Applying Evidential Reasoning to Prequalifying Construction Contractors[1]

M. Sönmez[2]; G. D. Holt[3]; J. B. Yang[4]; and G. Graham[5]

Abstract: A contractor prequalification process is a typical multiple criteria decision-making problem that embraces both quantitative and qualitative criteria. When facing such problems, a decision maker may need to provide uncertain, incomplete, or imprecise assessments due to a lack of information, time pressure and/or shortcomings in expertise. A multiple criteria decision-making method is then needed in order to deal with such assessments as well as for the meaningful and robust aggregation. This paper addresses these issues by applying an evidential reasoning approach to a contractor prequalification problem. The advantages and disadvantages of applying evidential reasoning to contractor prequalification problems in practice are also reported.

DOI: 10.1061/(ASCE)0742-597X(2002)18:3(111)

CE Database keywords: Contractors; Decision making; Construction industry.

Introduction

Multiple criteria decision-making (MCDM) problems that embrace both quantitative and qualitative criteria are very common in practice. When facing such MCDM problems, the literature and research show that the following difficulties may be encountered:

- Different types of assessments (e.g., numbers, linguistic terms, and/or stochastic values) depending on the characteristics of the decision criteria (Valls and Torra 2000),
- Imprecise and missing assessments due to the lack of data, shortcomings in expertise, time pressure and/or the decision maker (DM) is only willing or able to provide incomplete assessments (Kim and Ahn 1999), and
- Meaningful and robust aggregation of subjective and objective assessments made on multiple (decision) criteria.

In practice, a contractor selection problem can be described as a two-stage process. First, a large number of contractors are invited to tender and then a short list of contractors is drawn based on a set of predetermined criteria (prequalification stage). In the second stage, a contractor is selected from the short list to carry out the project (final contractor selection stage). A contractor prequalification problem (CPP) is a typical multiple criteria decision-making problem in which decision criteria are of both quantitative and qualitative natures and the aforementioned problems do occur. The aim of this paper is to present an application of the evidential reasoning (ER) approach to solve a CPP with uncertain, imprecise (incomplete), and/or missing information.

Following this introduction, the literature on contractor prequalification problems will be summarized. Then, a brief description of MCDM problems will be given. After that the ER approach will be explained and discussed with regard to why it is a suitable tool to deal with MCDM problems of above type. A hypothetical CPP illustrates application of the ER technique. The results and discussion will follow. The conclusion will include the advantages and disadvantages of the method in practice.

Contractor Prequalification

Contractor prequalification can be described as the screening of contractors by a construction owner (client or client's representative, etc.) based on a set of criteria, selected to determine the contractors' competence to perform the proposed construction contract (Russell and Skibniewski 1988). Contractor prequalification is, therefore, a multiple criteria decision process in which a wide range of criteria need to be considered and evaluated simultaneously. The prequalification decision-making process was fully described and explained in Russell and Skibniewski (1988).

Prequalification is essential for both contractors and clients (Bubshait and Al-Gobali 1996). On the client's side, it helps eliminate the incompetent, insufficiently financed, and inexperienced contractors from further consideration. On the contractor's side, it works as a form of external auditing of a contractor's ability. The process usually starts by establishing decision criteria, which will vary according to the selection scenario, such as type of project, type of client, time scale, etc. (Russell and Skibniewski 1988; Russell et al. 1992; Ng 1996). Several previous researches have studied prequalification criteria, mainly to determine which criteria to apply and their relative importance (Russell and Skibniewski 1990; Russell 1990; Holt et al. 1993; Holt et al. 1994a,b,c,d; Bubshait and Al-Gobali 1996; Ng 1996; Hatush and Skitmore 1997a,b). The next step in the prequalification process is to gather

---

[1] An earlier version of this paper was presented in *Second International Conference on Decision Making in Urban and Civil Engineering*, 20–22 November 2000, Lyon, France.

[2] Manchester School of Management, Univ. of Manchester, Institute of Science and Technology, P.O. Box 88, Manchester M60 1QD, UK. E-mail: mahmut.sonmez@stud.umist.ac.uk

[3] The Built Environment Research Unit, Univ. of Wolverhampton, West Midlands WV1 1SB, UK.

[4] Manchester School of Management, Univ. of Manchester, Institute of Science and Technology, P.O. Box 88, Manchester M60 1QD, UK.

[5] Dept. of Business Studies, Univ. of Salford, Salford M5 4WT, UK.

Note. Discussion open until December 1, 2002. Separate discussions must be submitted for individual papers. To extend the closing date by one month, a written request must be filed with the ASCE Managing Editor. The manuscript for this paper was submitted for review and possible publication on March 13, 2001; approved on September 5, 2001. This paper is part of the ***Journal of Management in Engineering***, Vol. 18, No. 3, July 1, 2002. ©ASCE, ISSN 0742-597X/2002/3-111–119/$8.00+$.50 per page.

data for each contractor to be prequalified—and evaluate contractors based on this (available) information—which may be both quantitative and qualitative in nature. Very often, there is also an inherent degree of uncertainty in these data.

Several strategies and approaches have been proposed to evaluate prequalification criteria in the construction literature. Russell and Skibniewski (1988) explained five such strategies. These were dimensional weighting, two-step prequalification, dimension-wide strategy, prequalification formulas, and subjective judgment:

- In dimensional weighting, decision makers are asked to evaluate contractors on a 1 to 10 scale, 1 being unsatisfactory and 10 being satisfactory. Then, a contractor's score is calculated as a weighted sum of ratings over all decision criteria (i.e., scores *x* weights);
- In two-step prequalification, the client has to identify the most important criteria for the first stage prequalification. If a contractor satisfies these criteria, it goes forward to the second stage. If not, the contractor is eliminated from further consideration;
- In a dimension-wide strategy, decision criteria are ranked in descending order from the most important to the least important. Then, one decision criterion (starting with the most important one) is considered at a time, and all contractors are evaluated only on this criterion. Any candidate contractor who fails to satisfy this criterion is eliminated. Successful contractors are then considered for the next most important criterion *ad infinitum*;
- The prequalification formula is used to reflect a client's requirements and objectives; and
- The strategy of subjective judgment is entirely based on decision makers' knowledge and experience in the industry including, for example, previous relationships with the alternative contractors. This strategy is subject to biases of the DM.

