An extension of the DS/AHP method by several levels of criteria

Abstract

An extension of the DS/AHP method is proposed in the paper. The extension takes into account the fact that the multi-criteria decision problem might have a number of levels of criteria. Moreover, expert judgments concerning the criteria are imprecise and incomplete. The proposed extension also uses groups of experts or decision makers for comparing decision alternatives and criteria. It does not require assigning favorability values for groups of decision alternatives and criteria. The computation procedure is reduced to solving a finite set of linear programming problems. Numerical examples explain and illustrate the proposed modifications.

Keywords: multi-criteria decision problem, analytic hierarchy process, Dempster-Shafer theory, pairwise comparison, linear programming.

1 Introduction

One of the most well-established and frequently used method for solving a multi-criteria decision problem is the *analytic hierarchy process* (AHP) proposed by Saaty [1]. In the AHP, the decision maker (DM) models a problem as a hierarchy of criteria and decision alternatives (DA's). After the hierarchy is constructed, the DM assesses the importance of each element at each level of the hierarchy. This is accomplished by generating entries in a pairwise comparison matrix where elements are compared to each other. For each pairwise comparison matrix, the DM uses a method to compute a priority vector that gives the relative weights of the elements at each level of the hierarchy. Weights across various levels of the hierarchy are then aggregated using the principle of hierarchic composition to produce a final weight for each alternative.

The strength of AHP is that it organizes various factors in a systematic way and provides a structured simple solution to decision making problems. However, additional to the fact that the AHP method must perform very complicated and numerous pairwise comparisons amongst alternatives, and it is also difficult to obtain a convincing consistency index with an increasing number of attributes or alternatives. Moreover, the method uses precise estimates of experts or the DM. This condition can not be satisfied in many applications because judgments elicited from experts are usually imprecise and unreliable due to the limited precision of human assessments.

In order to overcome these difficulties and to extend the AHP on a more real elicitation procedures, Beynon *et al* [2, 3] proposed a method using *Dempster-Shafer theory* (DST) and called the *DS/AHP method*. The method was developed for decision making problems with a single DM, and it applies the AHP for collecting the preferences from a DM and for modelling the problem as a hierarchical decision tree. A nice analysis of the DS/AHP method is given by Tervonen *et al* [4]. It should be noted that the main excellent idea underlying the DS/AHP method is not applying DST to the AHP. It is the comparison of groups of alternatives with a whole set of alternatives. The such type comparison is equivalent to the preference stated by the DM.

The DS/AHP method has many advantages. However, it does not allow us to apply the procedure of incomplete elicitation used for assessment of DA's on levels of criteria. At the same time, it is also difficult to obtain precise weights of criteria in many applications. The problem becomes more complicated when the number of levels of criteria is two or larger.

Therefore, we propose an extension of the DS/AHP method which generalize the method and partially overcome the above difficulties. The extension takes into account the fact that the multi-criteria decision problem might have a number of levels of criteria. Moreover, it is assumed that expert judgments concerning the criteria are imprecise and incomplete. The proposed extension also uses groups of experts or decision makers for comparing decision alternatives and criteria. It employs the fact that the belief and plausibility measures in the framework of DST can be regarded as lower and upper bounds for the probability of an event. A computation procedure realizing the extension is reduced to a number of rather simple linear programming problems.

2 Dempster-Shafer theory

Let U be a universal set under interest, usually referred to in evidence theory as the *frame of discernment*. Suppose N observations were made of an element $u \in U$, each of which resulted in an imprecise (non-specific) measurement given by a set A of values. Let c_i denote the number of occurrences of the set $A_i \subseteq U$, and $\mathcal{P}o(U)$ the set of all subsets of U (power set of U). A frequency function m, called *basic probability assignment* (BPA), can be defined such that [5, 6]:

$$m : \mathcal{P}o(U) \to [0,1],\ m(\varnothing) = 1,\ \sum_{A \in \mathcal{P}o(U)} m(A) = 1.$$

Note that the domain of BPA, $\mathcal{P}o(U)$, is different from the domain of a probability density function, which is U. According to [5], this function can be obtained as follows:

$$m(A_i) = c_i/N. \tag{1}$$

If $m(A_i) > 0$, i.e. A_i has occurred at least once, then A_i is called a *focal element*.

