PREDICTING CONSUMER INFORMATION SEARCH BENEFITS FOR PERSONALIZED ONLINE PRODUCT RANKING: A CONFIDENCE-BASED APPROACH

Mingyue Zhang, School of Economics and Management, Tsinghua University, Beijing, China, zhangmy.12@sem.tsinghua.edu.cn

Xunhua Guo, School of Economics and Management, Tsinghua University, Beijing, China, guoxh@sem.tsinghua.edu.cn

Guoqing Chen, School of Economics and Management, Tsinghua University, Beijing, China, chengq@sem.tsinghua.edu.cn

Qiang Wei, School of Economics and Management, Tsinghua University, Beijing, China, weiq@sem.tsinghua.edu.cn

Abstract:

Product ranking mechanism is an important service for e-commerce that facilitates consumers' decision-making process. This paper studies online product ranking under uncertainty. Different from previous studies that generally rank products merely based on predicted ratings, a new personalized product ranking method is proposed based on estimating consumer information search benefits and taking prediction uncertainty and confidence into consideration. Experiments using real data of movie ratings illustrate that the proposed method is advantageous over traditional point estimation methods, thus may help enhance customers' satisfaction with the decision-making process and choices through saving their time and efforts.

Keywords: Consumer search benefits, Ranking, Collaborative filtering, Prediction confidence

1 INTRODUCTION

Online shopping was enabled by the expansion of Internet and digital sources, and soon became popular on account of its convenience, low costs and fast processing. However, 'information overload' has been a serious problem due to the rapid growth in volume of products. Because of time and energy constraints, most customers are unable to inspect all available alternatives (Häubl & Trifts 2000). For instance, if we search for '*dress'* in Amazon.com, there are 1,148,021 results, as of January 2014. Even if the search keywords are extended to be more accurate, the amount of information is still beyond the processing capacity of people (462,041 results for '*dress for women*', 68,461 results for '*dress for women party*'). As a result, customers are most likely to end up browsing a small fraction of the products.

There is a rich literature on consumer search behavior which has been framed as a sequential process (McCall 1970; Nelson 1970; Lippman & McCall 1976; Ratchford & Srinivasan 1993; Zwick et al. 2003). Consumers inspect alternatives one at a time and each time he/she may accept the current product, recall a product that has passed over, or continue searching (Weitzman 1979; Adam 2001). Therefore, what matters most is the order that the products are presented to the specific consumer (Ghose et al. 2012). A ranked list of products which matches users' actual need would allow potential customers to make more informed decisions. In the existing literature, online product ranking methods are mostly developed in the perspective of 'Collaborative Filtering' (Adomavicius & Tuzhilin 2005). Most of the collaborative filtering approaches are aimed at rating predictions (Herlocker et al. 2000), and products are ordered with descending predicted ratings.

Nevertheless, it is argued that directly using predicted ratings as ranking scores may not be the best solution (McLaughlin & Herlocker 2004; Deshpande & Karypis 2004; Wang et al. 2008) because there is always uncertainty along with point prediction (Adam 2001). In this paper, taking uncertainty and prediction confidence into consideration, we propose a new personalized product ranking method which proceeds from estimating consumer information search (CIS) benefits. The remaining parts of the paper are organized as follows: Section 2 reviews related research in online product ranking and CIS benefits predicting methods. Section 3 presents the new method for ranking online products based on CIS benefits. Section 4 demonstrates an illustrative example. Section 5 presents and discusses the results from the experiments on real data, and section 6 concludes the paper.

2 LITERATURE REVIEW

2.1 Online Product Ranking

With the abundance and rapid growth of products available online, how to rank products effectively has been an urgent problem and attracts a number of scholars doing valuable and interesting research (Abdul-Muhmin 1999; Montgomery et al. 2004; Diehl & Zauberman 2005; Adomavicius & Tuzhilin 2005; Wang et al. 2008; Tian et al. 2009; Zhang et al. 2010). Related work in this area can be classified into two main streams (Abdul-Muhmin 1999): attribute-based ranking method where the order is only related to the features of products, such as price, quality, popularity and average rating; and personalized ranking method which incorporates the heterogeneity of customers apart from product features.

1) Attribute-based ranking

Widing and Talarzyk (1993) discussed how to best satisfy consumer needs through the design of product information ranking schema and they found that ordering alternatives by a linear weighted average of the product's attributes were superior to other ways. As intelligent decision-support systems are developed, e-retailers are willing to adopt them because they can reduce cognitive efforts and help buyers make wise decisions by ordering and filtering products online (Sproule & Archer 2000), and most of them have applied attribute-based ranking mechanisms (Montgomery et al. 2004). Recently, some other researchers also proposed ranking algorithms based on each considered product feature, using text analysis techniques to extract condensed information from massive customer reviews (Tian et al. 2009; Zhang et al. 2010).

2) Personalized ranking

Recommender systems are aimed at pushing some items that a user might be interested in (Adomavicius & Tuzhilin 2005). They either predict the absolute values of ratings that users would give to the unseen items and then order the products with descending ratings (Herlocker et al. 2000), or directly predict the relative preferences of users and output a ranked list of products (Jin et al. 2003). Wang et al. (2008) further framed the classic recommendation algorithm "Collaborative Filtering" as a *relevance* ranking problem, and integrated the *Probability Ranking Principle* of information retrieval into online product ranking, which resulted in an order according to products' probability of relevance to a user preference. Recently, Ghose et al. (2012) proposed a 'utility-preserving' ranking strategy from an economic perspective which took multi-dimensional preference and customer heterogeneity into consideration. However, although these studies have considered the characteristics of consumers to make ranking results personalized, the uncertainty of prediction has been neglected. To the best of our knowledge, some memory-based *CF* algorithms do not consider the uncertainty in the predicted ratings, which may potentially bias their ranking order, and the expectation criterion does not hold in deciding the order of products to be inspected (Adam 2001). In light of the above discussion, there is a need for a method for ranking online products such that it takes the uncertainty of rating predictions into account. The most closely related research to this ranking problem is the prediction of CIS benefits which combines the expected prediction value and the probability distribution of ratings (Wang et al. 2011).