Russell and Skibniewski (1990) developed a computer program called QUALIFIER-1 to aid decision makers in prequalification. This program was based on an aggregated weighing for each contractor obtained through the input rating for each decision criterion. Prequalification criteria, also called composite decision factors, are displayed hierarchically and each criterion is broken down into further criteria—decision factors—each of which has a different contribution to the evaluation of the associated upper level criterion. The advantages and disadvantages of this computer program were listed. The following two principal disadvantages of QUALIFIER-1 can be overcome by applying ER to prequalify contractors:

1. QUALIFIER-1 suffered from inability to adequately represent the risk profile of the DM and deal with the uncertainty associated with data collected on candidate contractors; and
2. Other algebraically defined formulations for the presented model could be investigated (Russell and Skibniewski 1990).

Russell et al. (1990) further developed QUALIFIER-1 by adding extra functions (such as a help function) thereby producing QUALIFIER-2. This program enables users to carry out sensitivity analysis and include heuristic knowledge in the analysis. However, QUALIFIER-2 still does not adequately deal with uncertainties associated with heuristic knowledge.

Holt et al. (1994b) classified the contractor selection process into three stages: (1) prequalification; (2) contractor evaluation; and (3) final selection. For each stage, three types of score were proposed (P1, P2, and P3, respectively). P1 scores represent the general organizational attributes of a contractor, and also provide an insight of specific contractor weakness(es). A multiattribute analysis technique was used to combine P2 scores (these representing the scores of more specific criteria) and P3 scores (representing bid amount) into a simple index. This index is determined by assigning a 40% weighting to the P2 scores and a 60% weighting to the P3 scores. Sensitivity analysis had revealed these percentages to best discriminate between contractors.

Holt et al. (1993, 1994a,b,c,d, 1995) provided example applications of multiattribute analysis to the evaluation of construction bidders. Holt et al. (1994d) developed a method to evaluate contractor prequalification criteria and provided guidelines for practitioners, highlighting areas to address when evaluating a contractor based on a particular criterion. Holt (1996) applied cluster analysis to reduce a large number of potential bidders, to identify only those most suitable to tender for a particular project. Ng (1996) studied different decision support systems for contractor prequalification and used a case-based reasoning approach to prequalify contractors. The prequalification criteria used and their relative importance were found to be different among type of clients, i.e., between public, private, architects, or engineers (Ng 1996). Hatush and Skitmore (1997b) applied a program evaluation and review technique approach to assess and evaluate contractor data against client goals, namely time, cost, and quality. Holt (1998) reviewed the use of different contractor selection methods and the following were identified as having been applied in this context: bespoke approaches, multiattribute analysis, multiattribute utility theory, cluster analysis, multiple regression, fuzzy set theory, and multivariate discriminant analysis. The advantages and disadvantages of these methods were also discussed. Despite this previous research, the problem of dealing with different types of assessments (qualitative, quantitative, and/or stochastic values); uncertain, incomplete, and vague assessments; and reconciling quantitative and qualitative CPP data remains. As Holt et al. (1993) concluded: "... [a contractor] selection process developed into an expert system would be welcomed by the UK construction industry." It is these aspects that this paper concentrates upon by applying the ER technique through the support of an intelligent decision system.

Multiple Criteria Decision Making

MCDM problems consist of multiple criteria, alternatives, and a DM or a group of decision makers (a single DM is assumed throughout the paper). The methodology of multiple criteria decision making can be divided into three steps: (1) structuring the decision problem, (2) formulating a preference model, and (3) evaluating and comparing alternatives (Ozernoy 1992). A MCDM problem with both qualitative and quantitative criteria is usually structured in a hierarchy. The goal of a typical MCDM problem is usually to select a best alternative, A_i, from a set of *n* alternatives ($A=\{A_1,A_2,\ldots,A_n\}$). Let us denote multiple criteria as $C_1,C_2,\ldots,C_m$. Each criterion may have different numbers of subcriteria denoted as $C_{11},C_{12},\ldots,C_{1k},C_{21},C_{22},\ldots,C_{2l},\ldots,C_{m1}, C_{m2},\ldots,C_{mp}$. A DM is first asked to determine the relative importance (i.e., weights) of main criteria such that $W_1,W_2,\ldots,W_m$ and subcriteria $W_{11},W_{12},\ldots,W_{1k},W_{21}, W_{22},\ldots, W_{2l},\ldots, W_{m1},W_{m2},\ldots,W_{mp}$. These weights are used for propagating lower level criteria assessments to respective upper levels and normalized so that $\Sigma_{i=1}^{m}W_i=1$. Then, the DM is required to evaluate *n* alternatives against the predetermined criteria.

There are several evaluation tools and methods proposed in the literature (see Hwang and Yoon 1981). The majority of these methods require certain and precise assessments from the DM. However, this may not be the case in real-world decision problems since most decision problems, if not all, are made under uncertainty and risk, time pressure, and in the presence of some information that is uncertain, incomplete, and/or missing.

Evidential Reasoning Approach

The ER approach (developed on the basis of decision theory and the Dempster-Shafer theory of evidence) offers a rational and reproducible methodology to aggregate uncertain, incomplete, and vague data (Yang and Sen 1994; Yang 2001). ER uses the concept of *degree of belief* to elicit a decision maker's preferences. The degree of belief can be described as the degree of expectation that an alternative will yield an anticipated outcome on a particular criterion. An individual's degree of belief depends on the knowledge of the subject and the experience. The use of belief functions can be justified by the fact that it may not always be reasonable or practical to expect or force individuals to make certain and precise assessments when evaluating decision criteria. This situation is due to several reasons. First, humans are not machines, so they tend to make judgments intuitively. Second, a DM may not always have adequate knowledge and/or experience to make certain assessments. Third, DMs may intuitively feel more comfortable providing their judgments in linguistic terms (rather than numerically), which, due to subjectivity, leads to ambiguity in human decision making (Poyhonen et al. 1997). Finally, information about decision criteria (and/or decision alternatives) may be incomplete, imprecise, or unavailable to the DM. ER can deal with all of these problems thereby making it particularly suitable for the CPP.

