Evidence Theory and Multiple Criteria Decision Analysis: The Evidential Reasoning Approach

Jian-Bo Yang [1], Dong-Ling Xu [1], Xinlian Xie [2], and Anil Kumar Maddulapalli [3]
[1] *Manchester Business School, The University of Manchester, Manchester M15 6PB, UK*
[2] *Transportation and Logistic College, Dalian Maritime University, Dalian, 116026, China*
[3] *India Science Lab, GM R&D, Bangalore, India*
Email: jian-bo.yang@mbs.ac.uk; ling.xu@mbs.ac.uk
Telephone: +44 (0)161 306 3427, Fax: +44 (0)161 306 3505

***Abstract*— The evidential reasoning (*ER*) approach has been developed to support multiple criteria decision analysis (*MCDA*). It is based on Dampster's combination rule for criteria aggregation and belief function for treating ignorance. In the original ER approach, however, alternative ranking depends on the accurate estimation of a value (utility) function, which may be difficult in certain decision environments. In this paper, the link and difference between the ER algorithm and Dampster's combination rule are analysed first. A new alternative ranking method is then investigated as an integrated part of the enhanced ER approach.**

Keywords: evidence theory, multiple criteria decision analysis, evidential reasoning, alternative ranking.

I. Basic Concepts of the Evidence Theory

The evidence theory was first investigated in 1960's (Dempster, 1967) and extended since 1970's (Shafer, 1976). It has found wide applications in many areas such as artificial intelligence, expert systems, pattern recognition, information fusion, database and knowledge discovery, multiple criteria decision analysis (MCDA), audit risk assessment, etc. (Yager, 2004; Yang and Singh, 1994; McClean and Scotney, 1997; Denoeux and Zouhal, 2001; Beynon, 2002; Yang, 2001; Yang et al., 2002, 2006; Xu and Yang, 2006). In this section, the basic concepts of the belief function and Dempster's combination rule are briefly introduced as a basis for introduction of the evidential reasoning approach in the next section.

Suppose $H = \{H_1, \cdots, H_N\}$ is a collectively exhaustive and mutually exclusive set of hypotheses, referred to as the frame of discernment. A basic probability assignment (bpa) is a mass function $m: \Theta \to [0,1]$, satisfying:

$$m(\Phi) = 0 \text{ and } \sum_{C \subseteq H} m(C) = 1 \tag{1}$$

with Φ being an empty set, C any subset of H, and Θ the power set of H, consisting of all the subsets of H, or

$$\Theta = \{\Phi, \{H_1\}, \cdots, \{H_N\}, \{H_1, H_2\}, \cdots, \{H_1, H_N\}, \cdots, \{H_1, \cdots, H_{N-1}\}, H\} \tag{2}$$

A probability mass $m(C)$ measures the degree of belief exactly assigned to C and represents how strongly C is supported by evidence. Probabilities assigned to all the subsets of H are summed to unity and there is no belief in the empty set. The probability assigned to H, or $m(H)$, is referred to as the degree of ignorance.

Associated with each bpa to C are a belief measure, $Bel(C)$, and a plausibility measure, $Pl(C)$, which are both mass functions: $\Theta \to [0,1]$, defined by the following equations:

$$Bel(C) = \sum_{B \subseteq C} m(B) \text{ and } Pl(C) = \sum_{C \cap B \neq \Phi} m(B) \tag{3}$$

$Bel(C)$ represents the exact support to C and its subsets, and $Pl(C)$ represents the possible support to C. The interval $[Bel(C), Pl(C)]$ can be seen as the lower and upper bounds of support to C. The two functions can be connected by the following equation

$$Pl(C) = 1 - Bel(\overline{C}) \tag{4}$$

where $\overline{C}$ denotes the complement of C. The difference between the belief and the plausibility of a set C describes the degree of ignorance of the assessment to C (Shafer, 1976).

