Effort and Accuracy Analysis of Choice Strategies for Electronic Product Catalogs

Jiyong Zhang
Human Computer Interaction Group
Swiss Federal Institute of Technology(EPFL)
CH-1015, Lausanne, Switzerland
+41 21 6936656
jiyong.zhang@epfl.ch

Pearl Pu
Human Computer Interaction Group
Swiss Federal Institute of Technology(EPFL)
CH-1015, Lausanne, Switzerland
+41 21 6936081
pearl.pu@epfl.ch

ABSTRACT

One crucial task for e-commerce systems is to help buyers find products that not only satisfy their preferences but also reduce their search effort. Usually the amount of available products is far beyond the upper limit that any individual could process by hand; thus product search tools are employed to generate target product(s) by eliciting the buyer's preferences and then executing some kind of choice strategies. We propose in this paper an extended effort–accuracy framework for measuring the performance of various choice strategies in terms of cognitive effort, elicitation effort and decision accuracy. The performance of a variety of basic choice strategies is further studied by theoretical analysis as well as empirical simulations. It shows that the performance of a given choice strategy is a tradeoff between choice accuracy and effort required from the users. The proposed framework also suggests a new efficient method of evaluating the user interfaces of e-commerce systems by analyzing the performance of the underlying choice strategies.

Categories and Subject Descriptors

H.5.2 [**Information Interfaces and Presentation**]: User Interfaces—*Theory and methods, Evaluation/methodology*; H.1.2 [**Models and Principles**]: User/Machine Systems—*Human factors, Human information processing*; J.8 [**Computer Applications**]: Internet Applications—*Electronic commerce*

General Terms

Performance, Measurement

Keywords

Electronic product catalog, preference elicitation, extended effort–accuracy framework, choice strategies, elicitation effort, elementary elicitation process

1. INTRODUCTION

With the rising prosperity of the World Wide Web (WWW), consumer e-commerce systems are becoming increasingly important in our daily lives. Unlike traditional commerce, where the activities are carried out directly between individuals or organizations, in e-commerce environments the buyer (or the decision maker) interacts with pre-designed computer systems to get information about products or services he/she wants. Usually the product or service information provided by the e-commerce system is far beyond any individual's cognitive effort to process without the help of search tools. Even when such tools are employed, search results are often inaccurate due to a lack of high precision in the preference model, and the end-user needs to examine many pages of possible choices for the final result. Studies from economics and psychology have shown that the individual has bounded rationality when making decisions due to his/her limited knowledge and computational capacity[13]. Therefore, one crucial task for e-commerce systems is to help buyers choose the products they prefer with a reasonable amount of effort and time.

We focus on a specific category of e-commerce systems called electronic product catalog (EPC)[5, 15], which provides a list of products for the buyer to select. Each of these products is represented by a number of attributes. The buyer needs to choose the product that most closely satisfies his/her preferences. In most cases these preferences cannot be fully satisfied and some tradeoffs have to be made between different attributes[8]. The process of choosing the most preferred product from the EPC can be formally described as solving a Multi-Attribute Decision Problem (MADP)[1] $\Psi = \langle \mathbf{X}, \mathbf{D}, \mathbf{O}, \mathbf{P} \rangle$, where $\mathbf{X} = \{X_1, \cdots, X_n\}$ is a finite set of attributes the product catalog has, $\mathbf{D} = D_1 \times \cdots \times D_n$ indicates the space of all possible products in the catalog (each $D_i (1 \leq i \leq n)$ is a set of possible domain values for attribute X_i), $\mathbf{O} = \{O_1, \cdots, O_m\}$ is a finite set of available products (also called alternatives or outcomes) that the EPC offers, and $\mathbf{P} = \{P_1, \cdots, P_t\}$ denotes a set of preferences that the decision maker may have. Each preference P_i may be identified in any form as required by the solution methods. The solution of a MADP is an alternative O most satisfying the decision maker's preferences.

---

[1]It is also known as multi-criteria decision making (MCDM) problem[3]. Our definition emphasizes on the term *attribute*, which is an objective aspect of products, not related to the decision maker's preferences.

