RC 22048 (98946) 30 April 2001 **Computer Science**

IBM Research Report

Design and Implementation of an Intelligent Decision Analysis System for E-Sourcing

Martin Bichler, Juhnyoung Lee, Chang Hyun Kim, Ho Soo Lee
IBM Research Division
Thomas J. Watson Research Center
P.O. Box 218
Yorktown Heights, NY 10598

Research Division
Almaden - Austin - Beijing - Haifa - T. J. Watson - Tokyo - Zurich

Design and Implementation of an Intelligent Decision Analysis System for E-Sourcing

Martin Bichler, Juhnyoung Lee, Chang Hyun Kim, Ho Soo Lee
IBM T.J. Watson Research Center
P.O. Box 218, Yorktown Heights, NY 10598
{bichler, jyl, changkim, leehs}@us.ibm.com

Abstract

Sourcing addresses critical decisions of what to buy, how much to buy, whom to buy from, and how to manage relationship with suppliers. Decision making in sourcing can involve a few hundred offerings each of which is described by several dozen attributes. When using traditional decision analysis techniques, sourcing specialists are often having difficulties in assigning appropriate weights to attributes and feel uncomfortable with decided results. In this paper, we present an innovative approach, where the decision makers only provide ordinal rankings over subsets of offerings. From the information implied by these ordinal rankings, the system derives a set of weights and an overall ranking of all the given offerings. With additional information from the decision maker, these results are iteratively refined. The paper describes the basic concepts and algorithms used as well as the implementation.

1. Introduction

Sourcing relates to the procurement of direct inputs used in the manufacture of a firm's primary outputs. E-sourcing is an internet-based business process for identifying, evaluating, negotiating, and configuring optimal groupings of buyers and suppliers into a supply chain that responds to changing market demands. Figure 1 depicts the sourcing process flow starting with the sourcing strategy and ending with acceptance or rejection of particular offers.

Efficient selection of bids and product offerings is a core activity in this sourcing process. Sourcing professionals have to make selections among hundreds of alternatives considering several dozens of criteria. Accountability is crucial when making such complex decisions particularly for business buyers whose decisions always need to be justified in terms of savings in time and cost (e.g., why certain suppliers won and others lost).

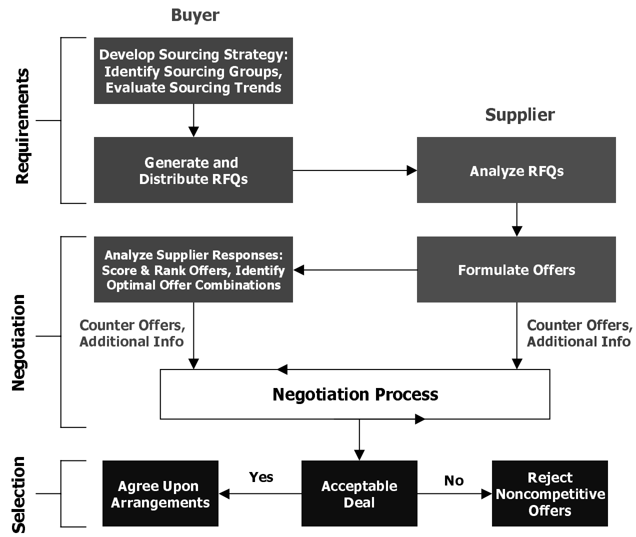

Figure 1: Sourcing Process Flow

Selection problems like this are challenging, because they require the balancing of multiple, often conflicting objectives. Traditional approaches to multi-attribute decision making (MADM) are often insufficient in situations with a large number of attributes and alternatives, which are typical in e-sourcing. In particular, traditional forms of weight elicitation do not work well under these circumstances.

ABSolute is an application framework providing buyer-side decision support for e-sourcing. The framework enhances traditional approaches to MADM by advanced visualization capabilities and WORA, a new weight assessment methodology based on ordinal rankings of alternatives. WORA was designed to alleviate the shortcomings of traditional weight assessment techniques.

