PROBABILISTIC RANKING OF MULTI-ATTRIBUTE ITEMS USING INDIFFERENCE CURVE

Xiaohui Gong* *H. Vicky Zhao** *Yan Lindsay Sun*†

* ECE Department, University of Alberta, Edmonton, AB Canada T6G 2V4
† ECBE Department, University of Rhode Island, Kingston, RI USA 02881
Email: gxiaohui@ualberta.ca, vzhao@ece.ualberta.ca, yansun@ele.uri.edu

ABSTRACT

This work proposes a novel probabilistic multi-attribute item ranking framework to estimate the probability of an item being a user's best choice and rank items accordingly. It uses indifference curve from microeconomics to model users' personal preference, and addresses the inter-attribute tradeoff and inter-item competition issues at the same time with little information loss. The proposed framework also considers the fact that a user can only compare a few items at the same time, and models the user's selection process as a two-step process, where the user first selects a few candidates, and then makes detailed comparison. Simulation results show that the proposed framework significantly outperforms existing multi-attribute ranking algorithms in terms of ranking quality.

1. INTRODUCTION

With the proliferation of social networks and online communities in the past decade, users are provided with massive amount of information and choices, and we are entering the "big data" era. To improve user experience and provide personalized services, personalized ranking plays an important role in web searching [1], database queries [2], and recommender systems [3] in eCommerce and social networks. Based on a user's personal preference and the user's explicit/implicit query, the algorithm ranks a set of items considering their relevance to the query [2], importance to the user [1], and match to the user's personal interest [3], etc. Then, the items are presented to the user according to their ranks. Personalized ranking helps the user quickly find the item/information that he/she needs.

Many personalized ranking algorithms have to consider multiple attributes of the items and address the *inter-attribute* tradeoff, for example, the tradeoff between reputation and price in eCommerce. There are currently two types of works to address this issue. The first tries to identify important attributes, and organize them into an importance hierarchy [4,5]. However, such methods may reduce the recommended items but cannot rank them. The second type uses attribute weighting [6] to quantify the relative importance of different attributes [7–9]. However, predetermined weights and normalization functions restrict the formation of personalized preferences, and they fail to address the *inter-item competition*, where a competitive item may reduce the chance of other items being selected.

To address the inter-item competition, the work in [10] proposed a personalized Multi-Attribute Probabilistic Selection framework (MAPS). In [10], each attribute is considered to be one dimension in a multi-dimensional space, and every item to compare is mapped to a point in the space. They use visual angle, the angle of the line connecting an item and the origin, to model users' personal preference, and to address the inter-attribute tradeoff and inter-item competition simultaneously. However, the visual angle approach reduces the dimensionality of the attributes by one, and causes information loss and incomplete description of users' preference.

In this work, we propose a novel probabilistic ranking framework using the concept of *indifference curve* from microeconomics, which offers a flexible way to model users' preference and addresses inter-attribute tradeoff and inter-item competition with little information loss. In addition, different from all prior works that assume users compare all items simultaneously, the proposed framework addresses the fact that a user have bounded rationality and can only compare a few items at the same time [11–13]. It models the user's decision making as a two-step selection process, where a user first selects a few candidates and then makes detailed comparison. Furthermore, the proposed framework outputs the probability that an item is the user's best choice, which provides important guidelines on appropriate pricing schemes, estimations of the market demand, and marketing strategies.

The rest of the paper is organized as follows. Section 2 introduces the proposed framework, and Section 3 provides detailed description of the proposed algorithms for user preference estimation and personalized ranking. Simulation results are shown in Section 4, and conclusions are drawn in Section 5.