According to [6], the *belief* $Bel(A)$ and *plausibility* $\mathrm{Pl}(A)$ measures of an event $A \subseteq \Omega$ can be defined as

$$\mathrm{Bel}(A) = \sum_{A_i : A_i \subseteq A} m(A_i),\quad \mathrm{Pl}(A) = \sum_{A_i : A_i \cap A \neq \varnothing} m(A_i). \tag{2}$$

As pointed out in [7], a belief function can formally be defined as a function satisfying axioms which can be viewed as a weakening of the Kolmogorov axioms that characterize probability functions. Therefore, it seems reasonable to understand a belief function as a generalized probability function [5] and the belief $\mathrm{Bel}(A)$ and plausibility $\mathrm{Pl}(A)$ measures can be regarded as lower and upper bounds for the probability of A, i.e., $\mathrm{Bel}(A) \leq \Pr(A) \leq \mathrm{Pl}(A)$.

3 The DS/AHP method

Suppose that there is a set of DA's $\mathbb{A} = \{A_1, ..., A_n\}$ consisting of n elements. Moreover, there is a set of criteria $\mathbb{C} = \{C_1, ..., C_r\}$ consisting of r elements. In the DS/AHP method, the DM chooses some subsets $B_k \in \mathcal{P}o(\mathbb{A})$ of DA's from the power set $\mathcal{P}o(\mathbb{A})$ in accordance with the certain criterion C_j from $\mathbb{C}$. The nice idea in [2, 3] is that instead of comparing DA's between each other, the DM has to identify favorable DA's from the set $\mathbb{A}$. This choice can be regarded as the comparison between the group or subset B_k of DA's and the whole set of DA's $\mathbb{A}$. In other words, in DS/AHP, all pairwise comparisons are made against the set $\mathbb{A}$. This is a very interesting and subtle approach. We should mention that only single DA's are compared in the AHP.

After the decision tree is set up, the weights of criteria have to be defined. They are obtained using the standard pairwise comparison method as in AHP. The DM also has to make pairwise comparisons between groups of DA's and the set $\mathbb{A}$, from which the so-called knowledge matrix [3] (a reduced matrix of pairwise comparisons with respect to each criterion) is formed for each criterion. After the comparisons are made, the knowledge matrices are multiplied in a specific way by the weights for criteria. Then priority values are obtained for groups of DA's and $\mathbb{A}$ using the eigenvector method. After the priority values have been obtained, they are combined using Dempster's rule of combination.

We illustrate the DS/AHP method by the following numerical example.

Example 1 *Let us study a decision problem where the DM has to choose which one of three types of transport to use. Three DA's (rail transport (A_1), motor transport (A_2), water transport (A_3)) are evaluated based on two criteria: reliability of delivery (C_1) and freight charge (C_2). The knowledge matrix for criterion C_1 is shown in Table 1. According to [3], a 6-point scale (1-6) is used for the pairwise comparisons instead of a 9-point scale (1-9) as in AHP. It can be seen from Table 1 that DA's A_2, A_3 are viewed as extremely favorable compared to the set $\mathbb{A} = \{A_1, A_2, A_3\}$. The zero's which appear in the knowledge matrix indicate no attempt to assert knowledge between groups of DA's, for instance, $\{A_1\}$ to $\{A_2, A_3\}$. This assertion can be made indirectly through knowledge of the favorability of A_1 to $\mathbb{A}$ and $\{A_2, A_3\}$ to $\mathbb{A}$ relatively. In Table 1, the indirect knowledge is that A_1 is not considered more favorable to $\{A_2, A_3\}$ in relation to $\mathbb{A}$. The knowledge matrix for criterion C_2 is shown in Table 2.*

The following rule for processing the knowledge matrices is proposed in [3]. If p is the weight for a criterion and x_{ij} is the favorability opinion for a particular group of DA's with respect to this criterion, then the resultant value is $p \cdot x_{ij}$ (the resultant change in the bottom row of the matrix is similarly $1/(p \cdot x_{ij})$). For instance, the knowledge matrix for freight charge

Table 1: The knowledge matrix for reliability of delivery

	$\{A_1\}$	$\{A_2, A_3\}$	$\{A_1, A_2, A_3\}$
$\{A_1\}$	1	0	4
$\{A_2, A_3\}$	0	1	6
$\{A_1, A_2, A_3\}$	1/4	1/6	1

Table 2: The knowledge matrix for freight charge

	$\{A_2\}$	$\{A_1, A_2, A_3\}$
$\{A_2\}$	1	1/2
$\{A_1, A_2, A_3\}$	2	1

can be rewritten by taking into account that the weight for C_2 *is* 0.4 *as shown in Table 3.*

Using the knowledge matrices for each of the criteria normalized knowledge vectors can be produced, following the traditional AHP method. The elements of the vectors can be regarded as the BPA's of groups of DA's. As a result, we get

$$m_1(\{A_1\} \mid \{C_1\}) = 0.398,\ m_1(\{A_2, A_3\} \mid \{C_1\}) = 0.457,\ m_1(\mathbb{A} \mid \{C_2\}) = 0.145,$$

$$m_2(\{A_2\} \mid \{C_2\}) = 0.56,\ m_2(\mathbb{A} \mid \{C_2\}) = 0.44.$$

By considering the criteria as independent pieces of evidence, these pieces of evidence can be combined by using Dempster's rule of combination. For brevity, we will not present the final results here. The interested reader should refer to [3].