In this paper we will focus on the basic decision analysis techniques implemented and the WORA methodology for weight assessment. The next section provides a brief overview of existing approaches to bid selection in e-sourcing. Then, section 3 surveys MAUT as well as some of the traditional approaches to weight assessment. Section 4 describes the core components of the ABSolute application framework, and section 5 explains the details of WORA and summarizes the results of a numerical simulation. Finally, section 6 concludes with some summarizing remarks.

2. Existing Approaches to Bid Selection in E-Sourcing

There are only a few bid analysis products currently available from companies such as Emptoris (www.emptoris.com), Frictionless Commerce (www.frictionless.com), Perfect (www.perfect.com), and Rapt (www.rapt.com). All of these products depend on a single decision analysis method and are limited in their capabilities for supporting the decision-making processes.

One approach used in commercial bid analysis products is optimization such as integer programming, and/or constraint programming. Bid analysis products from Emptoris and Rapt belong to this category. These products recommend a set of bids from multiple suppliers and optimize one or more objectives. A drawback of this approach is that its capability for recommending complex product offerings is limited. While this approach can be effective for simple objectives such as minimizing the total cost, it does not work well if the objectives involve complicated business rules over multiple attributes and are therefore only applicable in case of simple commodities.

Bid analysis products from Frictionless Commerce and Perfect are based on traditional decision analysis techniques, which have been actively studied in MADM, an area of operations research. The primary techniques in the field are Multi-Attribute Utility Theory (MAUT) [1], Simple Multi-Attribute Rating Technique (SMART) [2] and the Analytic Hierarchy Process (AHP) [3], all of them implemented in several software applications. The essence of all these widely used decision aids is breaking complicated decisions down into small pieces that can be dealt with individually and then recombined in an additive manner. The key difference among the various methods is the way the scores on individual attributes and their weights are assessed.

3. Traditional Decision Analysis

Most commercial bid selection tools such as Perfect and Frictionless Commerce are based on MAUT. They request a user to assign relative weights to individual attributes of alternatives (i.e., bids), and then use an additive value function in order to compute the scores of the alternatives. The systems then rank the alternative bids by score, and the user selects the winning bids among the top-rankers.

3.1. MAUT

The basic hypothesis of MAUT is that in any decision problem, there exists a real valued function U defined along the set of feasible alternatives, which the decision maker wishes to maximize. This function aggregates the criteria x_1 ... x_n. Besides, individual (single-measure) utility functions $U_1(x_1)$, ..., $U_n(x_n)$ are assumed for the n different attributes. The utility function translates the value of an attribute into ”utility units”. The overall utility for an alternative is given by the sum of all weighted utilities of the attributes. For an outcome that has levels x_1, …, x_n on the n attributes, the overall utility for an alternative is given by

$$U(x_1 ... x_n) = \sum_{i=1}^{n} w_i U(x_i)$$

The alternative with the largest overall utility is the most desirable under this rule. Each utility function $U(x_i)$ assigns values of *0* and *1* to the worst and best levels on that particular objective and

$$\sum_{i=1}^{n} w_i = 1,\ w_i >= 0 .$$

Consequently, the additive utility function assigns values of *0* and *1* to the worst and best conceivable outcomes, respectively. A basic precondition for the additive utility function is preferential independence of all attributes, which has been the topic of many debates on multi-attribute utility theory [2, p. 328]. Even in cases with interdependencies, the additive utility function is often used as a rough-cut approximation for a more complex non-linear utility function. MAUT is a theoretically widely accepted technique, and there have been many reported applications of multi-attribute utility theory in government and business decision-making.

3.2. Traditional Weight Assessment

The assessment of appropriate weights is key to MAUT and is what makes a "good" preference model. Assigning weights means determining trade-offs between different attributes of a decision. A fundamental weakness of the bid selection packages described in the last section is that they provide little guidance for weight assessment. Consequently, the resulting bid scores may not be reliable. When a user assigns weights to attributes for the first time, s/he might not understand their effect well. Assigning weights in the presence of several dozens of attributes, which is typically the case in procurement, is even more difficult.

Several techniques have been proposed to help users assign reasonable weights. One approach is called *pricing out* because it involves determining the value of one objective in terms of another (e.g. dollars). For example, one might say that 5 days faster delivery time is worth $400. The idea is to find the indifference point, i.e. determining the marginal rate of substitution between two attributes. Although this concept seems straightforward, it can be a difficult assessment to make.