The core of the evidence theory is Dempster's rule of combination by which evidence from different sources is combined or aggregated. The rule assumes that information sources are independent and uses the so-called orthogonal sum or intersection to combine multiple belief structures:

$$m = m_1 \oplus m_2 \oplus \cdots \oplus m_L \tag{5}$$

where $\oplus$ is the orthogonal sum operator. With two pieces of evidence m_1 and m_2, Dempster's rule of combination is defined as follows:

$$[m_1 \oplus m_2](D) = \begin{cases} 0, & D = \Phi \\ \dfrac{\sum_{C \cap B = D} m_1(C) m_2(B)}{1 - \sum_{C \cap B = \Phi} m_1(C) m_2(B)}, & D \neq \Phi \end{cases} \tag{6}$$

Note that Dempster's rule provides a non-compensatory process for aggregation of two pieces of evidence and can lead to irrational conclusions in the aggregation of multiple pieces of evidence in conflict (Murphy, 2000), in particular in cases where multiple pieces of evidence are compensatory in nature. On the other hand, the *ER* approach (Yang, et al., 1994, 2001, 2002, 2006) introduced in the next section provides a compensatory evidence aggregation process, which is different from Dempster's rule in that it treats basic probability assignments as weighted belief degrees, includes the concept of the degree of indecisiveness, and adopts a normalisation process for combined probability masses.

II. The main steps of the ER approach for MCDA

A *MCDA* problem can be modelled using the following belief decision matrix. Suppose M alternatives (A_l, l=1, …, M) are assessed on L criteria e_i (i=1, …, L) each on the basis of N common evaluation grades $H_n\ (n=1,\ldots,N)$, which are required to be mutually exclusive and collectively exhaustive. If alternative A_l is assessed to a grade H_n on a criterion e_i with a belief degree of $\beta_{n,i}(A_l)$, this assessment is denoted by $S_i(A_l)=S(e_i(A_l))=\{(H_n,\beta_{n,i}(A_l)),\quad n=1,\ldots,N\}$, which is a distributed assessment and referred to as a belief structure, where $\beta_{n,i}(A_l)\geq 0$ and $\sum\beta_{n,i}(A_l)\leq 1$. The individual assessments of all alternatives each on every criterion are represented by a belief decision matrix, defined as follows:

$$D_g=(S_i(A_l))_{L\times M} \tag{7}$$

Suppose ω_i is the weight of the i^{th} criterion, normalised by

$$0\leq\omega_i\leq 1 \text{ and } \sum_i\omega_i=1 \tag{8}$$

The *ER* approach has both the commutative and associative properties and as such can be used to combine assessments in any order. The *ER* aggregation process can be shown recursively, (Yang, 2001; Yang and Xu, 2002; Yang et al., 2006), summarised as the following main steps.

Step 1: *Assignment of the basic probability masses*

The basic probability masses for an assessment $S_f(A_l)$, denoted by f_n, are generated by:

$$f_n=\omega_f\beta_{n,f}(A_l) \text{ for } n=1,\ldots,N,\ f_H=\omega_f\beta_{H,f}(A_l)$$

$$f_\Theta=1-\omega_f\left(\sum_{n=1}^{N}\beta_{n,f}(A_l)+\beta_{H,f}(A_l)\right) \tag{9}$$

In the D-S theory, f_n may be interpreted as discounted belief. In *MCDA*, however, it is better interpreted as weighted belief as it means that in assessing an alternative A_l the f^{th} criterion only plays a limited role that is proportional to its weight. f_H represents the weighted ignorance in the assessment. f_Θ is referred as to the degree of indecisiveness left by $S_f(A_l)$, representing the amount of belief that is not yet assigned to any individual or subset of grades by $S_f(A_l)$ alone but needs to be jointly re-assigned in accordance with all other assessments.

Similarly, the basic probability masses g_n for another assessment $S_g(A_l)$ are generated by

$$g_n=\omega_g\beta_{n,g}(A_l) \text{ for } n=1,\ldots,N,\ g_H=\omega_g\beta_{H,g}(A_l)$$

$$g_\Theta=1-\omega_g\left(\sum_{n=1}^{N}\beta_{n,g}(A_l)+\beta_{H,g}(A_l)\right) \tag{10}$$

Step 2: *Combination of the basic probability masses*

f_n and g_n can be combined to generate aggregated probability masses using the following probabilistic combination equations:

$$\{H_n\}:\quad m_n=k\big(f_ng_n+f_n(g_H+g_\Theta)+(f_H+f_\Theta)g_n\big),\quad n=1,..,N \tag{11a}$$

$$\{H\}:\quad m_H=k\big(f_Hg_H+f_Hg_\Theta+f_\Theta g_H\big) \tag{11b}$$

$$\{\Theta\}:\quad m_\Theta=k\big(f_\Theta g_\Theta\big) \tag{11c}$$

$$k=\left(1-\sum_{n=1}^{N}\sum_{\substack{t=1\\t\neq n}}^{N}f_ng_t\right)^{-1} \tag{11d}$$

In the above equations, m_n and m_H measure the magnitudes of the joint beliefs to the grade H_n and the frame of discernment H, respectively, generated by combining the two assessments $S_f(A_l)$ and $S_g(A_l)$, with m_H thus representing the joint ignorance. m_Θ is the degree of indecisiveness left by both $S_f(A_l)$ and $S_g(A_l)$, representing the amount of belief that needs to be re-assigned back to all subsets of grades proportionally after the combination process is completed. k measures the degree of conflict between $S_f(A_l)$ and $S_g(A_l)$.

