Sentiment-Enhanced Explanation of Product Recommendations

Li Chen and Feng Wang
Department of Computer Science
Hong Kong Baptist University, China
{lichen, fwang}@comp.hkbu.edu.hk

ABSTRACT

Because of the important role of product reviews during users' decision process, we propose a novel explanation interface that particularly fuses the feature sentiments as extracted from reviews into explaining recommendations. Besides, it can explain multiple items altogether by revealing their similarity in respect of feature sentiments as well as static specifications, so as to support users' tradeoff making. Relative to existing works, we believe that this interface can be more effective, trustworthy, and persuasive.

Categories and Subject Descriptors

H.5.m **[Information interfaces and presentation (e.g. HCI)]**: Miscellaneous

Keywords

Recommendation explanation; sentiment analysis; e-commerce.

1. INTRODUCTION

How to effectively explain the recommended results has been recognized important for product recommender systems. In recent years, the focus has been mainly on increasing the system's *transparency* by means of explaining the underlying mechanism of collaborative filtering based or content-based algorithms. However, few works have attempted to improve the system's *effectiveness*, *trustworthiness* and *persuasiveness* via explanation [4]. In more detail, effectiveness refers to the system's ability in allowing users to make accurate decision. The trustworthiness represents the ability of inspiring users' trust in the system and furthermore their trusting intentions (like intention to return and intention to save efforts) [2]. Persuasiveness shows the system's ability to convince users to try or buy a proposed item. These three goals are indeed all crucial to any recommender systems in e-commerce environment.

Previously, we have developed a so called *preference-based organization interface* [2]. It differs from the traditional single-item explanation in that it does not emphasize explaining the recommended items one by one, but revealing a group of items' similarity in terms of their attribute values. Therefore, users might be aided to compare multiple items and make attribute tradeoffs among them. The user study revealed that our interface is capable of supporting users' decision making and inducing their trust.

Motivated by prior study, in the current work, we are interested in further improving this interface by fusing reviews' sentiment analysis results, so as to strengthen the above-mentioned three goals. Actually, product reviews are found taking important role in influencing users' evaluation of products and even their final choice [1]. That is, an active buyer is more likely to rely on other users' reviews to a product (e.g., opinions on the camera's image quality and ease of use) to judge its quality, instead of purely examining the product's static specifications. We hence propose a sentiment-enhanced organization interface, as the extension to our previous work. It is expected that this interface can be more *effective*, *trustworthy*, and *persuasive*, due to the incorporation of features' sentiment info. In other words, since product reviews are normally in the form of free texts, it should be meaningful to extract valuable info from them and present it in a way that can effectively assist users in making decisions.

2. INTERFACE DESIGN AND IMPLEMENTATION

Suppose the system recommends k items to a user during each cycle. Except the top-ranked item (that we call "top candidate"), the other items can be organized into m categories. The design principles that we derived from previous studies are referred [2]: 1) each category title acts as the explanation, to show the pros and cons of the contained products against the top candidate; 2) each category contains up to six products so as to avoid information overload; 3) the number of attributes accommodated in each explanation is controlled under five; 4) the explanations should be as diverse as possible since it is not informative to show two categories with similar titles. In addition to these general principles, we incorporate reviews' sentiment info into showing the similar properties of products (see Figure 1). For example, one category title is "*they have better values at price, screen size, and better opinion at resolution, but worse value at weight*", where "*better opinion at resolution*" reflects the average sentiment on "resolution" of these contained products given their reviews. In the following, we describe how we implement such interface.

