A Low-Effort Recommendation System with High Accuracy

A New Approach with Ranked Pareto-Fronts

In a simulation study, we demonstrate that recommendation systems using a choice-based conjoint analysis with hierarchical Bayes estimation require up to three times higher mental effort for the consumer than simple sorting mechanisms. However, consumers benefit from a choice-based conjoint analysis in terms of a significantly higher utility of the selected product. We further introduce the concept of a ranked Pareto-front which allows consumers to select a product with a better utility than they will select when using a choice-based conjoint analysis for the same low costs that using a simple sorting mechanism require.

DOI 10.1007/s12599-013-0295-z

The Authors

Dr. Jella Pfeiffer (✉)
Information Systems and Business Administration
Lehrstuhl für Wirtschaftsinformatik und BWL
Johannes Gutenberg-Universität Mainz
Jakob-Welder-Weg 9
55128 Mainz
Germany
jella.pfeiffer@uni-mainz.de

Prof. Dr. Michael Scholz
Juniorprofessur für Wirtschaftsinformatik mit Schwerpunkt E-Commerce
Universität Passau
Innstr. 43
94032 Passau
Germany
michael.scholz@uni-passau.de

Received: 2012-11-29
Accepted: 2013-05-08
Accepted after one revision by Prof. Dr. Hinz.

This article is also available in German in print and via http://www.wirtschaftsinformatik.de: Pfeiffer J, Scholz M (2013) Produktempfehlungssysteme mit minimalem Konsumentenaufwand und hoher Genauigkeit. Ein neuer Ansatz mit rangbasierter Pareto-Front. WIRTSCHAFTSINFORMATIK. doi: 10.1007/s11576-013-0388-9.

Electronic Supplementary Material
The online version of this article (doi: 10.1007/s12599-013-0295-z) contains supplementary material, which is available to authorized users.

1 Introduction

Online consumers today burrow through vast amounts of product information to find the best match to their preferences. This has boosted the popularity of recommendation systems promising to decrease consumers' search costs. Recent work has focused on collaborative-filtering recommendation systems and content-based recommendation systems (Adomavicius and Tuzhilin 2005). Such recommendation systems are, however, subject to several major problems that often reduce the quality of the recommendation (Ansari et al. 2000, pp. 364–365). One of the most important problems of both systems is the start-up problem: a buying profile consisting of several products is required for various consumers before a recommendation is possible. Collaborative-filtering systems also have a start-up problem when a product is new because no buying history for this product exists. Consequently, these systems can produce particularly inaccurate recommendations due to the missing data.

Utility-based recommendation systems aim to reduce the start-up problem by requiring the consumers to actively input their preferences into the system and elicit current and complete consumer preference profiles to compile a list of recommendations (Cao and Li 2007, p. 232; De Bruyn et al. 2008, p. 445; Huang 2011, p. 398; Scholz and Dorner 2012, p. 2; Xiao and Benbasat 2007, p. 139; Ansari et al. 2000, p. 365). Thus, neither historical data on the consumer nor the purchasing history of a product is needed. Recent utility-based recommendation systems have achieved a high level of accuracy by using sophisticated preference measurement approaches from marketing research that hedge individually elicited preferences with preferences from other consumers. Thus, in contrast to collaborative-filtering and content-based recommendation systems, there is no start-up problem with new consumers and a reduced start-up problem with new

Published online: 26 October 2013

product categories[1] when applying such sophisticated measurement approaches. Ansari et al. (2000, p. 365), for example, proposed a utility-based recommendation system that uses a choice-based conjoint analysis with a hierarchical Bayes estimation (Moore et al. 1998, p. 204). However, the high accuracy of utility-based recommendation systems is at the expense of higher consumer effort for inputting preferences.

Systems that require high consumer effort are usually not well adopted by consumers. This is reflected by the high diffusion of the low-effort content-based and collaborative-filtering recommendation systems in comparison to utility-based recommendation systems. Prominent collaborative-filtering and content-based examples can be found on Amazon.com, Last.fm or movie databases such as www.jinni.com or www.tastekid.com.

We introduce a new approach – ranked Pareto-fronts (RPF) – which requires low consumer effort, achieves high accuracy and does not suffer from the start-up problem. As the first step, it eliminates inferior alternatives by calculating the Pareto-front and, as the second step, it ranks the Pareto-efficient alternatives. Based on a cost model, we show in a simulation study that our new approach outperforms choice-based conjoint recommendation systems because it achieves the same accuracy with lower effort by consumers. In fact, it requires the same low effort as simple systems which, for example, only allow sorting products according to price.

This work does not only introduce a new methodology that is interesting for researchers, but is also highly relevant for e-shops. Indeed, we provide advice about the most suitable recommendation system for a certain type of consumer and our new easy-to-implement approach can help improve recommendation systems for a wide range of consumer types.

In the next section, we briefly introduce utility-based systems as systems that overcome several problems of common recommendation systems. We present the RPF as a substitute for utility-based recommendation systems with low consumer effort in Sect. 3. The utility and cost model used in our simulation study to measure effort and accuracy of the system is explained in Sect. 4. We present the design of our simulation together with a verification and validation in Sect. 5, and the results are discussed in Sect. 6. Finally the paper concludes with implications for practitioners and researchers, as well as limitations.

2 Recommendation Systems

2.1 Collaborative-Filtering and Content-Based Systems

Collaborative-filtering assumes consumers with a similar buying history have similar preferences. Thus, if consumer s is similar to consumer r, then products s has bought are recommendable for r (Adomavicius and Tuzhilin 2005, p. 737). Content-based recommendation systems, in contrast, use historical data of the same consumer to form recommendations (Adomavicius and Tuzhilin 2005, p. 735). These systems suffer from several major problems (Ansari et al. 2000, p. 365). First and foremost, they involve the start-up problem described in Sect. 1. Second, the buying profiles often include products that do not reflect the consumers' preferences, since these products have been, for example, bought as gifts. Third, the buying profile only consists of historical buying decisions and hence might not reflect a consumer's current preferences. And finally, collaborative-filtering systems do not provide any information as to why a particular product is recommendable.

2.2 Utility-Based Systems

Utility-based recommendation systems elicit current and complete consumer preference profiles to compile a list of recommendations (Huang 2011, p. 398). They estimate utility functions through explicit interaction with the consumer and are thus in contrast to content-based recommendation systems that like utility-based systems also use the product attributes to compile recommendations (Burke 2002, p. 334) which allow more sophisticated analyses, such as the computation of consumers' willingness-to-pay (Scholz and Dorner 2012). Utility-based recommendation systems are almost always based on the multiple attribute utility theory.[2] As the first step, utility-based recommendation systems estimate a consumer-specific utility function for each attribute that describes the product, such as price or brand. Based on these functions the system generates an ordered list of recommendable products as the second step. As the third step, the consumers walk through the list of products to find their preferred product.

Other, simpler utility-based recommendation systems only allow consumers to specify their most preferred attribute, i.e. they allow sorting products according to their most important attribute (Importance Sorting). Consequently, they assume that (i) the utility of a product is mainly determined by only one attribute and (ii) all consumers have the same preference order for this attribute (for instance, the lower the price, the better the utility). As price is often assumed to be the most important attribute, several recommendation systems only concede consumers to sort in a descending order by price (Price Sorting). More sophisticated systems that estimate a complete multi-attribute utility function are based on methods such as choice-based conjoint analysis (CBC). CBC requests their consumers to choose the preferred product among a small set of products in a number of choice tasks (Fritz et al. 2011, p. 272). This approach is easy for consumers because the choice tasks in a CBC resemble actual choices (Moore 2004, p. 300). However, methods such as CBC that are traditionally used to estimate utility functions for market segments are impractical for the application in recommendation systems because of the high consumer effort (De Bruyn et al. 2008, p. 446). Ansari et al. (2000, pp. 365–375) thus propose using hierarchical Bayes estimation as the core of utility-based recommendation systems. De Bruyn et al. (2008) seize this suggestion and develop a recommendation system that recommends products based on utility functions estimated by a CBC with hierarchical Bayes estimation. Consumers answer simple questions to reveal their preferences. These answers are then used to estimate utility functions. Although consumers only need to answer two simple questions to get recommendations, the provider of such a system has to conduct

[1] Such approaches, however, require data from other consumers in the same product category (roughly 100 other ratings in the same product category are sufficient).

