THE DS/AHP METHOD UNDER PARTIAL INFORMATION ABOUT CRITERIA AND ALTERNATIVES BY SEVERAL LEVELS OF CRITERIA

LEV V. UTKIN, NATALIA V. SIMANOVA

Abstract

An extension of the DS/AHP method is proposed in the paper. It takes into account the fact that the multi-criteria decision problem might have several 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. However, it does not require assigning favorability values for groups of decision alternatives and criteria. The computation procedure for processing and aggregating the incomplete information about criteria and decision alternatives is reduced to solving a finite set of linear programming problems. Numerical examples explain in detail and illustrate the proposed approach.

1 Introduction

One of the most well-established and frequently used methods for solving a multi-criteria decision problem is the *analytic hierarchy process* (AHP) proposed by Saaty[21]. 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 the 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, 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, 4] proposed a method using *Dempster-Shafer theory* (DST) and called by the *DS/AHP method*. The method was developed for decision making problems with a single DM, it exploited 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*[26]. It should be noted that the main excellent idea underlying the DS/AHP method is the comparison of groups of alternatives with a whole set of alternatives. This type of 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. 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.

Taking into account the above, we propose an extension of the DS/AHP method, which generalizes it and partially allows us to 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 assumes 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. It is shown in the paper that a computation procedure realizing the extension is reduced to a number of rather simple linear programming problems.

It should be noted that a lot of methods and approaches have been developed and proposed to solve multi-criteria decision making problems under imprecise and incomplete information about criteria and (or) decision alternatives and to model the preference information which is usually imprecise and incomplete in real situations. One of the pioneering works by Weber[31] provides a general framework for decision making with incomplete information, where the incomplete information about states of nature and utilities is formalized by means of probability intervals and linear inequalities, respectively. The proposed framework leads to solving the linear programming problems. Various extensions of the framework taking into account some peculiarities of eliciting the decision information have been provided by many authors. Danielson, Ekenberg *et al*[5, 9] proposed a class of second-order uncertainty models applied to decision making under incomplete information. An interesting robust classification model for the software defect prediction in the framework of the AHP was proposed by Peng, Kou *et al*[20].

One of the approaches for representing different types of imprecision is to consider a set of criterion weights produced by possible judgments provided by experts or by the DM[15, 16, 26]. In particular, the expert opinions in the form of the preference ratios have been studied by Salo and Hamalainen[22]. An interesting method for the analysis of incomplete preference information in hierarchical weighting models of the multi-criteria decision making leading to possibly non-convex sets of feasible attribute weights has been proposed by Salo and Punkka[23]. Park and Kim[19] formalized different kinds of judgments or statements by means of a number of linear inequalities for weights of criteria. Due to linearity, these inequalities form a convex polytope of vectors of weights and this fact gives the opportunity to use linear programming for computing a measure for ranking the DA's.

In order to model imprecision of judgments, the AHP method was modified by many authors by replacing the precise elements of comparison matrices with interval-valued or fuzzy elements[7, 8, 11, 12, 14, 17, 30]. At that, fuzzy comparison matrices are often transformed into interval comparison matrices using α-level sets. The fuzzy approach is a very useful and important extension of the AHP method. However, it requires to introduce additional assumptions concerning the corresponding membership functions. Moreover, most elicitation procedures in this case are based on standard pairwise comparisons which have some shortcomings.

The proposed method avoids the pairwise comparisons and uses the minimal initial information. Of course, it does not mean that we get in this case a better solution to the multi-criteria decision problem. On the contrary, we can not get a better solution by having a small amount of initial information. However, we get a more cautious solution and can choose between optimistic and pessimistic strategies.

In contrast to the above approaches, the proposed method first of all modifies the AHP method or the DS/AHP method. Second, it uses the most simple imprecise comparison judgments without eliciting additional information in the form of weights or probabilities. Experts even do not need to provide strong comparison judgments. They select only some subsets of DA's from the set of DA's in accordance with a certain criterion, and they select subsets of criteria from the set of criteria as favorable groups of criteria on every level. Third, the proposed method simultaneously models imprecise judgments on all levels of the hierarchy (DA's and all levels of criteria). Fourth, it is based on using the strong mathematical framework of DST.