2. THE PROPOSED FRAMEWORK

2.1. Problem Formulation

In this paper, we consider ranking items in an online shopping platform with two conflicting attributes, price and reputation, and our work can be extended to ranking items with more than two conflicting attributes. Consider a user query, which returns a list of matching items. For a matching item with price $P \in [P_{MIN}, P_{MAX}]$ and reputation $R \in [R_{MIN}, R_{MAX}]$, we normalize both reputation and price into the range $[0, 1]$ using a simple linear mapping function $p = (P_{MAX} - P)/(P_{MAX} - P_{MIN})$ and $r = (R - R_{MIN})/(R_{MAX} - R_{MIN})$. In this work, we use $P_{MAX} = 10^3$, $P_{MIN} = 10$, $R_{MAX} = 10^6$ and $R_{MIN} = 0$, and observe similar trends for other values and other normalization functions. After normalization, for both attributes, a larger value indicates a higher preference of the user. We use the term an *item's utility* $U(s)$ to quantify the user's personal level of satisfaction with the item s, and a larger utility means higher preference.

In this work, we consider rational and consistent user behavior with the following three assumptions [14]. The first is the *monotonicity assumption*, where we assume that an item's utility is higher when an attribute value is higher with the other(s) fixed. Second, we have the *diminishing value assumption*, which says users are assumed to have diminishing additional level of satisfaction with the increase of a certain attribute's value. That is, with the other attribute

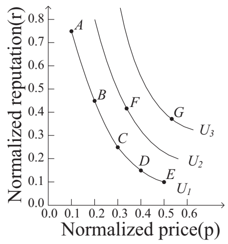

Fig. 1: An indifference map with $U(G) > U(F) > U(A) = \cdots = U(E)$.

values fixed, as one attribute value increases, the additional level of satisfaction that the user obtains decreases [14]. Last, we assume that users have *bounded rationality* [11, 12] and can only compare a few (usually 3 to 5) multi-attribute items at a time [13].

Given a set of items, an item is a *skyline item* if and only its attributes are not all worse (smaller) than those of any other items [15]. From the monotonicity assumption, a non-skyline item whose attributes are all smaller than those of a skyline item has a lower utility, and thus will never be picked by the user. Therefore, in our work, we consider skyline items only.

Denote the set of N skyline items as $S = \{s_i = \{p_i, r_i\}\}$, and without loss of generality, we sort them in the ascending order of normalized reputation with $r_1 < \cdots < r_N$ and $p_1 > \cdots > p_N$. We map all items into points in a two-dimensional space with the X and Y axes being the normalized price and reputation, respectively. In the following, we will use the two terms "point" and "item" interchangeably to represent the same concept.

Given N skyline items S, a user considers the tradeoff between price and reputation based on his/her personal preference, and chooses his/her personal best choice s_b. Due to the bounded rationality, we assume that a user first preselects n interested items/candidates from the set S. He/she then makes detailed comparison within the n items and finds the best choice s_b. For a given user, the goal of the proposed multi-attribute ranking algorithm is to understand the user's personal preference between the conflicting attributes and to rank the items accordingly such that the user's true best choice s_b is ranked as high as possible. In this work, we only consider the fixed-price buy-it-now market but not those requiring auctions. Also, we consider a dynamic market, where items can enter or exit the online market at any time, and where price and reputation change with time.

2.2. Indifference Curve and Marginal Rate of Substitution

In this work, we use the concept of *indifference curve* to model users' personal preference. An indifference curve is a graph showing different combinations of factors among which a user is indifferent, and points on the same indifference curve have the same utility value [14]. An indifference map is a collection of indifference curves with different utility values for a user, as shown in the example in Fig. 1. Indifference curves have three properties [14]: *First*, an indifference curve is always a non-increasing function. *Second*, different indifference curves do not intersect. *Third*, when comparing two indifference curves, the top one has a higher utility and is preferred by the user. With these properties, the problem of user preference modeling is changed to the estimation of the indifference curves, from which we can easily rank items. However, in real applications, we only have a few of the user's transaction (purchasing) records, and this limited information is insufficient to obtain the complete indifference map.

