$. For example, the five star scale (i.e. $N_r = 5$) such as the one used by Amazon.com, allows the users to give more detailed degrees of appreciation. The CBIS data set is defined as $\mathcal{D} = \{d^{(i)} =< u^{(i)}, e^{(i)}, \boldsymbol{v}^{(i)}, r^{(i)} > | e^{(i)} \in \mathcal{E}, \boldsymbol{v}^{(i)} \in \mathcal{V}, r^{(i)} \in \mathcal{R}, i = 1, \ldots, N\}$. Note that each observation $d^{(i)}$ is nothing else than an entry in the extended user-product matrix.

3.2 Using images to predict ratings

We consider the problem of CBIS as the maximization of a utility that ranks images for a user in a certain context. In this subsection, we exploit the power of images to define a more accurate utility which incorporates the information about both the rating and diversity.

Let $\mathcal{X} = \{x_1, x_2, \ldots, x_L\}$ be a list of L ranked images to recommend to a given user u in a context e where $x_t \in \mathcal{V}, t = 1, \ldots, L$, is the image at rank t in $\mathcal{X}$. The diversity of $\mathcal{X}$ imposes another condition that involves measuring dependencies (information redundancies) within subsets of products during the suggestion process. Therefore, the utility of the tth suggested product depends on both its rating and other products $\mathcal{X}_t = \{x_1, \ldots, x_{t-1}\}$ in the suggestion list that have been already consumed. The following utility function measures such compromise

$$x_t = \arg \max_{x \in \mathcal{V} - \mathcal{X}_t} s(x, u, e | \mathcal{X}_t) \tag{1}$$

To predict the ratings, we propose a generative model $p(u, e, x, r)$ which captures the joint probability (uncertainty) to observe a rating r for any entry (u, e, x). Note that one could predict probabilistically the rating using $p(r|u, e, x)$ obtained by conditioning $p(u, e, x, r)$ on (u, e, x, r). Based on product images, we consider similar users as those who have preferred similar images. For that end, we should first identify K user classes and M image classes from the observed data set $\mathcal{D}$. Then, two latent variables z and c label each data (u, e, x, r) with information about the user class and image class, respectively. We adopt the visual content flexible mixture model (VCC-FMM) [5]

$$p(u, e, x, r) = \sum_{c=1}^{M} \sum_{z=1}^{K} p(z)p(u|z)p(e|z)p(c)p(x|c)p(r|z, c) \tag{2}$$

The quantities $p(z)$ and $p(c)$ denote the *a priori* weights of user and image classes. $p(u|z)$ and $p(e|z)$ denote the likelihood of a user and context to belong respectively to the user's class z. $p(r|z, c)$ is the probability to generate a rating for a given user and image classes. Finally, $p(\boldsymbol{v}|c)$ is multi-dimensional continuous-valued generalized Dirichlet distribution (GD), parameterized by $2 \times d$-dimensional vector $\boldsymbol{\delta}_c$. We denote by Θ, the set of VCC-FMM parameters

$$\Theta = \big(p(z), p(c), p(u|z), p(e|z), \boldsymbol{\delta}_c, p(r|z, c)\big) \tag{3}$$

We train this model from the data $\mathcal{D}$ to identity the optimal parameters Θ_{ML} which maximize the log-likelihood of the data set $\log p(\mathcal{D})$

$$\Theta_{ML} = \arg\max_{\Theta} p(\mathcal{D}) = \prod_i p(u^{(i)}, e^{(i)}, x^{(i)}, r^{(i)}) \tag{4}$$

The numbers of user classes K and image classes M are unknown and their automatic identification from $\mathcal{D}$ is still a challenging problem in unsupervised learning. However, one could estimate automatically these numbers (M and K) from the data using minimum message length (MML) approach [5].