The paper is organized as follows. The main definitions of DST are proposed in Section 2. The main idea of the DS/AHP method and its illustration is given in Section 3. Problems of incomplete information about criteria and the ways for processing this information are discussed in Section 4. Section 5 considers the proposed extension applied to two levels of criteria and answers on the question how to unite the incomplete information about importance of criteria in this case. Two levels of criteria can be studied in the framework of group decision making. The generalization of the proposed extension to the arbitrary number of levels of criteria is studied in Section 6.

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[6, 25]:

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

Note that the domain of BPA, $\mathcal{P}o(U)$, is different from the domain of a probability density function, which is U. This function can be obtained as follows[6]:

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

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

The *belief* $Bel(A)$ and *plausibility* $\mathrm{Pl}(A)$ measures of an event $A \subseteq \Omega$ can be defined as[25]

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

As pointed out by Halpern and Fagin[10], 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[6] 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., $\mathrm{Bel}(A) \leq \Pr(A) \leq \mathrm{Pl}(A)$.

Let p_i be some unknown probability of the i-th element of the universal set $U = \{u_1, ..., u_m\}$. Then the probability distribution $p = (p_1, ..., p_m)$ satisfies the following inequalities for all focal elements A:

$$\mathrm{Bel}(A) \leq \sum_{i : u_i \in A} p_i \leq \mathrm{Pl}(A).$$

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 selects some subsets $B_k \subseteq \mathbb{A}$ of DA's from the set $\mathbb{A}$ in accordance with the certain criterion C_j from $\mathbb{C}$. The nice idea proposed by Beynon *et al*[2, 3, 4] is that the DM has to identify favorable DA's from the set $\mathbb{A}$ instead of comparing DA's between each other. 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. The DM also has to provide pairwise comparisons between groups of DA's and the set $\mathbb{A}$, from which the so-called knowledge matrix[4] (a reduced matrix of pairwise comparisons with respect to each criterion) is formed for each criterion.

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

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 corresponding hierarchical decision tree with one level of criteria is depicted in Fig. 1. The knowledge matrix for criterion C_1 is shown in Table 1. According to Beynon* et al*[4], a 6-point scale (1-6) is used for the pairwise comparisons instead of a 9-point scale (1-9) as in the 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 by Beynon* et al*[4]. 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 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 the paper by Beynon et al*[4].*

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

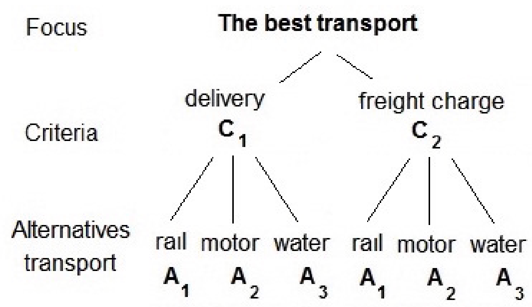

Figure 1: A hierarchical decision tree with one level of criteria

4 Incomplete information about criteria and DA's

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 the same 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.

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 select some subsets $B_k \subseteq \mathbb{A}$ of DA's from the set $\mathbb{A}$ in accordance with the certain criterion C_j from $\mathbb{C}$. Moreover, they select some subsets $D_i \subseteq \mathbb{C}$ from the set $\mathbb{C}$ 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 selects 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, providing the judgments, 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 selects 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 select 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 providing judgments in the form of subsets $B_1, ..., B_l$, respectively. This procedure is repeated 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. Suppose that 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 second rows).*