4 Incomplete information about criteria

The DS/AHP method is a powerful tool for solving multi-criteria decision problems. However, it has some disadvantages mentioned in the introductory section. First of all, it is difficult to assign a numerical value of the favorability opinion for a particular group of DA's. The second is that the standard procedure of the pairwise comparisons remains for criteria. Therefore, we propose to extend the DS/AHP method and to identify favorable criteria or groups of criteria from the set $\mathbb{C}$. Moreover, we propose to use only estimates like "preferable" or "not" by choosing the corresponding groups of DA's or criteria. We also suppose that there are many experts or DM's for evaluating DA's and criteria, and every expert judgment adds "1" to the corresponding preference.

Table 3: Updated knowledge matrix for freight charge

	$\{A_2\}$	$\{A_1, A_2, A_3\}$
$\{A_2\}$	1	1.25
$\{A_1, A_2, A_3\}$	0.8	1

Table 4: Expert preferences related to criteria

	$\{C_1\}$	$\{C_2\}$	$\{C_1C_2\}$
	D_1	D_2	D_3
c_k	6	4	5
$m(D_k)$	6/15	4/15	5/15

We again suppose that there is a set of DA's $\mathbb{A} = \{A_1, ..., A_n\}$ consisting of n elements and a set of criteria $\mathbb{C} = \{C_1, ..., C_r\}$ consisting of r elements. Experts choose some subsets $B_k \in \mathcal{P}o(\mathbb{A})$ of DA's from the power set $\mathcal{P}o(\mathbb{A})$ in accordance with the certain criterion C_j from $\mathbb{C}$. Moreover, they choose some subsets $D_i \in \mathcal{P}o(\mathbb{C})$ from the power set $\mathcal{P}o(\mathbb{C})$ as favorable groups of criteria.

In accordance with the introduced notation, expert's judgments can be represented in the form of preferences $B_k \succeq \mathbb{A}$, i.e., an expert chooses the subset B_k from the set of all DA's as the most preferable group of DA's. The preference $\mathbb{A} \succeq \mathbb{A}$ means that an expert meets difficulties in choosing some preferable subset of $\mathcal{P}o(\mathbb{A})$.

The expert elicitation and an assessment processing procedure can be represented by means of the two-step scheme.

At the *first step*, every expert picks out the most important or preferable group of criteria. If the number of experts, participating at this step, is N_C, then we can compute the BPA's $m(D_i) = c_i/N_C$ of all focal elements $D_i \subseteq \mathbb{C}$ (see Table 4), where $N_C = \sum_{i=1}^{2^r-1} c_i^{(k)}$.

At the *second step*, every expert chooses a subset $B_i \subseteq \mathbb{A}$ of DA's as the most preferable DA's from the set $\mathbb{A}$ with respect to the predefined criterion C_j. After all experts choose the subsets of DA's with respect to the j-th criterion, we have the set of integers $a_1^{(j)}, a_2^{(j)}, ..., a_l^{(j)}$ corresponding to the numbers of experts chosen subsets $B_1, ..., B_l$, respectively. This procedure is repeating r times for all $j = 1, ..., r$, i.e., for all criteria from the set $\mathbb{C}$. If we denote the total number of assessments related to DA's with respect to the j-th criterion $N_A^{(j)}$, then the conditional BPA of every subset B_i is computed as $m(B_i \mid C_j) = a_i^{(j)}/N_A^{(j)}$, $N_A^{(j)} = \sum_{i=1}^{2^n-1} a_i^{(j)}$ (see Table 5).

Example 2 *Let us return to Example 1. 15 experts provide preferences concerning criteria (see Table 4) and preferences concerning the DA's with respect to criteria C_1 and C_2 (see Table 5). The correspondences between subsets of criteria (DA's) and short notations D_k (B_k) are also represented in Tables 4 and 5.*

The next problem is to combine the above information for obtaining the weights of DA's. The main difficulty here is that in addition to judgments concerning single criteria, we have possible judgments concerning the groups of criteria. These judgments also have to be taken into account. The following approach can be proposed here for solving this problem.