2.2 Consumer Information Search Benefits

Consumer information search is an important part of purchase decision making (Ratchford & Srinivasan 1993; Hawkins et al. 2013). In an optimal search strategy, consumers only continue to inspect products if the incremental benefits of doing so outweigh the corresponding costs (Kim et al. 2010). Therefore, the critical step is estimating the benefits of CIS, which includes lower prices (Lippman & McCall 1976), higher utility, better quality, and so on. In the literature of consumer sequence search, consumers usually select only one item from a given set of available alternatives (Wright 1975), then the maximum value of utilities of all the inspected products can be used as the benefits of CIS (Diehl 2005; Moe 2006). As it is usually impossible to assess the real utility that consumers have experienced (Samuelson & Nordhaus 2001), consumers' ratings on purchased goods can be used to represent it (Adomavicius & Tuzhilin 2005). Wang et al. (2011) provided the definition of incremental benefits of CIS for heterogeneous consumers and this work lays the foundation of our research in this paper.

Suppose that r_{cs} denotes the rating for product s from consumer c, S_c and $\widetilde{S_c}$ denote the sets of products that have been inspected and not inspected, respectively, then the incremental benefits of inspecting $\widetilde{S_c}$ after inspecting S_c to consumer c can be expressed as (Wang et al. 2011):

$$B\left(c,\widetilde{S}_c \mid S_c\right)=B\left(c,\widetilde{S}_c \cup S_c\right)-B\left(c,S_c\right)=\max_{s\in\widetilde{S}_c\cup S_c}\left(r_{cs}\right)-\max_{s\in S_c}\left(r_{cs}\right) \tag{1}$$

Since the benefits for consumers are generated through inspecting products (Wu & Rangaswamy 2003), it is intuitive to rank products in an order where the expected benefits of inspecting *n* products increase along with the value of *n* at a decreasing rate (Diehl 2005). In this paper, we aim to improve the estimation method for CIS benefits by taking the prediction confidence information into consideration and apply this novel business intelligence algorithm in ranking online products. This ranking mechanism proves to be crucial to the accuracy of recommendation in our experiments.

3 AN OPTIMIZATION METHOD FOR RANKING BASED ON CONSUMER INFORMATION SEARCH BENEFITS

In light of the discussion in literature review, the prediction of incremental benefits after inspecting a number of items can be used as a factor to rank products, and the ideal order for products should be that, when a customer inspects products sequentially, his/her acquired benefits increase but the increasing rate decreases.

3.1 Framework of the Ranking Problem

The classical ranking method ("Collaborative Filtering") in which products are ordered with descending predicted ratings has some limitations. Firstly, they neglect the fact that a single product is just a part of the whole set containing inspected products. The net benefits of consumer information search are generated from all the inspected products, and therefore a newly inspected product cannot be isolated from the previous process and ranked independently. Secondly, the rating distribution on a product among all the consumers is not taken into account, and therefore the uncertainty in the prediction is not addressed. These issues can be conquered if we rank online products based on CIS benefits where the uncertainty of prediction are considered. Assume that we have an effective algorithm to predict the incremental benefits of CIS, then the problem of ranking can be framed as follows.

Suppose that the set of all products is I, and the ones that have been inspected by consumer c form the set S while the remaining ones form the set $\tilde{S}$. The variable y represents the sampled benefits, that is, $y = B(c, S) = \max\{r_{c,s_i} | s_i \in S\}$. Then the state of the system at any moment is described by $(\tilde{S}, y)$. Define $\psi(\tilde{S}, y)$ as the incremental benefits if the consumer continues to inspect the next one product in the current ranking, then for each step when selecting one product into the ordered set, we need to solve the following programming problem:

$$\begin{aligned} &\max \quad \psi\left(\tilde{S}, y\right) = B\left(c, s_i \mid S\right), s_i \in \tilde{S} \\ &s.t. \quad B\left(c, s_j \mid S + \left\{s_i\right\}\right) \le B\left(c, s_i \mid S\right), s_i \in \tilde{S}, s_j \in \tilde{S} - \left\{s_i\right\} \end{aligned} \tag{2}$$

This product ranking method based on CIS benefits also sheds some light on the classical problem in marketing---'when to stop searching'. If the products are ordered using our approach, the incremental benefits decreases while searching sequentially. Therefore, a consumer should stop searching when he/she finds out that the incremental benefits are less than the search costs. In this regard, the critical part of our ranking approach is to predict the incremental benefits of CIS accurately.

3.2 A Confidence-based Method for Estimating CIS Benefits

3.2.1 The Point and Distribution Estimation Method

As mentioned in Part 2.2, consumers' real ratings on a product can be used to represent the benefits after inspecting it. However, if we want to estimate the CIS benefits before the user inspects a product, the real ratings are not available. In most recommender systems, collaborative filtering (*CF*) methods are used to predict the ratings and provide simple single point prediction, which can be denoted by $\widehat{r_{cs}}$, and then Equation (1) can be replaced by:

$$E^{CF}\left(B\left(c, \tilde{S}_c \mid S_c\right)\right) = \max_{s \in \tilde{S}_c \cup S_c}\left(\widehat{r_{cs}}\right) - \max_{s \in S_c}\left(\widehat{r_{cs}}\right) \tag{3}$$

However, there is an apparent limitation if we use this so-called 'point-estimation' method to compute the incremental benefits of CIS. Since there is always uncertainty in a prediction, benefits that are derived only from expected ratings are not accurate. Therefore, the probability distribution of maximum rating $P(r_{cs} < r|s)$ which is widely used in economics (Weitzman 1979; Adam 2001; Zwick et al. 2003), should be considered. Formally, the distribution can be estimated as:

$$E\left(\max_{s \in S}\left(r_{cs}\right)\right) = \int_0^\infty P\left(\max_{s \in S}\left(r_{cs}\right) \ge r\right) dr = \sum_{r=1}^{K}\left[1 - \prod_{s \in S} P\left(r_{cs} < r \mid s\right)\right] \tag{4}$$

where K is the maximum rating in the domain of definition, given that the values of ratings are discrete.

