An Approach to Online Recommendation of Products with High Price-Performance Ratios Based on a Customized Price-Dominance Relationship

Hongfa Wang
Zhejiang University of Water Resources and Electric Power, Hangzhou, China
Email: hongfaw@163.com

Chen Xing
Zhejiang University of Water Resources and Electric Power, Hangzhou, China
Email: xingchen_82@hotmail.com

Abstract—Automatic and customized recommendation of products with high price-performance ratios is an important task in online shopping and booking in e-commerce. However, existing studies and techniques have not provided solutions to this problem. For example, existing recommendation techniques mainly recommend products which tend to be interesting for customers; existing skyline search techniques aim to find the best items (best in terms of each attribute) from candidates disregarding customized criteria. This paper introduces a customized price-dominance relationship for determining which product has a higher price-performance ratio in two given candidates, and proposes customized Skyline sets. Based on customized Skyline search, an approach is proposed for recommending products with the highest price-performance ratio to customers given their personalized parameters. The proposed approach is applied to hotel recommendations online. Experimental results demonstrate that the approach is capable of recommending hotels with the highest price-performance ratio to customers according to their measurement criteria. The proposed recommendation approach can help customers find cheap products of high qualities efficiently and effectively.

Index Terms—electronic commerce, customized recommendation, price-performance ratio, price dominance relationship

I. INTRODUCTION

With the development and applications of the techniques of internet networking and mobile computing, in particular electronic commerce (e-commerce), fast and convenient services of products online search and ordering have been offered on more and more e-commerce websites. There have been a variety of professional e-commerce websites available, such as shopping websites like Amazon, eBay, Taobao (China), tickets/hotels booking websites like expedia.com, booking.com and ctrip.com (China), and other e-commerce websites for particular categories of products or companies’ official e-commerce websites. Now more and more people choose to do shopping and/or booking online benefiting from advantages of e-commerce, such as fast services without restrictions on time and geographical locations, broad choices of products, and rich information. The statistics of 2009 annual online shopping issued by CNNIC (China Internet Network Information Center) showed that, the population of online-shopping customers reached 87.88 millions, and the total expense by Chinese online shoppers in the first half of the year 2009 reached 119.52 billions Chinese dollars, 2011 annual report from CNNIC showed that the number of online shoppers in 2011 increased by 48.6% than that in 2010. This indicates one of the fastest applications in terms of increase in customer numbers.

E-commerce has resulted in a revolution in the form of personal shopping and booking. However, similar to the traditional retailing, customers shopping/booking are conducted in two key steps.

- Step 1 (Candidates Selection) A customer choose candidates, denoted as $D_c$, from a set of products available, denoted as $D$, by filtering the products by his/her basic requirements/constraints on products attributes and prices according to his/her needs, preferences and budget. For online shopping, for example, when a customer buys a digital camera, his/her constrains include resolution>10mp, color="black" and $500<price<$800. In online booking, for example, a customer’s basic constrains for a hotel booking include check in on May 6, 2013, 1-3 star, single room, distance to city center<10km, price<$200.
- Step 2 (Candidates Comparison & Final Selection) In this step, the customer compares price-performance ratios among the candidates in $D_c$, and finally selects a product, $P$, with the highest
price-performance ratio normally (assuming the preferences have been specified in Step 1, i.e., the customers have no preferences for the candidates except price-performance ratios).

Technically speaking, Step 1 can be implemented by basic database query techniques. Consequently, such functions (product search/filtering and category navigation) have been available in most e-commerce websites. Customers can find candidates easily using these functions. In Step 2, a customer may manually compare price-performance ratios among candidates if \( Dc \) contains a small number of candidates, say several or a dozen. However, in real-life online shopping/booking, \( Dc \) may contain a large number of candidates, say dozens, hundreds, or even thousands, after filtering by customers’ constraints, as products could be from all over the world, rather than a couple of brick shops in the traditional retail situation. It may be unrealistic for the customer to compare all the candidates in \( Dc \), and find the one with the highest price-performance ratio. Therefore, it is demanding to provide such techniques and functions which are able to automatically recommend products with the highest price-performance ratios from the set of candidates to the customer in online shopping/booking.