Behavioral decision theory provides adequate knowledge describing people's decision behavior and presents typical approaches of solving decision problems in traditional environments where no computer aid is involved[6]. A classical effort–accuracy framework was also established that describes how people adapt their decision strategies by trading off accuracy and cognitive effort to the demands of the tasks they face. In the online electronic environment where the support of computer systems is pervasive, we believe that this adaptive decision behavior still exists, but that the classical framework needs to be adjusted. Though the decision maker's cognitive effort is still required, it can be significantly decreased by having computer programs carry out most of the calculation work automatically. However, the decision makers must expend some effort to explicitly state their preferences to the computer according to the requirements of the decision strategies employed. This process is usually called preference elicitation, and the effort that it requires is a new factor that may affect the decision maker's decision behavior.

In this paper we define a method of measuring the elicitation effort quantitatively and propose an extended effort–accuracy framework that measures the performance of a variety of choice strategies potentially employed in online product search tools in terms of cognitive effort, elicitation effort, and decision accuracy. Under this framework, we then investigate the relationship between the effort and accuracy of various choice strategies using theoretical analysis and simulation experiments. This paper is organized as follows: the related research work is reviewed in section 2. Section 3 describes the extended effort–accuracy framework. In the next section, a variety of choice strategies are studied and their elicitation effort is analyzed. Section 5 reports our main results from carrying out a set of simulation experiments to investigate the performance of various choice strategies in solving MADPs. Discussions of these results and conclusions are given in section 6 and 7 respectively.

2. RELATED WORK

Payne et al. proposed a classical effort–accuracy framework of describing how people select decision strategies when they face decision problems in traditional environments[6]. Typically, the decision maker navigates through all of the alternatives with a certain choice strategy and finally chooses the one that most satisfies his/her preferences. For example, when time permits and the decision maker's goal is to get an accurate result, he/she may use some strategies with a high level of cognitive effort such as the *weight additive(WADD)* strategy (choice strategies are described in section 4). However, in time-critical situations, the decision maker may adopt some strategies with a low degree of cognitive effort such as the *lexicographic* strategy, or the *elimination-by-aspect* strategy to speed the decision process. The performance of a set of choice strategies is measured in terms of cognitive effort and decision accuracy. By comparison, our extended effort–accuracy framework improves upon Payne's work by adding measurement to the effort involved in preference elicitation so that it is more suitable to model the performance of choice strategies in an online decision support environment.

Besides the basic choice strategies studied in the classical effort–accuracy framework, various advanced decision making approaches have been proposed to date in helping the decision maker to find the most preferred alternative within an EPC. For example, Stolze et al. [14] proposed the scoring tree method for building interactive e-commerce system based on multi-attribute utility theory (MAUT)[3]. Pu et al.[7] implemented the SmartClient approach based on a two-step decision process by first eliminating search space using constraint satisfaction techniques with little user involvement, and then guiding users to refine their decision accuracies using the WADD rule by performing a set of trade-off critiquings. Boutilier et al.[1] proposed the CP-nets approach based on conditional preference statements. Other technologies, such as Collaborative Filtering[10] and Analytic Hierarchy Process (AHP)[12] have also been proposed to help users make decisions. In this paper, we mainly investigate the performance of basic choice strategies. However, we believe that the performance of the above approaches can also be analyzed within the proposed framework. More importantly, currently the performance of each approach is generally evaluated by laborious empirical experiments with test users[9]. It is difficult to compare their performance since each approach is evaluated in a different environment by a different set of users. The extended effort–accuracy framework proposed here provides a new way of quantitatively measuring the performance of a given decision approach by effort and accuracy analysis.

3. THE EXTENDED EFFORT–ACCURACY FRAMEWORK

The classical effort–accuracy framework mainly investigated the relationship of decision accuracy and cognitive effort of processing information by different decision strategies in the offline situation. In the online decision support situation, however, the effort of eliciting preferences must be considered as well. We believe that the decision maker may select a certain strategy by trading off among 3 factors—cognitive effort, elicitation effort, and decision accuracy. Our framework extends the classical framework by adding measurement of elicitation effort. In the remainder of this section, we first recall the measurement of cognitive effort and accuracy in the classical framework, and then detail the method of measuring elicitation effort.

3.1 Measuring Cognitive Effort

Based upon the work of Newell and Simon[4], a decision approach can be seen as a sequence of elementary information processes (EIPs), such as reading the values of two alternatives on an attribute, comparing them, and so forth. By assuming that each EIP takes equal cognitive effort[2], the decision maker's cognitive effort is then measured in terms of the total number of EIPs. Conformed with the classical framework, a set of EIPs for the decision strategies is defined as: (1)READ: read an alternative's value on an attribute into short-term memory (STM), (2)COMPARE: compare two alternatives on an attribute, (3)ADD: add the values of two attributes in STM, (4)DIFFERENCE: calculate the size of the difference of two alternatives for an attribute, (5)PRODUCT: weight one value by another,

---

[2]Though this assumption is obviously imprecise, more studies by assigning different weighting of effort for the various EIPs show that the key relationships between the choice strategy and the decision environments were largely unchanged. See page 137 of [6].