Basic Concepts of Dempster and Shafer Theory

The idea of evidential reasoning was first introduced by Dempster (1967) and further extended by his student Shafer (1976). The ER algorithm is based on the theory of evidence developed by Dempster and Shafer (D-S theory) and a detailed explanation is given in (Yang and Sen 1994; Yang 2001). Therefore, it is the writers' intention to present only the operation of the combination rule of D-S theory, followed by an example.

The D-S theory uses a number between 0 and 1 to indicate the degree of belief (or degree of support) that a body of evidence provides for a proposition, which could represent a set of multiple hypotheses instead of a single hypothesis (Palacharla and Nelson 1994). An important feature of the D-S theory, as pointed out by Pearl (1990) is that "an expert [or a DM] may feel more comfortable describing the impact of an evidence in terms of weight assignment to classes rather than to individual points" (Murphy 2000, p. 1).

Suppose that a DM is asked to evaluate the past performance of a contractor against a number of grades (i.e., worst, bad, average, good, and excellent). The DM may evaluate one of the alternative contractors as 50% average, 40% good, and 10% excellent based on available information and evidence. A useful feature of the D-S theory is that belief in a hypothesis and its negation do not have to add to 1. In the previous example, the DM may have evaluated the contractor as 50% average and 40% good. The unassigned degree of belief (i.e., the remaining 10%) may be due to uncertainty caused by a lack of information or a shortage in expertise. Such a feature of D-S theory allows users to represent and manage uncertainty. As Smets (1999, p. 613) rightly put it, "a nice property of belief functions is that only what is known is used."

The *frame of discernment*, denoted by Θ, is a sample space in the D-S theory and is a finite nonempty set of propositions. A basic proposition is denoted by H_s, i.e., $H_s \subseteq \Theta$. In Θ, all propositions are required to be mutually exclusive and exhaustive. A probability *mass function* to every subset X of Θ $(X \subseteq \Theta)$ can be assigned, denoted by $m(X)$. Basic probability assignment (BPA) function is $m(X)$ where $m(X): 2^{\Theta} \rightarrow [0,1]$ such that $m(\phi) = 0$ and

$$\sum_{X \subseteq \Theta} m(X) = 1 \qquad (1)$$

where $0 \leq m(X) \leq 1$, for all $X \subseteq \Theta$ (Yang and Singh 1994; McErlean et al. 1999). Portion of the total belief exactly (i.e. 100%) committed to hypothesis X given a body of evidence is indicated by $m(X)$.

The quantity $m(\Theta)$ is a measure of the portion of the total belief that remains unassigned after the commitment of belief to all subsets of Θ. Ignorance or missing information in the D-S theory is $m(\Theta)$ (Vouros 2000; Smets 1999). If $m(X) = s(X \subseteq \Theta)$ and it is also known that no belief is assigned to other subsets of Θ, then $m(\Theta) = 1 - s$. Hence, the remaining belief is assigned to Θ, not to the negation of the proposition X (i.e., not the complement of X) (Yang and Singh 1994). The D-S theory also includes reasoning based on its rule of combination, which is subsequently defined. Given two BPAs such that $m_1(X)$ and $m_2(X)$, based on independent evidence, in Θ, then our task is to obtain a combined BPA denoted by $m_{12}(X)$, which is calculated according to Dempster's rule of combination as follows:

$$m_{12}(\phi) = 0 \qquad (2)$$

$$m_{12}(X) = \sum_{A \cap B = X} \frac{m_1(A) m_2(B)}{1 - K} \qquad (3)$$

$$K = \sum_{A \cap B = \phi} m_1(A) m_2(B) \qquad (4)$$

where $m_{12}(X)$ is computed from m_1 and m_2 by adding all products of the form $m_1(A) m_2(B)$ where A and B are selected from the subsets of Θ in all possible ways such that their intersection is X (Yang and Singh 1994). K is the mass that the combination assigned to null subset. It represents the contradictory evidence (Murphy 2000).

Suppose that the frame of discernment, Θ, is defined as {good, average, bad}. Then, an alternative, A_1, is assessed by using these grades with respect to two criteria, C_1 and C_2, which are assumed to have equal importance. Assume that the DM confirms that A_1 is 0.3 good, 0.5 average, and 0.2 bad in terms of the criterion C_1. In terms of criterion C_2, the DM confirms that A_1 is 0.7 good but unable to give a complete assessment [i.e., the remaining 0.3 is assigned to ignorance—$m(\Theta)$]. Table 1 shows the operation of the Dempster's combination rule.

When there are more than two rules/assessments, the first two rules are combined and then the third rule is combined with the combined result of the first two. The combination process continues until all the rules are combined in this fashion.

Evidential Reasoning Decision-Making Process

ER has increasingly been used in a diverse range of areas from engineering and management to safety, and has been applied to

Table 1. Operation of Dempster's Rule of Combination in Dempster and Shafer Theory

Evaluation grades	{G}	{A}	{B}	{Θ}	{ϕ}
Rules/assessments					
rule 1 C1	0.30	0.50	0.20	0	0
rule 2 C2					
{G} 0.7	{G}0.21	{ϕ}0.35	{ϕ}0.14	0	
{Θ} 0.3	{G}0.09	{A}0.15	{B}0.06	0	
Combined mass	0.30[a]	0.15	0.06	0	0.49[b]
Normalized mass	**0.59**[c]	**0.29**	**0.12**	0	0

[a] 0.21 + 0.09.

[b] 0.35 + 0.14.

[c] 0.30/(1 − 0.49).

different MCDM problems. Interested readers may refer to the following references for a full explanation of the method and associated algorithm: Yang and Singh (1994); Yang and Sen (1994); Wang et al. (1995); Wang et al. (1996); Yang and Sen (1997); and Yang (2001). The ER decision-making process is briefly described here in a stepwise manner:

1. Display a decision problem in a hierarchical structure;
2. Assign weights to each (main) problem criterion and also to their subcriteria (if any);
3. Choose a method for assessing a criterion either quantitatively or qualitatively;
4. Evaluate each alternative based on the lowest (i.e., bottom) level criteria in the hierarchic structure by one of the methods in (3);
5. Transform assessments between a main criterion and its associated subcriteria if they are assessed using different methods (i.e., quantitative and qualitative);
6. Quantify qualitative assessments at the top level if necessary and determine an aggregated value for each alternative; and
7. Rank alternatives and choose the one with the highest numerical value.