Unfortunately, to the best of our knowledge, little work has been done on this topic, and few e-commerce websites have offered such functions. Motivated by the demand and potential significance of this technique, we investigate this problem and aim to propose an efficient and effective approach for online recommendation of products with the highest price-performance ratios for different customers/customers according to personalized criteria and parameters given by them.

Automatic recommender has been extensively studied in the data mining community in the last decade [1-8]. Researchers have proposed a number of auto-recommender algorithms, and some algorithms have been successfully used in commercialized e-commerce websites. For example, when a user clicks on a book/paper in a website, relative books/papers may be recommended to the user with a prompt like “customers access this book/paper also click these books/papers ….”. Existing recommendation techniques [3-8] mainly focused on recommending products that are potentially interesting for customers who are accessing the website, rather than recommending products with the highest price-performance ratio.

Another technique related to this recommendation problem considered in this paper is Skyline search [9-15]. This technique is aimed to find an optimal subset, \( S \) (called a skyline set), from a candidate set, \( Dc \), of objects, which belongs to the same category, e.g., digital cameras, and single rooms in hotels at a particular city. A subset \( S \) is optimal if any object \( s \) in \( S \) satisfies: there exists no objects \( s’ \) in \( Dc \) such that \( s’ \) dominates \( s \), i.e., \( s’ \) has no attributes whose values are worse than the corresponding ones in \( s \), and \( s’ \) has at least one attribute whose value is better than that in \( s \). Whereas, Skyline search techniques are not suitable for finding products with the highest price-performance ratios as the techniques have not considered personal and real-time requirements on price-performance evaluations and calculations (details about the reasons have been addressed in Section II B).

This paper aims to propose an approach for online recommendation of products with the highest price-performance ratios. Firstly, a customized price-dominance relationship is introduced to compare the price-performance between two products, and customized Skyline sets are defined. Secondly, an approach was proposed for online recommendation of products with high price-performance ratios based on customized skyline search. A customized skyline set extracted from \( Dc \) is a recommendation set \( Dr \).

The remainder of the paper is organized as follows. Section II reviews and explores state of the art of two related techniques, auto recommendation and skyline search. Section III introduces a price dominance relationship. In Section IV, a customized skyline search algorithm is proposed for finding products with the highest price-performance ratios for different customers. Section V presents a real case study, the application of the proposed approach to customized recommendations for online booking of rooms with hotels. Section VI concludes the paper.

II. STATE OF THE ART OF RELATED TECHNIQUES

This section reviews two major closely related techniques, auto recommendation and Skyline search, and explores current status of applications. In exploring state of the art of the two techniques and current applications, we identify their relevance to the recommendation task considered in this paper, and explore why they are not applicable for this task.

A. Auto Recommendation Techniques

Auto recommendation has been extensively studied since the 1990s. Existing auto recommendation techniques include three major categories. The first category is collaborative filtering [5, 6], which recommends products to target customers based on the comments on the products given by previous customers. The second category is content-based recommendation [7,18], which recommends a user products which are similar to the ones that chosen by the user previously. Content-based recommendation first extracts content features of candidates, and then matches the features with preferences of customers, and finally recommends the products with high matching degrees to the customers. The third category is knowledge-based recommendation [8]. This technique utilizes efficacy knowledge in particular domains, i.e., the knowledge about how a product satisfies a particular user. State of the art of the auto recommendation techniques have been addressed in detail in two survey books [16, 17]. Readers are suggested to refer to [16, 17] for details about the recommendation techniques. Note that these auto recommendation techniques are aimed to find and recommending products that customers are likely interested in, rather than recommending products with the highest price-performance ratios.
B. Skyline Search Techniques

Skyline search focuses on extracting a skyline set from a given set of candidates effectively and efficiently. A number of approaches [9-15] have been proposed for skyline search. However, these approaches are not applicable for the recommendation problem considered in this paper. An example as shown in Table 1 is used to illustrate why they are not applicable. This example is based on two assumptions:

1) products are assumed to be single rooms in hotels, and
2) in the selection of the products, customers consider only two attributes of the hotels, star levels and distances from the city center.