When combining the point estimation (3) and distribution estimation (4), we can calculate the incremental benefits as follows which was proposed in previous research (Wang et al. 2011) and named as *NTP* (New Type of Prediction):

$$E^{NTP}\left(B\left(c,\widetilde{S}_c \mid S_c\right)\right)=\sum_{r=1}^{K}\left[1-\prod_{s\in \widetilde{S}_c\cup S_c} P\left(r_{cs}<r \mid \widehat{r_{cs}}\right)\right]-\sum_{r=1}^{K}\left[1-\prod_{s\in S_c} P\left(r_{cs}<r \mid \widehat{r_{cs}}\right)\right] \tag{5}$$

Although personalization has been considered compared to distribution estimation (Jones & Mendelson 2011), *NTP* still has some drawbacks. Using predicted rating $\widehat{r_{cs}}$ on behalf of $E(r_{cs}|z_c, z_s)$, where z_c, z_s are vectors representing the relevant information of consumer c and product s, the implicit assumption is that, when two distinct products have the same predicted rating for a customer, their real ratings have the same probability distribution. That is to say, the above method takes the heterogeneity of customers into account, but not that of products. The difference of confidence between two predictions has been neglected.

Many studies in collaborative filtering argue that the predictions of ratings cannot be treated equally because there is a confidence or reliability issue (Mazurowski 2013; Hernando et al. 2013), which is ignored in the above method. Confidence refers to the system's trust in its recommendations or predictions (Herlocker et al. 2000). The collaborative filtering algorithms, such as user-based *KNN*, predict ratings for unknown items based on previous ratings that have been assigned to this item by other similar users, who are called 'neighbors' (Mazurowski 2013). Intuitively, ratings from these 'neighbors' will affect the quality of prediction---the larger amounts and less disagreement among neighbors, the higher confidence is for the prediction. Therefore, consumers' search benefits will be different when the predicted ratings are the same but the prediction confidences are different. Therefore, we attempt to propose a new method to estimate consumers' search benefits which integrates the prediction confidence, in the hope that the results will be more accurate.

3.2.2 The Confidence-based Method

In a comprehensive review, Mazurowski (2013) summarized six different confidence estimation algorithms. Here we select and improve one of them (*Algorithm 3*: *Variability for Item*) to act as confidence estimation method because it is the best choice when both considering the performance and efficiency. Suppose that the set of consumers who have rated item s is C_s, where $C_s = \{c_1, c_2, c_3, \cdots c_n\}$. For a new consumer $c_i (c_i \notin C_s)$, if memory-based *CF* algorithms are used to make recommendations for him/her, then his/her 'neighbors' who have already rated item s become an important factor that will affect the confidence of rating prediction $r_{c_i s}$.

On one hand, 'confidence' is a continuous variable which ranges in the closed interval [0,1]. It is impossible to enumerate every value of the confidence level and its corresponding probability distribution $P(r_{cs} < r|\widehat{r_{cs}}, conf)$, and consequently some kinds of transformation are needed. As is analyzed above, the real rating distribution from neighbor consumers for the same item can be used as a simple measure of confidence for predictions in a recommender system. On the other hand, these 'neighbor' consumers are part of the whole consumer set C_s, and it might be controversial for the definition of 'neighbors' because there is disagreement for the lower bound of user similarity to be recognized as a 'neighbor'. Therefore, we can assume that all the customers who have the same predicted rating with target user c_i for this item are his/her neighbors. That is to say, the distribution of real ratings for the item being recommended can be used as a measure of confidence for predictions (McNee et al. 2003; McLaughlin & Herlocker 2004). In this way, the probability distribution under infinite confidence levels $P(r_{cs} < r|\widehat{r_{cs}}, conf), conf \in [0,1]$ has been transformed into distribution under finite different items $P(r_{cs} = i|\widehat{r_{cs}}, s), s \in S$. Obviously, there still exist other algorithms to estimate the confidence level. For example, the above method can be enriched by considering dynamic user ratings, which can be studied in future work.

From historical data, the probability of predicted rating $\widehat{r_{cs}}$ can be derived, when r_{cs} is given. Then according to Bayesian rule, the posterior distribution of r_{cs} can be calculated as:

$$P\left(r_{cs}=i \mid \widehat{r_{cs}}=j,s\right)=\frac{P\left(\widehat{r_{cs}}=j \mid r_{cs}=i\right)\cdot P\left(r_{cs}=i \mid s\right)}{\sum_{i=1}^{K} P\left(\widehat{r_{cs}}=j \mid r_{cs}=i\right)\cdot P\left(r_{cs}=i \mid s\right)} \tag{6}$$

where $P(\widehat{r_{cs}} = j | r_{cs} = i)$ is generated from the whole product set S, and the outcome $P(r_{cs} = i | \widehat{r_{cs}} = j, s)$ is specific to target product s, incorporating the confidence of prediction.

If we combine Equations (4) and (6), a personalized distribution can be calculated as follows (The new method is defined as 'Point Estimation Incorporating Uncertainty and Prediction Confidence' (*PUPC*)):

$$E^{PUPC}\left(\max_{s\in S}(r_{cs})\right) = \sum_{r=1}^{K}\left[1 - \prod_{s\in S} P\left(r_{cs} < r \mid \widehat{r_{cs}}, s\right)\right] \tag{7}$$

When using the posterior distribution to re-predict incremental benefits, the result will be:

$$E^{PUPC}\left(B\left(c, \tilde{S}_c \mid S_c\right)\right) = \sum_{r=1}^{K}\left[1 - \prod_{s\in \tilde{S}_c \cup S_c} P\left(r_{cs} < r \mid \widehat{r_{cs}}, s\right)\right] - \sum_{r=1}^{K}\left[1 - \prod_{s\in S_c} P\left(r_{cs} < r \mid \widehat{r_{cs}}, s\right)\right] \tag{8}$$

This means that the benefits after inspecting a set of items are not only related to the personalized predictions $\widehat{r_{cs}}$, but also the items. In other words, given a consumer c and two different products s_1 and s_2, the benefits of inspecting s_1 and s_2 are different even though they have the same predicted rating. This is because they are not identical products and the distributions of real ratings are different.