On one hand, by having BPA's $m(D_k)$ of subsets $D_k \subseteq \mathbb{C}$, the belief and plausibility

Table 5: Expert preferences related to DA's

	$\{A_1\}$	$\{A_2\}$	$\{A_3\}$	$\{A_1A_2\}$	$\{A_1A_3\}$	$\{A_2A_3\}$	$\{A_1A_2A_3\}$
	B_1	B_2	B_3	B_4	B_5	B_6	B_7
$a_i^{(1)}$	5	2	3	4	0	0	1
$a_i^{(2)}$	3	1	2	3	3	1	2
$m(B_i\|C_1)$	5/15	2/15	3/15	4/15	0	0	1/15
$m(B_i\|C_2)$	3/15	1/15	2/15	3/15	3/15	1/15	2/15

functions of D_k can be computed as

$$\mathrm{Bel}(D_k) = \sum_{i:D_i\subseteq D_k} m\left(D_i\right),$$

$$\mathrm{Pl}(D_k) = \sum_{i:D_i\cap D_k\neq\varnothing} m\left(D_i\right),\ k=1,...,2^r-1.$$

On the other hand, suppose that the j-th criteria is chosen by experts with some unknown probability p_j such that the condition $\sum_{j=1}^r p_j = 1$ is valid. Then the probabilities of criteria satisfy the following system of inequalities

$$\mathrm{Bel}(D_k) \leq \sum_{j:C_j\in D_k} p_j \leq \mathrm{Pl}(D_k),\ k=1,...,2^r-1. \tag{3}$$

By viewing the belief and plausibility functions as lower and upper probabilities, respectively, we can say that the set of inequalities (3) produces a set $\mathcal{P}$ of possible distributions $p=(p_1,...,p_r)$ satisfying all these inequalities. Let us fix a distribution p from $\mathcal{P}$. Then, by applying the total probability theorem, we can write the combined BPA of the subset B_k as follows:

$$m_p(B_k) = \sum_{j=1}^r m(B_k \mid C_j)\cdot p_j,\ p\in\mathcal{P}.$$

It should be noted that the obtained BPA depends on the probability distribution $p\in\mathcal{P}$. Therefore, the belief and plausibility functions of B_k also depend on the fixed probability distribution $p\in\mathcal{P}$ and are

$$\mathrm{Bel}_p(B_k) = \sum_{i:B_i\subseteq B_k} m_p(B_i) = \sum_{j=1}^r p_j\cdot\left(\sum_{i:B_i\subseteq B_k} m(B_i \mid C_j)\right),$$

$$\mathrm{Pl}_p(B_k) = \sum_{i:B_i\cap B_k\neq\varnothing} m_p(B_i) = \sum_{j=1}^r p_j\cdot\left(\sum_{i:B_i\cap B_k\neq\varnothing} m(B_i \mid C_j)\right).$$

The obtained belief and plausibility functions linearly depend on p. Consequently, we can find the lower belief and upper plausibility functions by solving the following linear programming problems:

$$\mathrm{Bel}(B_k) = \inf_{p\in\mathcal{P}} \mathrm{Bel}_p(B_k) = \inf_{p\in\mathcal{P}} \sum_{j=1}^{r} p_j \cdot \left(\sum_{i:B_i\subseteq B_k} m(B_i \mid C_j) \right),$$

$$\mathrm{Pl}(B_k) = \sup_{p\in\mathcal{P}} \mathrm{Pl}_p(B_k) = \sup_{p\in\mathcal{P}} \sum_{j=1}^{r} p_j \cdot \left(\sum_{i:B_i\cap B_k\neq\varnothing} m(B_i \mid C_j) \right)$$

subject to $\sum_{j=1}^{r} p_j = 1$ and (3).

When we do not have information about criteria at all, then the set of constraints to the above linear programming problems are reduced to one constraint $\sum_{j=1}^{r} p_j = 1$. Note that the optimal solutions to the linear programming problem can be found at one of the extreme points of the convex sets $\mathcal{P}$ of distributions produced by the linear constraints. Since we remain only one constraint $\sum_{j=1}^{r} p_j = 1$ which forms the unit simplex, then its extreme points have the form

$$(1, 0, ..., 0),\ (0, 1, ..., 0), ..., (0, 0, ..., 1).$$

Hence, it is obvious that the optimal belief and plausibility functions of the DA B_k can be computed as follows:

$$\mathrm{Bel}(B_k) = \min_{j=1,\dots,r} \sum_{i:B_i\subseteq B_k} m(B_i \mid C_j), \tag{4}$$

$$\mathrm{Pl}(B_k) = \max_{j=1,\dots,r} \sum_{i:B_i\cap B_k\neq\varnothing} m(B_i \mid C_j). \tag{5}$$

It is interesting to note that the belief function of the optimal DA in the case of prior ignorance about criteria is computed by using the "maximin" technique, i.e., we first compute the minimal "combined" belief function of every DA over all criteria in accordance with (4). Then we compute the maximal belief function among the obtained "combined" belief functions. The plausibility function of the optimal DA is computed by using the "maximax" technique in accordance with (5).