Step 3: *Generation of the combined belief degrees and distributed assessment*

If there are more assessments than two, step 2 can be repeated with each new assessment combined with m_n, m_H and m_Θ. After all assessments are combined in this recursive fashion, the finally combined probability masses need be normalised to generate the combined belief degrees β_n and β_H by proportionally re-assigning m_Θ back to all subsets of grades as follows:

$$\{H_n\}:\quad \beta_n^l=\frac{m_n}{1-m_\Theta},\ n=1,..,N \tag{12a}$$

$$\{H\}:\quad \beta_H^l=\frac{m_H}{1-m_\Theta} \tag{12b}$$

The combined assessment of the voice A_l is then given by

$$S(A_l)=\{(H_1,\beta_1^l),(H_2,\beta_2^l),\cdots,(H_N,\beta_N^l),(H,\beta_H^l)\} \tag{13}$$

The above belief distribution provides a panoramic view about the combined assessment of the alternative A_l with the degrees of strength and weakness explicitly measured by the belief degrees.

III. Comparison of Dempster's rule and the *ER* algorithm by example

The *ER* algorithm has at least the following features: 1> taking into account the relative importance of evidence; 2> taking account of indecisiveness explicitly; 3> modelling ignorance explicitly by breaking unassigned probability mass into two parts and treating them accordidngly; 4> generating rational conclusions in the combination of multiple pieces of evidence of a compensatory nature. To show these features, examine the following two pieces of evidence in conflict (Wang et al., 2006):

$m_1(A)=0.99$, $m_1(B)=0.01$, $m_1(C)=0$;

$m_2(A)=0$, $m_2(B)=0.01$, $m_2(C)=0.99$.

Before combination, the two pieces of evidence show that B is almost an unlikely event as only a tiny probability of 1% is assigned to it. After implementing Dempster's combination rule, however, B becomes a certain event and is assigned a probability of 100%, which seems irrational in terms of compensatory evidence aggregation, although this result may still be meaningful in terms of non-compensatory aggregation. Table 1 shows the results generated using Dempster's combination rule.

Table 1 Combination of conflict evidence by Dempster's combination rule

Belief Structure	{A}	{B}	{C}
m_1	0.99	0.01	0
m_2	0	0.01	0.99
$m_1 \otimes m_2$ (before normalization)	0	0.0001	0
$m_1 \otimes m_2$ (after normalization)	0	1	0

On the other hand, the *ER* algorithm treats original evidence by a belief structure, assigns a relative weight to each piece of evidence, takes into account its indecisiveness explicitly, and normalises aggregated evidence to generate a combined belief structure. Suppose the relative weights of the two pieces of evidence are given by w_1 and w_2 respectively, with $w_1+w_2=1$. Table 2 shows the *ER*-aggregated results which seem rational if the two pieces of evidence are compensatory in nature.

Table 2 The ER combined belief degrees under different sets of relative weights

Relative weights		$(m_1 \otimes m_2)$ (A)	$(m_1 \otimes m_2)$ (B)	$(m_1 \otimes m_2)$ (C)
w_1=0,	w_2=1	0	0.01	0.99
w_1=0.2,	w_2=0.8	0.058	0.01	0.932
w_1=0.4,	w_2=0.6	0.305	0.01	0.685
w_1=0.5,	w_2=0.5	0.495	0.01	0.495
w_1=0.6,	w_2=0.4	0.685	0.01	0.305
w_1=0.8,	w_2=0.2	0.932	0.01	0.058
w_1=1,	w_2=0	0.99	0.01	0

IV. A NEW RANKING METHOD FOR THE ER APPROACH

While a distribution is useful to show the diversity of an assessment, it is not convenient for ranking alternatives. On the other hand, the mean value of a distribution is usually used as a mean evaluation rating of an alternative. However, the mean value of a distribution depends on how the assessment grades are quantified. Suppose the grade H_n is quantified by $u(H_n)$. In a most general case, $u(H_n)$ should not be a single fixed value but a variable interval, denoted by $[a_n, b_n]$ with $b_n \geq a_n$ and a_n and b_n not fixed. Suppose the grade H_{n+1} is also quantified by a variable interval $[a_{n+1}, b_{n+1}]$. Then, H_{n+1} is said to be preferred to H_n if and only if $a_{n+1} > b_n$ for n = 1,…, N-1. Without loss of generality, suppose $u(H_n)$ is normalised so that $a_1 = 0$ and $b_N = 1$.