3.3 The New Ranking Approach based on CIS Benefits

This confidence-based method for estimating CIS benefits can be applied into the framework of ranking problem which was described in Part 3.1. Afterwards, our ranking approach will make up the drawbacks of traditional ranking method (*CF*) to some extent, that is, the uncertainty of *CF* predictions for ratings can be released by incorporating confidence information. The pseudo-code of the new ranking approach is shown in Figure 1.

RankByPUPC

Input: Data Object Set $D = \left\{\left(c_1, s_1, r_{c_1 s_1}\right), \left(c_1, s_2, r_{c_1 s_2}\right), \cdots, \left(c_i, s_j, r_{c_i s_j}\right), \cdots\right\}$

Output: Rank Lists for All Customers $Rank = \left\{R_{c_1}, R_{c_2}, \cdots, R_{c_i}, \cdots\right\}$

where R_{c_i} is the rank result for customer c_i.

// Select a CF method to predict r_{cs}
for (each consumer-product pair (c, s) where r_{cs} is known in the data object set D) {
  use *CF* to generate $\widehat{r_{cs}}$ as the prediction of r_{cs};
}
// Estimate the probability distribution in our approach
for ($k = 1,2, \cdots, K$) {
  for ($s = s_1, s_2, \cdots s_j, \cdots$) {
    for ($r = 1,2, \cdots, K$) {
      calculate the probability distribution of $P(r_{cs} < r | \widehat{r_{cs}} = k, s)$; } }
}
// Rank the products for each customer based on CIS benefits
for (each customer $c = c_1, c_2, \cdots, c_i, \cdots$) {
  do {
    // Predict incremental benefits of CIS when selecting one product into the order
    for (each unsorted product of customer c, $s = s_1, s_2, \cdots s_j, \cdots$) {
      get the set of products that already ranked S_c;
      calculate CIS incremental benefits $E^{PUPC}\left(B(c, s | S_c)\right)$ with Equation 8;
    }
    // get the maximum benefits
    add the product with maximum incremental benefits into the rank list R_c;
  } while (the set of unsorted products is not empty)
}
output rank results for all customers $Rank = \left\{R_{c_1}, R_{c_2}, \cdots, R_{c_i}, \cdots\right\}$

Figure 1. Pseudo-code of RankByPUPC

4 ILLUSTRATIVE EXAMPLE

Suppose that there are 3 different products (s_1, s_2, s_3) and 5 customers (c_1, c_2, c_3, c_4, c_5) in the system. For a target customer c_1, we want to demonstrate a personalized ranking list of the 3 items. A qualified *CF* is used to predict the ratings while the real ones can only be obtained after the item is purchased. Table 1 shows the predicted and real ratings of 3 items to 5 customers.

Items	c_1	c_2	c_3	c_4	c_5
s_1 Predicted ratings	5	5	5	5	4
s_1 Real ratings	4	5	4	5	5
s_2 Predicted ratings	4	1	2	3	1
s_2 Real ratings	5	1	2	3	2
s_3 Predicted ratings	4	4	3	5	3
s_3 Real ratings	4	4	4	5	2

Table 1. Historical Rating Data

In order to generate a ranking list for a specific customer, the first step is to estimate his or her incremental benefits of CIS accurately. From these historical data shown above, we can calculate the probability distribution needed in each CIS benefits estimation method (*CF*, *NTP*, *PUPC*) and then give their estimated results.

4.1 Incremental Benefits Predicted by Different Methods

Consider the scenario that after consumer c_1 has inspected product s_1, he/she wants to continue inspecting s_2 and s_3. Now let's predict his/her incremental benefits using 3 different methods: *CF*, *NTP* (Wang et al. 2011), and *PUPC* (our method).

1) The actual value of incremental benefits for c_1 is: 5-4=1
2) *CF*: $E^{CF}\left(B\left(c, s_2, s_3 \mid s_1\right)\right) = \max_{s \in s_1 \cup s_2 \cup s_3}\left(\widehat{r_{cs}}\right) - \max_{s \in s_1}\left(\widehat{r_{cs}}\right) = 5 - 5 = 0$
3) *NTP*: After taking uncertainty into account, the probability distribution $P(r_{cs} < r|\widehat{r_{cs}})$ is obtained from historical data, as shown in Table 2.

r	1	2	3	4	5
$\widehat{r_{cs}}$ =1	0	0.5	1	1	1
$\widehat{r_{cs}}$ =2	0	0	1	1	1
$\widehat{r_{cs}}$ =3	0	0	1/3	2/3	1
$\widehat{r_{cs}}$ =4	0	0	0	0	0.5
$\widehat{r_{cs}}$ =5	0	0	0	0	0.4

Table 2. $P(r_{cs} < r|\widehat{r_{cs}})$ Used in NTP

Then the incremental benefits predicted by *NTP* are: (Equation 5)

$$E^{NTP}\left(B\left(c, s_2, s_3 \mid s_1\right)\right) = \left(5 - 0.4 \times 0.5 \times 0.5\right) - \left(5 - 0.4\right) = 0.3$$

4) *PUPC*: After incorporating the confidence information, the probability distribution $P(r_{cs} < r|\widehat{r_{cs}}, s)$ in our method can be calculated from historical data and a part of it is presented in Table3:

r	1	2	3	4	5
$s_1, \widehat{r_{c_1 s_1}} = 5, Conf_1$	0	0	0	0	0.5
$s_2, \widehat{r_{c_1 s_2}} = 4, Conf_2$	0	0	0	0	0
$s_3, \widehat{r_{c_1 s_3}} = 4, Conf_3$	0	0	0	0	1

($Conf_1, Conf_2, Conf_3$ denote the different confidence levels of prediction.)