Prequalifying Construction Contractors

In this section, the steps of the ER decision-making process will be applied to a multiple criteria CPP. In this problem, there is a similar set of decision criteria applied to those advocated by Holt et al. (1994d) and for simplicity, the same set of criteria weights were used (note that the weights are rounded). The decision criteria are hierarchically displayed and weights are shown in Fig. 1.

When an alternative is evaluated on the criterion Contractor's Organization, for example, there are subcriteria (attributes) such as age, size, image, quality control policy, health and safety policy, and litigation tendency, which could be used as the basis of the evaluation (Fig. 2). Some of these subattributes may only be assessable using subjective judgments while the remainder might be assessed numerically. The subcriterion image is a qualitative attribute requiring subjective assessment, e.g., against a number of grades that could be used for this purpose. The following grades might apply: {none, poor, average, and good}.

A DM may then state that an alternative's image (e.g., Contractor K) is 30% average and 60% good represented by {average (0.3), good (0.6)}. In this statement, average and good are the two distinctive grades and the numbers 30 and 60 are called the degrees of belief of the DM. Since the degrees of belief indicated here do not total 1 or 100% (0.3+0.6<1 or 30%+60% <100%), the assessment is said to be incomplete. These incomplete assessments are likely to occur in real life problems because of a lack of data and evidence, or because the DM is unable (or not prepared) to make assessments due to a lack of expertise on such criteria. Individuals are required to support their assessments with evidence and justify the degree of belief that they assign to each grade. In the previous assessment, for example, grades are to be defined by the DM.

Contractor Evaluation

Five alternative contractors were considered in order to simplify the calculations. The contractors have been assessed by a DM

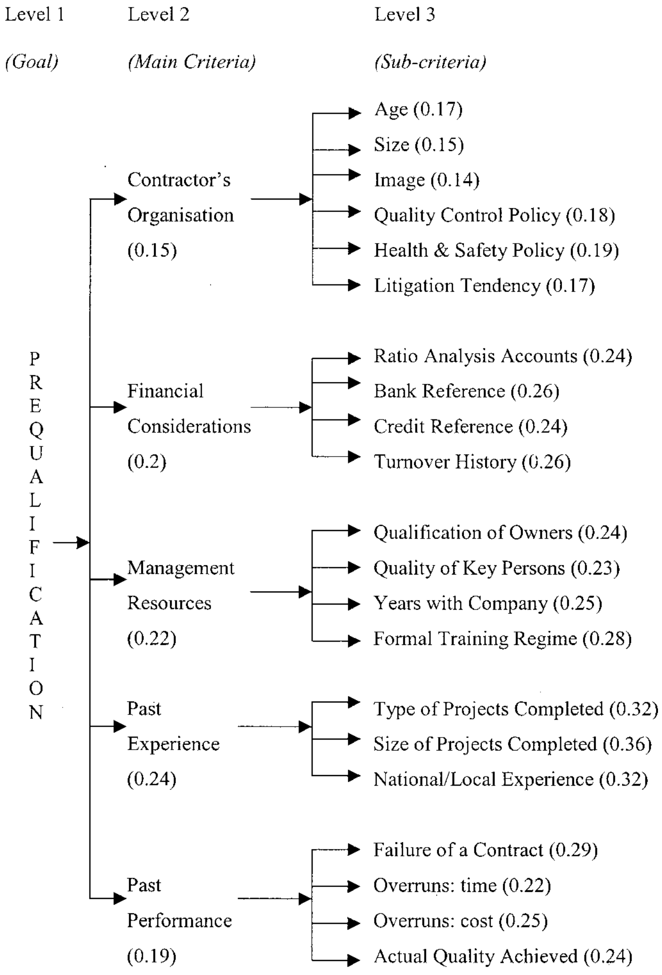
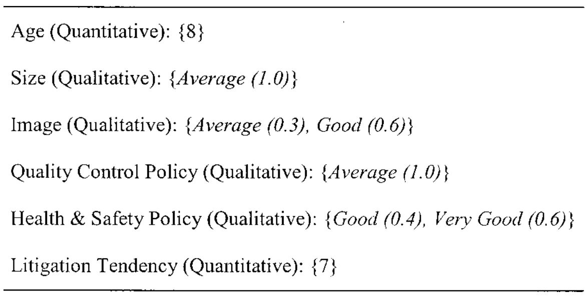


Fig. 1. Hierarchical display of prequalification of construction contractors problem (Holt et al. 1994d)


Age (Quantitative): {8}

Size (Qualitative): {*Average (1.0)*}

Image (Qualitative): {*Average (0.3), Good (0.6)*}

Quality Control Policy (Qualitative): {*Average (1.0)*}

Health & Safety Policy (Qualitative): {*Good (0.4), Very Good (0.6)*}

Litigation Tendency (Quantitative): {7}

Fig. 2. Assessments made on subcriteria of contractor's organization

(i.e., client) based on available information and expertise as shown in Table 2. A set of definitions of decision criteria and areas to address when assessing alternative contractors were given by Holt et al. (1994d) and these were used for undertaking this assessment. Table 2 shows that the DM used different types of grades when evaluating contractors. The use of different grades facilitates data collection and allows capture of the DM's preferences, experience, intuition, or beliefs and also implies that the DM is not manipulated either by the method or decision analyst who may help during the decision process. This is because the DM use his or her own expressions to evaluate decision criteria. Although this may increase ambiguity, uncertainty, or imprecision in the data, the ER approach facilitates this through rule and utility based knowledge transformation, which will be explained later.