[2] For an introduction into multiple attribute utility theory see Wallenius et al. (2008).

Table 1 Example of dominated products

Product	Photo resolution	Zoom factor	Price	Dominated by	Overall utility (U_p)
A	u_{ph} (11.6 MP) = 0.6	u_{zf} (11.5×) = 0.3	u_{pr} (610 EUR) = 0.2	B, D	1.1
B	u_{ph} (13.8 MP) = 0.8	u_{zf} (15.0×) = 0.4	u_{pr} (470 EUR) = 0.4	None	1.6
C	u_{ph} (10.5 MP) = 0.5	u_{zf} (18.5×) = 0.5	u_{pr} (540 EUR) = 0.3	None	1.3
D	u_{ph} (12.7 MP) = 0.7	u_{zf} (15.0×) = 0.4	u_{pr} (470 EUR) = 0.4	B	1.5
E	u_{ph} (9.4 MP) = 0.4	u_{zf} (11.5×) = 0.3	u_{pr} (260 EUR) = 0.7	None	1.4

a CBC for each product category that is offered to the consumers.

2.3 Pareto-Front Approach

CBC is more accurate than Importance Sorting (see Sect. 2.2) in providing recommendations that fit the consumer's preferences but also costs the consumer much more effort. We aim to provide an approach with high quality recommendations without requiring the consumer to input preference information. This approach shows the consumer only non-dominated products. A product dominates another product if all attributes are at least as good and, for at least one attribute, it is strictly better. Assume for example the following set of digital cameras which are described by utility values for photo resolution (u_{ph}), zoom factor (u_{zf}) and price (u_{pr}) (see **Table 1**).

Although product *D* has the second highest overall utility value, as well as the second best photo resolution, zoom factor, and price and might, therefore, be recommendable, it is dominated by product *B* and will not be considered by any rationally thinking consumer. A product that is not dominated by any other product is called Pareto-efficient. The set of all Pareto-efficient products finally forms the Pareto-front (in the example, products *B*, *C* and *E* are Pareto-efficient). Narrowing the set of available products to the Pareto-front offers two advantages. First, the number of products in the recommendation list can be decreased without any consumer input. And second, it increases the objective decision quality (Aksoy et al. 2011 p. 113; Häubl and Trifts 2000, p. 8; Payne et al. 1993, p. 34).

Recommendation systems that use the concept of Pareto-fronts rely on less consumer input, because instead of eliciting utility functions, only the preference order over each attribute must be known (the same assumption that we formulated for Importance Sorting). For determining the Pareto-front in the example, we do not need to know the utility values but we can compare the product attributes directly with another: Product *D* is dominated by B because 12.7 MP < 13.8 MP and the price and the zoom factor are the same. For most attributes the preference order for the different attributes is homogeneous among consumers. For example, we can assume that all consumers would prefer digital cameras with lower prices and higher zoom factors, given everything else is equal. Yet, some attributes will have a heterogeneous preference order like color or design. For these ones, the consumers need to reveal the preference order, by indicating which colors, designs, etc. they prefer over others. In summary, for determining which product dominates other products, we only need to judge whether a product attribute and not to what extent it is better than another one. Thus, an ordinal preference order for each attribute is sufficient and, in contrast to sophisticated utility-based systems like CBC, we do not need to know utility values.

3 A New Approach: The Ranked Pareto-Front

Products that are Pareto-efficient are handled as equivalent when using the Pareto-front concept. However, some Pareto-efficient products have a higher probability to be the designated best product, because they are dominant over other Pareto-efficient products in several attributes. Consider, for example, products *B* and *E* from **Table 1**. *B* dominates *E* in two attributes (photo resolution and zoom factor) and is dominated in only one attribute (price) by *E*.

In our new approach, we suggest computing a rank for all Pareto-efficient products that is based on the number of dominant attributes compared to all other Pareto-efficient products. We compare a particular product p_1 with all other products on the Pareto-front and increment the rank of p_1 by one for each attribute *i* if its attribute level (x_{ip1}) is better than that of p_2. **Figure 1** demonstrates the entire algorithm.

Since products *B*, *C* and *E* are Pareto-efficient, our proposed algorithm runs on these products only. Product *B* dominates both products *C* and *E* in two attributes each, because it has a better photo resolution than *C* and *E* (13.8 mega pixel is better than 10.5 mega pixel is better than 9.4 mega pixel), a better zoom factor than *E* (15× is better than 11.5×), and a better price than *C* (470 Euro is better than 540 Euro). It therefore

Algorithm Ranked Pareto-front

```
for a = 1 → Length(paretoEfficientProducts) do
  p1 ← paretoEfficientProducts[a]
  for b = a + 1 → Length(paretoEfficientProducts) do
    p2 ← paretoEfficientProducts[b]
    for i = 1 → Length(attributes) do
      x_ip1 ← p1[i]
      x_ip2 ← p2[i]
      if (isPositive(i) and x_ip1 > x_ip2) or (isNegative(i) and x_ip1 < x_ip2)
      then
        ranks[a] ← ranks[a] + 1
      else if (isPositive(i) and x_ip1 < x_ip2) or (isNegative(i) and x_ip1 >
      x_ip2) then
        ranks[b] ← ranks[b] + 1
      end if
    end for
  end for
end for
```

Fig. 1 Pseudo-code of the ranked Pareto-front algorithm gets a rank of 4. Product C dominates B in one and E in two attributes resulting in a rank of 3. Finally, product E dominates both B and C in one attribute each and is, with a rank of 2, the worst Pareto-efficient product. In this example product B is most likely to be the best product (having the highest rank) and will be recommended first.

In our example, all three Pareto-efficient products have a different rank. When applying the RPF approach to a larger set of products, some products might have the same rank. We suggest sorting products with paired ranks in random order because they should be equally likely to be chosen.

Let us analyze different types of decision-making strategies to verify that RPF is a promising recommendation system approach and has a higher probability to be chosen by consumers.

As B dominates E in two attributes, but E dominates B in only one attribute, B has a higher probability to be the best product when using a lexicographic decision strategy. The first step of a lexicographic strategy is to build an importance ranking for all attributes. Thereafter the consumers select the product that best fulfills the most important attribute. Assuming that each attribute is equally likely to be ranked as the most important one by a consumer, in our example B has a chance of 66.67 % to be preferred to E when using a lexicographic strategy because two out of three attributes are better than E.

We have a similar argument for consumers that eliminate products with inappropriate attribute levels. B has a higher chance than product E of not being eliminated due to higher levels in two out of three attributes.

Consumers who additively accumulate weighted attribute utilities and thus use a utility-maximizing strategy will also benefit from an RPF since the lower attribute utility of price for product B can be compensated by a higher attribute utility for two attributes (resolution and zoom factor), whereas E has a higher attribute utility for only one attribute. We therefore assume more accurate results for different decision strategies when using the RPF concept instead of the simple Pareto-front concept which does not sort alternatives according to their rank.