(6)ELIMINATE: eliminate an alternative from consideration, (7)MOVE: move to next element of the external environment, and (8)CHOOSE: choose the preferred alternative and end the process.

3.2 Measuring Accuracy

Accuracy of choice can be defined in many ways. For example, Grether et al.[2] use the frequency of selection of dominated alternatives as a measure of decision quality. According to Pu et al.[9], the accuracy of choice is measured by the number of alternatives chosen correctly. The classical effort–accuracy framework uses the maximization of expected value (EV) as the criterion to measure the accuracy of decision strategies. One advantage of this measure is that it can distinguish not only that an error has occurred but also the severity of the error.

Since no risk or uncertainty is involved in the MADPs and in the case of certainty the criterion of EV is equivalent to the utility value of each alternative, here we define the accuracy of a strategy S as the utility value $V(O_S)$, where V is the value function determined in the WADD strategy, and O_S is the alternative selected by strategy S.

In practice, we expect that the WADD strategy could achieve 100% accuracy, and the strategy of random choice (selecting an alternative randomly from the alternative set) will get zero accuracy. By doing so we define the *relative accuracy* of a strategy S as:

$$RA_S = \frac{V(O_S) - V(O_{RAND})}{V(O_{WADD}) - V(O_{RAND})} \qquad (1)$$

where O_{RAND} is the alternative determined by random choice.

3.3 Measuring Elicitation Effort

In computer–aided decision environments, a considerable amount of decision effort goes into preference elicitation since people need to "tell" their preferences explicitly to the computer system. So far no formal method has been given to measure the preference elicitation effort. An elicitation process can be decomposed into a series of basic interactions between the user and the computer, such as selecting an option from a list, filling in a blank, answering a question, etc. We call these basic interaction actions elementary elicitation processes (EEPs). In our analysis, we define the set of EEPs as follows: (1)SELECT: select an item from a menu or a dropdown list, (2)FILLIN: fill in a value to an edit box, (3)ANSWER: answer a basic question, (4)CLICK: click a button to execute an action[3].

It is obvious that different EEPs require different elicitation effort (for instance, the EEP of one CLICK would be much easier than an EEP of FILLIN a weight value for a given attribute). For the sake of simplification, we currently assume that each EEP requires an equal amount of effort from the user. Therefore, given a specific decision approach, elicitation effort is measured by the total amount of EEPs it may require.

This elicitation effort is a new factor for the online environment. The main difference between cognitive effort and elicitation effort lies in the fact that the cognitive effort is

a description of the mental activities in processing information, while the elicitation effort is about the interaction effort between the decision maker and the computer system through pre-designed user interfaces. Even though the decision makers already have clear preferences in their mind, they must still state their preferences in a way that the computer can "understand". With the help of computer systems, the decision maker is able to reduce the cognitive effort by compensating with a certain degree of elicitation effort.

Let's consider a simple decision problem with 3 attributes and 4 alternatives. When a computer system is not involved, the cognitive effort of solving this problem by the WADD strategy will be 24 READS, 8 ADDS, 12 PRODUCTS, 3 COMPARES, and 1 CHOOSE. The total number of EIPs is therefore 48[4]. However, with the aid of a computer system, the decision maker could get the same result by spending 2 units of elicitation effort (FILLIN the weight value of first 2 attributes) and 1 unit of cognitive effort (CHOOSE the final result).

4. CHOICE STRATEGIES

As mentioned earlier, a variety of choice strategies could be adopted to help decision makers find the preferred solution(s) for a given decision problem. Each choice strategy can be thought of as a method (or a sequence of operation) for searching through all available alternatives. In this section we introduce a set of basic choice strategies for solving decision problems, and then analyze their cognitive and elicitation effort in online situations. Their detailed performances will be shown in the next section by simulation experiments.