In the presence of ignorance, the lower and upper bounds of the belief degree to which an alternative A_l is assessed to H_n is given by β_n^l and $(\beta_n^l + \beta_H^l)$ respectively. Hence, the mean value of a distribution is characterised by a maximum and minimum value given as follows:

$$u_{\min}(S(A_l)) = (\beta_1^l + \beta_H^l)u(H_1) + \sum_{n=2}^{N} \beta_n^l u(H_n) \quad (14)$$

$$u_{\max}(S(A_l)) = \sum_{n=1}^{N-1} \beta_n^l u(H_n) + (\beta_N^l + \beta_H^l)u(H_N) \quad (15)$$

Then, an alternative A_l is preferred to another alternative A_k if $u_{\min}(S(A_l)) \geq u_{\max}(S(A_k))$.

The pairwise comparison of A_l and A_k can be made directly using equations (14) and (15) if the utilities of the grades are given precisely by the decision maker, or fixed to $u(H_n)= a_n = b_n$. Otherwise, we can construct the following linear programming model for comparing A_l with A_k:

$$\begin{aligned} \max \quad & \sigma_{kl} = u_{\max}(S(A_k)) - u_{\min}(S(A_l)) \\ s.t. \quad & \mathbf{X} = [(a_1, u(H_1), b_1), \cdots, (a_N, u(H_N), b_N)]^T \in \Omega_d \end{aligned} \quad (16a)$$

$$\Omega_d = \left\{ \mathbf{X} \left| \begin{array}{ll} a_n \leq u(H_n) \leq b_n & n = 1, \cdots, N \\ b_n \geq a_n & n = 1, \cdots, N \\ a_{n+1} - b_n \geq \delta & n = 1, \cdots, N-1 \\ a_1 = 0,\ b_N = 1 & \end{array} \right. \right\} \quad (16b)$$

In (16b), δ is a small non-negative real number that should be large enough to make the difference between two adjacent grades significant and meaningful, with $0 \leq \delta \leq 1/(N-1)$. If the optimal value of formulation (16a) for a sufficiently small δ is negative, or $\sigma_{kl} < 0$, it will mean that there is always $u_{\min}(S(A_l)) > u_{\max}(S(A_k))$ in any permissible ways that the assessment grades are quantified, as bounded by Ω_d, so that $S(A_l)$ is least favoured and $S(A_k)$ is most favoured. Hence, $\sigma_{kl} < 0$ means that A_l is evidentially preferred to A_k. Otherwise, there will exist no definite relation between A_l and A_k, and assumptions would have to be made or preference conditions would need to be added to (16b) in order to differentiate between A_l and A_k, which may be referred to as assumption-based or preference-based comparison. Note that from the evidential preference relations between alternatives a robust partial ranking can be generated.

In general, the larger the value of δ, the more powerful formulate (16a) is to differentiate alternatives. For example, formulation (16a) will have a unique solution of $u(H_n) = a_n = b_n = (n-1)/(N-1)$ if δ is assumed to take the maximum permissible value of $1/(N-1)$, which is equivalent to assign a single fixed value to each grade so that grades are evenly distributed in the utility space. On the other hand, a grade may be quantified to a fixed utility interval rather than to a fixed single utility. For example, each grade may be quantified to an equal utility interval so that $(a_n - b_n) = 1/(N-1)$, which is equivalent to assign δ =0, $a_1 = 0$, $a_{n+1} = b_n = (n-1)/(N-1)$ for $n = 1, \cdots, N-1$, and $b_N = 1$. Anyway, if it is necessary to make such assumptions for ranking alternatives, they should be made adequately to suit specific decision situations.