By having the belief and plausibility functions of all subsets B_k, $k = 1, ..., 2^n - 1$, we can determine the "best" DA. The choice of the "best" DA is based on comparison of intervals produced by the belief and plausibility functions. There exist a lot of methods for comparison. We propose to use the most justified method based on the so-called caution parameter [8, 9] or the parameter of pessimism $\eta \in [0, 1]$ which has the same meaning as the optimism parameter in Hurwicz criterion [10]. According to this method, the "best" DA from all possible ones should be chosen in such a way that makes the convex combination $\eta \cdot Bel(B) + (1 - \eta)Pl(B)$ achieve its maximum. If $\eta = 1$, then we analyze only belief functions and make pessimistic decision. This type of decision is very often used [11, 12]. If $\eta = 0$, then we analyze only plausibility functions and make optimistic decision.

Example 3 *Let us return to Example 2 and find the belief and plausibility functions of subsets* D_1, D_2, D_3:

$$\mathrm{Bel}(D_1) = m(D_1) = 6/15,\ \mathrm{Pl}(D_1) = m(D_1) + m(D_3) = 11/15,$$

$$\mathrm{Bel}(D_2) = m(D_2) = 4/15,\ \mathrm{Pl}(D_2) = m(D_2) + m(D_3) = 9/15,$$

$$\mathrm{Bel}(D_3) = \mathrm{Pl}(D_3) = 1.$$

Let us compute the belief and plausibility functions of DA's A_1, A_2, A_3. *The linear programming problem for computing the belief function of the first DA* A_1 *is of the form:*

$$\begin{aligned}\mathrm{Bel}(A_1) &= \inf_{p\in\mathcal{P}} \left(p_1 \cdot m(A_1|C_1) + p_2 \cdot m(A_1|C_2)\right) \\ &= \inf_{p\in\mathcal{P}} \left(p_1 \cdot 5/15 + p_2 \cdot 3/15\right)\end{aligned}$$

subject to $p_1 + p_2 = 1$ *and*

$$6/15 \leq p_1 \leq 11/15,\ 4/15 \leq p_2 \leq 9/15.$$

The optimal solution is $p_1 = 2/5$, $p_2 = 3/5$. *Hence* $\mathrm{Bel}(A_1) = 0.253$. *The linear programming problem for computing the plausibility function of* A_1 *has the same constraints and the objective function*

$$\begin{aligned}\mathrm{Pl}(A_1) &= \sup_{p\in\mathcal{P}} \left(\sum_{i=1}^{2} p_i \cdot \left(m(B_1|C_i) + m(B_4|C_i) + m(B_5|C_i) + m(B_7|C_i)\right)\right) \\ &= \sup_{p\in\mathcal{P}} \left(p_1 \cdot 10/15 + p_2 \cdot 11/15\right).\end{aligned}$$

The optimal solution is $p_1 = 2/5$, $p_2 = 3/5$. *Hence* $\mathrm{Pl}(A_1) = 0.707$. *The belief and plausibility function of other DA's can be computed in the same way:* $\mathrm{Bel}(A_2) = 0.093$, $\mathrm{Pl}(A_2) = 0.467$, $\mathrm{Bel}(A_3) = 0.16$, $\mathrm{Pl}(A_3) = 0.427$. *It can be seen from the results that the first DA is optimal by arbitrary values of* η *due to the inequalities* $\mathrm{Bel}(A_1) \geq \mathrm{Bel}(A_3) \geq \mathrm{Bel}(A_2)$ *and* $\mathrm{Pl}(A_1) \geq \mathrm{Pl}(A_2) \geq \mathrm{Pl}(A_3)$.

If we would not have information about importance of criteria, then

$$\mathrm{Bel}(A_1) = 3/15,\ \mathrm{Pl}(A_1) = 11/15,$$

$$\mathrm{Bel}(A_2) = 1/15,\ \mathrm{Pl}(A_2) = 7/15,$$

$$\mathrm{Bel}(A_3) = 2/15,\ \mathrm{Pl}(A_3) = 8/15.$$

5 Two levels of criteria

Let us consider a case when there are two levels of criteria. The first (highest) level contains t criteria from the set $\mathbb{C} = \{C_1, ..., C_t\}$. Every criterion of the first level has the number k_1, where $k_1 = 1, ..., t$. For the criterion of the first level with the number k_1, there are r criteria from the set $\mathbb{C}_2(k_1) = \{C_1(k_1), ..., C_r(k_1)\}$ on the second level[1]. Every criterion of the second level has the number (k_1, k_2). For example, the third criterion of the second level with respect to the second criterion of the first level has the number $(2, 3)$. Experts choose some subsets $D_i \subseteq \mathbb{C}$ from the power set $\mathcal{P}o(\mathbb{C})$ as favorable groups of criteria on the first level. Experts also choose some subsets $D_k(k_1) \subseteq \mathbb{C}_2(k_1)$ from the power set $\mathcal{P}o(\mathbb{C}_2(k_1))$ as favorable groups of criteria on the second level with respect to the criterion of the first level with the number k_1.