4.1 The Weighted Additive Strategy

The weighted additive (WADD) strategy is based on the multi-attribute utility theory (MAUT)[3]. Let the symbol $\succsim$ denote the decision maker's preference order, e.g. $A \succsim B$ means "A is preferred or indifferent to B". According to the Utility Theory[17][5], for a given MADP, there exists a function $V : \mathbf{O} \to \Re$, called a utility function (or value function as the MADP considers the case of certainty only), that for any two possible alternatives O and $\bar{O} \in \mathbf{O}$,

$$O \succsim \bar{O} \Leftrightarrow V(O) \geq V(\bar{O}) \qquad (2)$$

More specifically, an alternative O can be represented by a set of attribute values $\langle X_1 = x_1, \cdots, X_n = x_n \rangle$ (in short as $\langle x_1, \cdots, x_n \rangle$), thus the above formula can be rewritten as

$$\langle x_1, \cdots, x_n \rangle \succsim \langle \bar{x}_1, \cdots, \bar{x}_n \rangle$$
$$\Longleftrightarrow \qquad (3)$$
$$V(\langle x_1, \cdots, x_n \rangle) \geq V(\langle \bar{x}_1, \cdots, \bar{x}_n \rangle)$$

Usually the value function V is scaled from zero to one. If the attributes $X_1, \cdots, X_n$ are *mutually preferentially inde-*

---

[3]We assume that the actions are in their basic forms only. For example, the FILLIN operation is not allowed to elicit more than one value or even an expression, otherwise the usability issue will arise.

[4]The detail analysis is given at page 80-81 of [6]. This example assumes the values of all attributes are numeric and consistent with the decision maker's preferences.

[5]We follow the general assumptions in utility theory that the preferences order $\succsim$ is complete, transitive, continuous, and independent[17].

$pendent^6$, then the value function is in the additive form:

$$V(\langle x_1, \cdots, x_n \rangle) = \sum_{i=1}^{n} w_i v_i(x_i) \qquad (4)$$

where v_i is a component value function of X_i ranged in $[0, 1]$, and w_i is the weight value of X_i satisfying

$$\sum_{i=1}^{n} w_i = 1, \qquad w_i > 0 \qquad (5)$$

The WADD strategy evaluates the value of each alternative by formula (4), and the alternative with the highest overall evaluation value is chosen as the optimal solution.

To use the WADD strategy, the weight and component value function for each attribute must be elicited. Theoretically, if the attributes are mutually preferentially independent, and the required parameters are determined precisely, the WADD strategy could be 100% accurate. In practice, however, this strategy may produce errors since user's specific preferences cannot be fully elicited and represented in such a specific form. Nevertheless, compared to the heuristic strategies introduced in the following subsection, the WADD strategy could gain higher accuracy and is used as the baseline strategy in our analysis of simulation experiments.

4.2 Basic Heuristic Strategies

Besides the WADD strategy, decision makers also use a variety of heuristic strategies in solving the decision problems they face. In this subsection we recall these strategies briefly, and further detail is available from Payne[6].

The equal weight (EQW) strategy. This strategy is a simplified version of the WADD strategy which ignores information about the relative importance (weight) of each attribute. An overall value for each alternative is obtained by simply summing the values for all of its attributes, and the alternative with the highest overall value is selected as the final solution.

The elimination-by-aspects (EBA) strategy. This strategy begins by determining the most important attribute. The cutoff value for that attribute is retrieved, and all alternatives with values for that attribute below the cutoff are eliminated. The process continues with the second most important attributes, then the third, and so on, until only one alternative remains. This strategy was first described by Tversky[16].

The majority of confirming dimensions (MCD) strategy. Described by Russo and Dosher[11], this strategy involves processing pairs of alternatives. The values for each of the two alternatives are compared on each attribute, and the alternative with a majority of winning (better) attribute values is selected. The retained alternative is then compared with the next one among the set of alternatives. This comparison process repeats until all alternatives have been evaluated and the final winning alternative has been identified.