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

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

5 Processing and aggregating the incomplete information

A method for aggregating and processing the above incomplete information totally depends on the criterion of decision making. Roughly speaking, a large part of decision methods consists of aggregating the different local criteria from the set $\mathbb{C}$ into a function called a global criteria, which has to be maximized. According to these methods, global criteria by a finite set $\mathbb{A}$ of DA's can be represented as follows:

$$F(\mathbf{w},\mathbf{u}_k) \to \max_{\mathbb{A}}. \tag{3}$$

Here $\mathbf{w} = (w_1, ..., w_r)$ is the vector of "weights" or importance measures of criteria; $\mathbf{u}_k = (u_{1k}, ..., u_{rk})$, $k = 1, ..., n$, is the vector of "weights" or utilities of the k-th DA with respect to every criterion from $\{C_1, ..., C_r\}$; F is some function allowing us to combine the "weights" of criteria and DA's in order to get a final measure of "optimality" of every DA. In particular, one of the most widely applied criteria is the linear function F, i.e.,

$$F(\mathbf{w},\mathbf{u}_k) = \sum_{i=1}^{r} (w_i \cdot u_{ik}). \tag{4}$$

However, it is obvious that, by having only partial information about $\mathbf{w}$ and $\mathbf{u}_k$, we can get only partial information about F. Suppose that the vectors $\mathbf{w}$ and $\mathbf{u}_k$ take values from the sets $\mathcal{W}$ and $\mathcal{U}_k$, respectively. Here $\mathcal{U}_k = \mathcal{U}_{1k} \times ... \times \mathcal{U}_{rk}$ is the Cartesian product of r intervals $\mathcal{U}_{ik}$. Then, by accepting that $\mathbf{w} \in \mathcal{W}$ and $\mathbf{u}_k \in \mathcal{U}_k$, we can say that the function F belongs to some interval or a set of intervals $\mathcal{F}$. Moreover, all elements of $\mathcal{F}$ are equivalent in the sense that we can not choose a more preferable element or a subset of elements from the set because all elements in $\mathcal{W}$ and $\mathcal{U}_k$ are equivalent in the same sense.

Suppose that every constraint for $\mathbf{w}$ is linear. Then the set $\mathcal{W}$ is convex. Given the convex set $\mathcal{W}$ and the Cartesian product $\mathcal{U}_k$, we will prove that $\mathcal{F}$ is a convex interval and, therefore, has the lower $\underline{F}$ and upper $\overline{F}$ bounds.

Let us fix a vector $\mathbf{u}_k \in \mathcal{U}_k$. Since $\mathcal{W}$ is convex, then there are some lower $\underline{F}(\mathbf{u}_k)$ and upper $\overline{F}(\mathbf{u}_k)$ bounds for $F(\mathbf{u}_k)$ by the fixed vector $\mathbf{u}_k$. The function $F(\mathbf{u}_k)$ is non-decreasing with every u_{ik}, $i = 1, ..., r$, because there holds $w_i \geq 0$. This implies that the lower bound for $\underline{F}(\mathbf{u}_k)$ is achieved at point $u_{ik} = \inf \mathcal{U}_{ik}$ and the upper bound for $\overline{F}(\mathbf{u}_k)$ is achieved at point $u_{ik} = \sup \mathcal{U}_{ik}$. Finally, we can conclude that the lower bound for $F(\mathbf{w},\mathbf{u}_k)$ over the set $\mathcal{W} \times \mathcal{U}_k$ is $\underline{F}(\underline{\mathbf{u}}_k)$ and the upper bound over the same set is $\overline{F}(\overline{\mathbf{u}}_k)$. Here $\underline{\mathbf{u}}_k$ and $\overline{\mathbf{u}}_k$ are vectors with elements $\inf \mathcal{U}_{ik}$ and $\sup \mathcal{U}_{ik}$, respectively.