<table>
<thead>
<tr>
<th>Single room</th>
<th>Star level of hotel</th>
<th>Distance to city center (km)</th>
<th>Price ($)</th>
</tr>
</thead>
<tbody>
<tr>
<td>P_1</td>
<td>2</td>
<td>9.5</td>
<td>280</td>
</tr>
<tr>
<td>P_2</td>
<td>2</td>
<td>7.2</td>
<td>291</td>
</tr>
<tr>
<td>P_3</td>
<td>2</td>
<td>4.3</td>
<td>330</td>
</tr>
<tr>
<td>P_4</td>
<td>3</td>
<td>7.0</td>
<td>395</td>
</tr>
<tr>
<td>P_5</td>
<td>3</td>
<td>4.5</td>
<td>440</td>
</tr>
<tr>
<td>P_6</td>
<td>3</td>
<td>6.5</td>
<td>425</td>
</tr>
</tbody>
</table>

Skyline search techniques have three major limitations for the recommendation problem in this paper.

First, results found by skyline search are absolutely dominate/best products. In this example, a dominate product, h, is a hotel (a single room) that is no worse than any hotel h’ in the candidate set, Dc, and is better than h’ in at least one attribute. Consequently, skyline search techniques are not able to compare price-performance ratios between products in different levels, e.g., P_3 and P_5 in Table 1. P_3 has a star level higher than P_5 by one, and P_3 and P_5 have similar distances from the city center. The price of P_3 is higher than that of P_5 by $100. The dominance relationship defined in skyline search is not suitable for comparing the price-performance between P_3 and P_5.

Second, the dominance relationship used in skyline search uses a common criterion without considerations of customized criteria, and thus it is not able to assess price-performance ratios of products for different customers. For example, for P_3 and P_5, a customer C_1 may think it is “cheap” to pay $100 more for a single room in a hotel with one-star higher, and thus P_3 has a higher price-performance ratio than P_5 in his/her mind. While, another customer C_2 may not think so, and thus P_5 has a lower price-performance ratio than P_3 in his/her mind.

Third, most existing skyline search techniques are unable to meet the high efficiency requirement for real-time recommendations in online shopping/booking. Skyline search normally takes much time, especially when the number of candidates is large and/or a large number of attributes are considered [11]. Consequently, existing skyline search techniques are not applicable for online recommendation of products with high price-performance ratios.

This paper aims to propose an approach for recommending products with the highest price-performance ratios to customers according to their personalized criteria and parameters. The three limitations have been overcome in this approach.

C. Applications

Now, most existing e-commerce systems and websites have employed auto recommendation systems. For example, in Amazon, one of the largest e-retailing websites, the system records details about customers’ activities in the website, such as what products the customers browsed, what products they purchased, and their comments on the products. Its auto recommender system is able to discover customers’ preferences by exploring historical navigating and browsing behavior of every user in the website, and thus it can recommend products which are potentially interesting for different customers. Most of the existing recommender systems focus on identifying products which are likely interesting for customers, in order to recommend right products to target customers. However, these recommender systems are not able to identify and recommend products with high price-performance ratios from a same category of products.

III. CUSTOMIZED PRICE-DOMINANCE RELATIONSHIP

To recommend products with the highest price-performance ratios to customers, a customized price dominance relationship is introduced first to evaluate and compare price-performance ratios between two products. Different people may have different criteria and concerned attributes in assessing price-performance ratios. This means different people may have different psychological prices for a same product. Therefore, in this paper, a customized price function is introduced to obtain a psychological price for every customer, and then a customized price-dominance relationship and a customized skyline set are introduced.