Many preference conditions may be added to formulation (16b) for differentiating alternatives. For example, if an improvement from a low grade is more appreciated than from any better grade, then the following conditions could be added to formulation (16b)

$$u(H_{n+1}) - u(H_n) \geq u(H_{n+2}) - u(H_{n+1})$$

for n=1, …, N-2 (17)

Adversely, if improvement from a high grade is more appreciated than from any lower grade, then the following conditions could be added to formulation (16)

$$u(H_{n+2}) - u(H_{n+1}) \geq u(H_{n+1}) - u(H_n)$$

for n = 1, …, N-2 (18)

On the other hand, if the degree of improvement from a grade to an adjacent better grade is bounded, the following conditions may be added to formulation (16b), with τ being a bounding index and $\tau \geq 1$,

$$\frac{1}{\tau} \leq \frac{u(H_{n+2}) - u(H_{n+1})}{u(H_{n+1}) - u(H_n)} \leq \tau \quad \text{for } n = 1, \ldots, N\text{-}2 \qquad (19)$$

V. A CASE STUDY TO SHOW THE NEW RANKING METHOD

A case study in new product development is presented to demonstrate the application of the *ER* approach coupled with the new ranking method. In this case study, thousands of data sets generated from various surveys were used to formulate a number of decision models for assessing over 160 customer voices for a world-leading car manufacturer. A customer voice is a description, stated in the customer's own words, of the benefit to be fulfilled by a product or service. It is not appropriate to present all the models in detail in this paper due to the limited space and the confidentiality. In this section, the generic structure and a typical small scale model will be discussed to illustrate the *ER* approach and the new ranking method for the prioritisation of customer voices.

A. Problem description

The hierarchy for assessing four voices using evidence generated from two surveys is shown in Figure 1, which is the main window of the decision support system named *IDS* developed to implement the *ER* approach. This main window includes a bottom-right tree view for displaying a hierarchy of assessment criteria, a bottom-left list view for listing all voices to be assessed, a menu bar for listing all the *IDS* functions, and a shortcut bar for quick access to frequently used *IDS* functions. In Figure 1, there are seven criteria at the bottom of the hierarchy, which are derived from the evidence generated from Survey 1 and Survey 2. Survey 1 has five bottom-level criteria and Survey 2 has two bottom-level criteria.

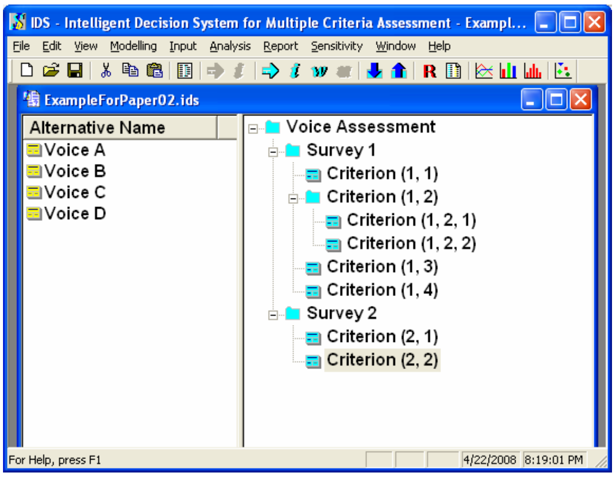

Figure 1 *IDS* main window for voice assessment

Given the nature of the two surveys, weights are suggested by the company staff in such a way that Survey 1 is regarded to be more important than Survey 2. The weight for Survey 1 was initially judged to be twice as large as that for Survey 2. The weights for Survey 1 and Survey 2 are normalised to ⅔ and ⅓ respectively, which are subject to sensitivity analysis. The other criteria in the same groups of the hierarchy are all equally weighted in this case study.

The original data sets generated from the two surveys were pre-processed into the distributions on the survey scales. The pre-processed data are then transformed to the distributed assessments on the common priority scale {NO, LOW, AVERAGE, HIGH, TOP}, which are used as evidence for assessing the four voices on the seven bottom level criteria, summarised in a belief decision matrix as shown in Table 3. The numbers in the brackets are belief degrees associated with the evaluation grades on the common scale in the order of {NO, LOW, AVERAGE, HIGH, TOP}.

The "Unassigned Priority" (or Unknown) is not explicitly given in Table 3 and can be calculated by one minus the sum of the five numbers in an assessment. In the assessment of Voice B on "Criterion (2, 2)", for example, the total belief degree is 0.96, which means that 4% of the responses are missing from the original Survey 2 data. The missing information (ignorance) of 4% recorded in the belief decision matrix is preserved in the *ER* aggregation process.