In summary, we claim that a product with a higher rank is more likely to be purchased than a product with a lower rank, because products with a higher rank (i) have a higher chance to provide the best level of the most important attribute, (ii) have a higher probability to fulfill consumer-specific aspiration levels (i.e. minimal and maximal acceptable attribute levels), and (iii) have a higher probability to offer a high utility value that is accumulated based on attribute utility values. It has to be noted that the RPF is a multi-criteria decision approach and does not compute a utility value for each product, but only a probability of being the best product.

We introduce a utility and cost model for measuring the accuracy and effort of the discussed recommendation systems in the next section. These models have shown evidence in several recent investigations.

4 A Utility and Cost Model

4.1 Utility Model

The utility U_p of a product p is usually the sum of weighted single-attribute utility functions $u_i(x_{ip})$ (Butler et al. 2008, p. 751; Montgomery et al. 2004, p. 193)

$$U_p = \sum_{i=1}^{m} w_i u_i(x_{ip}), \tag{1}$$

where m is the number of attributes, w_i is the weighting of attribute i, $\sum_{i=1}^{m} w_i = 1$, reflecting how important the attribute is for the consumer, and x_{ip} is the attribute level that describes attribute i of product p and $u_i(x_{ip})$ is the attribute level utility.

Following other works, we express the single-attribute utility function as an exponential function with diminishing or increasing but monotonous utilities (Butler et al. 2001, p. 805; van Ittersum and Pennings 2012, p. 95)

$$u_i(x_{ip}) = a_i - b_i e^{c_i x_i}, \tag{2}$$

where a_i, b_i and c_i are scaling constants.

4.2 Cost Model

As explained in Sect. 2, the interaction with a utility-based recommendation system typically comprises three steps ((1) estimation of utility functions, (2) generation of a recommendation list and (3) consumers' choice from the recommendation list). Consequently, from the consumer's viewpoint, the costs for the complete choice process, C_{choice}, consist of the costs C_{pref} for measuring preferences as the first step, the costs C_{wait} for waiting for the system to respond as the second step, and the costs C_{select} for selecting an alternative from the recommendation set as the third step. The costs C_{wait} will not be modeled. Because of low expected run times, this term should be negligibly low and will hardly be perceived as waiting time by the consumer.

In summary, C_{choice} consists of mental costs for measuring preferences (e.g., conducting a conjoint analysis) and of mental costs for evaluating the set of recommended products and making a final decision:

$$C_{choice} = C_{pref} + C_{select}. \tag{3}$$

C_{pref} will differ with respect to which preference measurement method is used and C_{select} depends on the quality of the recommendation list. The better the order of products in the list, the earlier the consumer will be satisfied with a highly recommended product and stop the selection process. Furthermore, both cost blocks depend on the consumer's individual decision behavior.

For determining the individual's mental costs, we follow the effort-accuracy framework by Johnson and Payne (1985, p. 399). This framework is based on the theory that decision makers trade-off between effort and accuracy when choosing which decision strategy to best apply in the given decision environment. Johnson and Payne propose the following list of elementary information processes (EIPs) for determining the effort for applying a strategy. Each EIP describes a mental step that a decision maker might apply: *ADD*, *COMPARE*, *DIFFERENCE*, *ELIMINATE*, *MOVE*, *PRODUCT* and *READ*. **Table 2** describes the different steps.

C_{pref} and C_{select} can be determined by summing up the EIPs which each consumer needs to provide during the preference measurement and for selecting alternatives. Estimates exist for weighting the effort of EIPs in order to reflect that some EIP are more effortful than others (Bettman et al. 1990, p. 131; Johnson and Payne 1985, p. 406; Lohse and Johnson 1996, p. 31). We take the estimates by Johnson and Payne (1985) because Bettman et al. (1990) report problems with increased values caused by effortful consumer's mouse movements. The estimate for READ is from Lohse and Johnson (1996) because they use a more sophisticated method (eye-tracking) than

Table 2 Elementary Information Processes

EIP	Effort	Description
ADD	0.9	Add two values
COMPARE	0.3	Two pieces of information are compared
DIFFERENCE	n.a.	The difference of two values is calculated
ELIMINATE	0.3	Remove an alternative or an attribute from consideration
MOVE	0.23	Go to another piece of information
PRODUCT	1.2	An attribute level utility is multiplied with an attribute weight
READ	0.23	One piece of information is read into the short term memory

```
Algorithm  WADD
for p = 1 to n do
  {Compute the attribute level utilities for each attribute.}
  for i = 1 to m do
    READ attr_i                          } mn
    READ x_ip
    if i == 1 then
      PRODUCT U_p = w_i * u_i(x_ip)      } n
    else
      PRODUCT prod = w_i * u_i(x_ip)     } (m-1)n
      ADD U_p+ = prod
    end if
  end for
  {Keep the alternative with the higher utility.}
  if p > 1 then
    ELIMINATE alt_p with lower U         } n-1
  end if
end for
```

Fig. 2 EIP cost model for consumers using a weighted additive decision strategy

Johnson and Payne to estimate the effort for READ. The effort for DIFFERENCE is not used in this paper, because none of our strategies use DIFFERENCE.

In order to account for individual differences of human-decision behavior, about fifteen decision-strategies have been identified so far in literature (Pfeiffer 2012, p. 20; Svenson 1979, p. 89). One example is the normative weighted additive rule (WADD). It assumes that a decision maker computes the utility of each product, U_p, and chooses the product with the highest utility. For determining C_{select} of a consumer that additively accumulates weighted attribute utilities, we suggest to use pseudo-code notation to note the sequence of EIPs that is needed as is shown in **Fig. 2**. For each of the n alternatives and the m attributes, two *READs* are needed for reading in the attribute and the attribute level. Computing the weighted additive utility function costs m-times *PRODUCT* and $m-1$ times *ADD* for each of the n alternatives. Finally, for finding the alternative with the highest utility value, the consumer needs $n-1$-times *ELIMINATE*. Thus, whenever the consumer applying a WADD-strategy has to choose an alternative from n alternatives that are described by m attributes, he will need $C^{WADD}_{select} = 2mnREAD + mnPRODUCT + (m-1)nADD + (n-1)ELIMINATE$. Similar analyses for determining C_{select} can be done for other strategies (see **Fig. A-1**, **Fig. A-2** and **Fig. A-3**).

The number of alternatives which the consumer considers from the recommendation list depends on the quality of the ordering of the products and the consumer's stopping behavior. A large number of studies investigated when people stop the search for a better alternative and identified stopping rules (Hey 1982, p. 65). The bounce rule, for example, assumes that the search is stopped at p_{t+l}, if $U(p_t) \geq U(p_{t+l})$, $\vee l$ where $l = 1$ is the *one bounce rule* and $l = 2$ is the *two bounce rule*, etc. (Hey 1982, p. 73). For example, a consumer applying a two bounce rule would stop her search in the recommendation list as soon as the utility values of three consecutive products are monotonously decreasing. For l-bounce stopping rules holds that the better the recommendation list is sorted decreasingly according to the consumer's utility, the earlier the consumer stops the search and thus the fewer alternatives he considers.

The approach used above for C_{select} can be easily transferred to determine C_{pref} for a CBC-based recommendation system. In a CBC analysis, a consumer is confronted with ct choice tasks, where each choice task has n_{ct} alternatives. Thus the total number of alternatives considered for a CBC analysis will be $n = ct \times n_{ct}$ and C_{pref} for the CBC analysis is determined as is C_{select} because for modeling decision behavior in choice task the same decision strategies can be used as for modeling the choice of an alternative from a recommendation list. Because alternatives in choice tasks are not sorted, we do not assume that consumers apply a stopping rule but rather consider all n_{ct} alternatives in each choice task. In CBC analysis, n_{ct} is usually low (e.g., three or four).