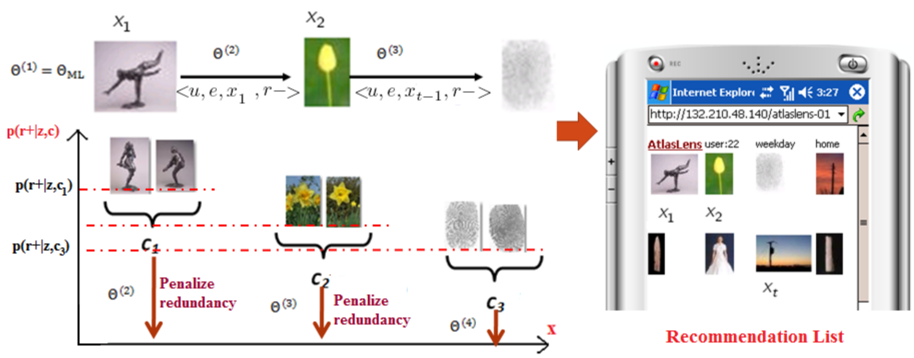


Fig. 3. The principle of our diversity-based ranking for suggesting highly rated and diversified images. Horizontal axis outlines the similarity of images while vertical axis defines the probability for an image to get a high rating. Images of the class c_1 (similar to x_1) are penalized by setting $p(r^+|z, c_1) \simeq 0$.

3.3 Using images in to rank by diversity

The diversity of the suggestion lists $\mathcal{X}$ can be measured as the degree of dissimilarity between all images in the list. Based on the visual information of products, we maximize both the diversity and the rating by an appropriate design of the ranking function $s(x, u, e|\mathcal{X}_t^{ue})$. Since consumers make binary purchase decisions (buy or not), we employ a binary scale $\{r^+, r^-\}$ for ratings. The products are ranked probabilistically according to a utility which favors those with high ratings as follows

$$s(x, u, e) = \log \frac{p(r^+|x, u, e)}{p(r^-|x, u, e)} \tag{5}$$

where $p(r^-|x, u, e) = \sum_{r=1}^{T_r} p(r|x, u, e)$, $p(r^+|x, u, e) = 1 - p(r^-|x, u, e)$ and T_r is a threshold used to separate positive and negative ratings.

Now, the principle of our diversity-ranking strategy is to recommend only "highly rated" products which belong to "different classes". Given that we have already recommended $\mathcal{X}_t$ products, we select the current one such that it is "visually" dissimilar from those in $\mathcal{X}_t$ by assuming previous products *"irrelevant"*. This assumption is implemented by generating negative ratings for the consumed products $\{< u, e, x_{t'}, r^- >, t' = 1, \ldots, t\}$. In order to take into account the new information about the irrelevance of $\mathcal{X}_t$, the parameters of the model $\Theta^{(t)}$ are successively updated from each observation $< u, e, x_{t'}, r- >$. Let $s(x, u, e; \Theta)$ be the utility (5) computed using a certain model Θ given by equation (3). Then, to promote the diversity, x_t^{ue} is selected according to the utility (1) and having the form (5) with $s(x, u, e; \Theta^{(t)})$ except that the parameters are updated with diversity information:

$$s(x, u, e|\mathcal{X}_t^{ue}) = s(x, u, e; \Theta^{(t)}) \tag{6}$$

The general scheme of our algorithm is given as follows. Initially, we set $\mathcal{X}_1^{ue} = \emptyset$ and $\Theta^{(1)} = \Theta_{ML}$ given by Eq. (4). Then, each time an image x_{t-1} of class c_{t-1}^* is suggested, we use a cost-effective online learning since an offline relearning is not a reasonable solution. The probability of positive ratings for images of the same class $p(r^+|z, c_{t-1}^*)$ are updated effectively as [4]

$$\begin{aligned} &p(r^+|z, c_{t-1}^*) = 0, \\ &c_{t-1}^* = \arg\max_c p(c|u, e, x_{t-1}, r^+) = \frac{p(c, u, e, x_{t-1}, r^+)}{p(u, e, x_{t-1}, r^+)} \end{aligned} \tag{7}$$