Table 3A Belief Decision Matrix for Assessing Voices

Criterion	Voice			
	A	B	C	D
Criterion (1, 1)	{0.01, 0.01, 0.05, 0.25, 0.69}	{0, 0.01, 0.04, 0.29, 0.66}	{0, 0.03, 0.13, 0.35, 0.49}	{0.00, 0.01, 0.13, 0.46, 0.39}
Criterion (1, 2, 1)	{0, 1, 0, 0, 0}	{0, 0, 0, 0, 1}	{0, 0, 0, 0, 1}	{0, 0, 0, 0, 1}
Criterion (1, 2, 2)	{0, 1, 0, 0, 0}	{1, 0, 0, 0, 0}	{0, 0, 0, 0, 1}	{0, 0, 0, 0, 1}
Criterion (1, 3)	{0, 0, 0, 0, 1}	{0, 0, 0, 0, 1}	{0, 0, 0, 0, 1}	{0, 0, 0, 0, 1}
Criterion (1, 4)	{0.15, 0.06, 0.26, 0.06, 0.47}	{0.14, 0.06, 0.27, 0.06, 0.47}	{0.32, 0.07, 0.35, 0.07, 0.19}	{0.16, 0.07, 0.32, 0.07, 0.38}
Criterion (2, 1)	{0.91, 0, 0, 0, 0}	{0, 0, 0, 0, 0.98}	{0, 0, 0, 0, 0.98}	{0.97, 0, 0, 0, 0}
Criterion (2, 2)	{0.01, 0.05, 0.13, 0.21, 0.51}	{0.01, 0.04, 0.10, 0.17, 0.64}	{0.04, 0.07, 0.13, 0.18, 0.56}	{0.02, 0.05, 0.13, 0.21, 0.56}

B. Voice ranking and sensitivity analysis

In this section, we use the model as discussed in the previous sub-section to demonstrate how to apply the *ER* approach and the *IDS* software for aggregating assessments from the bottom level criteria of the hierarchy progressively to the top level voice assessment, ranking the four voices, and conducting sensitivity analysis. The following figures were all generated using the *IDS* software. In the case study, the original survey datasets were provided in Excel files. The data were transformed to the common priority scale using Matlab programmes. The generated data were then read into the *IDS* software through its data input interface.

For each voice, the *ER* approach is employed to first combine the assessments on the two bottom-level criteria "Criterion (1, 2, 1)" and "Criterion (1, 2, 2)", resulting in an aggregated assessment on the higher level criterion "Criterion (1, 2)". The assessments on Criterion (1, 1), Criterion (1, 2), Criterion (1, 3) and Criterion (1, 4) are then combined to generate an aggregated assessment on Survey 1. Similarly, the assessments on Criterion (2, 1) and Criterion (2, 2) are combined to generate an aggregated assessment on Survey 2.

The initially aggregated assessments on the two surveys are then further combined to generate an assessment for each voice on the overall criterion "Voice assessment" as the following overall distributed assessments:

$$S(A) = \{(H_1, 0.1164), (H_2, 0.1856), (H_3, 0.0655), (H_4, 0.0721), (H_5, 0.5444), (H, 0.0161)\}$$

$$S(B) = \{(H_1, 0.0865), (H_2, 0.0109), (H_3, 0.0483), (H_4, 0.0574), (H_5, 0.7938), (H, 0.0031)\}$$

$$S(C) = \{(H_1, 0.0442), (H_2, 0.0183), (H_3, 0.0731), (H_4, 0.0707), (H_5, 0.7913), (H, 0.0024)\}$$

$$S(D) = \{(H_1, 0.1210), (H_2, 0.0168), (H_3, 0.0820), (H_4, 0.1012), (H_5, 0.6737), (H, 0.0054)\}$$

If the five grades are quantified so that a change from a low grade is more appreciated than from a higher grade, then we will have a pseudo-concave utility function for the qualitative grades, for example as shown in Figure 2. Given the *fixed* pseudo-concave utility function as shown in Figure 2, the four voices are ranked as shown in Figure 3, or Voice $C \succ$ Voice $B \succ$ Voice $D \succ$ Voice A, where $\succ$ means "is preferred to".