To address this issue, we use another concept from microeconomics, the *marginal rate of substitution* (MRS), which is the maximum amount of the attribute on the Y axis (normalized reputation in this work) that a user is willing to give up to obtain one additional unit of the attribute on the X axis (normalized price). To simplify the analysis, we assume in this work that the indifference curves are continuously differentiable. Then, MRS at a given point is the magnitude of the slope of the indifference curve evaluated at that point. In this work, we use MRS as partial knowledge of the indifference curve to help model users' personalized preference.

MRS has an important *diminishing property* that can help us extract information of the indifference curves and users' preference from a finite number of purchasing records. From the diminishing value assumption, as the price (reputation) increases, the additional satisfaction that user gains with one more unit increment of price (reputation) decreases. Consequently, with user's satisfaction fixed, as the price (reputation) increases, the user is willing to give up less on reputation (price) to gain an additional unit of price (reputation). Therefore, MRS decreases as the price increases along the curve, and the continuously differentiable indifference curves are convex.

Indeed, with only a few purchasing records, we cannot extract perfect information of MSRs of the complete price-reputation plane, but can only estimate their ranges at a few points. Still, as will be demonstrated later, these estimated slope ranges can help capture users' preference and offer good ranking quality.

Based on the above, given a few user's purchasing records, our proposed framework first extracts user's personal preference and estimates the ranges of the slopes of the indifference curves at different points in the 2D price-reputation plane. Then given a new set of skyline items, for each item, we use the estimated slope ranges to estimate the probability that it is the user's best choice, and then rank them based on these estimated probabilities.

2.3. Performance Evaluation

Given the top-down ranking list of the N skyline items in the set S, let v_b be the ranking position of the (known) user's best choice s_b. $v_b = 1$ when s_b is considered to have the highest probability of being purchased by the user, and $v_b = N$ when it is considered to be the least favorable item for the user. We use *ranking quality* $rq = (N - v_b)/(N-1)$, the percentage of items ranked worse than s_b with larger ranking positions than v_b, to evaluate the performance [10]. A larger value of ranking quality indicates better accuracy, where $rq = 1$ when the best choice is accurately ranked the first, and $rq = 0$ when s_b is ranked the last.

3. PROBABILISTIC RANKING WITH SLOPE RANGE ESTIMATION

3.1. Slope Range Estimation

For a point s in the 2D price-reputation plane, let k_s denote the true slope of the indifference curve at s. From their properties, indifference curves have non-positive slopes with $k_s \leq 0$. Given a set of the user's historical purchasing records, we study in the following how to narrow down the slope range.

3.1.1 Estimation from One Transaction: We first consider one single record where among a set of skyline items S with $r_1 < \cdots < r_N$ and $p_1 > \cdots > p_N$, the user purchases s_i as his/her best choice. Note that item s_i has normalized reputation r_i and price p_i.

Given the best choice $s_b = s_i$, we first divide the whole set S into two subsets: the first includes all points above the best choice s_b with $S_+ = \{s_{i+1}, \cdots, s_N\}$; and the second includes all points below s_b with $S_- = \{s_1, \cdots, s_{i-1}\}$. We study these points separately and have Theorem 1. The proof is in [16].

Theorem 1. *For an item $s_j \neq s_b$, let k_{jb} be the slope of the line connecting s_j and the best choice s_b. For all $s_j \in S_+$, we have $k_{s_j} \leq k_{jb}$; and for all $s_j \in S_-$, we have $k_{jb} \leq k_{s_j} \leq 0$.*

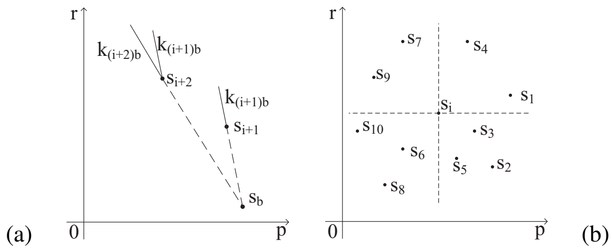

Fig. 2: Refinement of estimated slope range with (a) one transaction record, and (b) multiple transactions.