The satisficing (SAT) strategy. Satisficing is one of the oldest heuristics identified in decision making literature [13]. With this strategy, alternatives are considered one at a time,

---

^{6}The attributes $X_1, \cdots, X_n$ are *mutually preferentially independent* if every subset Y of these attributes is preferentially independent of its complementary set of attributes. Details about this definition and the existence of an additive value function can be found in page 111 of [3].
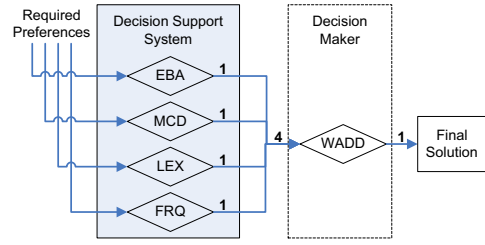

Figure 1: The *C4* decision strategy

in the order they occur in the set. Each attribute of an alternative is compared to a predefined cutoff value, which is often known as the *aspiration level*. If any attribute value is below the aspiration level, then that alternative is rejected. The first alternative which passes the cutoffs for all attributes is chosen. A choice can therefore be made before all alternatives have been evaluated. In the case where no alternative passes all the cutoffs, the cutoff can be relaxed and the process repeated, or an alternative can be randomly selected.

The lexicographic (LEX) strategy. For this strategy, the most important attribute is determined, the values of all the alternatives on that attribute are examined, and the alternative with the best value on that attribute is selected. If two alternatives have equal values, the second most important attribute is examined. This continues until the tie is broken.

The frequency of good and bad features (FRQ) strategy. Alba and Marmorstein[1] suggested that decision makers may evaluate or choose alternatives based simply upon counts of the good or bad features the alternatives possess. To implement this strategy, the decision maker needs to develop cutoffs for specifying good and bad features, and then to count the number of such features. This strategy could be viewed as the application of a voting rule to multi-attribute choice, where the attributes can be viewed as voters.

4.3 Hybrid Strategies

The basic heuristic strategies introduced above are widely adopted when time is critical in offline decision environments to diminish the decision maker's cognitive effort. In an online context, these strategies also have the advantage of saving the decision maker's elicitation effort by solving the decision problem with only partial preference information.

However, in the case when time is not critical, people are willing to make some extra effort to achieve accurate solutions. Some *hybrid strategies* which combine several strategies together may meet this requirement. For example, instead of executing only 1 strategy and showing only 1 alternative as the solution, the decision support system may execute several strategies and show all of their results to the user at the same time. The user can then spend some extra cognitive effort to choose the most preferred alternative. By doing so the final decision would reach higher accuracy.

Here we propose a specific hybrid strategy called *C4*, which is a combination of four basic heuristic strategies: EBA, MCD, LEX, and FRQ. The decision procedure is illustrated in Figure 1: First the decision maker inputs his/her preferences to the system according to the requirements of the four strategies. Then the decision support system executes

Table 1: elicitation effort analysis of choice strategies

Strategy	Parameters required to be elicited
WADD	weights, component value functions
EQW	component value functions
EBA	importance order, cutoff values
MCD	*none*
SAT	cutoff values
LEX	importance order
FRQ	cutoff values for good and bad features
C4	cutoff values, importance order

the four basic strategies simultaneously and produces up to 4 different alternatives for the decision maker. Finally the decision maker spends a certain amount of cognitive effort to select the final alternative using the WADD strategy. As the WADD strategy is completed by the decision maker, it requires no elicitation effort. The elicitation effort for *C4* would be counted by the total parameters that the four heuristic strategies require. We expect that the *C4* strategy could gain much higher decision accuracy than using the underlying basic strategies individually.

4.4 Analysis of Cognitive and Elicitation Effort

With the support of computer systems, the cognitive effort for WADD as well as the basic heuristic strategies is quite low. The decision maker inputs his/her preferences, and the decision support system executes that strategy and shows one alternative. Then the decision maker chooses this alternative and the decision process is ended. Thus the cognitive effort is equal to 1 EIP: CHOOSE the final alternative and exit the process. For the *C4* strategy, the cognitive effort of solving a MADP with n attributes and m alternatives is equal to that of solving a problem with n attributes and *4* alternatives in the traditional situation, the cognitive effort of which has been studied in [6].

According to their definitions, various choice strategies require that preferences with different parameters be elicited. For example, for the WADD strategy, the component value function and the weight for each attribute must be obtained. While for the EBA strategy, the importance order and cutoff value for each attribute are required. The required parameters for each strategy are shown in table 1.

For each parameter in the strategies aforementioned, a certain amount of elicitation effort is required. This elicitation effort may vary with different implementations of the user interface. For example, to elicit the weight value of an attribute, the user can just FILLIN the value to an edit box, or the user can ANSWER several questions to approximate the weight value. In our analysis and the following simulation experiments, we follow the "at least" rule: the elicitation effort is determined by the least number of EEP(s). In the above example, the elicitation effort for obtaining a weight value is measured as 1 EEP.