5 Simulation

We have developed a new IT artifact in form of a method that is relevant for solving the recommendation problem that many e-commerce companies face. We have also shown research rigor in the constructions of this method by relying on well-established theoretical foundations (Hevner et al. 2004, p. 88): We extend a well-established approach (Pareto-front) with respect to the theory of consumer decision behavior (our approach is designed to rank Pareto-efficient products according to the probability of being the best product for several consumer decision strategies). Another crucial step of the design science cycle, which we address in this section, is the design evaluation with rigorous methods (Hevner et al. 2004, pp. 82–87). As evaluation criteria, we determine both effort and accuracy of the compared systems with the effort-accuracy framework by Johnson and Payne (1985), since it is well accepted as theoretical framework in the field of information systems (Todd and Benbasat 1992, p. 376, 1994, p. 538).

We will evaluate the different recommendation systems with an agent-based model instead of, for example, a laboratory experiment, an analytical approach or a field study, because of several reasons and thus will follow the approach of other works with comparable research questions (e.g. Hinz and Eckert 2010; Hostler et al. 2011). Agent-based models allow for modeling and analyzing heterogeneous behavior and can capture the fact that individuals have different utility functions and apply different decision strategies. In contrast, obtaining results that allow implications for particular decision strategies in empirical experiments is difficult because it is hard to observe which exact decision strategy a consumer has applied (Bröder and Schiffer 2003, p. 196; Rieskamp and Hoffrage 2008, p. 262). Moreover, we have to investigate a sequence of interdependent decisions due to the answers an individual provides in the preference measurement part determining the order of recommendations and with it the individual's final choice. Such complex and dynamic interactions would be difficult to show in an analytical model. Therefore, we conduct a simulation in which we model 1000 consumers who make a buying decision while using different recommendation systems. This allows us to make clear statements about different consumer types. Each consumer is described by (i) a utility function, (ii) a decision strategy and (iii) a stopping rule. Consumers first interact with the recommendation system and thereafter evaluate recommended products until the stopping rule terminates the evaluation. At the end each consumer selects a product.

5.1 Simulation Design

We conducted a pre-study in order to be able to determine realistic parameters for our agents. In the pre-study 50 participants revealed their utility functions and aspiration levels for buying a new digital camera. Utility functions were elicited using a self-explicated approach that consisted of two steps. In a first step, the consumers revealed their aspiration levels for each attribute and in the second step they specified the attribute weightings as well as the parameters for exponential single-attribute utility function (see (2)). Finally, consumers rated the ten most recommendable cameras. An R^2 of 0.51 as well as an RMSE of 1.23 indicate that the sequence of the presented recommendable cameras was perceived as rather accurate by the consumers. This indicates that the utility functions were reliably and validly estimated.

5.1.1 Utility Function

We estimated distributions for each exponential single-attribute utility function parameter (a_i, b_i and c_i) and each attribute weight w_i from the results of the above mentioned laboratory experiment. Attribute weights were measured as a discrete value on an 11-point scale. To estimate the parameters a_i, b_i and c_i we assumed that the best attribute level has a single-attribute utility value of 1 whereas the worst attribute level has a utility value of 0. The participants in the experiment specified the utility value for the average level (i.e. the attribute level that is exactly between the best and the worst) of each attribute using an 11-point scale. The value specified for the average attribute level was than transformed into [0; 1]. These utility values allowed us to compute the parameters a_i, b_i and c_i. We used digital cameras described by photo resolution, zoom factor, size, display size, video resolution, number of settings, light sensitivity, and price in both the laboratory experiment and the simulation. Each attribute is described by six (for photosensitivity) to 117 (for size) different attribute levels.

Aspiration levels were directly specified by the experimental participants. We used these data to (i) estimate for how many attributes consumers specify aspiration levels and (ii) how restrictive consumers specify aspiration levels. A complete list of all parameter distributions estimated based on the laboratory experiment and used in our simulation can be found in **Table A-1**.

5.1.2 Decision Strategies

We choose the most common decision strategies (Wang and Benbasat 2009, p. 3; Yee et al. 2007, p. 534) which we evenly distribute over our 1000 simulated consumers.

- WADD (Weighted Additive Rule): This is the normative rule in the decision making literature. It assumes that a decision maker computes the utility for each product, U_p.
- EBA (Elimination by Aspect Strategy): Decision makers sort the attributes i according to their weight w_i. Starting with the attribute with the highest weight, they iteratively remove products if the value of the ith attribute does not meet the aspiration level for this attribute. All attribute levels that fulfill the aspiration level are acceptable for the decision maker. The strategy stops if there is only one alternative left or all attributes are considered. We follow Yee et al. (2007, p. 534) and assume a deterministic process where the order in which attributes are considered is decreasing according to w_i.
- LEX (Lexicographic Heuristic): Decision makers consider the attribute i with the highest w_i and select the product whose attribute level x_{ip} has the highest value: $u_i(x_{ip})$. If this returns more than one alternative, they iteratively compare the remaining alternative across the next most important attribute until there is only one product left.
- CONJ (Conjunctive Strategy): A decision maker removes a product if at least one of its attribute levels violates an aspiration level. If all considered products satisfy the aspiration levels, the user chooses randomly among them.

5.1.3 Stopping Rule

We assume that consumers will not consider all alternatives in the third step of the recommendation system. Rather, we assume consumers to rationally and consistently use the bounce rule to make a final selection because the system saves them the effort to consider all alternatives. Consumers of type EBA using a two bounce rule, for example, would use EBA to compare the three alternatives that are ranked highest. If the one with the highest rank wins, she stops the search. Otherwise, she compares the winning alternatives with alternatives four and five from the recommendation list with EBA, etc.

5.1.4 Recommendation Systems

We implement a benchmark system that presents products in a random order to the consumers (Random Sorting). As a simple utility-based recommendation system we implement a system that only allows sorting products according to their price and a system that allows sorting products according to the consumers' most important attribute. We further implement a CBC-based recommendation system with hierarchical Bayes estimator by using the R-package "bayesm" (Rossi et al. 2005). Our estimator is a hybrid Gibbs sampler with a random walk Metropolis step for the multinomial logit coefficients. The hierarchical multinomial logit model is specified with mixture of normals heterogeneity to overcome the strict assumptions of a normal multinomial logit model (Fiebig et al. 2010, p. 397). We generated 20,000 Markov Chain Monte Carlo draws from which

Table 3 Overview of the compared recommendation systems

System	Consumer interaction (C_{pref})	System computation (C_{wait})	Result (C_{select})
Random sorting	None	None	Select a Product from the Recommendation List
Price sorting	None	Sort products according to their price	
Importance sorting	Enter most important attribute[a]	Sort products according to the consumer's most important attribute	
Pareto-front	None[a]	Compute Pareto-front	
CBC	Choose the preferred product in several choice tasks	Compute multi-attribute utility functions based on a multinomial regression with hierarchical Bayes estimator	
Ranked Pareto-front	None[a]	Compute Pareto-front and probabilities to be the best product	

[a]Importance Sorting and the two Pareto-front approaches only require consumer input for heterogeneous preference orders

the first 10,000 were used for burn-in and the last 10,000 were used for estimating the utility functions. The consumers of the CBC-based recommendation system have to choose one product out of three or take the no choice option in 25 choice tasks. We use 20 of the choice tasks for estimating the consumers' utility functions and 5 choice tasks (hold-out tasks) for computing the validity of the estimated functions. The stimuli for the choice tasks were generated using a D-optimal design[3] (Kanninen 2002).