Suppose that the k-th criterion on the first level is chosen by experts with some unknown probability q_k such that the condition $\sum_{k=1}^{t} q_k = 1$ is valid. Then the probabilities of the criteria satisfy the following system of inequalities

$$\mathrm{Bel}(D_k) \leq \sum_{j:C_j \in D_k} q_j \leq \mathrm{Pl}(D_k),\ k = 1, ..., 2^t - 1. \tag{6}$$

Let $\mathcal{Q}$ be the set of probability distributions produced by all constraints (6).

Suppose that the j-th criterion on the second level with respect to the k-th criterion of the first level is chosen by experts with some unknown probability $q_j(k)$ such that the condition $\sum_{j=1}^{r} q_j(k) = 1$ for every $k = 1, ..., t$ is valid. Then the probabilities of the criteria satisfy the following system of inequalities

$$\mathrm{Bel}(D_l(k)) \leq \sum_{j:C_j(k) \in D_l(k)} q_j(k) \leq \mathrm{Pl}(D_l(k)),\ l = 1, ..., 2^r - 1,\ k = 1, ..., t. \tag{7}$$

Let $\mathcal{Q}(k)$ be the set of probability distributions produced by all constraints (7) by a fixed value of k.

Denote

$$a_{jl}(k) = \sum_{i:B_i \subseteq B_l} m(B_i \mid C_j(k)),\ b_{jl}(k) = \sum_{i:B_i \cap B_l \neq \varnothing} m(B_i \mid C_j(k)).$$

Here the index j corresponds to the j-th criterion of the second level chosen with respect to the k-th criterion of the first level. The index l means the number of subset B_l chosen for computing its belief and plausibility functions.

Let us fix the probability distributions $q = (q_1, ..., q_t)$ and $q(k) = (q_1(k), ..., q_r(k))$, $k = 1, ..., t$. Now we can write the conditional belief $\mathrm{Bel}_{q,q(k)}(B_l)$ and plausibility $\mathrm{Pl}_{q,q(k)}(B_l)$ functions of B_l under conditions of fixed distributions q and $q(k)$, $k = 1, ..., t$,

$$\mathrm{Bel}_{q,q(k)}(B_l) = \sum_{i:B_i \subseteq B_l} m_{q,q(k)}(B_i) = \sum_{k=1}^{t} q_k \sum_{j=1}^{r} q_j(k) \cdot a_{jl}(k), \tag{8}$$

[1] We assume that the sets of criteria on the second level corresponding to every criterion of the first level are identical, i.e., $\mathbb{C}_2(i) = \mathbb{C}_2(k) = \mathbb{C}_2$ for $i \neq k$.

$$\mathrm{Pl}_{q,q(k)}(B_l) = \sum_{i:B_i \cap B_l \neq \varnothing} m_{q,q(k)}(B_i) = \sum_{k=1}^{t} q_k \sum_{j=1}^{r} q_j(k) \cdot b_{jl}(k). \tag{9}$$

By minimizing the belief function and by maximizing the plausibility function over all distributions $q \in \mathcal{Q}$ and $q(k) \in \mathcal{Q}(k)$, $k = 1, ..., t$, we can get the unconditional lower belief and upper plausibility functions of B_l. This can be carried out by solving the optimization problems

$$\mathrm{Bel}(B_l) = \min_{q,q(k)} \mathrm{Bel}_{q,q(k)}(B_l), \tag{10}$$

$$\mathrm{Pl}(B_l) = \max_{q,q(k)} \mathrm{Pl}_{q,q(k)}(B_l) \tag{11}$$

subject to (6) and (7).

At first sight, this is typical quadratic programming problems having linear constraints and nonlinear objective functions. However, we can show that every optimization problem from the above can be solved by considering a set of $t + 1$ linear programming problems.