5. SIMULATION EXPERIMENTS AND RESULTS

Our simulation experiments are based on the Monte Carlo method. The scale of a MADP is determined by two factors:
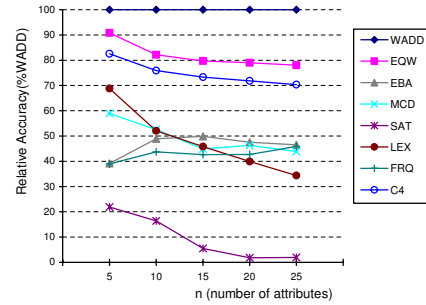
Figure 2: The relative accuracy of MADPs based on different number of attributes, where m(number of alternatives) = *1,000*

the number of attributes n, and the number of alternatives m. We classify the decision problems into *20* categories according to the value of n and m: n has five values (5, 10, 15, 20, and 25), and m has four (10, 100, 1,000 and 10,000). Each category contains 500 randomly generated MADPs, whose domain values for a given attribute are also determined randomly: the lower bound of each attribute is set to 0, and the upper bound is determined randomly from the range of 2 to 100. Formally speaking, for each attribute x_i, we define $D_i = [0, z_i]$, where $z_i \in [2, 100]$.

As shown in table 1, each choice strategy (except MCD) requires the elicitation of some specific parameters to represent the user's preferences. To simulate the preference elicitation process of each MADP, we assume that the component value function for each attribute is a polynomial form and is approximated by 3 mid-value points that are randomly generated[7]. Thus each component value function requires 3 units of EEPs. Other required parameters such as weight and cutoff value (each requires 1 unit of EEP) for each attribute are also simulated by the random generation process. The order of importance is determined by the weight order of the attributes for consistency.

In our simulation experiments, the WADD strategy is appointed as the baseline strategy, and the relative accuracy of a strategy is calculated according to equation (1) and is averaged over each category. The elicitation effort is measured in terms of the total number of EEPs required by the specific strategy, and the cognitive effort is measured by the required units of EIPs. Since the relationship between accuracy and cognitive effort has already been studied and analyzed by Payne et al.[6], in this section we focus on the performance of each strategy in terms of choice accuracy and elicitation effort.

Figure 2 shows the changes in *relative accuracy* for the listed choice strategies as *the number of attributes* increases in the case that each MADP has 1,000 alternatives. Increasing the number of attributes leads to a monotonic decrease in accuracy of the EQW, LEX, SAT and *C4* strategies. Besides the baseline WADD strategy, the EQW strategy achieves the highest accuracy, while the SAT strategy has the lowest accuracy. The four basic heuristic strategies EBA, MCD, LEX and FRQ are within the middle-level accuracy range. Of particular interest is that the proposed *C4* strategy, which combines the above four basic strategies, could achieve a

---

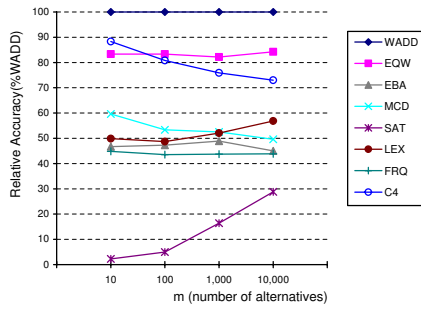[7]The procedure of assessing component value functions with mid-value points is introduced in page 120 of [3].

Figure 3: the relative accuracy of MADPs based on different number of alternatives, where n(number of attributes) $= 10$
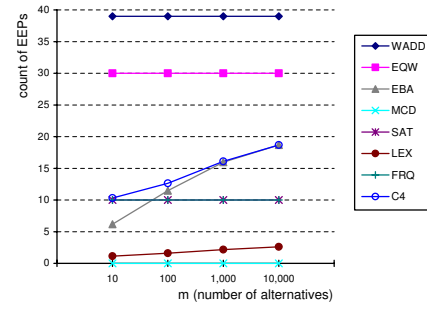

Figure 5: the elicitation effort of MADPs based on different number of alternatives, where n(number of attributes) $= 10$
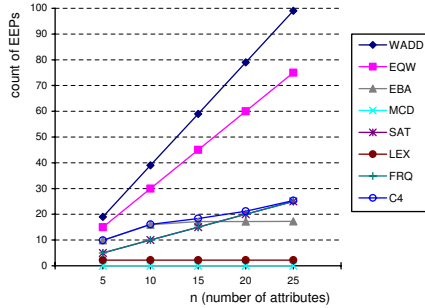