Table 2. Assessment Scores of Contractors based on Subcriteria

		Contractors				
Criteria	Subcriteria	K	L	M	N	O
Contractor's organization	Age	8	5	3	7	10
	Size	Average (1.0)	Average (0.2); large (0.6)	Short (1.0)	Average (0.3); large (0.6)	Large (0.4); very large (0.6)
	Image	Average (0.3); good (0.6)	Poor (0.25); average (0.75)	None (0.5); poor (0.4)	Average (0.2); good (0.7)	Poor (0.5); average (0.5)
	Quality control policy	Available (1.0)	Intention to have (1.0)	None (1.0)	Intention to have (1.0)	Available (1.0)
	Health and safety policy	Good (0.6); very good (0.4)	Average (0.4); good (0.5)	Very poor (0.2); poor (0.7)	No information	Average (0.4); good (0.4)
	Litigation tendency	7	No information	2	5	3
Financial considerations	Ratio analysis accounts	Low (0.4); average (0.4)	Very low (0.1); low (0.8)	High (0.6); very high (0.3)	Average (0.4); high (0.5)	High (0.75); very high (0.25)
	Bank reference	Poor (0.6); average (0.3)	Very poor (0.85); poor (0.15)	No information	Average (1.0)	Good (0.2); very good (0.7)
	Credit reference	No information	Poor (0.6); average (0.3)	Good (0.8); very good (0.1)	Average (0.6); good (0.3)	Good (0.5); very good (0.4)
	Turnover history	Average (0.5); good (0.4)	Poor (0.2); average (0.7)	Good (0.4); excellent (0.5)	Average (0.6); good (0.4)	Good (0.5); excellent (0.4)
Management resources	Qualification of owners	Poorly qualified's (0.3); average (0.6)	Poorly qualified (0.5); average (0.5)	Highly qualified (0.6); excellent (0.4)	Highly qualified (0.8); excellent (0.2)	Highly qualified (0.35); excellent (0.65)
	Qualification of key persons	Average (0.2); Highly qualified (0.8)	Poorly qualified (0.3); average (0.7)	Average (0.8); highly qualified (0.2)	Highly qualified (0.5); excellent (0.5)	Highly qualified (0.6); excellent (0.4)
	Years with company	5	2	2	5	10
	Formal training regime	None (0.5); poor (0.4)	None (0.3); poor (0.5)	Poor (0.5); satisfactory (0.4)	Satisfactory (0.3); good (0.6)	No information
Past experience	Type of projects completed	No evidence (0.2); little evidence (0.7)	Little evidence (0.6); sufficient evidence (0.3)	Sufficient evidence (0.5); clear evidence (0.4)	Sufficient evidence (0.6); clear evidence (0.3)	Clear evidence (0.7); very clear evidence (0.2)
	Size of projects completed	Very small (0.45); small (0.55)	Small (0.5); average (0.4)	Very small (0.4); small (0.6)	Average (0.3); big (0.6)	Big (0.6); very big (0.3)
	National/local experience	Few (0.4); a few (0.5)	Average (0.65); some (0.35)	Few (0.5); a few (0.4)	Some (0.75); many (0.25)	Average (0.85); some (0.15)
Past performance	Failure of a contract	3	5	4	2	1
	Overruns: time	Very late (1.0)	Late (1.0)	Late (0.5); on time (0.5)	On time (1.0)	On time (1.0)
	Overruns: cost	Very much (0.5); much (0.5)	Much (0.6); little (0.4)	Little (0.3); very little (0.7)	Very little (0.65); none (0.35)	Very little (0.6); none (0.4)
	Actual quality achieved	Very low (0.35); low (0.65)	Average (0.55); high (0.45)	High (0.65); very high (0.35)	Average (0.25); high (0.75)	Very high (0.55); outstanding (0.45)

Table 3. Transformation of Lowest Level Criteria Assessments to Upper Levels

Upper level criterion	Contractor's organization				
Subcriteria/assessment grades	Very bad	Bad	Average	Good	Very good
Age (years)	1	3	5	9	12
Size	Very small	Small	Average	Large	Very large
Image	None; poor (0.1)	Poor (0.9)	Average (0.85)	Average (0.15); good (0.8)	Good (0.2)
Quality control policy	None	—	Intention to have	—	Available
Health and safety records	Very poor	Poor	Average	Good	Very good
Litigation tendency (x times)	8	6	4	2	0

The evaluation scores in Table 2 can be categorized as (1) subjective-quantitative, (2) certain/uncertain, and (3) complete/incomplete. Take the criterion ratio analysis of accounts for example. Contractor K was assessed as 40% low and 40% average in Table 2. This assessment is said to be both subjective and incomplete, as the degrees of belief of the DM do not add up to one. The remaining 20% may represent a lack of evidence or missing information in the contractor's prequalification submission. Contractor M has not provided any information regarding the criterion bank reference, so a detailed assessment cannot be made. The incomplete assessments can be due to uncertainty and the associated doubts of the DM. The missing assessments may be unavoidable because there is a lack of information for some alternatives for a particular criterion or the DM may not have enough knowledge or expertise to make an assessment on a particular criterion. However, the ER approach has the capability to deal with such a set of assessments through the support of an intelligent decision system (IDS) developed by Yang and Xu (2000).

Assessment Transformation

Suppose that a DM has made assessments on Contractor K regarding those subcriteria of contractor's organization as shown in Fig. 2. This assessment is a mix of quantitative and qualitative evaluation. It needs to be combined and transformed to the associated upper level so that a single and aggregated evaluation (index) can be found for this upper-level criterion, contractor's organization. This transformation can be either rule or utility-based depending on the decision maker's preference. An example of the transformation process can be seen in Table 3. The criterion contractor's organization consists of six subcriteria, two of which are of quantitative nature. Since contractor's organization is evaluated against five verbal grades, all subcriteria assessments need to be transformed to these grades. If the subcriterion is evaluated against the same number of verbal grades as the associated upper-level criterion, then the transformation is straightforward. However, if the number of verbal grades is different, then a rule-based transformation is necessary. For example, a contractor's image is assessed using four verbal grades and this needs to be transformed to the five grades. The following rule is given by the DM based on his knowledge and subjective experience. A contractor with no image is thought to have a very bad organization as far as the subcriterion image is concerned while if a contractor has a poor image, then this contractor's organization is said to be, to a small extent, very bad (0.1) and, to a large extent, bad (0.9). If one looks at the quantitative criterion litigation tendency for example, a contractor with a high litigation tendency record (eight times) in the past is considered to have a very bad organization in this respect. A contractor with no current legal actions (zero litigation tendency record) is considered to be a very good organization. It should be noted that no data was lost during this transformation process (Yang 2001).