Denote

$$\mathbb{E}_{q(k)} a_l(k) = \sum_{j=1}^{r} q_j(k) \cdot a_{jl}(k)$$

Note that the multiplier $\mathbb{E}_{q(k)} a_l(k)$ in (8) depends only on the probability distributions from the set $\mathcal{Q}(k)$ and does not depend on the distributions from $\mathcal{Q}$ and $\mathcal{Q}(i)$, $i \neq k$. The same can be said about all the multipliers of the above form. This implies that under condition $q_k \geq 0$, $k = 1, ..., t$, there hold

$$\min_{q,q(k)} \mathrm{Bel}_{q,q(k)}(B_l) = \min_{q \in \mathcal{Q}} \sum_{k=1}^{t} q_k \left(\min_{q(k) \in \mathcal{Q}(k)} \mathbb{E}_{q(k)} a_l(k) \right),$$

$$\max_{q,q(k)} \mathrm{Pl}_{q,q(k)}(B_l) = \max_{q \in \mathcal{Q}} \sum_{k=1}^{t} q_k \left(\max_{q(k) \in \mathcal{Q}(k)} \mathbb{E}_{q(k)} a_l(k) \right).$$

Hence, for computing the belief function, we get the set of t simple linear programming problems

$$\underline{\mathbb{E}} a_l(k) = \min_{q(k)} \mathbb{E}_{q(k)} a_l(k)$$

under constraints (7) or $q(k) \in \mathcal{Q}(k)$ and the linear programming problem

$$\mathrm{Bel}(B_l) = \min_{q} \sum_{k=1}^{t} q_k \cdot \underline{\mathbb{E}} a_l(k) \tag{12}$$

under constraints (6) or $q \in \mathcal{Q}$.

The same can be said about computing the plausibility function, i.e.,

$$\mathrm{Pl}(B_l) = \max_{q} \sum_{k=1}^{t} q_k \cdot \overline{\mathbb{E}} a_l(k) \tag{13}$$

under constraints (6) or $q \in \mathcal{Q}$, where $\overline{\mathbb{E}}a_l(k)$, $k = 1, ..., t$, is obtained by solving t simple linear programming problems

$$\overline{\mathbb{E}}a_l(k) = \max_{q(k)} \mathbb{E}_{q(k)} a_l(k)$$

under constraints (7) or $q(k) \in \mathcal{Q}(k)$.

Example 4 *Let us return to Example 1 and suppose that there are two transport firms. Every firm offers the freight services, but have different levels of the delivery reliability and freight charge. Two experts prefer the first firm and three experts prefer both the firms. Hence the BPA's of the subsets D_1, D_2, D_3 are 0.4, 0, 0.6, respectively. The preferences of experts on the second level of criteria with respect to the first and second firms are shown in Table 6 and in Table 7, respectively. These tables also contain the BPA's $m(D_l(k))$ of all subsets of $\mathbb{C}_2(k)$. The expert judgments about DA's with respect to the first and second criteria of the second level are given in Table 5. Here we assume that the weights of DA's of identical criteria of the second level are identical, i.e., experts do not recognize or do not "see" the first level of criteria and estimate DA's with respect to the set $\mathbb{C}_2(k)$. This implies that $m(B_i \mid C_l(k)) = m(B_i \mid C_l(j))$ for all possible i, l, k, j.*

First of all, we find the values of $\underline{\mathbb{E}}a_l(k)$ and $\overline{\mathbb{E}}b_l(k)$ for $k = 1, 2$. For instance, there holds for $l = 1$ ($B_1 = \{A_1\}$), $k = 1$,

$$\begin{aligned}\mathbb{E}a_1(1) &= q_1(1) \cdot a_{11}(1) + q_2(1) \cdot a_{21}(1) \\ &= q_1(1) \cdot m(B_1 \mid C_1(1)) + q_2(1) \cdot m(B_1 \mid C_2(1)) \\ &= q_1(1) \cdot 5/15 + q_2(1) \cdot 3/15.\end{aligned}$$

$$\begin{aligned}\mathbb{E}b_1(1) &= q_1(1) \cdot b_{11}(1) + q_2(1) \cdot b_{21}(1) \\ &= q_1(1) \cdot (m(B_1 \mid C_1(1)) + m(B_4 \mid C_1(1)) + m(B_5 \mid C_1(1)) + m(B_7 \mid C_1(1))) \\ &\quad + q_2(1) \cdot (m(B_1 \mid C_2(1)) + m(B_4 \mid C_2(1)) + m(B_5 \mid C_2(1)) + m(B_7 \mid C_2(1))) \\ &= q_1(1) \cdot 10/15 + q_2(1) \cdot 11/15.\end{aligned}$$

Constraints are of the form (see (7)):

$$\begin{aligned}0.2 &\le q_1(1) \le 0.6, \\ 0.4 &\le q_2(1) \le 0.6, \\ 1 &= q_1(1) + q_2(1).\end{aligned}$$

By solving the linear programming problems with the above constraints, we get

$$\underline{\mathbb{E}}a_1(1) = 0.4 \cdot 5/15 + 0.6 \cdot 3/15 = 0.253,$$

$$\overline{\mathbb{E}}b_1(1) = 0.6 \cdot 10/15 + 0.4 \cdot 11/15 = 0.693.$$

In the same way, we can find all values of $\underline{\mathbb{E}}a_l(k)$ and $\overline{\mathbb{E}}b_l(k)$, which are represented in Table 8.