Figure 4: the elicitation effort of MADPs based on different number of attributes, where m(number of alternatives) $= 1,000$
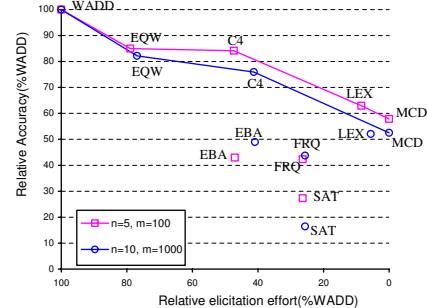

Figure 6: Elicitation effort/Accuracy tradeoffs for various choice strategies

much higher accuracy than when they perform separately (For instance, when there are 25 attributes, the accuracy of $C4$ strategy is about 65% higher than the average accuracy of the four basic strategies).

Figure 3 shows the relationship between *relative accuracy* and *the number of alternatives* for the listed choice strategies. The accuracy of EQW, MCD, DBA, and FRQ strategies remains stable as the number of alternatives increases. The LEX and SAT strategies, however, have a trend of gaining higher accuracy when there are more alternatives in the decision problem. The $C4$ strategy, though its accuracy decreases when the number of alternatives increases, could still gain a considerably higher accuracy than that of the EBA, MCD, LEX, and FRQ strategies.

The effect of the number of attributes on elicitation effort for each strategy is shown in figure 4. The elicitation effort required to use heuristic strategies increases more slowly than the effort required to use the WADD strategy as the number of attributes increases. For instance, when the number of attributes is 20, the elicitation effort for the $C4$ strategy is only about 25% of that of WADD strategy. The FRQ and SAT strategies require the same level of elicitation effort. The LEX strategy requires the least elicitation effort except for the MCD strategy, which requires *zero* elicitation effort in all cases.

Figure 5 shows the relationship between *elicitation effort* and *the number of alternatives* for each strategy. As the number of alternatives increases exponentially, the level of

elicitation effort for WADD, EQW, MCD, SAT, and FRQ strategies remains unchanged, and it increases slowly for the LEX, EBA and $C4$ strategies. As a whole, the elicitation effort of the studied choice strategies is quite robust to the number of alternatives that a decision problem has.

A combined study of figure 2 to 5 leads to some interesting conclusions. For each category of MADPs, some choice strategies, such as WADD and EQW, could gain relatively high decision accuracy with proportionally high elicitation effort. Other choice strategies, especially $C4$, MCD, EBA, FRQ, and LEX, could achieve a reasonably high accuracy with much lower elicitation effort compared with the baseline WADD strategy. Figure 6 illustrates the relationship between elicitation effort and choice accuracy for various strategies when solving different scales of decision problems. For the MADPs with 10 attributes and 1,000 alternatives, the MCD strategy could achieve above 50% relative accuracy without any elicitation effort. The $C4$ strategy in particular could achieve over 75% accuracy while only spending about 40% elicitation effort.

We say that a choice strategy, say A, is dominated if and only if there exists another strategy B which has higher relative accuracy and lower cognitive effort and elicitation effort than A in the same situation. Figure 6 shows that, for the MADPs with 5 attributes and 100 alternatives, the WADD, EQW, $C4$, EBA, and MCD are non-dominated approaches and have the possibility of being the optimal strategy according to the user's decision goal. For example, if the user's goal is to make decisions as accurately as possible, WADD is

be the best strategy among the listed strategies, and if the decision maker's goal is to have reasonable accuracy with certain elicitation effort, then the *C4* strategy may be the best choice.

6. DISCUSSION

These experimental results suggest that the designers of decision support systems could tradeoff decision accuracy with elicitation effort by implementing various choice strategies. Since a non-dominated decision strategy is better than a dominated strategy, in practice, decision support systems are required to provide as many non-dominated strategies as possible to satisfy decision makers' various decision goals.

We studied the performance of the *C4* strategy, which can achieve a higher accuracy while requiring users to expend some extra cognitive and elicitation effort than the basic strategies it contains. The *C4* strategy demonstrates how the performance of a new decision strategy could be measured quantitatively in terms of cognitive effort, elicitation effort, and accuracy within the extended effort–accuracy framework.