Table 3. $P(r_{cs} < r|\widehat{r_{cs}}, s)$ *Used in PUPC*

It can be seen intuitively that, the confidence level of prediction for item $s_2, \widehat{r_{c_1 s_2}} = 4$ is lower than that for item $s_3, \widehat{r_{c_1 s_3}} = 4$ even though they have the same value of predicted rating. The incremental benefits predicted by *PUPC* are: $E^{PUPC}\left(B\left(c, s_2, s_3 \mid s_1\right)\right) = 5 - \left(5 - 0.5\right) = 0.5$

The comparison between the three methods is illustrated in Table 4. It is clear that the prediction of incremental benefits by *PUPC* is the closest to the true value.

	$\max\left(r_{cs} \mid s \in \{s_1, s_2, s_3\}\right)$	$\max\left(r_{cs} \mid s \in \{s_1\}\right)$	$B\left(c, s_2, s_3 \mid s_1\right)$
True value	5	4	1
CF	5	5	0
NTP	4.9	4.6	0.3
PUPC	5	4.5	0.5

Table 4. Incremental Benefits Predicted by CF, NTP and PUPC

4.2 Ranking Results

For the specific customer c_1, it is obvious that his/her optimal ranking should be $s_2 \to s_1 \to s_3$ or $s_2 \to s_3 \to s_1$ according to the real ratings in Table 1. After applying the ranking algorithm (‘*RankByPUPC’*), we can get the ranking result of *PUPC* method.

1) *CF*: The products are ranked with descending predicted ratings, and the result would be:

$$s_1 \to s_2 \to s_3 \quad \text{Or} \quad s_1 \to s_3 \to s_2$$

2) *NTP*: To select the first product in the order, we need to calculate their inspected benefits respectively:

$$B^{NTP}\left(s_1\right) = 4.6\,, \quad B^{NTP}\left(s_2\right) = 4.5\,, \quad B^{NTP}\left(s_3\right) = 4.5$$

Therefore, s_1 will be in the first position. Applying Equation (2), ranking result would be:

$$s_1 \to s_2 \to s_3 \quad \text{Or} \quad s_1 \to s_3 \to s_2$$

It is worth noting that if we use incremental benefits predicted by *NTP* to rank the products, it will always have the same result with *CF* because the probability $P(r_{cs} < r|\widehat{r_{cs}})$ is non-incremental with the increase of $\widehat{r_{cs}}$. That is why this approach is excluded in the experiment part.

3) *PUPC*: The three products’ inspected benefits predicted by our method are shown as follows:

$$B^{PUPC}\left(s_1\right) = 4.5\,, \quad B^{PUPC}\left(s_2\right) = 5\,, \quad B^{PUPC}\left(s_3\right) = 4$$

Thus, s_2 will be in the first position. According to algorithm *RankByPUPC*, the result would be:

$$s_2 \to s_1 \to s_3 \quad \text{Or} \quad s_2 \to s_3 \to s_1$$

It is clear that the result generated by *PUPC* is consistent with the real optimal ranking.

5 EXPERIMENTS ON REAL DATA

Among different kinds of methods for estimating benefits and ranking products, the experiments in this paper are mainly conducted on two types of methods as mentioned above: *CF* and *PUPC*. This is because 'Collaborative Filtering' is the most frequently used point estimation method by online retailers. Here we select one of the most classical *CF* algorithms as an instance: user-based K-Nearest-Neighbor (*KNN*) method in which the number of nearest neighbors is set to 20 and the similarity between users are calculated with *Pearson* correlation. The results are compared between *CF* and its corresponding *PUPC* method.

The experiments consist of two parts. Firstly, in the search scenario, we use the two types of methods to estimate the consumer search benefits on the well-known dataset of *MovieLens* and then compare their performances. The measure used for comparison is 'Root Mean Squared Error' (*RMSE)*. Secondly, in the recommendation scenario, we compare the ranking results of the two different methods. A benchmark ranking can also be acquired with real ratings given by users.

5.1 Settings

1) Data:

We select the data of *MovieLens* which is offered by Grouplens project (*http://grouplens.org/datasets/*) to compare performances on estimating benefits and ranking products. Considering the fact that our method is sensitive to dataset sparsity (e.g. the computation of probability), in data pre-processing we deleted the users whose number of co-rated items with every neighbor is less than 20. The dataset was then randomly divided into two subsets: 2/3 are used as training set and 1/3 are used as test set.

2) Metrics to compare performances:

In the first part, because the purpose of the experiments is to predict search benefits, we use the classical measure Root Mean Squared Error *(RMSE)* to compare different methods, which is denoted as follows:

$$RMSE = \sqrt{\frac{\sum_{c \in C}\left[E\left(\hat{B}\left(c, \tilde{S} \mid S\right)\right) - B\left(c, \tilde{S} \mid S\right)\right]^2}{|C|}}$$

In the second part, various metrics are used to evaluate the ranking performances in order to cross-validate the results. Firstly, if we hypothesize that customers will inspect all the items which are recommended (top-N products in the ranking list), then traditional metrics which are commonly used in information retrieval can be applied, because we only care about the fraction of products in the recommendation set that will be purchased by the customer. Here products with the highest real rating by a customer are considered to be purchased/relevant. As a result, *precision*, *recall*, and *F-score* are calculated. Secondly, for systems that return ranked items, it is desirable to also consider their order. Two metrics are used to validate the performance results: Mean Average Precision (*MAP*) and Normalized Discounted Cumulative Gain (*nDCG*) (Järvelin 2002; Abbassi et al. 2009).