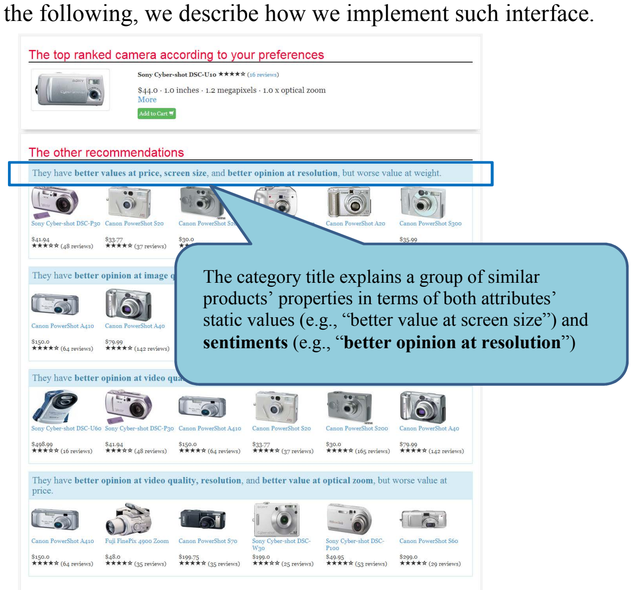

Figure 1. Explanation incorporated with feature sentiments.

WWW'14 Companion, April 7–11, 2014, Seoul, Korea.
ACM 978-1-4503-2745-9/14/04.
http://dx.doi.org/10.1145/2567948.2577276

Feature-based sentiment analysis. Being different from classical document-based sentiment classification, we conduct feature-based sentiment analysis so as to know review writers' opinions on specific features. For this purpose, we first use a part-of-speech tagger to extract frequent nouns and noun phrases from reviews as feature candidates. We then identify the associated opinion(s) by looking for the nearby adjective words or phrases. The sentiment score (in the range of [1, 5]) is concretely determined via a dictionary SentiWordNet. Moreover, we consider the appearance of intensifier and negations words in adjusting the score. The synonymous features are then grouped and mapped to attributes through computing the lexical similarity with WordNet. As a result, each product is denoted as $\{(f_i, os_i, x_i)_{1:m}, (f_j, os_j)_{m+1:n}\}$, where os_i is the sentiment on attribute f_i which also has static value x_i, and f_j is the opinion feature like "ease of use" which only has sentiment os_j.

Modeling of user preferences. Grounded on the Multi-Attribute Utility Theory, we model the active user's preferences over all products as a weighted additive form of value functions:

$$U(p) = \sum_{i=1}^{m} w_i * [\alpha * v_i(x_i(p)) + (1-\alpha) * v_{senti}(os_i(p))] + \sum_{j=m+1}^{n} w_j * v_{senti}(os_j(p))$$

In this formula, $v_i(x_i(p))$ is the value function of the i-th attribute w.r.t. product p, $v_{senti}(os_i(p))$ is the function over the attribute's sentiment (and $v_{senti}(os_j(p)$ is the opinion feature's function). w_i is the attribute's weight (in [1, 5]). Its default value is 3 and the default value on α (that is used to control the relative importance between an attribute's static value and its sentiment) is 0.50. We also assign default value function to un-stated attributes (e.g., "the cheaper, the better" for price, "the lower, the better" for weight, "the higher, the better" for all sentiments). Then, according to this preference model, all products can be ranked by their utilities, and the top k ones will form the set of recommendations.

Generation of category candidates. Among the k items, except the ranked 1st one left as the top candidate, each of the others is converted into a *tradeoff vector* $\{(attribute, tradeoff)\}$, where the *tradeoff* is either *improved* (↑) or *compromised* (↓), indicating the *attribute's* value or sentiment is better or worse than the one of the top candidate. The formula below shows how the tradeoff vector is determined regarding each item.

If p's sentiment on f_i is negative $(os_i(p) < 3)$

$$tradepff(f_i, p', p) = \begin{cases} \uparrow_o \ , & if\ 1 \le i \le m,\ x_i(p') \succcurlyeq x_i(p)\ and\ os_i(p') > os_i(p) \\ \uparrow_o \ , & if\ m < i \le n,\ os_i(p') > os_i(p) \\ \downarrow \ , & if\ 1 \le i \le m,\ x_i(p') \prec x_i(p) \end{cases}$$

Else $(os_i(p) \ge 3)$

$$tradeoff(f_i, p', p) = \begin{cases} \uparrow_v & if\ 1 \le i \le m, x_i(p') \succ x_i(p)\ and\ os_i(p') \ge os_i(p) \\ \downarrow & if\ 1 \le i \le m, x_i(p') \prec x_i(p) \end{cases}$$

where p is the top candidate, and p' is the compared item. It can be seen that when the top candidate's sentiment on one attribute f_i is negative, the main focus is on judging whether the compared item has better sentiment (i.e., $\uparrow_o$) on f_i, while if the sentiment is positive, it emphasizes showing the better static value ($\uparrow_v$) of f_i.