The simulation results need to be interpreted with some caution. First of all, the elicitation effort is measured by approximation. As mentioned earlier, we assumed that each EEP requires an equal amount of effort from the users. Currently it is unknown whether this obviously imprecise assumption would affect the simulation results largely. In addition, when measuring the decision accuracy, the WADD strategy is chosen as the baseline, assuming that it contains no error. However, this is not the case in reality. Moreover, as the MADPs in the simulation experiments are generated randomly, there is a potential gap between the simulated MADPs and those in real applications.

Despite these limitations, the simulation work provides useful insights into the performance of choice strategies. In e-commerce systems, the user interface can be determined largely by the underlying choice strategies, the performance of which can be further measured within the extended effort–accuracy framework. In this sense our work provides a new method of evaluating and comparing the performance of different user interfaces for e-commerce systems.

7. CONCLUSION AND FUTURE WORK

In this paper we proposed an extended effort–accuracy framework for modeling the performance of different decision strategies in the decision support environments. The method of measuring the effort of preference elicitation was given and a variety of decision strategies were then evaluated through simulation experiments. The proposed framework also provides a new method of evaluating different user interfaces used in e-commerce systems automatically by analyzing their underlying choice strategies.

In the future, we plan to improve this framework by taking into account the aforementioned shortcomings, and verifying the model's predictions with traditional user studies. So far, we have only studied the performance of a set of basic decision strategies, most of which are heuristic and intuitive. Our ongoing research also involves the use of the currently developed effort–accuracy framework to evaluate the performance of more advanced and hybrid approaches such as those based on collaborative filtering, soft-CSP, or CP-nets.

8. ACKNOWLEDGMENTS

9. REFERENCES

[1] C. Boutilier, R. Brafman, C. Domshlak, H. Hoos, and D. Poole. CP-nets: A tool for representing and reasoning with conditional ceteris paribus preference statements. *Journal of Artificial Intelligence Research*, 21:135–191, 2004.

[2] D. M. Grether and L. L. Wilde. Consumer choice and information: New experimental evidence. *Information Economics and Policy*, 1:115–144, 1983.

[3] R. Keeney and H. Raiffa. *Decisions with Multiple Objectives: Preferences and Value Tradeoffs.* Cambridge University Press, 1993.

[4] A. Newell and H. Simon. *Human Problem Solving.* Printice-Hall, Englewood Cliffs, NJ, 1972.

[5] J. W. Palmer. Retailing on the WWW: the use of electronic product catalogs. *International Journal of Electronic Markets*, 7(3):6–9, 1997.

[6] J. Payne, J. Bettman, and E. Johnson. *The Adaptive Decision Maker.* Cambridge University Press, 1993.

[7] P. Pu and B. Faltings. Enriching buyers' experiences: the smartclient approach. In *SIGCHI conference on Human factors in computing systems*, pages 289–296. ACM Press New York, NY, USA, 2000.

[8] P. Pu and B. Faltings. Decision tradeoff using example-critiquing and constraint programming. *Constraints: An International Journal*, 9(4), 2004.

[9] P. Pu and P. Kumar. Evaluating example-based search tools. In *ACM Conference on Electronic Commerce (EC'04)*, pages 208–217, New York, USA, 2004.

[10] P. Resnick, N. Iacovou, M. Suchak, P. Bergstorm, and J. Riedl. GroupLens: An Open Architecture for Collaborative Filtering of Netnews. In *Proceedings of ACM 1994 Conference on Computer Supported Cooperative Work*, pages 175–186, Chapel Hill, North Carolina, 1994. ACM.

[11] J. Russo and B. Dosher. Strategies for multiattribute binary choice. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 9:676–696, 1983.

[12] T. Saaty. *The Analytic Hierarchy Process.* New York: McGraw-Hill, Inc., 1980.

[13] H. A. Simon. A behavioral model of rational choice. *Quarterly Journal of Economics*, 69:99–118, 1955.

[14] M. Stolze. Soft navigation in electronic product catalogs. *international Journal on Digital Libraries*, 3(1):60–66, July 2000.

[15] M. Torrens. *Scalable Intelligent Electronic Catalogs.* Phd. thesis no. 2690, Swiss Federal Institute of Technology (EPFL), Lausanne (Switzerland), December 2002.

[16] A. Tversky. Elimination by aspects: A theory of choice. *Psychological Review*, 79:281–299, 1972.

[17] J. von Neumann and O. Morgenstern. *Theory of Games and Economic Behavior.* 2nd Edition, Princeton University Press, Princeton.