Table 6: Expert preferences related to criteria on the second level

	$D_1(1)$	$D_2(1)$	$D_3(1)$
c_l	2	4	4
$m(D_l(k))$	0.2	0.4	0.4

Table 7: Expert preferences related to criteria on the second level

	$D_1(2)$	$D_2(2)$	$D_3(2)$
c_l	3	2	5
$m(D_l(k))$	0.3	0.2	0.5

Now we can compute the lower unconditional belief function $\mathrm{Bel}(B_l)$ *from (12) by solving the linear programming problem with objective function* $q_1 \cdot \underline{\mathbb{E}}a_l(1) + q_2 \cdot \underline{\mathbb{E}}a_l(2)$ *and constraints (6):*

$$0.4 \leq q_1 \leq 1,$$
$$0 \leq q_2 \leq 0.6,$$
$$1 = q_1 + q_2.$$

In the same way, we can find the upper unconditional plausibility function $\mathrm{Pl}(B_l)$ *from (13) by solving the linear programming problem with objective function* $q_1 \cdot \overline{\mathbb{E}}b_l(1) + q_2 \cdot \overline{\mathbb{E}}b_l(2)$ *and the same constraints.*

The corresponding computation results are shown in Table 9. It can be seen from the results that the first DA is "optimal".

References

[1] T.L. Saaty. *Multicriteria Decision Making: The Analytic Hierarchy Process.* McGraw Hill, New York, 1980.

[2] M. Beynon. DS/AHP method: A mathematical analysis, including an understanding of uncertainty. *European Journal of Operational Research*, 140:148–164, 2002.

[3] M. Beynon, B. Curry, and P. Morgan. The Dempster-Shafer theory of evidence: An alternative approach to multicriteria decision modelling. *Omega*, 28:37–50, 2000.

Table 8: Intermediate results of computing the belief and plausibility functions

	B_1	B_2	B_3	B_4	B_5	B_6	B_7
$\underline{\mathbb{E}}a_l(1)$	0.253	0.093	0.16	0.573	0.533	0.293	1
$\overline{\mathbb{E}}b_l(1)$	0.706	0.467	0.427	0.84	0.907	0.747	1
$\underline{\mathbb{E}}a_l(2)$	0.24	0.087	0.153	0.547	0.533	0.287	1
$\overline{\mathbb{E}}b_l(2)$	0.713	0.467	0.453	0.847	0.913	0.76	1

Table 9: Unconditional belief and plausibility functions

	B_1	B_2	B_3	B_4	B_5	B_6	B_7
$\text{Bel}(B_l)$	0.245	0.089	0.156	0.557	0.533	0.289	1
$\text{Pl}(B_l)$	0.71	0.467	0.443	0.844	0.91	0.755	1

[4] T. Tervonen, R. Lahdelma, and P. Salminen. A method for elicitating and combining group preferences for stochastic multicriteria acceptability analysis. TUCS Technical Report 638, Turku Centre for Computer Science, Turku, Finland, November 2004.

[5] A.P. Dempster. Upper and lower probabilities induced by a multi-valued mapping. *Annales of Mathematical Statistics*, 38:325–339, 1967.

[6] G. Shafer. *A Mathematical Theory of Evidence*. Princeton University Press, 1976.

[7] J.Y. Halpern and R. Fagin. Two views of belief: Belief as generalized probability and belief as evidence. *Artificial Intelligence*, 54:275–317, 1992.

[8] J. Schubert. On ρ in a decision-theoretic apparatus of Dempster-Shafer theory. *International Journal of Approximate Reasoning*, 13:185–200, 1995.

[9] K. Weichselberger. *Elementare Grundbegriffe einer allgemeineren Wahrscheinlichkeitsrechnung*, volume I Intervallwahrscheinlichkeit als umfassendes Konzept. Physika, Heidelberg, 2001.

[10] R.D. Luce and H. Raiffa. *Games and decisions*. Wiley, New York, 1957.

[11] T. Augustin. Expected utility within a generalized concept of probability - a comprehensive framework for decision making under ambiguity. *Statistical Papers*, 43:5–22, 2002.

[12] K.-M. Osei-Bryson. Supporting knowledge elicitation and consensus building for Dempster-Shafer decision models. *International Journal of Intelligent Systems*, 18:129–148, 2003.

[13] P. Walley. *Statistical Reasoning with Imprecise Probabilities*. Chapman and Hall, London, 1991.