The *swing-weighting approach* requires the decision maker to compare individual attributes directly by imagining hypothetical outcomes. Starting with a hypothetical alternative that has the worst outcome in all attributes, the decision maker writes down other hypothetical alternatives, which have the best outcome in only one of the attributes. The various hypothetical alternatives are then ranked. The worst alternative gets 0 points, the best alternative gets 100 points. From this one can compute the weights by dividing the points by the sum of all points.

Since many decision makers feel unable to provide exact weights, some of the more recent approaches only ask for uncertain estimates. For example, methods from fuzzy decision analysis use fuzzy sets for weights and individual scoring functions and fuzzy operators for the aggregation of those fuzzy sets [4].

AHP uses a different approach to weight determination. Cognitive psychology has found that people are poor at assimilating large quantities of information on problems. A principle used in AHP is that comparative judgments are applied to construct a symmetric matrix of pairwise comparisons of all combinations of attributes. The method is based on the mathematical structure of consistent matrices and their associated right-eigenvector's ability to generate true or approximate weights. The right eigenvector of the matrix results in the weights for the different objectives.

A variation of AHP is the Geometric Mean Technique. Several researchers pointed out that the geometric mean is more appropriate in order to obtain the relative importance of the elements being compared [5] [6]. REMBRANDT is a software package implementing this technique. The advantage of both AHP and the Geometric Mean Technique is that pairwise comparisons are relatively easy to understand for the decision maker. The disadvantage is the exponential increase of pairwise comparisons in case of larger numbers of attributes and alternatives.

When it comes to hundreds of alternatives and dozens of attributes, most traditional forms of MADM and weight elicitation do not work well. Procurement specialists are often facing difficulties in assigning appropriate weights, and feel uncomfortable with the decided results.

4. ABSolute Sourcing Framework

ABSolute is an application framework providing buyer-side decision support for e-sourcing. It provides an integrated approach to support complex purchasing decisions and an opportunity to satisfy analysis capabilities required in large-scale purchasing processes. ABSolute integrate methods from multi-attribute decision analysis with advanced visualization techniques. The core components of the ABSolute decision analysis framework are

- a user interface for visual analysis,
- MAUT – a traditional and widely used decision aid, and
- WORA – a new methodology designed to determine weights in the presence of a large number of criteria.

The first steps in the ABSolute sourcing process are the development of a sourcing strategy as well as RFQ creation (see Figure 1). RFQ creation involves also the modeling of preferences, i.e. the relevant attributes as well as individual utility functions for them need to be chosen. For example, considering price a purchasing specialist might conclude that a lower price is preferable to a higher price and that the range of acceptable prices lies between $40 and $60 per unit. Based on the RFQ bids are solicited from the suppliers.

The following section we will concentrate on the next step in the sourcing process flow, namely the analysis of supplier responses, i.e. bids. This step consists of a number of finer grained process steps involving visual analysis of bids, dominance analysis, and finally MAUT enhanced by the WORA technique.

4.1. Interactive Visual Analysis

The ABSolute user interface combines the features of different analysis methods, and allows the users to selectively utilize alternative methods for different analysis needs. The visual analysis provides an effective means to navigating through the information space, intuitively understanding the properties of given options, and perceiving interesting patterns in the subject data set. In several ways, the visual analysis mechanism is designed to assist users in the RFQ/bid analysis process: It presents the entire information space of submitted RFQs/bids in a single page. This compact display makes it easy to navigate through the information space and visually compare all the RFQs/bids s/he is interested. It also helps users by visually showing them concrete comparison data that might otherwise be only abstractly in their heads or on paper in non-visual form.

4.2. Dominance Analysis

When many alternatives are present, it is common to reduce the choice set to a more manageable size by first eliminating ”inferior” alternatives. Domination and the conjunctive rule are both mechanisms for the identification of ”good” alternatives. In ABSolute it is easy to determine alternatives that do not meet the aspiration levels preset in the RFQ in at least one of the dimensions. This feature is also called *conjunctive procedure*. A drawback of the conjunctive rule is that it is non-compensatory. If an alternative barely fails in a given dimension, it cannot be compensated for with surplus elsewhere.