Table 3 shows that a contractor with 6 litigation tendencies indicates a bad organization. This raises the question: How does one assess a contractor with intermediate values, e.g., 1, 3, 5, etc.? Let us consider 5, which is half-way between 6 (equivalent of grade bad) and 4 (equivalent of grade average). The contractor's organization is assumed to be 50% bad and 50% average. Here, DMs are assisted by computer software developed based on the evidential reasoning approach when converting lower level attributes' assessments to the respective upper level.

In some cases, individuals may want to use utility as a means of transformation. Suppose that the DM has assigned the following utility values to each outcome of the subcriterion age (age of a contractor refers to the length of time that it has been in the business): $u(1)=0$, $u(3)=0.25$, $u(5)=0.5$, $u(9)=0.75$, and $u(12)=1$. If a contractor is assessed by any intermediate value, the utility can be found by using a piece-wise linear function (Yang and Sen 1996). Assuming that a contractor has been in the construction market for 10 years, its utility can be found using the following piece-wise linear function formula:

$$u(10)=\frac{u(12)-u(9)}{12-9}(10-9)+u(9) \qquad (5)$$

This contractor therefore, has a utility of 0.833 with regard to the subcriterion age. Once all subcriteria assessments are converted to their associated upper levels, the aggregated assessment under upper-level criteria are required to be propagated to the top level as can be seen in Table 4.

A contractor's organization (one of the five main criteria) that is very bad, is thought to be the worst alternative while another one with a very good organization is said to be an excellent prequalified alternative. Such transformations need to be carried out for all other criteria simultaneously. That is, the ER algorithm aggregates all the degrees of belief of the DM from the bottom to the top levels.

Results and Discussion

The evaluations given by the DM in Table 2 were fed into the computer program IDS via evidential reasoning. The IDS has

Table 4. Example of Transforming Upper-Level Criterion Assessments to Top Level

Select the best prequalified contractors	Contractor's organization
Worst	Very bad
Bad	Bad
Average	Average
Good	Good
Excellent	Very good

Table 5. Combined Assessments of Contractors at Main Criteria Level

	Contractors				
Main criteria	K	L	M	N	O
Contractor's organizations	Very bad (0.08); Bad (0.11) Average (0.29) Good (0.23) Very good (0.28)	Very bad (0.16) Bad (0.13) Average (0.47) Good (0.19) Very good (0.00)	Very bad (0.36) Bad (0.46) Average (0.00) Good (0.15) Very good (0.00)	Very bad (0.00) Bad (0.13) Average (0.45) Good (0.19) Very good (0.02)	Very bad (0.01) Bad (0.06) Average (0.30) Good (0.27) Very good (0.33)
Financial considerations	Worst (0.00) Very poor (0.12) Poor (0.15) Average (0.29) Strong (0.07) Very strong (0.03) Excellent (0.00)	Worst (0.27) Very poor (0.21) Poor (0.23) Average (0.21) Strong (0.00) Very strong (0.00) Excellent (0.00)	Worst (0.00) Very poor (0.00) Poor (0.00) Average (0.00) Strong (0.24) Very strong (0.28) Excellent (0.14)	Worst (0.00) Very poor (0.00) Poor (0.03) Average (0.67) Strong (0.15) Very strong (0.11) Excellent (0.00)	Worst (0.00) Very poor (0.00) Poor (0.00) Average (0.00) Strong (0.26) Very strong (0.27) Excellent (0.39)
Management resources	None (0.25) Poor (0.54) Average (0.16) Good (0.00)	None (0.48) Poor (0.47) Average (0.00) Good (0.00)	None (0.18) Poor (0.39) Average (0.31) Good (0.09)	None (0.05) Poor (0.18) Average (0.41) Good (0.34)	None (0.00) Poor (0.00) Average (0.49) Good (0.23)
Past experience	Very low (0.21) Low (0.60) Average (0.13) High (0.00) Very high (0.00)	Very low (0.00) Low (0.36) Average (0.47) High (0.10) Very high (0.00)	Very low (0.32) Low (0.37) Average (0.14) High (0.11) Very high (0.00)	Very low (0.00) Low (0.00) Average (0.27) High (0.59) Very high (0.07)	Very low (0.00) Low (0.00) Average (0.25) High (0.53) Very high (0.16)
Past performance	Very poor (0.48) Poor (0.24) Average (0.14) High (0.14) Very high (0.00)	Very poor (0.00) Poor (0.32) Average (0.65) High (0.03) Very high (0.00)	Very poor (0.00) Poor (0.00) Average (0.62) High (0.27) Very high (0.11)	Very poor (0.00) Poor (0.02) Average (0.13) High (0.57) Very high (0.28)	Very poor (0.00) Poor (0.00) Average (0.00) High (0.38) Very high (0.62)

produced the following combined and distributed assessment results at the main criteria level as shown in Table 5 by using the combination rule explained earlier. A demonstration version of IDS with an example model of the contractor prequalification problem presented in this paper can be obtained from J. B. Yang through e-mail request at jian-bo.yang@umist.ac.uk.

In Table 5, the numbers in brackets are the aggregated degrees of belief of the DM and due to missing or incomplete assessments given in Table 2, the total degree of belief for each alternative on each criterion does not total 1. Table 5 is useful in that it gives an initial idea of how contractors perform on each main decision criterion and allows the client to see the strong and weak points of each candidate contractor. It may also be displayed in a graphical format allowing the DM to see the performance of contractors at a glance. At this stage, however, it is difficult for the DM to rank five contractors in order of preference by looking at the results presented in Table 5. The distributed assessments in Table 5 are therefore propagated to the top level through rule and/or utility based transformation and by considering the relative importance of decision criteria. The overall scores produced by IDS are shown in Table 6.