A fifth system presents products that are Pareto-efficient in a random order. We used LESS (Linear Elimination Sort for Skyline) as fast exact algorithm to compute all Pareto-efficient products (Godfrey et al. 2006). 79 out of 130 cameras were identified as Pareto-efficient. These 79 cameras were presented in a random order. We further implemented our proposed algorithm of an RPF in a sixth recommendation system. This system uses the algorithm presented in **Fig. 1** to sort the Pareto-efficient products. Our simulated consumers use all the recommendation systems presented above to find an appropriate digital camera out of 130 cameras. Each consumer evaluates the final recommendation set of each system and selects an appropriate product based on the stopping rule discussed above. When using Importance Sorting and the CBC-based recommendation system, it additionally requires effort from our consumers to interact with the system ($C_{pref} > 0$). If two products are assigned the same rank with the RPF system, then they are displayed in random order. From the 130 products, only 8 Pareto-efficient solutions shared the rank with one other product.

Table 3 gives an overview of the recommendation systems that will be compared in this work. CBC and Importance Sorting require consumer interaction before they can compute the recommendation list. The costs for measuring preferences with a CBC and for selecting a product out of the recommendation list of any system are computed as described in Sect. 4.2. Importance Sorting furthermore requires selecting the most important attribute. Each consumer therefore reads in each attribute (m-times READ with each 0.23 EIP) and compares them (($m-1$)-times COMPARE with each 0.3 EIP) to find the most important attribute, resulting in 3.94 EIPs for the eight attributes of the digital cameras. After the consumer-interaction part (if existent), the recommendation list needs to be generated. The two Pareto-front approaches compute the set of Pareto-efficient solutions first before they can start sorting products. The RPF approach, in addition, needs to determine the ranks (with algorithm 1, see **Fig. 1**). The CBC needs to compute the multi-attribute utility function based on the consumer input in the first step (see Sect. 4.1). The Pareto-front approach presents the Pareto-efficient products in random order on the recommendation list. The others sort products according to some system-specific criterion (prices, utilities, ranks).

5.2 Verification and Validation

For ensuring the rigor of our simulation model, we follow the guidelines by Rand and Rust (2011, pp. 7–10) for verification and validation. For validating that our implemented model corresponds to reality, we follow four steps: micro-face validation, macro-face validation, empirical input validation and empirical output validation.

We ensure that the input parameters used to define the model correspond in a meaningful way to real-world individuals in the micro-face validation by defining buying decision parameters that have been extensively investigated in recent research. First, we model utility functions that form the theoretical foundations of several utility estimation approaches such as self-explication approaches and conjoint analysis (Green et al. 2001, p. 60). We furthermore parameterize these functions based on the findings of the laboratory experiment (see Sect. 5.1). Second, agents are characterized by decision strategies that have been found to describe actual human-decision behavior (Biggs et al. 1985, p. 975; Olshavsky 1979, p. 306).

The macro-face validation ensures that the aggregate patterns correspond to real-world patterns. We warrant a macro-face valid model by implementing four different strategies that cover the variety of human decision behavior. Since there is no consumer interaction between consumers but only between the recommendation system and the consumer, we validate aggregate patterns by observing that different consumer types as defined on the micro-level lead to a heterogeneous population of consumers on the macro level. On the one extreme with the WADD strategy, we incorporate individuals into the model who invest high effort

[3]A D-optimal design seeks to minimize the set of stimuli necessary to investigate the main effects (here attributes) of the variance of an observed variable (here product choice).

Table 4 Mean of consideration set sizes for different stopping rules. Standard deviations are in parentheses

System	1 bounce rule	2 bounce rule	3 bounce rule
Random sorting	5.04 (14.14)	6.92 (14.79)	14.91 (14.92)
Price sorting	8.73 (20.41)	10.83 (20.26)	15.13 (19.94)
Importance sorting	6.12 (17.25)	7.70 (17.14)	9.83 (17.22)
Pareto-front	4.89 (9.48)	6.60 (9.41)	9.30 (9.49)
CBC	3.96 (13.04)	5.96 (13.95)	6.99 (12.93)
RPF	3.88 (8.63)	5.24 (8.55)	6.89 (8.64)

Table 5 Validity of the recommendation systems without selecting a product

System	EBA	CONJ	LEX	WADD
Random sorting	−0.003	0.004	−0.006	−0.007
Price sorting	−0.387	−0.367	−0.380	−0.389
Importance sorting	0.303	0.281	0.322	0.251
Pareto-front	0.012	0.013	0.012	0.011
CBC	0.702	0.450	0.547	0.869
RPF	0.868	0.860	0.865	0.864

in their decision during the decision process: A behavior that can be found in real-world decision-making contexts for consumers with, for example, high product involvement (Denstadli and Lines 2007, p. 126). We model individuals who invest little effort in their decision with the LEX strategy, a behavior that is typical for low product involvement (Hoyer 1984, p. 823). Consumers with high expertise typically select products with a CONJ strategy rather than a WADD strategy (Denstadli and Lines 2007, p. 125).

In order to ensure that our output corresponds to the real world, we compare the number of products that the agents consider in the selection phase (consideration set size) between the different systems. In accordance with the findings of Häubl and Trifts (2000, p. 15), the consideration set size decreases when a recommendation system is used. We furthermore compare the consideration set size of our simulation with the consideration set size of the participants in our pre-study. In our pre-study, we found that the consideration set consisted of minimally one, maximally ten and 5.62 products on average (SD = 2.78). As shown in **Table 4**, the results when using a 2 bounce rule match well with those of the advanced recommendation systems (CBC, RPF) and for the Random Sorting and the Pareto-front approach both a 1 and a 2 bounce rule fit well to our observed data. For the two simple sorting recommendation systems, a 1 bounce rule leads to better results than a 2 bounce rule. In summary, our simulation results for the consideration set size fit well to the results gathered from our pre-study with real respondents.

We also use the results for calibrating the parameter l of the bounce rule. We decided to use the two bounce rule, because by allowing the agents to consider more alternatives, it allows the simple recommendation systems to achieve a higher accuracy. Furthermore, the preference measurement method in our pre-study resembles more a CBC than any of the other recommendation system approaches, making the CBC the preferred approach for calibrating the parameter l for the bounce rule. The consideration set size can be found in the appendix for each strategy.

We discuss the results of our simulation with regard to our output variables cost and accuracy (see Sect. 5.2) in the next section. A detailed presentation of all results can be found in the online Appendix.

6 Results

6.1 Validity

We computed the validity of our CBC-based recommendation systems as the correlation of the predicted and rank based on the real utility function of the agents. **Table 5** indicates that (i) consumers using WADD especially benefit from an outstanding validity (0.869), and (ii) and the validity is moderately good also for consumers using other strategies (between 0.450 and 0.702).

The computed validity of the CBC-based recommendation system expresses the accuracy of the recommendation list. We also estimate the accuracy of the recommendation lists for all other systems as the correlation between the systems' predicted product rank and the product rank based on the true utilities. **Table 5** indicates that the recommendation list based on a random sorting or the Pareto-front approach is uncorrelated with the true product utilities (the coefficients are close to zero). Importance sorting produces a significantly better recommendation list than random sorting or the Pareto-front approach ($p < 0.001$), but a significantly worse list compared to a CBC-based recommendation system ($p < 0.001$). The proposed RPF approach produces the best recommendation lists for all investigated decision strategies (the CBC validity is comparable only for WADD: 0.869 to 0.864). This underlines the fact that the RPF computes the probability for any product of being the best product independent of the consumers' decision strategy. Price sorting is interestingly worse than random sorting.