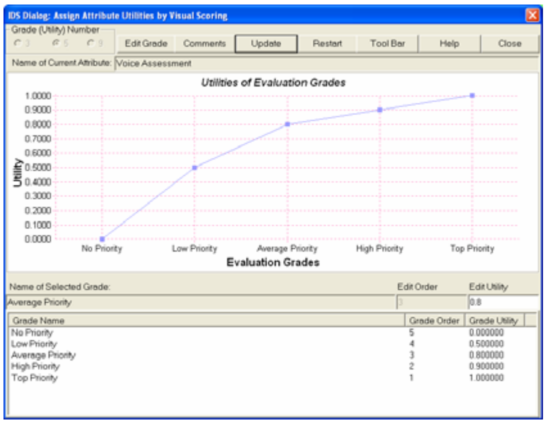

Figure 2 Pseudo-concave utility function

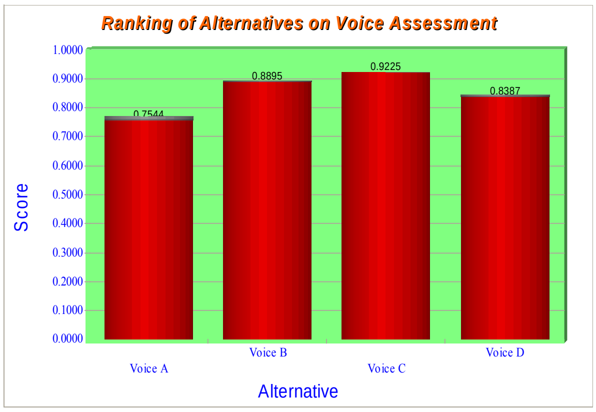

Figure 3 Pseudo-concave ranking

In the above ranking, we assumed that the precise utilities of all the grades had been estimated. Such an assumption is questionable in that the robustness of the ranking generated on the basis of the assumption needs to be examined. In general, as discussed in the previous section, the utility of a qualitative grade should be a variable interval but not fixed to a single value, unless the decision makers can provide sufficient and precise preference information to estimate a unique utility function for all assessment grades. In what follows, we employ the generic ranking method proposed in the previous section to generate the ranking of the four voices and investigate its robustness without making the above strict assumptions unnecessarily.

Take the pairwise comparison of Voice *C* and Voice *A* for example. From their overall assessments, the minimum mean utility of Voice *C* and the maximum mean utility of Voice *A* are given as follows:

$$u_{\min}(S(C)) = 0.0466u(H_1) + 0.0183u(H_2) + 0.0731u(H_3) + 0.0707u(H_4) + 0.7913u(H_5)$$

$$u_{\max}(S(A)) = 0.1164u(H_1) + 0.1856u(H_2) + 0.0655u(H_3) + 0.0721u(H_4) + 0.5605u(H_5)$$

The model for comparing Voice *C* with Voice *A* is then given by

$$\max \quad \sigma_{\mathbf{AC}} = u_{\max}(S(A)) - u_{\min}(S(C))$$
$$s.t. \quad \mathbf{X} = [(a_1, u(H_1), b_1), \cdots, (a_N, u(H_N), b_N)]^T \in \Omega_d$$

Set δ to be a sufficiently small positive number that makes differences between the grades meaningful, for example $\delta = 0.1 \times (1/(5-1)) = 0.025$. Then, the optimal value of the linear programming problem is given by $\sigma_{\mathbf{AC}}$=-0.0192, which means that it is evidentially true that Voice *C* $\succ$ Voice *A*.

Similarly, we can construct a linear programming model for comparing each pair of the four voices. The results are summarised in Table 4. We can then generate the evidential preference relations between certain pairs of voices as shown in Table 5, resulting in the following partial evidential ranking of the voices: Voice *C* ~ Voice *B* $\succ$ Voice *D* $\succ$ Voice *A,* where ~ means "is equivalent to".