Domination procedures entail the identification of alternatives that are equal to or worse than some other alternative on every single dimension. Let's analyze two bids *A* and *B* and the associated consequences

$$\boldsymbol{x}' = (x_1', ..., x_i', ..., x_n') \text{ and}$$
$$\boldsymbol{x}'' = (x_1'', ..., x_i'', ..., x_n'')$$

$\boldsymbol{x}'$ dominates $\boldsymbol{x}''$ whenever $x_i' \geq x_i''$, for all i and $x_i' > x_i''$ for some i. The set of consequences that is not dominated is called the *efficient frontier* or *Pareto optimal set*. ABSolute provides functions and a graphical user interface to easily determine dominance relationships in the bids.

4.3. Decision Analysis

Many complex decision problems involve multiple conflicting objectives. It is often true that no dominant alternative that is better than all other alternatives will exist in terms of all of these objectives. In fact, simulation shows that the Pareto optimal set increases with the number of attributes [7].

One cannot maximize benefits and at the same time minimize costs. In essence, the decision maker is faced with a problem of trading off the achievement of one objective against another objective. Consequently, after sorting out the inferior alternatives, preference models must be established, i.e. utility functions in order to find the best alternative for a subject. These utility functions should allow explicit comparisons between alternatives differing in many ways.

ABSolute uses MAUT (see section 3.1) as the basic decision analysis technique to rank the set of alternatives. Although MAUT was used with large numbers of attributes, previous analyses suggest that the predictive validity of multi-attribute models was adequate only when there were fewer than five attributes [8]. The main problem in the presence of many attributes is the assessment of appropriate weights. Hierarchical structuring of attributes provides one mechanism to allow the decision maker focusing on specific subsets of these attributes sequentially, but it doesn't solve the problem of getting reasonable weights.

Often purchasing managers have a natural preference of one alternative over another, but they are having difficulties expressing these preferences as precise weights. Several MADM experts also called for consideration of the ability of multi-criteria methods to aid decision maker learning [9]. The contention is that decision makers may often start off with only an initial impression of what they want from alternatives. As the analysis proceeds, decision makers can more accurately form their preferences. MAUT, AHP and REMBRANDT do not directly consider learning. They assume a clear underlying preference function to be elicited from the decision maker [10, p. 170].

WORA (Weight determination based on Ordinal Rankings of Alternatives) is a new method, which utilizes the information contained in ordinal rankings of bids in order to estimate the decision maker's true weights. It provides an interactive weight elicitation procedure, which helps the user learn and continuously refine his/her own preferences in MAUT. The following section describes the principles of the WORA technique.

5. The WORA Technique

The visual interface of ABSolute helps users examine and select subsets of offers, and create and arrange ordinal rankings over these subsets. From the information implied by these ordinal rankings, the system derives a set of weights of attributes, and then an overall ranking of all the given offers using optimization techniques. With additional information from the decision maker, these results can be refined. This basic process is illustrated in Figure 3.

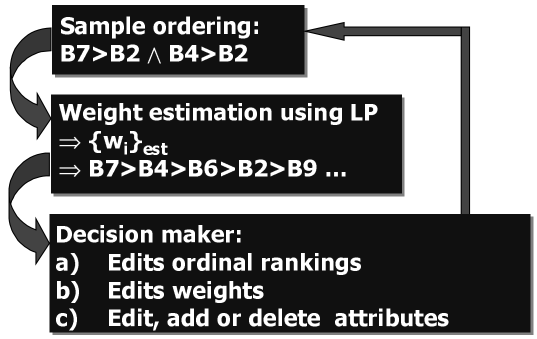

Figure 2: The WORA Process

In a first step the decision maker provides some sample rankings of bids of the type of $B_1 \succ B_2 \succ B_3$. ABSolute checks these sample rankings for *intransitive preferences* and *dominance violations*. Intransitive preferences describe situations where a decision maker prefers bids $B_1 \succ B_2 \succ B_3 \succ B_1$. A utility function for these preferences would have to consist of $U(B_1) > U(B_2) > U(B_3) > U(B_1)$ which is impossible. It turns out that detecting intransitive preferences in multiple sample rankings is similar to the problem of detecting cycles in a directed graph, which can be solved using algorithms such as a depth first search. The sample rankings are also checked for dominance violations. In these cases a user provides $B_1 \succ B_2$, although bid B_2 dominates B_1. Both types of inconsistencies are displayed to the user in order to be resolved. The sample rankings provided by the decision maker are then transformed into constraints to the following linear program (LP), which generates estimates for the decision maker's weights.