5.2 Results

5.2.1 Search Scenario

In the search scenario, we need to estimate CIS benefits and there are 3 parameters: the order of the products (*Order*), length of the products set (*N*), and percentage of already inspected items (*Per*), where $Order \in$ {random order, decreasing order by predicted ratings}, $N \in \{5,10,20,30,40\}$ and $Per \in \{0.1,0.2,0.3,0.4,0.5,0.6,0.7,0.8,0.9\}$. This problem can be described as: there are N products $(s_1, s_2, \cdots s_N)$ in the set where products are presented by a specific *Order*, after a customer c $(c \in C)$ has inspected $N \times Per$ products sequentially, we need to estimate the incremental benefits if he/she continues to inspect the remaining. Since there are $2 \times 5 \times 9 = 90$ cases (2 values of *Order*, 5 values of *N*, 9 values of *Per*), we compare their average performances with varying *N* and *Per*.

1) The average *RMSE* of algorithm *PUPC* and *CF* in cases with varying *N* ($N \in \{5,10,20,30,40\}$) are calculated respectively and demonstrated in Figure 2.

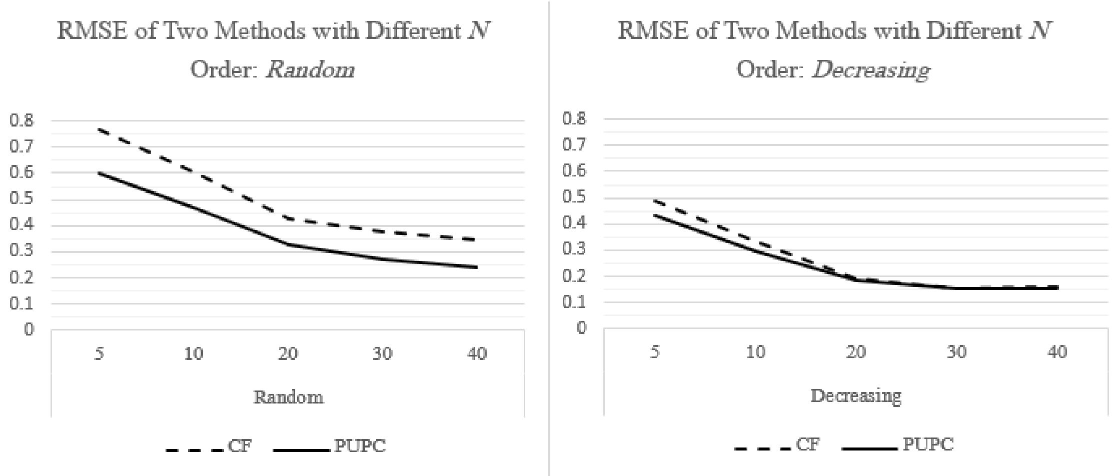

Figure 2. *Average RMSE of Two Methods with Different N and Order*

It can be seen that *PUPC* method outperforms *CF* in all above cases since smaller *RMSE* means better estimation. In order to better understand the performances of the two methods, we calculate the average improvements of *PUPC* over *CF* with different *N* which are illustrated in Table 5.

Order	N	IMPR(*PUPC*,*CF*)	Order	N	IMPR(*PUPC*,*CF*)
Random	5	21.81%	Decreasing	5	11.79%
	10	22.76%		10	9.59%
	20	23.87%		20	4.64%
	30	28.51%		30	1.25%
	40	30.27%		40	1.90%

Table 5. *Average Improvements of PUPC over CF with Different N and Order*

2) Similar to the above, the average *RMSE* of algorithm *PUPC* and *CF* in cases with varying *Per* ($Per \in \{0.1,0.2,0.3,0.4,0.5,0.6,0.7,0.8,0.9\}$) are also calculated and demonstrated in Figure 3.

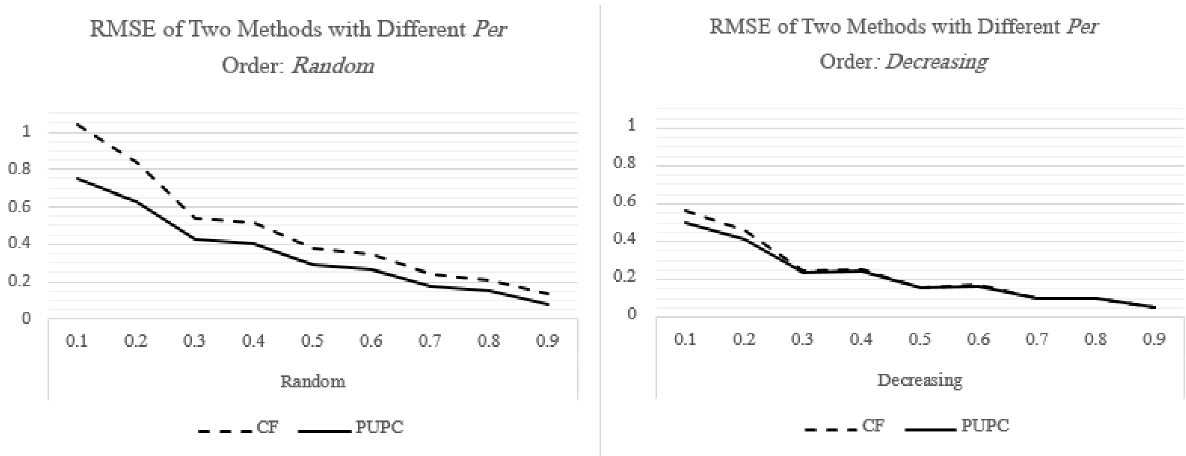

Figure 3. *Average RMSE of Two Methods with Different Per and Order*

Similarly, we calculate the average improvements of *PUPC* over *CF* with different *Per* one step further and illustrate the results numerically in Table 6.