As presented in **Table 6**, price is highly correlated with most other attributes. In case the single-attribute utility functions of two attributes have different signs of slopes (one is monotonously increasing (e.g., the more settings the better), the other decreasing (the higher the price, the worse)), a high correlation means that one attribute is in trade off to the other attribute. For price this is the case for most other attributes (zoom factor, screen resolution, video resolution, settings and photosensitivity). Considering only the price to sort products therefore leads to low utilities for those five attributes. Although price is typically one of the most important attributes, sorting only for price means neglecting other important attributes which leads to a negative correlation between the single-attribute utilities for price and the overall product utilities. Price sorting is thus inefficient and will scare off consumers.

The results in **Table 5** also indicate that the RPF outperforms other approaches

Table 6 Correlations between the product attributes based on the attribute levels (not the attribute level utilities)

	A1	A2	A3	A4	A5	A6	A7	A8
Resolution (A1)	–	0.039	−0.198	−0.212	0.019	0.036	−0.252	−0.041
Zoom factor (A2)	0.039	–	0.813	0.329	0.178	0.320	0.453	0.143
Size (A3)	−0.198	0.813	–	0.558	0.148	0.199	0.611	0.298
Price (A4)	−0.212	0.329	0.558	–	0.408	0.484	0.643	0.289
Screen resolution (A5)	0.019	0.178	0.148	0.408	–	0.458	0.236	0.132
Video resolution (A6)	0.036	0.320	0.199	0.484	0.458	–	0.243	0.253
Settings (A7)	−0.252	0.453	0.611	0.643	0.236	0.243	–	0.354
Photosensitivity (A8)	−0.041	0.143	0.298	0.289	0.132	0.253	0.354	–

in respect to the best ordering of products in the recommendation list. This result is independent of consumers' decision strategies or stopping rules. In the next subsection, we compare the complete search process (including the selection of a product) across all investigated systems.

6.2 Effort and Accuracy

To evaluate the different recommendation systems, we use two measures: the consumer's effort and the accuracy of the recommendation technique. The effort is expressed as described in the cost model (see Sect. 3) and for the accuracy (or quality) we use the utility value of the selected product normed in [0; 1] indicating the distance between the utility of the selected product and the utility-maximizing product.

Figure 3 presents the results. The further left and the further up a system is plotted in the diagram, the better it is. In summary, the RPF clearly outperforms the competing systems. Only for consumers applying a WADD strategy can a CBC system achieve a slightly better accuracy but at the expense of much more effort. A detailed table of our results can be found in the online Appendix in **Table A-2**.

Our results illustrate the trade-off between accuracy and effort when using a CBC-based recommendation system. Other utility-based systems such as Price Sorting or Importance Sorting systems are inferior to a CBC-based system in terms of accuracy, but superior in terms of effort. We also found again that a system that allows sorting products according to their price has the lowest accuracy. Presenting consumers an unsorted list of products (Random Sorting) leads to higher accuracy and allows selecting a product with comparable costs.

Between 80 and 90 % of the effort a consumer has when using a CBC-based recommendation system is caused by the choice tasks (C_{pref}). Although it might be possible to slightly reduce the number of choice tasks, the effort for such systems would be still higher compared to other systems.

Recommendation systems that present Pareto-efficient products (see Pareto Front in **Fig. 3**) are characterized by costs that are comparable to (EBA, CONJ, LEX) or significantly higher (WADD) than those when using an Importance Sorting system.[4] We also found significantly higher accuracy when using an Importance Sorting system compared to a system that presents randomly sorted Pareto-efficient products for consumers using an EBA or CONJ strategy, but significantly lower accuracy for consumers using WADD. A recommendation system presenting randomly ordered Pareto-efficient products is therefore only recommendable for consumers using a WADD strategy.

As **Fig. 3** demonstrates, our proposed RPF approach allows consumers to select an appropriate product at costs comparable to those of using an Importance Sorting system. However, the accuracy is significantly higher than (for EBA, CONJ, LEX) or comparable to that of a CBC-based recommendation system (for WADD). We therefore can conclude that our proposed approach is superior to Importance Sorting systems as well as CBC-based systems. Furthermore, our results strongly support our claim that a product with a better rank is more likely to be purchased than a product with a worse rank (see Sect. 3) because the quality of the ranking can be evaluated by the accuracy of the recommendation process of the system. The higher the accuracy, the better was the rank of the chosen product. First, our approach yielded better accuracy than other systems for consumers using EBA and CONJ, which both rely on consumer-specific aspiration levels. Second, our approach was better than others for consumers using LEX, which considers the most important attribute. Third, also for consumers using WADD, thus a strategy that chooses the product with highest utility value; our system was comparable to or better than other systems (CBC) in terms of accuracy.

Our proposed approach furthermore guarantees to select a Pareto-efficient product which avoids regret (Loomes and Sugden 1982, p. 805). 7.4 % of our consumers selected a product that is not Pareto-efficient when using the CBC-recommendation system. We interestingly found that all consumers selected a Pareto-efficient product when using the Importance Sorting system.

7 Discussion

7.1 Research Implications

We analyzed recommendation systems that do not need historical data and do not, therefore, suffer from the start-up problem. Utility-based recommendation systems, that fall into this category, require the consumer to explicitly input preferences which demand a high consumer effort. Based on the framework by Johnson and Payne (1985, p. 399), we developed a cost model to quantify the effort a consumer experiences for the product choice. As expected, our simulation showed that a utility-based CBC-recommendation system provides

[4] We conducted ANOVAs with Tukey's honest significance test to compare our recommendation systems in terms of effort and accuracy. We used an α-level of 0.05 to test for statistical significance.