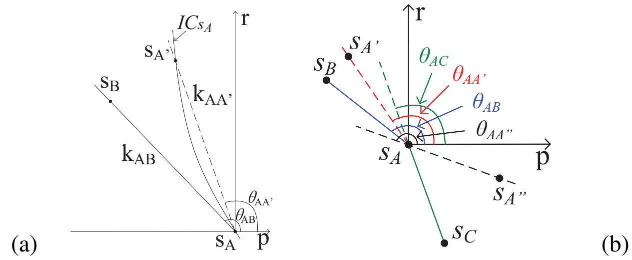

Fig. 3: (a) Two-item comparison, and (b) three-item comparison.

It is worth mentioning that from Theorem 1, with a single purchasing record $\{S, s_b\}$, we can only update the slope ranges of non-selected items, but not that of the best choice s_b.

Given the above initial slope range estimation, we can further refine the estimation using the diminishing MRS property in Section 2.2. Consider the example in Fig. 2a, there are two items s_{i+1} and s_{i+2} above the best choice $s_b = s_i$. From Theorem 1, initially, we have $k_{s_{i+2}} \leq k_{(i+2)b}$ and $k_{s_{i+1}} \leq k_{(i+1)b}$ with $k_{(i+2)b} > k_{(i+1)b}$. Note that from the diminishing property of MRS, we have $k_{s_{i+2}} \leq k_{s_{i+1}}$. Therefore, we can update the range of $k_{s_{i+2}}$ to $k_{s_{i+2}} \leq k_{s_{i+1}} \leq k_{(i+1)b}$. Similarly, we can update the ranges for all items in S_- accordingly.

3.1.2 Refinement with Multiple Transactions: From Theorem 1, with one purchasing record, we can update the upper bounds of the slope ranges for all items above the best choice and the lower bounds of the slope ranges for all items below the best choice. Now, we consider slope range refinement with multiple transactions.

For a given item s_i, we divide the remaining items from all previously known transactions into four subsets: $S_i^I = \{s_j : p_j > p_i, r_j > r_i\}$, $S_i^{II} = \{s_j : p_j \leq p_i, r_j > r_i\}$, $S_i^{III} = \{s_j : p_j \leq p_i, r_j \leq r_i\}$ and $S_i^{IV} = \{s_j : p_j > p_i, r_j \leq r_i\}$, where each s_j has its own estimated slope range $[\underline{k}_{s_j}, \bar{k}_{s_j}]$. In the example in Fig. 2b, $S_i^I = \{s_1, s_4\}$, $S_i^{II} = \{s_7, s_9\}$, $S_i^{III} = \{s_6, s_8, s_{10}\}$ and $S_i^{IV} = \{s_2, s_3, s_5\}$.

To further refine the estimation results from multiple records, we again use the diminishing property of MRS, and use items in S_i^{II} to update the lower bound of s_i's slope range $\underline{k}_{s_i}$, and use items in S_i^{IV} to update the upper bound of s_i's slope range $\bar{k}_{s_i}$. The diminishing property of MRS says $k_{s_i} \leq k_{s_j}$ for all $s_j \in S_i^{IV}$ and $k_{s_i} \geq k_{s_j}$ for all $s_j \in S_i^{II}$. Therefore, we update $\bar{k}_{s_i} = \min_{s_j \in S_i^{IV} \cup s_i} \{\bar{k}_{s_j}\}$ and $\underline{k}_{s_i} = \max_{s_j \in S_i^{II} \cup s_i} \{\underline{k}_{s_j}\}$. The above estimation and refinement enable us to convert the historical purchasing records into *personalized records* in the price-reputation plane. Define the set of personalized records as $H = \{h_i = \{s_i, [\underline{k}_{s_i}, \bar{k}_{s_i}]\}\}$.