Subsequently, the Apriori algorithm (a popular tool for retrieving frequent patterns) is performed over all tradeoff vectors, in order to discover the frequently occurring subsets of (*attribute*, *tradeoff*) pairs. Each retrieved subset indicates a category candidate (e.g., "{(price, $\uparrow_v$), (screen size, $\uparrow_v$), (resolution, $\uparrow_o$), (weight, ↓)}" that represents a group of products). Note that a product might belong to more than one category in the case that it has different subsets of tradeoff properties shared by different groups of products.

Selection of categories. Each category candidate is additionally computed with a score, suggesting its gains versus losses relative to the top candidate, as well as the matching degree with the user's current preferences and the diversity degree with other already selected categories:

$$\text{Score(C)} = \left(\sum_{i}^{|C|} w_i \times tradeoff_i\right) \times \left(\frac{1}{|SR(C)|} \sum_{p \in SR(C)} U(p)\right) \times Diveristy(C, SC)$$

where *C* denotes the currently concerned category candidate with a set of $(attribute, tradeoff)$ pairs, *SR(C)* denotes z products in *C* (ranked by their utilities), and *SC* denotes the set of categories selected so far. The tradeoff value $tradeoff_i$ is default set as 0.75 if "$\uparrow_v$" (or "$\uparrow_o$"), or 0.25 if "↓". The diversity degree $Diveristy(C, SC)$ is defined by both the categories' titles and their contained products [2]. The selection process ends when there are *m* categories decided.

3. EVALUATION PLAN

We have implemented the explanation interface in two product domains: digital camera (194 items, 6 attributes, and 3 opinion features) and laptop (139 items, 7 attributes, and 4 opinion features). The parameters k (the number of recommendations), *m* (the number of categories), and z (the number of items in each category) were set as 25, 4, and 6, respectively, according to the design principles noted before. For the next step, we plan to perform a user study to compare it with our previous version. We are setting up an online experiment website that can be convenient for participants to take part in the study. Each user's actions will be automatically recorded in a log file for our analysis. Moreover, we will obtain her/his subjective perceptions with the interface, based on the evaluation framework for recommender systems [3]. Specifically, we aim to validate the following three hypotheses through the experiment:

Hypothesis 1: the new interface (shorted as Senti-ORG) would be more ***effective*** than the original design (ORG [2]) in terms of aiding users to make accurate and confident decisions;

Hypothesis 2: Senti-ORG would be more ***trustworthy*** than ORG, so that users are more inclined to return to use it;

Hypothesis 3: Senti-ORG would be more ***persuasive***, given that more users would be prepared to buy product chosen from it.

4. CONCLUSION

In this paper, we present the design and implementation of a new explanation interface for product recommenders. We believe this interface could be more effective, trustworthy, and persuasive, than related ones, due to the incorporation of feature sentiments as extracted from product reviews. In the future, we will not only test it through user study, but also plug such technique into different types of recommender systems for ideally aiding users' decisions.

5. ACKNOWLEDGMENTS

We thank grants ECS/HKBU211912 and NSFC/61272365.

6. REFERENCES

[1] Chatterjee, P. 2001. Online Reviews: Do Consumers Use Them? In *Advances in Consumer Research* 28, 129-133.

[2] Chen, L. and Pu, P. 2010. Experiments on the Preference-based Organization Interface in Recommender Systems. In *TOCHI* 17, 1, 1-33.

[3] Pu, P., Chen, L. and Hu, R. 2011. A User-Centric Evaluation Framework for Recommender Systems. In *RecSys'11*, 157-164.

[4] Tintarev, N. and Masthoff, J. 2011. Designing and Evaluating Explanations for Recommender Systems. In *Recommender Systems Handbook*, 479-510.