with $p(c, u, e, x, r) = \sum_z p(z)p(u|z)p(e|z)p(c)p(x|c)p(r|z, c)$. Intuitively, Eq. (7) allows the selection of image class representatives with the highest predicted ratings. Therefore, the proposed diversity-ranking strategy seeks for "novel" products with high-ratings as illustrated in Fig. 3. The first product to suggest comes from the class "piece of art" in left since it has the highest rating. Once it is selected its probability of high-rating is reduced to zero. The second product to suggest will necessarily come from the class with the second highest rating (flowers). This process is repeated until the suggestion list is filled. Note that a rating-based ranking strategy can be implemented straightforwardly by considering constant parameters, i.e. $\Theta^{(t)} = \Theta_{ML}, \forall t$

$$s(x, u, e|\mathcal{X}_t^{ue}) = s(x, u, e; \Theta_{ML}) \tag{8}$$

4 Experiments

The aim of this experiment is to measure the contribution of the visual information in making accurate recommendations comparatively with state-of-art methods. We make comparisons with some representative algorithms used for rating prediction that are the Aspect model [11], Pearson Correlation (PCC)[18], Flexible Mixture Model (FMM) [19], the Decoupled Model (Decoupled) [12] and the User Rating Profile (URP)[14]. For CF approaches, we consider images as a categorical variable. To investigate the usefulness of contextual information, we evaluate the V-FMM which is the variant of VCC-FMM with only one (homogeneous) context information, i.e. $\mathcal{E} = \{1\}$. We measure the performance of algorithms in terms of both accuracy of predicting ratings.

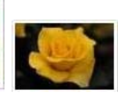
Classes	**c = 2**	**c = 11**	**c = 14**	**c = 28**
Images				
$\mathbf{z = 1}$	0.21	0.08	0.74	0.17
$\mathbf{z = 2}$	0.01	0.84	0.07	0.32
$\mathbf{z = 3}$	0.88	0.32	0.92	0.06

Table 1. Sample visual content class representatives and the estimated probability of high rating $p(r^+|z, c)$, for a selected set of user classes.

Table 2. Averaged MAE with standard deviations over 10 runs of the different algorithms on $\mathcal{D}$. The relative improvement rates are computed by comparing MAe of each algorithm with that of PCC.

	PCC(baseline)	Aspect	FMM	URP	Decoupled	V-FMM	VCC-FMM
Avg MAE	1.327	1.201	1.145	1.116	1.095	0.890	0.646
Std Deviation	0.040	0.051	0.036	0.042	0.037	0.034	0.014
Improvement	0.00%	9.49%	13.71%	15.90%	17.48%	32.94%	55.84%

4.1 Data Set

We present experimental results conducted on a collected from 27 subjects who participated in the experiment (i.e. $N_u = 27$) during a period of three months. The participating subjects are graduate students in faculty of science (computer science, mathematics, biology, and chemistry). Subjects received periodically (twice a day) a list of three images on which they assign relevance degrees expressed on a five star rating scale (i.e. $N_r = 5$). A data set $\mathcal{D}$ of 13446 ratings is collected ($N = 13446$). We have used a general-purpose collection of 4775 images collected in part from Washington University and another part from collections of free photographs. The image collection which we experiment here contains both man-made and natural images and categorized into 41 categories. To represent images, we have employed both local and global descriptors. For local descriptors, we use the 128-dimensional Scale Invariant Feature Transform (SIFT) to represent image patches. We employ vector quantization to SIFT descriptors and we build a histogram for each image ("bag of visual words"). The size of the visual vocabulary is 100. For global descriptors, we used the color correlogram for image texture representation, and the edge orientation histogram. An image descriptor is a 140-dimensional

4.2 Prediction accuracy

Experiment protocol We divide the data set $\mathcal{D}$ into two halves: one for training VCC-FMM and the remaining part for validation. We measure the accuracy of the prediction using the Mean Absolute Error (MAE) which is the average of the absolute deviation between the ratings $r_{\boldsymbol{v}}^{ue}$ in the validation data $\mathcal{D}_{test}$ and the predicted ones $\hat{r}_{\boldsymbol{v}}^{ue} = \sum_r r(p(u,e,\boldsymbol{v},r)/\sum_r p(u,e,\boldsymbol{v},r)$

$$MAE = \frac{1}{|\mathcal{D}_{test}|} \sum_{d_i \in \mathcal{D}_{test}} |r_{\boldsymbol{v}^{(i)}}^{u^{(i)}e^{(i)}} - \hat{r}_{\boldsymbol{v}^{(i)}}^{u^{(i)}e^{(i)}}| \tag{9}$$