The last step is to check the consistency of the estimated slope range at each point. For point s_i, its estimated slope range $[\underline{k}_{s_i}, \bar{k}_{s_i}]$ should satisfy $\underline{k}_{s_i} \leq \bar{k}_{s_i}$. If $\underline{k}_{s_i} > \bar{k}_{s_i}$, it means that the user shows inconsistent behavior in the historical records, and the corresponding personalized record h_i should be discarded from H to ensure accurate information collection.

3.2. Ranking

3.2.1 Slope Range Estimation in the New Market: Following the above steps, we can estimate the slope ranges of the indifference curves at a few points where there were corresponding items in the historical records. However, given a new query in a dynamically changing market, it is possible that there are new items that we have no prior information about their slope ranges.

For a new item s_{i_0}, we need to estimate the upper and lower bounds of its slope range. To estimate the upper bound, we first search all items in the historical data and find a set of its M closest neighbors S_u whose upper bounds are non-zero with $\bar{k}_{s_j} < 0$ for all $s_j \in S_u$. We then estimate the upper bound of $k_{s_{i_0}}$ using weighted sum $\bar{k}_{s_{i_0}} = \sum_{s_j \in S_u} w_j \bar{k}_{s_j}$, where the weight $w_j = (d_{i_0 j})^{-1} / \sum_{s_j \in S_u} [(d_{i_0 j})^{-1}]$ is inversely proportional to the distance $d_{i_0 j}$ between s_{i_0} and s_j. Similarly, to estimate the lower bound of $k_{s_{i_0}}$, we find a set of its M closest neighbors S_l among all items in the historical data whose lower bounds are finite, and estimate $\underline{k}_{s_{i_0}}$ using $\underline{k}_{s_{i_0}} = \sum_{s_j \in S_l} w_j \underline{k}_{s_j}$. We use $M = 3$ in this work and observe similar results for other values of M.

3.2.2 Two-Step Ranking: We consider the scenario where a user first preselects n candidates that he/she might be interested in, and then makes detailed comparison within the n items. To find the probability that an item s_i is the best choice, we first need to find all possible preselected candidate sets S_{IS_j} that includes s_i, and then for each such set S_{IS_j}, find the probability that s_i has the largest utility compared to all other items in S_{IS_j}.

Mathematically, let $\mathcal{P}[S_{IS_j}]$ be the probability that S_{IS_j} is the preselected candidate set, and let $\mathcal{P}[s_i = \text{best}|S_{IS_j}]$ denote the probability that s_i is the preferred item among all in S_{IS_j}. Then, the probability that s_i is the user's best choice is

$$\mathcal{P}_{s_i} = \sum_{S_{IS_j} : s_i \in S_{IS_j}} \mathcal{P}[s_i = \text{best}|S_{IS_j}] \mathcal{P}[S_{IS_j}]. \quad (1)$$

To model the candidate pre-selection process, we adopt the *visual angle* model in [10]. In particular, a preference density function $f(\psi)$ is used to model the probability that a user is interested in items at angle ψ in the price-reputation plane, and we use the same method as in [10] to estimate $f(\psi)$. For skyline item s_i, define its visual angle as $\psi_i = \arctan(r_i/p_i)$. Given the estimated $f(\psi)$, from the derivation in [16], the probability that the pre-selected candidate list is $S_{IS_j} = \{s_{j-1}, s_j, s_{j+1}, s_{j+2}\}$ is

$$\mathcal{P}[S_{IS_j}] = \int_{\psi_j}^{\psi_{j+1}} f(\psi) \mathrm{d}\psi. \quad (2)$$