5.1. LP Formulation

Suppose we have a subset of bids $B = \{B_1, B_2, \ldots, B_k\}$ which we can rank. Assume that $B_1 \succeq B_2 \succeq B_3 \succeq \ldots B_k$, that is, in the set B, the object B_1 is most preferable; B_2 is next and so on. The score S_i of each bid B_i is computed as

$$S_i = \sum_j w_j f(a_{ij}) \text{ for } i = 1, \ldots, k$$

where the weights w_j are unknown, and j is the number of an attribute. These weights should satisfy the relationship $B_1 \succeq B_2 \succeq B_3 \succeq \dots B_k$. We formulate the following LP:

Maximize 0
Subject to:

$$S_1 \geq S_2$$
$$S_2 \geq S_3$$
$$.$$
$$.$$
$$S_{k-1} \geq S_k$$
$$\sum_j w_j = 1$$
$$w_j \geq 0 \text{ for each } j$$

Since the weights should be non-negative and add up to 1, we add the non-negativity constraint ($w_j \geq 0$), and the normalization constraint ($\sum_j w_j = 1$). A feasible solution to all these constraints can be obtained by solving the LP.

In this formulation we have not given any objective function to the LP, since any feasible solution will satisfy our requirements. But it is certainly possible to specify any linear objective function.[1] If the decision maker can rank several subsets B, we have to write the constraints for each such subset. This way, the solution to the LP is guaranteed to satisfy all the information we know about such subsets.

5.2. Numerical Experimentation

A basic research question in this context is, how many sample rankings it takes to get to a good estimate of the decision makers weights. In a computer simulation we analyzed this question as well as potential strategies, which help reduce the number of sample rankings that need to be provided by the decision maker.

In the simulation we assume the decision maker's "true" weights to be known in advance. This means there is no decision maker learning involved and the task is to elicit these "hidden" preferences. Based on the "true" weights the simulation calculated the decision maker's "true" ranking of bids B_{true},. Before every simulation round a set of bids was generated with random attribute values between 0 and 1. In an iterative procedure the simulation revealed an additional binary sample ranking of the type

Loop 1: $B_1 \succeq B_2$,
Loop 2: $B_1 \succeq B_2 \succeq B_3$,
Loop 3: $B_1 \succeq B_2 \succeq B_3 \succeq B_4$
…

In every loop a set of estimated weights $\{w_i\}_{est}$ was calculated, and based on these weights an estimated ranking of all bids, B_{est}. The simulation calculated the Bravais Pearson correlation coefficient of B_{true} and B_{est}, which was used as an indicator for the quality of the estimated weights.

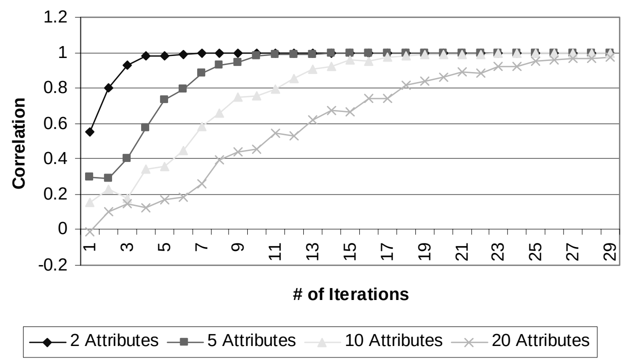

Figure 4: Speed of Convergence

Figure 4 shows the average correlation coefficients after 40 simulation runs given 70 random bids and different numbers of attributes. The number of revealed binary rankings needed to get a very good estimate (i.e. a correlation coefficient >0.9) was a little higher than the number of attributes. Each iteration revealed additional information and the estimate was getting gradually better. If more binary rankings were revealed in each iteration, then of course, less iterations were needed to get a high correlation. Additional simulations showed that the number of bids does not have a significant impact on the speed of convergence.