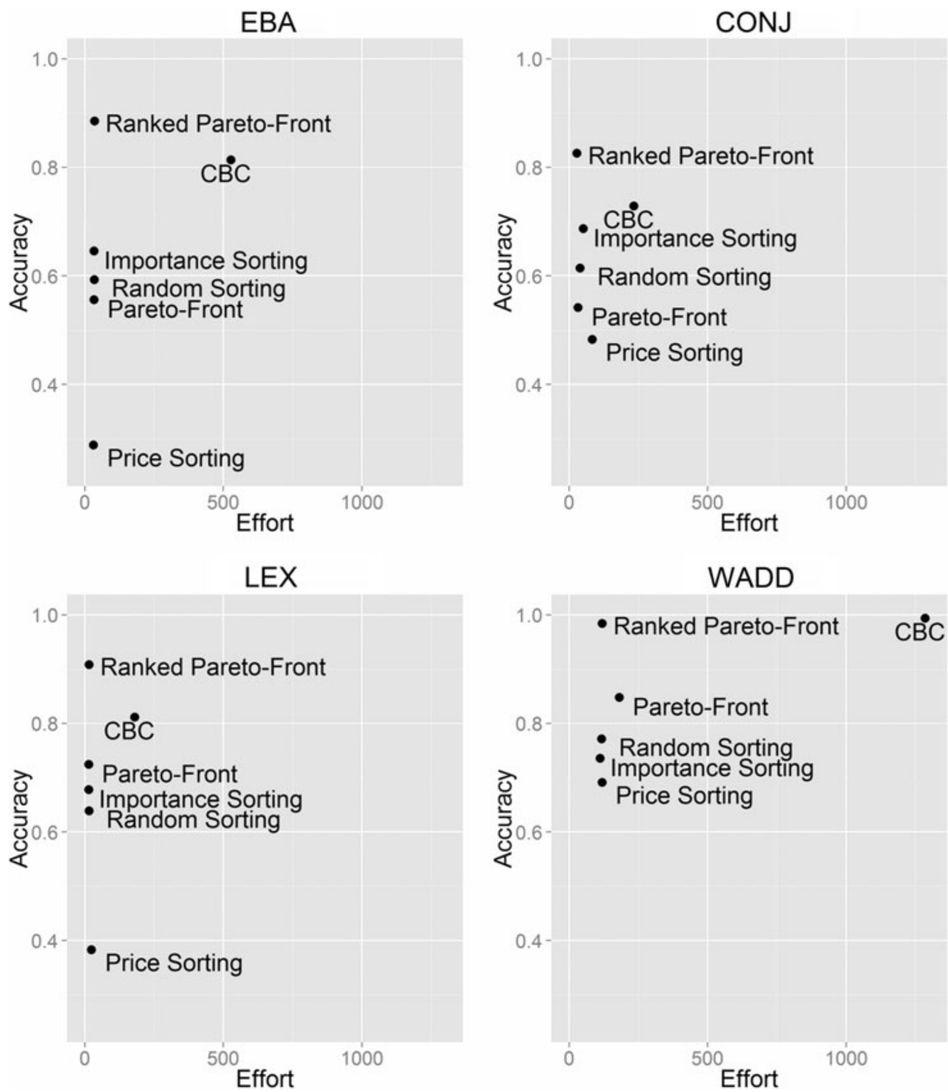

Fig. 3 Results across all decision strategies when using a two-bounce rule

very accurate recommendations. Yet, depending on the applied decision strategy, it requires roughly between 1.5 to 3 times more consumer effort than for easy importance-sorting according to the most preferred attribute. Between 80 % and 90 % of the cost of CBC-recommendation systems results from the preference measurement step which is why we suggested a new approach that does not require any input of consumer's preferences: the RPF approach. We recommend this approach because it outperforms CBC-recommendation systems in terms of accuracy in particular for consumer types that do not follow a utility-maximizing strategy, while requiring the same low costs as a simple sorting mechanism. Furthermore, fast algorithms for computing the Pareto-front exist and are easy to implement and the RPF approach does not suffer from the start up problem. Yet, CBC-recommendation systems should be chosen when market shares (Jedidi et al. 1996), optimal product configurations (Dorner and Scholz 2013, p. 10) or the individual's willingness's-to-pay (Gensler et al. 2012; Schlereth et al. 2012) are to be estimated in addition to providing recommendations.

7.2 Managerial Implications

Many e-shops sort products according to their price by default. We cannot recommend restricting sorting to a prespecified attribute, because it is clearly outperformed by allowing the consumer to choose the sorting criteria individually.

Sorting products according to their price is especially counterproductive because price is typically highly correlated with beneficial product attributes (e.g. the number of settings of a digital camera) but has in contrast to other product attributes an upside-down effect on the overall utility. Sorting products according to their prices in ascending order is therefore equal to sorting beneficial product attributes in ascending order. The higher the number of beneficial attributes a consumer considers, the more likely it is that products with low overall utility value are ranked highly when sorting according to price because of the trade-off between price and beneficial attributes.

Instead of offering sorting according to an attribute, we instead recommend using the RPF for compiling recommendation lists with little consumer effort. Our results indicate that consumers who did not apply a CONJ strategy considered only between three and six products on average when using the RPF and assuming consumers to use a two bounce rule. The RPF might, therefore, also be applicable if the number of products is much larger than in our simulation. In a dataset of 2778 apartments, for example, we found only 601 to be Pareto-efficient.

Most existing research studies focus on recommending products of only one category. Yet, in several situations, consumers prefer buying product sets. Consider a consumer who is going to start with digital photography. The consumer might require at least a camera body and a lens, while some consumers start with a more sophisticated set including a camera body, several lenses and an external flash. Applying the RPF can help to narrow down the overwhelming number of possible product sets at virtually no cost. In other problem domains (e.g., IT project portfolio management systems) the RPF might furthermore contribute to improving other decision support systems that identify valuable Pareto-efficient alternatives.

7.3 Limitations

Our approach is subject to three main limitations. First, the RPF requires that the products can be compared across a common set of attributes (i.e., they all belong to the same product category) and that the ordering on each attribute is equal for each individual. This might be true for attributes such as price or camera resolution (i.e., the cheaper, the better or the higher the resolution, the better), but not for nominal attributes such as color. An easy way to solve this problem would be to ask the consumer for her preference order. This can be achieved with little additional cost. If we assume that eight attributes are nominal and each attribute consists of ten levels, a consumer needs 141 EIPs to generate a preference order for all eight attributes when using an efficient sorting algorithm such as quicksort, introsort or bubblesort.[5] Taking into account that the difference between a CBC and an RPF in terms of effort is much higher than 141 EIPs (see Fig. 3), we still can claim that a CBC requires more consumer effort than an RPF.

Second, since the RPF approach presents a system in which no consumer input is required, the consumer might experience low control (Kamis et al. 2008, p. 171) and low social presence (Kumar and Benbasat 2006, p. 437). We therefore recommend allowing the consumer to interact with the recommendation system in the third step, for example, in form of filters that permit setting aspiration levels. This is particularly useful for EBA or CONJ consumers. Moreover, this approach would have the positive side-effect to even further decrease the consumer effort for selecting a product from the recommendation list.

Third, in contrast to laboratory experiments and field studies, in simulations the construction of external validity is harder because rigid assumptions might limit the ability to generalize the results. In our case, a non-representative sample of agents is a main thread to external validity (Wohlin et al. 2012, p. 110). We assume our agents to be purists in that they do not deviate from the process that the strategies prescribe. However, because of the very heterogeneous and highly adaptive nature of human decision-making, testing for all possible combinations and errors that decision-makers might make would be impossible. We thus decided to test four very different decision-types and one stopping rule which allow analyzing the systems across different decision contexts systematically. We decided to use the bounce rule in our simulation instead of other stopping rules (Hey 1982, p. 65) because other stopping rules would have forced us to make strong assumptions about the parameters (for example, the reservation price rule assumes that each decision maker stops the search once the products falls below her personal reservation price). Furthermore, the RPF already showed the best accuracy before the selection stage and, therefore, our results strongly support the superior performance of the RPF independent from the stopping rule.

To the best of our knowledge, this article is the first which adapts the Pareto-front concept in order to allow sorting Pareto-efficient products. Although we were not able to test all conceivable decision strategies and combinations of them, our cost model is designed in a way that makes it easily adaptable to further strategies.

In future work, we would like to test the systems that performed best in our study, namely the CBC-recommendation system, the RPF and the Importance Sorting, in an experiment with real consumers. While the simulation study was able to systematically analyze which system performs best for which consumer type by using objective measurements for accuracy and effort, an experiment would reveal consumer's perceived effort and perceived accuracy as well as the performance of the systems for many more kinds of different decision behavior that might occur in real-world settings.

References

Adomavicius G, Tuzhilin A (2005) Toward the next generation of recommender systems: a survey of the state-of-the-art and possible extensions. IEEE Trans Knowl Data Eng 17(6):734–749

Aksoy L, Cooil B, Lurie NH (2011) Decision quality measures in recommendation agents research. J Interact Mark 25(2):110–122

Ansari A, Essegaier S, Kohli R (2000) Internet recommendation systems. J Mark Res 37(3):363–375

Bettman JR, Johnson EJ, Payne JW (1990) A componential analysis of cognitive effort in choice. Organ Behav Hum Decis Process 45(1):111–129

Biggs SF, Bedard JC, Gaber BG, Linsmeier TJ (1985) The effect of task size and similarity on the decision behavior of bank loan officers. Manage Sci 31(8):970–987

Bröder A, Schiffer S (2003) Bayesian strategy assessment in multi-attribute decision research. J Behav Decis Mak 16(3):193–213

Burke R (2002) Hybrid recommender systems: survey and experiments. User Model User-Adapt Interact 12(4):331–370

Butler JC, Dyer JS, Jia J, Tomak K (2008) Enabling e-transactions with multi-attribute preference models. Eur J Oper Res 186(2):748–765