The next step is to compute $\mathcal{P}[s_i = \text{best}|S_{IS_j}]$, the probability that s_i is the preferred item in S_{IS_j}, we first consider a simple scenario of comparing two items s_A and s_B where s_A is below s_B, as shown in Figure 3a. To determine the probability that s_A is preferred to s_B, we first consider the indifference curve IC_{s_A} that passes through s_A. Since all points on IC_{s_A} have the same utility as s_A, we can compare s_B with any point on IC_{s_A}. In this work, we choose the point $s_{A'}$ whose distance to s_A is the same as that between s_A and s_B. Let k_{AB} denote the slope of the line connecting s_A and s_B, and $k_{AA'}$ be the slope of the line connecting $s_{A'}$ and s_A. We define function $\theta(k) = \pi + \arctan(k)$ to convert slope k to angle θ, and we have $\theta_{AB} = \theta(k_{AB})$ and $\theta_{AA'} = \theta(k_{AA'})$, as shown in Fig. 3a. Since $s_{A'}$ and s_B have the same distance to s_A, comparing their positions is equivalent to comparing the two angles θ_{AB} and $\theta_{AA'}$. From Fig. 3a, it is easy to see that when $\theta_{AB} > \theta_{AA'}$, $s_{A'}$ is above s_B and $U(s_A) > U(s_B)$, and vice versa.

Table 1: Simulation results of the ranking quality of different multi-attribute ranking algorithms.

		Type 1	Type 2	Type 3	Type 4	Type 5
Indifference curve		96.55%	94.63%	97.01%	99.87%	99.41%
MAPS		90.71%	86.06%	93.56%	98.79%	99.14%
Weighted sum	Max	86.56%	76.25%	92.87%	99.99%	100%
	Min	15.04%	25.16%	9.97%	0%	1.57%
	Average	68.20%	61.61%	72.47%	48.0%	75.83%

To compare θ_{AB} and $\theta_{AA'}$, note that the indifference curve IC_{s_A} is convex, and we have $-\infty < k_{AA'} \leq k_A \leq \bar{k}_A$ where k_A is the true slope of IC_{s_A} at point s_A, and $\bar{k}_A$ is the estimated upper bound of k_A. Define $\bar{\theta}_A = \theta(\bar{k}_A)$ and $\theta_A = \theta(k_A)$. Without any prior knowledge of $\theta_{AA'}$ or the position of $s_{A'}$, we assume that $\theta_{AA'}$ is uniformly distributed in the range $[\pi/2, \bar{\theta}_A]$. Note that if $\theta_{AB} > \bar{\theta}_A$, we have $\theta_{AB} > \theta_{AA'}$, and thus s_A is always preferred to s_B. Therefore, the probability that s_A is preferred to s_B is

$$\mathcal{P}[U(s_A) \geq U(s_B)] = \begin{cases} \frac{\theta_{AB} - \pi/2}{\bar{\theta}_A - \pi/2} & \text{if } \pi/2 \leq \theta_{AB} \leq \bar{\theta}_A, \\ 1 & \text{if } \bar{\theta}_A < \theta_{AB}. \end{cases} \quad (3)$$

Now we consider the scenario where we compare three or more items at the same time. We first consider the case where there are three items s_A, s_B, s_C where s_A is below s_B but above s_C as shown in Fig. 3b. Same as Fig. 3a, we consider the indifference curve IC_{s_A} that passes through s_A and find the point $s_{A'}$ on IC_{s_A} that has the same distance to s_A as s_B. The definitions of $\theta_{AA'}$ and θ_{AB} are the same as above. Similarly, we find another point $s_{A''}$ on IC_{s_A} that has the same distance to s_A as s_C. Let k_{AC} be the slope of the line connecting s_A and s_C, and define $\theta_{AC} = \theta(k_{AC})$. Let $k_{AA''}$ be the line connecting s_A and $s_{A''}$ and define $\theta_{AA''} = \theta(k_{AA''})$. Define $\underline{\theta}_A = \theta(\underline{k}_A)$. Note that the indifference curve IC_{s_A} is convex. Thus, we have $k_{AA'} < k_{AA''}$ and $\theta_{AA'} < \theta_{AA''}$. So the probability that s_A has the largest utility among the three is equivalent to the probability that $\theta_{AA'} \leq \theta_{AB}$ and $\theta_{AA''} \geq \theta_{AC}$ under the constraint that $\theta_{AA'} < \theta_{AA''}$. Without any prior knowledge of the positions of $s_{A'}$ and $s_{A''}$, we assume that $\theta_{AA'}$ and $\theta_{AA''}$ are independent. Following the same analysis as (3), from [16], we have