One particularity of the above simulation is that all ranked samples are chained, i.e. one bid of sample i is also contained in sample i+1. The chained sample ranking $B_1 \succeq B_2 \succeq B_3$, reveals also the transitive preference $B_1 \succeq B_3$,. In fact, when we

[1] Numerical experimentations such as the ones described in section 5 have shown that the feasible region in most cases is very small and the type of objective function does not have a significant impact.

compare the revelation of connected and unconnected samples it can be shown that connected samples achieve high correlations faster than unconnected ones. Figure 5 shows the results of two simulation runs, each with 10 attributes. In every iteration a sample ranking of 4 bids is revealed.

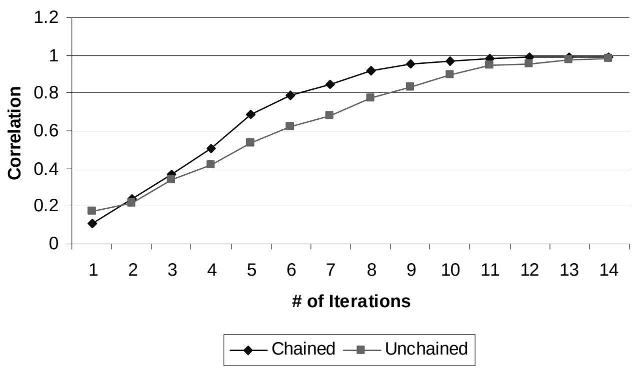

Figure 5: Chained and Unchained Samples

Further simulation has shown that revealing information from dominance relationships can be used to reduce the number of iterations. These strategies can be used in order to suggest samples to the decision maker.

6. Conclusions

ABSolute provides advanced decision analysis capabilities for bid selection. The interactive visualization enables to easily navigate through a large information space. MAUT provides the strength of a traditional decision analysis technique, which has already been applied to numerous real-world situations.

The traditional weight assessment techniques used with MAUT are not sufficient in the presence of large numbers of criteria. WORA provides an innovative approach to utilize the information contained in ordinal rankings of alternatives. In an iterative process it aids the decision maker in learning about their own preferences and assessing reasonable weights. This way, WORA makes the thought process explicit and leads to more informed decisions. In purchasing situations with many alternatives and many attributes WORA can be an important complement to conventional techniques.

References

[1] R. L. Keeny and H. Raiffa, *Decision Making with Multiple Objectives: Preferences and Value Tradeoffs*. Cambridge, UK: Cambridge University Press, 1993.

[2] W. Edwards, “How to use multiattribute utility measurement for social decisionmaking,” *IEEE Transactions on Systems, Man, and Cybernetics SMC*, vol. 7, pp. 326-340, 1977.

[3] T. L. Saaty, *The Analytic Hierarchy Process*. New York, USA: McGraw Hill, 1980.

[4] R. A. Ribeiro, “Fuzzy multiple attribut decision making: A review and new preference elicitation techniques,” *Fuzzy Sets and Systems*, vol. 78, pp. 155-181, 1996.

[5] F. A. Lootsma, “Multi-criterial analysis and budget reallocation in long-term research planning,” *European Journal of Operational Research*, vol. 47, pp. 293-305, 1991.

[6] V. Belton and A. E. Gear, “On a shortcoming of Saaty's method of analytical hierarchies,” *Omega*, vol. 11, pp. 227-230, 1983.

[7] H.-J. Zimmermann and L. Gutsche, *Multi-Criteria Analyse - Einfuehrung in die Theorie der Entscheidungen bei Mehrfachzielsetzungen*. Berlin; Heidelberg: Springer-Verlag, 1991.

[8] G. W. Fisher, “Utility models for multiple objective decisions: To they accurately represent human preferences,” *Decision Sciences*, vol. 10, pp. 451-479, 1979.

[9] D. Vanderpooten, “The interactive approach in MCDA: A technical framework and some basic conceptions,” *Mathematical and Computer Modelling*, vol. 12, pp. 1213-1220, 1990.

[10] D. L. Olson, *Decision Aids for Selection Problems*. New York, et al.: Springer, 1995.