Table 4Pairwise comparison indices of the voices

Voice A	Voice B	Voice C	Voice D
$\sigma_{\mathbf{AA}}$=0.000	$\sigma_{\mathbf{AB}}$=-0.017	$\sigma_{\mathbf{AC}}$=-0.019	$\sigma_{\mathbf{AD}}$=-0.001
$\sigma_{\mathbf{BA}}$=0.246	$\sigma_{\mathbf{BB}}$=0.000	$\sigma_{\mathbf{BC}}$=0.003	$\sigma_{\mathbf{BD}}$=0.118
$\sigma_{\mathbf{CA}}$=0.251	$\sigma_{\mathbf{CB}}$=0.043	$\sigma_{\mathbf{CC}}$=0.000	$\sigma_{\mathbf{CD}}$=0.117
$\sigma_{\mathbf{DA}}$=0.175	$\sigma_{\mathbf{DB}}$=-0.006	$\sigma_{\mathbf{DC}}$=-0.009	$\sigma_{\mathbf{DD}}$=0.000

Table 5Evidential partial ranking of the voices

Voice A	Voice B	Voice C	Voice D
—	B$\succ$A	C$\succ$A	D$\succ$A
—	—	—	—
—	—	—	—
—	B$\succ$D	C$\succ$D	—

To differentiate Voice *B* from Voice *C*, additional preference information need be provided. For instance, if the utility function of the assessment grades is assumed to be generically pseudo-concave, then the following linear programme can be constructed:

$$\max \quad \sigma_{\mathbf{BC}} = u_{\max}(S(B)) - u_{\min}(S(C))$$
$$s.t. \quad \mathbf{X} = \in \Omega_d \oplus (17)$$

The optimal value of the above problem is given by $\sigma_{\mathbf{BC}}$=-0.0019, which means that Voice *C* $\succ$ Voice *B* given the assumption of a *generic* pseudo-concave utility function. If $\sigma_{\mathbf{BC}}$ were still non-negative, then more strict preference information would be needed. If precise grade utilities are estimated, a complete ranking of the voices can always be generated. For this case study, however, such strict preference information was not needed.

VI. CONCLUSION

The *ER* approach provides another intersection process for compensatory evidence combination, complementary to Dempster's rule of non-compensatory combination. The *ER* algorithm is different from Dempster's rule in that it treats basic probability assignments as weighted belief degrees, includes the concept of the degree of indecisiveness, and adopts a normalisation process for combining probability masses. The proposed ranking method provides a generic process to support multiple criteria decision analysis.

ACKNOWLEDGMENT

This work was partially supported by the UK Engineering and Physical Science Research Council under the Grant No: EP/F024606/1 and the Natural Science Foundation of China under the Grant No: 60736026

REFERENCES

[1]. Beynon, M., 2002. DS/AHP method: A mathematical analysis, including an understanding of uncertainty. European Journal of Operational Research 140(1), 148-164
[2]. Dempster, A.P., 1967. Upper and lower probabilities induced by a multi-valued mapping. Annals of Mathematical Statistics 38, 325-339.
[3]. Denoeux, T., Zouhal, L.M., 2001. Handling possibilistic labels in pattern classification using evidential reasoning. Fuzzy Sets and Systems 122(3), 409-424
[4]. McClean, S., Scotney, B., 1997. Using evidence theory for knowledge discovery and extraction in distributed databases. International Journal of Intelligent Systems 12, 763-776.
[5]. Murphy, C.K., 2000. Combining belief functions when evidence conflicts. Decision Support Systems 29, 1-9.
[6]. Shafer, G., 1976. A Mathematical Theory of Evidence. Princeton University Press, Princeton.
[7]. Wang, Y. M. Yang, J. B. and Xu, D. L.. The evidential reasoning approach for multiple attribute decision analysis using interval belief degrees. European Journal of Operational Research 175(1), 35-66.
[8]. Xu, D.L., Yang, J.B., Wang, Y.M., 2006. The ER approach for multi-attribute decision analysis under interval uncertainties. European Journal of Operational Research 174(3), 1914-1943.
[9]. Yager, R.R., 2004. Decision making using minimization of regret. International Journal of Approximate Reasoning 36, 109-128.
[10]. Yang, J. B., Singh, M. G., 1994. An evidential reasoning approach for multiple attribute decision making with uncertainty. IEEE Transactions on Systems, Man, and Cybernetics 24(1), 1-18.
[11]. Yang, J. B., 2001. Rule and utility based evidential reasoning approach for multiattribute decision analysis under uncertainties. European Journal of Operational Research 131, 31-61.
[12]. Yang, J. B., Xu, D. L., 2002. On the evidential reasoning algorithm for multiattribute decision analysis under uncertainty", IEEE Transactions on Systems, Man, and Cybernetics Part A: Systems and Humans. 32(3), 289-304.
[13]. Yang, J.B., Wang, Y.M., Xu, D.L., 2006. The Evidential reasoning approach for MADA under both probabilistic and fuzzy uncertainties. European Journal of Operational Research 171(1), 309-343.