$$\begin{aligned} &\mathcal{P}[U(s_A) \geq U(s_B), U(s_A) \geq U(s_C)] \\ = &\begin{cases} \frac{-2(\theta_{AB} - \pi/2)(\pi - \theta_{AC})}{(\bar{\theta}_A - \pi)^2 + (\pi/2 - \underline{\theta}_A)^2 - \pi^2/4} & \text{if } k_{AB} < k_{AC}, \\ \frac{\pi^2 - \theta_{AC}\pi - 2\theta_{AB}\pi + \theta_{AB}{}^2 + \theta_{AC}{}^2}{(\bar{\theta}_A - \pi)^2 + (\pi/2 - \underline{\theta}_A)^2 - \pi^2/4} & \text{if } k_{AB} > k_{AC}. \end{cases} \end{aligned} \quad (4)$$

Using the same method, given s_A, s_B and s_C as in Fig. 3b, we can also calculate the probabilities that s_B and s_C are preferred among the three, respectively. Detailed derivations can be found in [16]. The comparison among four or more items is similar.

In summary, for each skyline item in a new market, we use (1) - (4) to compute the probability that it is the user's best choice. We then rank all items in the descending order of $\mathcal{P}_{s_i}$.

4. SIMULATION RESULTS

To validate the performance of the proposed personalized ranking framework, we use synthetic markets to simulate practical online shopping queries and compare the proposed framework with prior works. Same as in [10], for each query, we generate a synthetic market with N skyline items, where N is a randomly chosen number in $[20, 100]$. Then, we generate N price values in the range $[10, 10^3]$ and N reputation values in the range $[0, 10^6]$, both following the power law distribution. Without loss of generality, we order the reputation and price values from low to high with $P_1 < \cdots < P_N$ and $R_1 < \cdots < R_N$. Then, N skyline items are generated with (P_i, R_i) being the ith item. We follow Section 2.1 and normalize all prices and reputations.

To simulate users' selection behavior, we use the widely used Cobb-Douglas model in economics [14] to model the user's level of satisfaction with an item. For an item with normalized price p and normalized reputation r, the item's utility function is $U(p, r) = p^\alpha \cdot r^\beta$ with $\alpha \geq 0$ and $\beta \geq 0$ being the parameters quantifying the importance of price and reputation to the user, respectively. We consider five different categories of users, summarized as follows:

Type 1: Users in this category consider price and reputation to be equally important, and we use $\alpha = \beta = 1$ as an example.

Type 2: Users in this category consider price to be more important than reputation, and we choose $\alpha = 2$ and $\beta = 1$ as an example. For other values of α and β with $\alpha > \beta$, we observe the same trend.

Type 3: These users consider that reputation is more important than price and we use $\alpha = 1$ and $\beta = 2$ as an example.

Type 4: Users in this category consider price only and always choose the cheapest item. We use $\alpha = 1$ and $\beta = 0$ as an example.

Type 5: Users in this category consider reputation only and always choose the item with the highest reputation in the market. We use $\alpha = 0$ and $\beta = 1$ as an example.

For a user, his/her utility function is used as ground truth to choose his/her best choice with the largest utility. For each type of users, we repeat the simulations 30,000 times and average the results.

Table 1 compares the proposed indifference curve (IC) based method with MAPS [10] and the weighted sum approach [7–9] where the utility function $V(s_i) = \gamma r_i + (1 - \gamma) p_i$ is used to estimate users' personal preference. It shows the ranking quality defined in Section 2.3. The IC based method and MAPS estimate users' personal preference based on their previous $L = 5$ purchasing records, and each user in the IC based method preselects $n = 4$ items for detailed comparison. Since the performance of the weighted sum approach is very sensitive to the selected weight parameter γ, we show its best, worst, and average performance in Table 1. We can see that the proposed method outperforms both the weighted sum approach and MAPS for almost all five types of users. Even though for Type-4 (price only) and Type-5 (reputation only) users, the weighted sum approach with the optimum weight gives approximately the same performance as MAPS and the indifference curve based method, its average and worst performances are much worse, and therefore, the proposed method is preferred considering the overall performance.