First of all, customized parameters are introduced to specify customized criteria on price-performance ratios for different customers. For example, assume a user considers only two attributes (assessment attributes), star level and distance to city center, and the user’s criteria in assessing and comparing the price-performance ratios between single rooms in two hotels are: if the values of other attributes are at the same level, the user thinks it is worthwhile to pay $100 more for an increase of 1 in star-level, and it is worthwhile to pay $5 more for an decrease of 1km in distance to city center. We call the increases in price like $100 and $5 ‘acceptable increases in price’.

The customized parameters for customers are defined formally as follows.

Definition 1 (Customized parameters)

• Assessment attribute \( A=[a_1, a_2, \ldots, a_m] \) \( (m>0) \) is a set of attributes considered by a user for
assessing price-performance ratios in product selection. For example, if the user considers two attributes, star-level and distance to city center, then \( A = \{ \text{star-level, distance to city center} \} \). In this paper, we assume that \( a_i \) (\( i = 1, 2, \ldots, m \)) only takes numerical values, i.e., levels of the attribute. Discrete values can be converted to numerical ones according to the preferences of the user, for example, if the user considers the attribute color, and he/she has three preferences, black, white, blue (in descending order of preference), the three values (black, white, blue) can be converted to \( \{ 1, 2, 3 \} \) respectively.

- **Acceptable increases in price** 
  \( E = \{ e_1, e_2, \ldots, e_m \} \) (\( m > 0 \)) indicates that the user is willing to pay extra \( S e_i \) for a unit increase in the value of \( a_i \), e.g., one star-level increase, and a decrease of 1 km in distance to the city center. Note that \( e_i \) may take negative values. For example, if \( a_i \) is distance to city center, then price decreases when \( a_i \) increases. In this case, \( e_i \) takes negative values.

- **Threshold of price difference** 
  \( \varepsilon \) (\( 0 < \varepsilon < 1 \)) is a maximal threshold to restrict the degree of price difference.

**Definition 2 (Customized price function)** Given customized parameters, \( A = \{ a_1, a_2, \ldots, a_m \} \), and \( E = \{ e_1, e_2, \ldots, e_m \} \), where \( a_i \) takes value \( x_i \) (\( i = 1, 2, \ldots, m \)), a customized price function for the user is defined as

\[
q(x_1, x_2, \ldots, x_m) = \sum_{i=1}^{m} e_i x_i + b \quad (1)
\]

where \( b \) is a constant to be determined, which can be seen as a base price.

Assume \( n \) products with actual prices \( p_1, p_2, \ldots, p_n \) are provided to select. Customized prices \( q_1, q_2, \ldots, q_n \) can be obtained by Eq. (1). The value of \( b \) can be determined by minimizing mean square deviation between customized prices and actual prices, i.e., minimizing

\[
f(b) = \sum_{k=1}^{n} (q_k - p_k)^2 = \sum_{k=1}^{n} \left( \sum_{i=1}^{m} e_i x_{ki} + b - p_k \right)^2 \quad (2)
\]

Let \( f(b) = 0 \), we have

\[
b = \frac{\sum_{k=1}^{n} (p_k - \sum_{i=1}^{m} e_i x_{ki})}{n}
\]

In here, \( b \) can be seen as a base price for the products. It is reasonable to take it as the base price as it is an average level in the domain.

A psychological price \( q \) for a user can be obtained by Eq. (1) with a value of \( b \). The product is ‘expensive’ (i.e., the price-performance ratio is low) for the user if its actual price \( p \) is higher than the user’s psychological price, \( q \), otherwise, the product is ‘cheap’, i.e., the price-performance ratio is high.

To compare price-performance ratios between different products for different customers, we introduce a customized price-dominance relationship.

**Definition 3 (Customized price-dominance relationship, \( \succ \))** Given two products \( P_1 \) and \( P_2 \) with actual prices \( p_1 \) and \( p_2 \), and customized parameters \( A, E, \) and \( e \). For user \( U \), \( P_2 \) price dominates \( P_1 \), denoted as \( P_2 \succ P_1 \), if

\[
\frac{\delta(q_2, p_2) - \delta(q_1, p_1)}{(\delta(q_2, p_2) + \delta(q_1, p_1)) / 2} > \varepsilon \quad (3)
\]

where, \( \delta(q_1, p_1) = \frac{q_1 - p_1}{p_1} \), \( \delta(q_2, p_2) = \frac{q_2 - p_2}{p_2} \).