Butler JC, Morrice DJ, Mullarkey PW (2001) A multiple attribute utility theory approach to ranking and selection. Manage Sci 47(6):800–816

Cao Y, Li Y (2007) An intelligent fuzzy-based recommendation system for consumer electronic products. Expert Syst Appl 33(1):230–240

De Bruyn A, Liechty JC, Huizingh EK, Lilien GL (2008) Offering online recommendations with minimum customer input through conjoint-based decision aids. Mark Sci 27(3):443–460

Denstadli JM, Lines R (2007) Conjoint respondents as adaptive decision makers. Int J Mark Res 49(1):117–132

Dorner V, Scholz M (2013) Predicting and economically exploiting utility thresholds with utility-based recommendation systems. In: Proc of the 21st eur conf inform syst

Fiebig DG, Keane MP, Louviere J, Wasi N (2010) The generalized multinomial logit model: accounting for scale and coefficient heterogeneity. Mark Sci 29(3):393–421

Fritz M, Schlereth C, Figge S (2011) Empirical evaluation of fair use flat rate strategies for mobile Internet. Bus Inf Syst Eng 3(5):269–277

Gensler S, Hinz O, Skiera B, Theysohn S (2012) Willingness-to-pay estimation with choice-based conjoint analysis: addressing extreme response behavior with individually adapted designs. Eur J Oper Res 219(2):368–378

Godfrey P, Shipley R, Gryz J (2006) Algorithms and analyses for maximal vector computation. VLDB J 16(1):5–28

[5] Sorting a set of α attribute levels costs $\alpha ld(\alpha)$ READ and COMPARE operations.

Abstract

Jella Pfeiffer, Michael Scholz

A Low-Effort Recommendation System with High Accuracy

A New Approach with Ranked Pareto-Fronts

In recent studies on recommendation systems, the choice-based conjoint analysis has been suggested as a method for measuring consumer preferences. This approach achieves high recommendation accuracy and does not suffer from the start-up problem because it is also applicable for recommendations for new consumers or of new products. However, this method requires massive consumer input, which causes consumer reluctance. In a simulation study, we demonstrate the high accuracy, but also the high user's effort for using a utility-based recommendation system using a choice-based conjoint analysis with hierarchical Bayes estimation. In order to reduce the conflict between consumer effort and recommendation accuracy, we develop a novel approach that only shows Pareto-efficient alternatives and ranks them according to the number of dominated attributes. We demonstrate that, in terms of the decision accuracy of the recommended products, the ranked Pareto-front approach performs better than a recommendation system that employs choice-based conjoint analysis. Furthermore, the consumer's effort is kept low and comparable to that of simple systems that require little consumer input.

Keywords: Recommendation systems, Preference measurement, Pareto-front, Effort, Accuracy, Simulation

Green PE, Krieger AM, Winder Y (2001) Thirty years of conjoint analysis: reflections and prospects. Interfaces 31(3):56–73

Häubl G, Trifts V (2000) Consumer decision making in online shopping environments: the effects of interactive decision aids. Mark Sci 19(1):4–21

Hevner AR, March ST, Park J, Ram S (2004) Design science in information systems research. MIS Q 28(1):75–105

Hey JD (1982) Search for rules for search. J Econ Behav Organ 3:65–81

Hinz O, Eckert J (2010) The impact of search and recommendation systems on sales in electronic commerce. Bus Inf Syst Eng 2(2):66–77

Hostler RE, Yoon VY, Guo Z, Guimaraes T, Forgionne G (2011) Assessing the impact of recommender agents on on-line consumer unplanned purchase behavior. Inf Manag 48(8):336–343

Hoyer WD (1984) An examination of consumer decision making for a common repeat purchase product. J Consum Res 11(3):822–829

Huang S (2011) Designing utility-based recommender systems for e-commerce: evaluation of preference-elicitation methods. Electron Commer Res Appl 10(4):398–407

Jedidi K, Kohli R, DeSarbo WS (1996) Consideration sets in conjoint analysis. J Mark Res 33(3):364–372

Johnson EJ, Payne JW (1985) Effort and accuracy in choice. Manage Sci 31(4):395–414

Kamis A, Koufaris M, Stern T (2008) Using an attribute-based decision support system for user-customized products online: an experimental investigation. MIS Q 32(1):159–177

Kanninen BJ (2002) Optimal design for multinomial choice experiments. J Mark Res 39(2):214–227

Kumar N, Benbasat I (2006) The influence of recommendations and consumer reviews on evaluations of websites. Inf Syst Res 17(4):425–439

Lohse GL, Johnson EJ (1996) A comparison of two process tracing methods for choice tasks. Organ Behav Hum Decis Process 68(1):28–43

Loomes G, Sugden R (1982) Regret theory: an alternative theory of rational choice under uncertainty. Econ J 92(368):805–824

Montgomery A, Hosanagar K, Krishnan R, Clay K (2004) Designing a better shopbot. Manage Sci 50:189–206

Moore WL (2004) A cross-validity comparison of rating-based and choice-based conjoint analysis models. Int J Res Mark 21(3):299–312

Moore WL, Gray-Lee J, Louviere JJ (1998) A cross-validity comparison of conjoint analysis and choice models at different levels of aggregation. Mark Lett 9(2):195–207

Olshavsky RW (1979) Task complexity and contingent processing in decision making: a replication and extension. Organ Behav Hum Perform 24:300–316

Payne JW, Bettman JR, Johnson EJ (1993) The adaptive decision maker. Cambridge University Press, Cambridge

Pfeiffer J (2012) Interactive decision aids in e-commerce. Springer, Berlin

Rand W, Rust RT (2011) Agent-based modeling in marketing: guidelines for rigor. Int J Mark Res 28(3):181–193

Rieskamp J, Hoffrage U (2008) Inferences under time pressure: how opportunity costs affect strategy selection. Acta Psychol 127(2):258–276

Rossi PE, Allenby GM, McCulloch R (2005) Bayesian statistics and marketing. Wiley, West Sussex

Schlereth C, Eckert C, Skiera B (2012) Using discrete choice experiments to estimate willingness-to-pay intervales. Mark Lett 23(3):761–776

Scholz M, Dorner V (2012) Estimating optimal recommendation set sizes for individual consumers. In: Proc of the 33rd int conf inform syst

Svenson O (1979) Process descriptions of decision making. Organ Behav Hum Perform 23:86–112

Todd PA, Benbasat I (1992) An experimental investigation of the impact of computer based decision aids on processing effort. MIS Q 16(3):373–393

Todd PA, Benbasat I (1994) The influence of decision aids on choice strategies and conditions of high cognitive load. IEEE Trans Syst Man Cybern 24(4):537–547

van Ittersum K, Pennings J (2012) Attribute-value functions as global interpretations of attribute importance. Organ Behav Hum Decis Process 119(1):89–102

Wallenius J, Dyer JS, Fishburn PC, Steuer RE, Zionts S, Deb K (2008) Multiple criteria decision making, multiattribute utility theory: recent accomplishments and what lies ahead. Manage Sci 54(7):1336–1349

Wang W, Benbasat I (2009) Interactive decision aids for consumer decision making in e-commerce: the influence of perceived strategy restrictiveness. MIS Q 33(2):293–320

Wohlin C, Runeson P, Höst M, Ohlsson MC, Regnell B, Wesslén A (2012) Experimentation in software engineering. Springer, Berlin

Xiao B, Benbasat I (2007) E-commerce product recommendation agents: use, characteristics, and impact. MIS Q 31(1):137–209

Yee M, Dahan E, Hauser JR, Orlin J (2007) Greedoid-based noncompensatory inference. Mark Sci 26(4):532–549