5. CONCLUSION

In this work, we proposed a novel multi-attribute probabilistic ranking framework, which uses indifference curves to address multi-attribute tradeoff and to model users' personal preference with little information loss. The proposed framework contains two parts: estimation of the slope ranges of the indifference curves from users' historical purchasing records, and two-step ranking that estimates the probability of each item being the user's best choice and ranks items accordingly. Simulation results show that the proposed scheme outperforms existing multi-attribute ranking algorithms in terms of ranking quality.

6. REFERENCES

[1] L. Page, S. Brin, R. Motwani, and T. Winograd, "The PageRank citation ranking: bringing order to the web," *Stanford InfoLab Technical Report*, 1999.

[2] Z. Chen and T. Li, "Addressing diverse user preferences in SQL-query-result navigation," pp. 641–652, 2007.

[3] G. Adomavicius and A. Tuzhilin, "Toward the next generation of recommender systems: A survey of the state-of-the-art and possible extensions," *IEEE Trans. Knowl. Data Eng.*, vol. 17, no. 6, pp. 734–749, 2005.

[4] W. Kiebling, "Foundations of preferences in database systems," *ACM Int. Conf. on Very Large Databases*, pp. 311–322, 2002.

[5] Mindolin D and J. Chomicki, "Discovering relative importance of skyline attributes," *ACM Int. Conf. on Very Large Databases*, vol. 2, no. 1, pp. 610–621, 2009.

[6] J. Wallenius, J. S. Dyer, P. C. Fishburn, R. E. Steuer, S. Zionts, and K. Deb, "Multiple criteria decision making, multiattribute utility theory: Recent accomplishments and what lies ahead," *Management Science*, vol. 54, no. 7, pp. 1336–1349, 2008.

[7] U. Junker, "Preference-based search and multi-criteria optimization," *National Conf. Artificial Intelligentce*, pp. 34–40, 2002.

[8] P. Viappiani, B. Faltings, and P. Pu, "The lookahead principle for preference elicitation: Experimental results," *Flexible Query Answering Systems*, pp. 378–389, 2006.

[9] J. C. Butler, J. S. Dyer, J. Jia, and K. Tomak, "Enabling e-transactions with multi-attribute preference models," *European Journal of Operational Research*, vol. 186, no. 2, pp. 748–765, 2008.

[10] Q. Feng, L. Liu, Y. Sun, T. Yu, and Y. Dai, "Enhancing personalized ranking quality through multidimensional modeling of inter-item competition," *IEEE Int. Conf. Collaborative Computing: Networking, Applications and Worksharing*, pp. 1–10, 2010.

[11] H. A. Simon, *Models of Man: Social and Rational*, Wiley, 1957.

[12] G. Gigerenzer and R. Selten, *Bounded rationality: The adaptative toolbox*, MIT Press, 2002.

[13] G. Miller, "The magical number seven, plus or minus two: some limits on our capacity for processing information," *Psychological review*, vol. 63, no. 2, pp. 81, 1956.

[14] R. S. Pindyck and D. L. Rubinfeld, *Microeconomics*, Prentice Hall, 2005.

[15] S. Borzsony, Donald D. Kossmann, and K. Stocker, "The skyline operator," *IEEE Int. Conf. Data Engr.*, pp. 421–430, 2001.

[16] Xiaohui Gong, H. Vicky Zhao, and Yan Lindsay Sun, "Probabilistic ranking of multi-attribute items using indifference curve," *Technical Report*, http://www.ece.ualberta.ca/~vzhao/PRankingICReport.pdf.