Order	Per	IMPR($PUPC$,CF)	Order	Per	IMPR($PUPC$,CF)
Random	0.1	27.57%	Decreasing	0.1	11.62%
	0.2	25.21%		0.2	11.04%
	0.3	21.87%		0.3	2.58%
	0.4	21.61%		0.4	5.05%
	0.5	23.29%		0.5	0.93%
	0.6	22.29%		0.6	2.19%
	0.7	27.77%		0.7	1.01%
	0.8	27.59%		0.8	0.83%
	0.9	39.35%		0.9	0.11%

Table 6. Average Improvements of PUPC over CF with Different Per and Order

In conclusion, when estimating the CIS incremental benefits in a search scenario, the performances of *PUPC* are better than *CF* in all kinds of cases. But it is noteworthy that the improvements of *PUPC* over *CF* are varying with different parameters. That is to say, the improvements change with *Order*, *N* and *Per*.

5.2.2 Recommendation Scenario

In the recommendation scenario, a ranking list can be generated for each user by two different types of methods. There is one parameter when comparing the accuracy of recommendation without considering the product order: size of recommendation set (N). Here we set $N \in \{5,6,7,8,9,10,11,12,13,14,15,16,17,18,19,20\}$. After calculating the *precision*, *recall* and *F-score* in each case, we can get the PR curves (Precision-Recall Curves) to compare the performances between different methods. The curves are shown in Figure 4 and we can see that *PUPC* outperforms *CF* since greater values of *precision* and *recall* mean better recommendation.

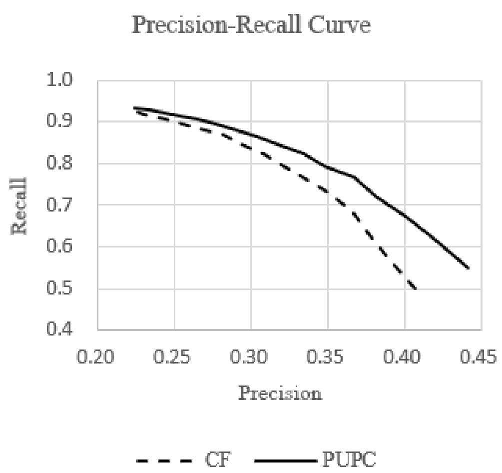

Figure 4. Precision-Recall Curves under Various Recommendation Sets

On the other hand, as we mentioned earlier, customers usually cannot inspect all the items in the recommendation set because of limited time and energy. Therefore, the order in which products are presented is also essential when items are inspected sequentially. Table 7 illustrates the ranking performances of the two types of algorithms by different metrics. It is clear that *PUPC* outperforms traditional *CF* method because of incorporating the uncertainty and prediction confidence. Additionally, we have conducted paired *t*-tests on the above data experiments, and the statistical results show that, both the incremental benefits estimation results and the ranking results of *PUPC* method are significantly better than those of *CF* method.

Algorithms	*MAP*	*nDCG*
CF	0.536779859071	0.669146006168
$PUPC$	0.569188209941	0.680894623861
IMPR($PUPC$, CF)	6.04%	1.76%

Table 7. Ranking Performances between Different Methods

6 CONCLUSION AND FUTURE WORK

There has been extensive research about how to generate intelligent ranking order of all available product items to help buyers make preferred decisions and support sellers to achieve higher profits. In the literature of recommender systems, most of the recommendations are listed based on predicted ratings from the highest to lowest. While uncertainty exists along with prediction, confidence information should be considered and merely using predicted ratings as criteria may not be the best ranking. This paper has proposed a new product ranking method based on consumer information search benefits which takes uncertainty and prediction confidence into consideration. In this way, the order of products is more closely relate to the real optimal ranking, which can save the customers' time and efforts, and consequently enhance their satisfaction with the decision-making process and choices.

It is worth mentioning that there are also other algorithms to estimate prediction confidence. In this paper we only choose the real rating distribution as a representation. Therefore, incorporating other confidence estimation algorithms could be a future direction of this research. Another interesting perspective is that our ranking approach could also be used for evaluating the performances of confidence estimation algorithms. Besides, future studies could also use other collaborative filtering methods and more real datasets to obtain more consolidated understanding of the proposed approach.

Acknowledgments

This work was partly supported by the National Natural Science Foundation of China (grant number 71110107027) and the MOE Project of Key Research Institute of Humanities and Social Sciences at Universities of China (grant number 12JJD630001).

References

Abbassi, Z., Amer-Yahia, S., Lakshmanan, L.V., et al. (2009). Getting recommender systems to think outside the box. In Proceedings of the third ACM conference on Recommender systems. ACM, 285-288.

Abdul-Muhmin, A.G. (1999). Contingent decision behavior: effect of number of alternatives to be selected on consumers' decision processes. Journal of Consumer Psychology, 8 (1), 91-111.

Adam, K. (2001). Learning while searching for the best alternative. Journal of Economic Theory, 101 (1), 252-280.

Adomavicius, G. and Tuzhilin, A. (2005). Toward the next generation of recommender systems: A survey of the state-of-the-art and possible extensions. Knowledge and Data Engineering, IEEE Transactions on, 17 (6), 734-749.

Browne, G.J., Pitts, M.G. and Wetherbe, J.C. (2007). Cognitive stopping rules for terminating information search in online tasks. MIS Quarterly, 31 (1), 89-104.

Chun, Y.H. (2000). Sequential Search and Selection Problem under Uncertainty. Decision Sciences, 31 (3), 627-648.

Deshpande, M. and Karypis, G. (2004). Item-based top-n recommendation algorithms. ACM Transactions on Information Systems, 22 (1), 143-177.

Diehl, K. (2005). When Two Rights Make a Wrong: Searching Too Much in Ordered Environments. Journal of Marketing Research, 42 (3), 313-322.

Diehl, K., and Zauberman, G. (2005). Searching Ordered Sets: Evaluations from Sequences under Search. Journal of Consumer Research, 31 (4), 824-832.