In Eq. (3), \( \delta(q_2, p_2) \) represents the difference between the psychological price \( q_2 \) and the actual price \( p_2 \).

Equation 3 indicates that the difference between \( \delta(q_2, p_2) \) and \( \delta(q_1, p_1) \) is greater than the threshold \( \varepsilon \).

Based on the customized price-dominance relationship, a product set to be recommended, \( D_r \), can be defined as a customized skyline set from a candidate set \( D_c \), i.e., a subset \( S \) of \( D_c \), such that there exist no other product \( Q \) in \( D_c \) such that \( Q \succ P \). The customized skyline set is defined formally as follows.

**Definition 4 (Customized skyline set)** Given a candidate set \( D_c \), user \( U \) and the user’s customized parameters, the customized skyline set \( S \) for \( U \) is defined as

\[
S = \{ P \mid P \in D_c \land \exists Q \in D_c \quad s.t. Q \succ P \} \quad (4)
\]

The customized skyline set for user \( U \) is essentially a set of products with the highest price-performance ratio for \( U \).

**IV. CUSTOMIZED SKYLINE SEARCH**

Since the customized skyline set in Definition 4 is a set of products with the highest price-performance ratios to be recommended, the customized recommendation approach proposed in this paper is customized skyline search in \( D_c \). The customized skyline search approach is described as follows.

**Input:** Candidate set \( D_c = \{ p_1, p_2, \ldots, p_n \} \), user's customized parameters.

**Output:** Customized skyline set \( S \)

**Step 1:** Obtain the value of \( b \).

**Step 2:** Obtain customized prices of the \( n \) products, \( q_1, q_2, \ldots, q_n \).

**Step 3:** Find product \( P_k \) with the maximal value of \( \delta(q_k, p_k) \) (\( i = 1, 2, \ldots, n \)), and insert \( P_k \) into \( S \).

**Step 4:** Compare \( P_i (i \neq k) \) with \( P_k \) for all \( i = 1,2, \ldots, n \), and insert \( P_i \) into \( S \) if \( P_i \succ P_k \) does not hold.

**Step 5:** Output \( S \).

The computation complexity of the approach is \( O(|D_c|) \), which is linear. Therefore, it can satisfy the real-time requirement easily. This will be validated in a real case study in Section V.

The following example is used to illustrate how a customized skyline set, i.e., a set of products with the highest price-performance ratios, is obtained by the proposed approach.
Example 1 Given $D_c$, a set of six rooms in Table 1, and customized parameters given by user $U$, $A=\{\text{star level, distance to city center}\}$, $e_1=100$, $e_2=5$ and $\varepsilon=0.1\%$. $S$ is obtained by the customized skyline search approach.

In Step 1, the value of $\delta$, 142.67 is obtained. In Step 2, customized prices $(295.17, 306.67, 321.17, 407.67, 420.17, 410.17)$ of the six products are obtained. In Step 3, the product with the maximal value of $\delta$, $P_1$, is found. In Step 4, $P_1$ is compared with $P_i$ ($i=2, 3, ..., 6$). Finally, $S=\{P_1, P_2\}$ is obtained.

V. A CASE STUDY: ONLINE RECOMMENDATION OF HOTEL

This section applies the proposed algorithm to customized recommendations for online room booking with hotels.

We implemented the algorithm by programming in Java, and built a system for personalized recommendation. The system contains two interfaces. One is for accepting parameters given by customers, and the other is for accepting databases of candidates for recommendation.

To test the effectiveness and efficiency of the recommendation system, we applied the system to hotel recommendation. We obtained a database of hotels from ctrip.com, which is a professional and representative booking website with around 90 million members. This website provides services of booking hotels, flights and train tickets and vacation packages. The system was used to recommend rooms of the highest price-performance ratios for different customers given their personalized parameters.