Each number under each grade in Table 6 again indicates the aggregated degrees of belief of the DM. The results in Table 6 can be described as distributed because each contractor has been, to some extent, assessed to more than one evaluation grade. Contractor K, for example, is assessed to be 22% worst, 36% bad, 23% average, 7% good and 4% excellent. By looking at the results presented in Table 6, one could rank the alternatives. However, in some cases alternatives may have very close assessments and it may be very hard to rank them. One solution to this problem is to quantify the grades at the top level—the goal of the problem. There are several ways of doing this. One of them is to use multiple attribute utility theory (Hwang and Yoon 1981) in the form of lottery type questions. Another is to use goal-programming techniques as suggested in Yang and Sen (1996). The former appeared the most appropriate for the construction

Table 6. Overall Assessment of Alternative Contractors

Alternative contractor	Grades				
	Worst	Bad	Average	Good	Excellent
K	0.22	0.36	0.23	0.07	0.04
L	0.18	0.32	0.39	0.05	0.00
M	0.17	0.23	0.28	0.19	0.05
N	0.01	0.05	0.42	0.37	0.09
O	0.00	0.01	0.23	0.36	0.30

Table 7. Expected Utilities of Alternative Contractors

Alternative contractors	Minimum utility	Maximum utility	Average utility
K	0.41	0.49	0.45
L	0.44	0.50	0.47
M	0.50	0.58	0.54
N	0.72	0.78	0.75
O	0.77	0.87	0.82

Table 8. Aggregated Results for Contractor K from Table 6

Grades	Scores	Utility
Worst	0.22	0
Bad	0.36	0.4
Average	0.23	0.7
Good	0.07	0.85
Excellent	0.04	1.0
Unassigned degree of belief	0.08	—

Note: Min $u(K)=[(0.22+0.08)\times u(\text{worst})]+[0.36\times u(\text{bad})]+[0.23\times u(\text{average})]+[0.07\times u(\text{good})]+[0.04\times u(\text{excellent})]$

Min $u(K)=[0.30\times 0]+[0.36\times 0.4]+[0.23\times 0.7]+[0.07\times 0.85]+[0.04\times 1]=\mathbf{0.41}$

Max $u(K)=[0.22\times u(\text{worst})]+[0.36\times u(\text{bad})]+[0.23\times u(\text{average})]+[0.07\times u(\text{good})]+[(0.04+0.08)\times u(\text{excellent})]$

Max $u(K)=[0.22\times 0]+[0.36\times 0.4]+[0.23\times 0.7]+[0.07\times 0.85]+[0.12\times 1]=\mathbf{0.49}$

Average $u(K)=[\text{Min } u(K)+\text{Max } u(K)]/2=(0.41+0.49)/2=\mathbf{0.45}$.

sector and the following utility values for each grade were assigned at the top level: worst $=0$, bad $=0.4$, average $=0.7$, good $=0.85$, and excellent $=1$. The concept of utility is used as a tool to quantify the verbal evaluation grades so that the contractors can be ranked in order of preference based on expected utility.

In Table 6, the total degree of belief assigned to alternative Contractor M is 92%. There is a 8% unassigned degree of belief. Due to the unassigned degrees of belief, the IDS yields two expected utility values for each alternative: a minimum and a maximum expected utility. In reality, the unassigned degree of belief can fall into any grades. However, the ER algorithm considers the two most extreme cases where the unassigned degree of belief either goes to the least preferred grade with minimum utility or to the most preferred grade with maximum utility. The minimum and maximum utilities shown in Table 7 were generated by the IDS based on the given utility values for each grade above. The minimum and maximum utilities of Contractor K, for example, are found in Table 8. The expected utility of an alternative is defined as the total degree of belief assigned to each grade times the utility of each grade.

The alternatives may be ranked based on the average utility but this may be misleading. In order to say with confidence that one alternative dominates another, the preferred alternative's minimum utility must be equal to or greater than the dominated alternative's maximum utility. For example, based on an average utility alternative, Contractor O is preferred to alternative Contractor N, but this comparison may differ if it is based on the maximum and minimum utilities. There is a small possibility that alternative N may be preferred to alternative O if the missing assessments for N go to the most preferred grade and those for O go to the least preferred grade. Therefore, it would not be said for sure that contractor O dominates contractor N because N's maximum utility is greater than the minimum utility of O (i.e., $0.78 > 0.77$). To precisely differentiate alternatives O and N, the quality of the original assessments related to the two alternatives given in Table 2 need to be improved. In response to the decision maker's request for simplicity, average utility is used to rank alternatives. The ranking of alternatives is as follows: $\text{O}>\text{N}>\text{M}>\text{L}>\text{K}$. Suppose that the DM introduced a cut-off point for prequalification as follows: the contractors whose average utilities equals less than 0.5 will not be considered for the second stage (i.e., final contractor selection). Therefore, the Contractors K and L (Table 7) are to be eliminated from further consideration based on this statement.

Conclusion

This paper introduced an application of evidential reasoning to solve a multiple criteria contractor prequalification problems with uncertain, incomplete (imprecise), and/or missing information. The process of building a MCDM model of a hierarchical structure was presented, in which both quantitative and qualitative information was represented in a unified manner through equivalent knowledge transformation. It is shown that the ER approach can handle both quantitative and qualitative data, which may be vague and/or incomplete. It can be used as a MCDM method enabling a DM to give a judgment according to their knowledge, expertise, and available information at the time a decision is made. It is important to obtain the decision maker's true preferences in a decision-making problem in order to ensure that a rational decision can be made based on the real preferences of the DM. The ER approach provides this by using the concept of degree of belief.

It must be noted that the writers do not claim that the ER approach is superior to other MCDM approaches such as multiple-attribute utility theory and analytic-hierarchy process. The aim was to introduce this approach to practitioners as an alternative MCDM approach, which may also be used as a sensitivity analysis tool to verify the results obtained by other construction procurement methods. The method presented in this paper and the decision-support tool could also be used as a database for monitoring the progress of hundreds of contractors and/or suppliers in the construction industry. Whenever a piece information becomes available about one contractor, the previous evaluation on this contractor can be upgraded immediately. The accuracy and reliability of the assessments given by a particular assessor can also be checked by others since all the evaluations are stored in the decision-support tool.

References

Bubshait, A. A., and Al-Gobali, K. H. (1996). "Contractor prequalification in Saudi Arabia." *J. Manage. Eng.*, 12(2), 50–54.

Dempster, A. P. (1967). "Upper and lower probabilities induced by a multivalued mapping." *Ann. Math. Stat.*, 38, 325–339.

Hatush, Z., and Skitmore, M. (1997a). "Criteria for contractor selection." *Constr. Manage. Econom.*, 15(1), 19–38.

Hatush, Z., and Skitmore, M. (1997b). "Assessment and evaluation of contractor data against client goals using PERT approach." *Constr. Manage. Econom.*, 15(4), 327–340.