Results The first five columns of table 2 show clearly the added value of the visual content comparatively with pure CF techniques. For instance, the improvement in the rating's prediction reported by V-FMM is 22.27% and 19.81% comparatively with the recent CF approaches FMM and URP, respectively. VCC-FMM which takes into account the context information has also improved the accuracy of the prediction comparatively with the others (at least an additional 15.28%). From consumer psychology [3], this fact outlines clearly the influence of the contextual situation on user long-term needs.

4.3 Usefulness of suggestion lists

Experiment protocol This evaluation measures the effectiveness of rating-based and diversity-based ranking strategies in terms of user satisfactions. In each experiment run, we initialize $\mathcal{X} = \emptyset$ and we put $T_r = 3$ to separate negative (r^-) and positive (r^+) ratings. We collect satisfaction indicators from the human subjects who participated in the generation of that data set. Each subject is recommended eight images on each of which he/she attributes a binary relevance degree: "0" for not-relevant and "1" for relevant. Then, we evaluate quantitatively the usefulness of the suggestion using the precision computed as the proportion of relevant images in the the list.

Results Figure 4 shows that for rating-based ranking, the higher the size of suggestion lists, the lower the value of average precision. Also, the diversity-based ranking reaches better user's satisfaction (18.06% in average) than rating-based one. Indeed, by removing "visual redundancy", we improve the usefulness of suggestion lists which conforms with consumer psychology researches [15]. Finally, it is shown that the "optimal" size of suggestion lists, i.e. highest average precision, are four and eight images for rating-based and diversity-based suggestions, respectively.

5 Conclusions

In this paper, we have studied the contribution of the visual and contextual information in the improvement of the usefulness of recommender systems. The proposed model predicts the outcome of the user's decision making in each context based the preferences of other users with similar interests on product images. Experiments showed that images helped significantly in increasing the accuracy of rating prediction and usefulness of suggestion lists.

Acknowledgements

The completion of this research was made possible thanks to Natural Sciences and Engineering Research Council of Canada (NSERC) and Bell Canada's support.

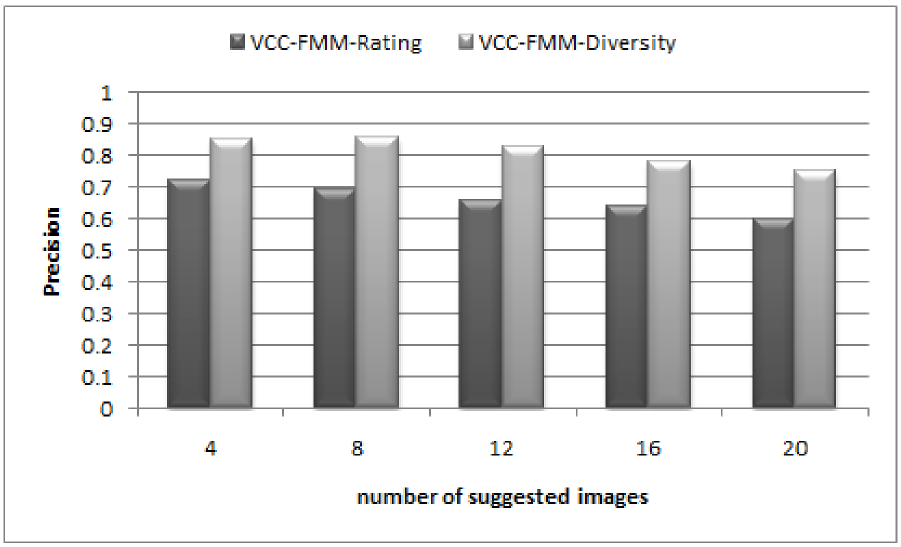


Fig. 4. Average precision reported by two ranking methods for building suggestion lists.