Assume four customers want to book a room with hotels in Beijing, China on ctrip.com. The customers prefer hotels closer to the city center. Basic constraints given by the user are: check-in date, October 10, 2012, double bed room, 3-4 star levels, $200<\text{price}<500$ and internet available. We searched on ctrip.com with the constraints and obtained a set of 368 candidates that satisfy the constraints. Statistics about the 368 hotels include average price=$364.37$, average distance to city center=16.12 km. Among the 368 hotels, there were 253 three-star hotels with an average price of $334.19$, and 115 four-star hotels with an average price of $430.75$.

The four customers need to book four rooms in the 368 hotels. They consider two attributes, star level and distance to city center when they assess price-performance ratios. The customized parameters in $E$ for the four customers are listed in Table III.

The proposed approach was used to recommend rooms to the four customers according to their customized personalized parameters. Table II lists the hotels recommended to the customers.

Table III lists the customized parameters and customized skyline sets for each user. Each hotel in a skyline set is represented by its ID. For $e_1$ and $e_2$ in Table III, assume corresponding attributes are $a_1=\text{star level}, a_2=\text{distance to city center}$. Comparing $U_1$ and $U_2$ in Table III, we see that $U_1$ and $U_2$ have a same criterion on star levels, and $U_1$ is willing to pay $9.5$ more than $U_2$ for a decrease of 1 km in distance. Hotels 138 and 140 have an advantage in terms of distance for $U_1$, and are recommended to $U_1$. Comparing $U_1$ and $U_3$, we see that they share a same criterion on distance, and $U_3$ is concerned more on star levels than $U_1$. Consequently, hotel 2 is recommended to $U_3$ with an advantage on star level. Comparing $U_1$ and $U_4$, we see that $U_4$ requires a smaller threshold than $U_1$. As a result, hotel #358 is not recommended to $U_4$.

In addition, the approach has a linear computation complexity. The real test showed that average run time of the approach was around 0.1 second. Consequently, the approach is able to satisfy the real-time requirement in online customized recommendations.

VI. CONCLUSION AND FUTURE WORK

In order to help customers find desired products with the highest price-performance ratios from a large number of products within the same category in online booking/shopping, this paper investigated auto recommendations of products with the highest price-performance ratios. We first introduced a customized
price-dominance relationship for comparing price-performance ratios between products. Then we formulated a set of products with the highest price-performance ratios as a customized skyline set. We finally developed a recommendation approach based on customized skyline search.

We applied the proposed approach to hotel recommendations in online accommodation booking. Experimental results in a real case study on room booking at ctrip.com demonstrated that the proposed approach is capable of recommending products with the highest price-performance ratios to customers according to the customers’ different criteria and parameters. Moreover, the proposed approach has an advantage of high-time-efficiency with a linear computation complexity.

The proposed approach is able to provide customers with customized recommendations of products with the highest price-performance ratios in online booking/shopping, so that the customers may find their desired products with the highest price-performance ratios straightforwardly.

We are going to improve the proposed approach in future to enhance its effectiveness. For example, the system only can only process numerical values. However, in real applications, some attributes’ values may be categorical. For instance, the service quality of hotels can be evaluated as ‘excellent’, ‘good’, ‘general’ and ‘bad’. A simple way is to convert the categorical values to numerical values, say 1, 2, 3, 4, and the distance to city center, 1km and 5km, the converted numerical values are rough without a common standard, and different people may have different understandings. To achieve a precise evaluation of price-performance ratios for different customers, a strategy is needed for assessing the non-numerical values of attributes precisely for different customers.

ACKNOWLEDGMENT

This work was partially supported by Natural Science Foundation of Zhejiang Province of China (grant no. Y12F030062).

REFERENCES


Hongfa Wang professor from Zhejiang University of Water Resources and Electric Power. His major research interest is intelligent computing.

Chen Xing lecturer from Zhejiang University of Water Resources and Electric Power. His major research interest is social computing, data mining.