Holt, G. D. (1996). "Applying cluster analysis to construction contractor classification." *Build. Environ.*, 31(6), 557–568.

Holt, G. D. (1998). "Which contractor selection methodology?" *Int. J. Proj. Manage.*, 16(3), 153–164.

Holt, G. D., Olomolaiye, P. O., and Harris, F. C. (1993). "A conceptual alternative to current tendering practice." *Build. Res. Inf.*, 21(3), 167–172.

Holt, G. D., Olomolaiye, P. O., and Harris, F. C. (1994a). "Factors influencing U.K. construction clients' choice of contractor." *Build. Environ.*, 29(2), 241–248.

Holt, G. D., Olomolaiye, P. O., and Harris, F. C. (1994b). "Applying multi-attribute analysis to contractor selection decisions." *Eur. J. Purchasing Supply Manage.*, 1(3), 139–148.

Holt, G. D., Olomolaiye, P. O., and Harris, F. C. (1994c). "Incorporating project specific criteria and client utility into the evaluation of construction tenderers." *Build. Res. Inf.*, 22(4), 214–221.

Holt, G. D., Olomolaiye, P. O., and Harris, F. C. (1994d). "Evaluating pre-qualification criteria in contractor selection." *Build. Environ.*, 29(4), 437–448.

Holt, G. D., Olomolaiye, P. O., and Harris, F. C. (1995). "Application of an alternative contractor selection model." *Build. Res. Inf.*, 23(5), 255–264.

Hwang, C. L., and Yoon, K. (1981). *Multiple attribute decision making methods and applications: A state-of-the-art survey*, Springer, New York.

Kim, S. H., and Ahn, B. S. (1999). ''Interactive group decision making procedure under incomplete information.'' *Eur. J. Oper. Res.,* 116(3), 498–507.

McErlean, F. J., Bell, D. A., and Guan, J. W. (1999). ''Modification of belief in evidential causal networks.'' *Inf. Software Technol.,* 41(9), 597–603.

Murphy, C. K. (2000). ''Combining belief functions when evidence conflicts.'' *Decision Support Syst.,* 29(1), 1–9.

Ng, S. T. (1996). ''Case-based reasoning decision support for contractor pre-qualification.'' PhD thesis, Univ. of Manchester, Institute of Science and Technology, Manchester, U.K.

Ozernoy, V. M. (1992). ''Choosing the 'best' multiple criteria decision-making method.'' *INFOR,* 30(2), 159–171.

Palacharla, P., and Nelson, P. (1994). ''Understanding relations between fuzzy logic and evidential reasoning methods.'' *Proc., 3rd IEEE Int. Conf. on Fuzzy Systems,* Vol. 2, IEEE, New York, 1933–1938.

Pearl, J. (1990). ''Reasoning with belief functions: An analysis of compatibility.'' *Int. J. Approx. Reasoning,* 4(5/6), 363–389.

Poyhonen, M. A., Hamalainen, R. P., and Salo, A. A. (1997). ''An experiment on the numerical modelling of verbal ratio statements.'' *J. Multi-Criteria Decision Anal.,* 6(1), 1–10.

Russell, J. S. (1990). ''Model for owner prequalification of contractors.'' *J. Manage. Eng.,* 6(1), 59–75.

Russell, J. S., Hancher, D. E., and Skibniewski, M. J. (1992). ''Contractor prequalification data for construction owners.'' *Constr. Manage. Econom.,* 10(2), 117–135.

Russell, J. S., and Skibniewski, M. J. (1988). ''Decision criteria in contractor prequalification.'' *J. Manage. Eng.,* 4(2), 148–164.

Russell, J. S., and Skibniewski, M. J. (1990). ''QUALIFIER-1: Contractor prequalification model.'' *J. Comput. Civ. Eng.,* 4(1), 77–90.

Russell, J. S., Skibniewski, M. J., and Cozier, D. R. (1990). ''QUALIFIER-2: Knowledge-based system for contractor prequalification.'' *J. Constr. Eng. Manage.,* 116(1), 157–171.

Shafer, G. (1976). *A mathematical theory of evidence*, Princeton University Press, Princeton, N.J.

Smets, P. (1999). ''Practical uses of belief functions.'' *Proc. 15th Conf. on Uncertainty in Artificial Intelligence*, K. B. Laskey and H. Prade, eds., Morgan Kaufmann, San Francisco, 612–621.

Valls, A., and Torra, V. (2000). ''Using classification as an aggregation tool in MCDM.'' *Fuzzy Sets Syst.,* 115(1), 159–168.

Vouros, G. (2000). ''Representing, adapting and reasoning with uncertain, imprecise and vague information.'' *Expert Syst. Applic.,* 19(3), 167–192.

Wang, J., Yang, J. B., and Sen, P. (1995). ''Safety analysis and synthesis using fuzzy sets and evidential reasoning.'' *Reliability Eng. Syst. Safety,* 47(2), 103–118.

Wang, J., Yang, J. B., and Sen, P. (1996). ''Multi-person and multi-attribute design evaluations using evidential reasoning based on subjective safety and cost analysis.'' *Reliability Eng. Syst. Safety,* 52(2), 113–128.

Yang, J. B. (2001). ''Rule and utility based evidential reasoning approach for multiattribute decision analysis under uncertainties.'' *Eur. J. Oper. Res.,* 131(1), 31–61.

Yang, J. B., and Sen, P. (1994). ''A general multi-level evaluation process for hybrid MADM with uncertainty.'' *IEEE Trans. Syst. Man Cybern.,* 34(10), 1458–1473.

Yang, J. B., and Sen, P. (1996). ''Preference modelling by estimating local utility functions for multiobjective optimization.'' *Eur. J. Oper. Res.,* 95(1), 115–138.

Yang, J. B., and Sen, P. (1997). ''Multiple attribute design evaluation of complex engineering products using the evidential reasoning approach.'' *J. Eng. Design,* 8(3), 211–230.

Yang, J. B., and Singh, M. (1994). ''An evidential reasoning approach for multiple attribute decision making with uncertainty.'' *IEEE Trans. Syst. Man Cybern.,* 24(1), 1–18.

Yang, J. B., and Xu, D. L. (2000). ''Intelligent decision system via evidential reasoning.'' *IDS Version 1.1*, IDSL, Sale, Cheshire, U.K.