References

1. G. Adomavicius, R. Sankaranarayanan, , S. Sen, and A. Tuzhilin. Incorporating contextual information in recommender systems using a multidimensional approach. *ACM TOIS*, 23(1):103–145, 2005.
2. G. Adomavicius and A. Tuzhilin. Toward the Next Generation of Recommender Systems: A Survey of the State-of-the-Art and Possible Extensions. *IEEE TKDE*, 17(6):734–749, 2005.
3. R. Belk. Situational Variables and Consumer Behavior. *Journal of Consumer Research*, 2:157–164, 1975.

4. S. Boutemedjet and D. Ziou. A Graphical Model for Context-Aware Visual Content Recommendation. *IEEE Transactions on Multimedia*, 10(1):52–62, 2008.
5. S. Boutemedjet, D. Ziou, and N. Bouguila. Unsupervised Feature Selection for Accurate Recommendation of High-Dimensional Image Data. In *Proc. of NIPS*, 2007.
6. I. Cantador, A. Bellogín, and P. Castells. Ontology-based personalised and context-aware recommendations of news items. In *Proc. of IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology*, pages 562–565, 2008.
7. H. Chen and D. Karger. Less is More:Probabilistic Models for Retrieving Fewer Relevant Documents. In *Proc. of SIGIR 29*, pages 429–436, 2006.
8. M. Creusen and J. Schoormans. The Different Roles of Product Appearance in Consumer Choice. *Journal of Product Innovation Management*, 22(1):63–81, 2005.
9. M. Fernández, D. Vallet, and P. Castells. Using historical data to enhance rank aggregation. In *Proceedings of the 29th annual international ACM SIGIR conference on Research and development in information retrieval*, page 644. ACM, 2006.
10. A. Fiore, S. Lee, and G. Kunz. Individual differences, motivations, and willingness to use a mass customization option for fashion products. *European Journal of Marketing*, 38(7):835–849, 2004.
11. T. Hofmann. Latent Semantic Models for Collaborative Filtering. *ACM TOIS*, 22(1):89–115, 2004.
12. R. Jin, L. Si, and C. Zhai. A Study of Mixture Models for Collaborative Filtering. *Journal of Information Retrieval*, 9:357–382, 2006.
13. H. Lowe, I. Antipov, W. Hersh, and C. Smith. Towards knowledge-based retrieval of medical images. The role of semantic indexing, image content representation and knowledge-based retrieval. In *Proceedings of the AMIA Symposium*, page 882. American Medical Informatics Association, 1998.
14. B. Marlin. Modeling User Rating Profiles For Collaborative Filtering. In *Proc. of Advances in Neural Information Processing Systems 16 (NIPS)*, 2003.
15. L. McAlister and E. Pessemier. Variety Seeking Behavior: An Interdisciplinary Review. *The Journal of Consumer Research*, 9(3):311–322, 1982.
16. P. Messaris. *Visual Persuasion: The Role of Images in Advertising*. Sage Pubns, 1997.
17. R. Mooney and L. Roy. Content-Based Book Recommending Using Learning for Text Categorization. In *Proc. of ACM DL*, 2000.
18. P. Resnick, N. Iacovou, M. Suchak, P. Bergstrom, and J. Riedl. Grouplens: An Open Architecture for Collaborative Filtering of Netnews. In *ACM CSCW*, 1994.
19. L. Si and R. Jin. Flexible Mixture Model for Collaborative Filtering. In *Proc. of ICML*, pages 704–711, 2003.
20. L. Si and R. Jin. Unified Filtering by Combining Collaborative Filtering and Content-Based Filtering via Mixture Model and Exponential Model. In *Proc. of CIKM*, pages 156 – 157, 2004.
21. A. Yeung, N. Gibbins, and N. Shadbolt Contextualising tags in collaborative tagging systems. In *Proceedings of the 20th ACM conference on Hypertext and hypermedia*, pages 251–260, 2009.
22. C. Zhai, W. W. Cohen, and J. Lafferty. Beyond Independent Relevance: Methods and Evaluation Metrics for Subtopic Retrieval. In *Proc. SIGIR*, pages 10–17, 2003.