Ghose, A., Ipeirotis, P.G. and Li, B. (2012). Designing ranking systems for hotels on travel search engines by mining user-generated and crowdsourced content. Marketing Science, 31 (3), 493-520.
Häubl, G., and Trifts, V. (2000). Consumer decision making in online shopping environments: The effects of interactive decision aids. Marketing science, 19 (1), 4-21.
Hauser, J.R. and Wernerfelt, B. (1990). An Evaluation Cost Model of Consideration Sets. Journal of Consumer Research, 16 (4), 393-408.
Hawkins, D.I., Mothersbaugh, D.L. and Best, R.J. (2013). Consumer Behavior: Building Marketing Strategy. McGraw-Hill Irwin.
Herlocker, J.L., Konstan, J.A. and Riedl, J. (2000). Explaining collaborative filtering recommendations. In Proceedings of the 2000 ACM conference on Computer supported cooperative work. ACM, 241-250.
Hernando, A., Bobadilla, J., Ortega, F. and Tejedor, J. (2013). Incorporating reliability measurements into the predictions of a recommender system. Information Sciences, 218, 1-16.
Järvelin, K. and Kekäläinen, J. (2002). Cumulated gain-based evaluation of IR techniques. ACM Transactions on Information Systems, 20 (4), 422-446.
Jin, R., Si, L., Zhai, C. and Callan, J. (2003). Collaborative filtering with decoupled models for preferences and ratings. In Proceedings of the twelfth international conference on Information and knowledge management, ACM, 309-316.
Jones, R. and Mendelson, H. (2011). Information Goods vs. Industrial Goods: Cost Structure and Competition. Management Science, 57 (1), 164-176.
Kim, J.B., Albuquerque, P. and Bronnenberg, B.J. (2010). Online Demand under Limited Consumer Search. Marketing Science, 29 (6), 1001-1023.
Kim, S.M., Pantel, P., Chklovski, T. and Pennacchiotti, M. (2006). Automatically assessing review helpfulness. In Proceedings of the 2006 Conference on Empirical Methods in Natural Language Processing, Association for Computational Linguistics, 423-430.
Lippman, S.A. and McCall, J. (1976). The economics of job search: A survey. Economic inquiry, 14 (2), 155-189.
Mazurowski, M.A. (2013). Estimating confidence of individual rating predictions in collaborative filtering recommender systems. Expert Systems with Applications, 40 (10), 3847-3857.
McCall, J. (1970). Economics of information and job search. Quarterly Journal of Economics, 84 (1), 113-126.
McLaughlin, M.R. and Herlocker, J.L. (2004). A collaborative filtering algorithm and evaluation metric that accurately model the user experience. In Proceedings of the 27th annual international ACM SIGIR conference on Research and development in information retrieval. ACM, 329-336.
McNee, S.M., Lam, S.K., Guetzlaff, C., et al. (2003). Confidence displays and training in recommender systems. In Proceedings of INTERACT'03 IFIP TC13 International Conference on Human-Computer Interaction, 176-183.
Moe, W.W. (2006). An empirical two-stage choice model with varying decision rules applied to internet clickstream data. Journal of Marketing Research, 43 (4), 680-692.
Montgomery, A.L., Hosanagar, K., Krishnan, R. and Clay, K.B. (2004). Designing a better shopbot. Management Science, 50 (2), 189-206.
Nelson, P. (1970). Information and Consumer Behavior. The Journal of Political Economy, 78 (2), 311-329.
Ratchford, B.T. (1982). Cost-benefit models for explaining consumer choice and information seeking behavior. Management Science, 28 (2), 197-212.
Ratchford, B.T and Srinivasan, N. (1993). An Empirical Investigation of Returns to Search. Marketing Science, 12 (1), 73-87.
Samuelson, P.A and Nordhaus, W.D. (2001). Economics (Mcgraw-Hill International Editions). Samuelson, P.A., editor. Boston, MA: McGraw-Hill.
Shani, G. and Gunawardana, A. (2011). Evaluating recommendation systems. Recommender systems handbook. Springer US, 257-297.
Sproule, S. and Archer, N. (2000). A buyer behaviour framework for the development and design of software agents in e-commerce. Internet Research, 10 (5), 396-405.
Stigler, G.J. (1961). The economics of information. The journal of political economy, 69 (3), 213-225.

Tian, P., Liu, Y., Liu, M. and Zhu, S. (2009). Research of product ranking technology based on opinion mining, In Intelligent Computation Technology and Automation, 2009. ICICTA'09. Second International Conference on. IEEE, 4, 239-243.
Wang, H., Guo, X., Wei, Q. and Chen, G. (2011). Providing a Service for Interactive Online Decision Aids through Estimating Consumers' Incremental Search Benefits. ICIS Proceedings, Paper 10.
Wang, J., Robertson, S., de Vries, A.P. and Reinders, M.J. (2008). Probabilistic relevance ranking for collaborative filtering. Information Retrieval, 11 (6), 477-497.
Wang, J. and Zhu, J. (2009). Portfolio theory of information retrieval. In Proceedings of the 32nd international ACM SIGIR conference on Research and development in information retrieval. ACM, 115-122.
Weitzman, M.L. (1979). Optimal search for the best alternative. Econometrica: Journal of the Econometric Society, 641-654.
Widing, R.E. and Talarzyk, W.W. (1993). Electronic information systems for consumers: an evaluation of computer-assisted formats in multiple decision environments. Journal of Marketing Research, 125-141.
Wright, P. (1975). Consumer choice strategies: Simplifying vs. optimizing. Journal of Marketing Research, 12 (1), 60-67.
Wu, J. and Rangaswamy, A. (2003). A fuzzy set model of search and consideration with an application to an online market. Marketing Science, 22 (3), 411-434.
Zhang, K., Narayanan, R. and Choudhary, A. (2010). Voice of the customers: mining online customer reviews for product feature-based ranking. In Proceedings of the 3rd conference on Online social networks, USENIX Association, 11-11
Zhang, R. and Tran, T. (2011). An information gain-based approach for recommending useful product reviews. Knowledge and Information Systems, 26 (3), 419-434.
Zwick, R., Rapoport, A., Lo, A.K.C and Muthukrishnan, A.V. (2003). Consumer sequential search: not enough or too much? Marketing Science, 22 (4), 503-519.