The next question is how to change the global criterion in (3) when we have a set of possible functions F. Suppose that w_i and u_{ik} belong to closed intervals for all $i = 1, ..., r$ and $k = 1, ..., n$. Then the function

F also belongs to a closed interval. Then the choice of the "best" DA can be based on comparison of intervals of F. There exist a lot of methods for comparison. We propose to use the most justified method based on the so-called caution parameter[24, 32, 27, 28] or the parameter of pessimism $\eta \in [0,1]$ which has the same meaning as the optimism parameter in Hurwicz criterion[13]. According to this method, the "best" DA from all possible ones should be chosen in such a way that makes the convex combination $\eta \cdot \inf F + (1-\eta) \sup F$ achieve its maximum. If $\eta = 1$, then we analyze only lower bounds for F for all DA's and make pessimistic decision. This type of decision is very often used[1, 18]. If $\eta = 0$, then we analyze only upper bounds for F for all DA's and make optimistic decision.

Therefore, the next problems are how to find $\mathbf{w}$ and $\mathbf{u}_k$, how to interpret $\mathbf{w}$ and $\mathbf{u}_k$, how to prove that w_i and u_{ik} belong to closed intervals, how to compute the lower and upper bounds for the functions F from $\mathcal{F}$. These problems arise due to the fact that we do not have complete information about weights of criteria and DA's. The following approach can be proposed here for solving the above problems.

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

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

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

On the other hand, suppose that the j-th criterion is selected 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:

$$\mathrm{Bel}(D_k) \leq \sum_{j:C_j \in D_k} p_j \leq \mathrm{Pl}(D_k), \ k = 1, ..., 2^r - 1. \tag{5}$$

Here p_j can be regarded as the weight w_j of the j-th criterion, $j = 1, ..., r$.

By viewing the belief and plausibility functions as lower and upper probabilities, respectively, we can say that the set of inequalities (5) 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:

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

In fact, we apply here the linear function of weights $p_1, ..., p_r$ of the local criteria $C_1, ..., C_r$. The BPA $m(B_k \mid C_j)$ can be regarded as the weight u_{jk} of the k-th DA or the sum of weights of the k-th group of DA's with respect to the j-th criterion.

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:

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

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

subject to $\sum_{j=1}^{r} p_j = 1$ and (5).

It should be noted here that the lower belief function is the lower bound for F in (4) and the upper plausibility function is the upper bound for F in (4).

When we do not have information about criteria at all, then the set of constraints to the above linear programming problems is 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's B_k can be computed as follows:

$$\mathrm{Bel}(B_k) = \min_{j=1,\dots,r} \sum_{i:B_i\subseteq B_k} m(B_i \mid C_j), \tag{6}$$

$$\mathrm{Pl}(B_k) = \max_{j=1,\dots,r} \sum_{i:B_i\cap B_k\neq\varnothing} m(B_i \mid C_j). \tag{7}$$

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 smallest "combined" belief function of every DA over all criteria in accordance with (6). Then we compute the largest 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 (7).

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 with the parameter of pessimism $\eta \in [0,1]$. 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.

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

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

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

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

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:

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

subject to $p_1 + p_2 = 1$ and

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

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*

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

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*

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

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

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

6 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 select some subsets $D_i \subseteq \mathbb{C}$ from the set $\mathbb{C}$ as favorable groups of criteria on the first level. Experts also select some subsets $D_k(k_1) \subseteq \mathbb{C}_2(k_1)$ from the set $\mathbb{C}_2(k_1)$ as favorable groups of criteria on the second level with respect to the criterion of the first level having the number k_1.

Suppose that the k-th criterion on the first level is selected 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:

$$\mathrm{Bel}(D_k) \leq \sum_{j:C_j\in D_k} q_j \leq \mathrm{Pl}(D_k),\ k = 1, ..., 2^t - 1. \tag{8}$$

Let $\mathcal{Q}$ be the set of probability distributions produced by all constraints (8).

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

$$\mathrm{Bel}(D_l(k)) \leq \sum_{j:C_j(k)\in D_l(k)} q_j(k) \leq \mathrm{Pl}(D_l(k)),$$
$$l = 1, ..., 2^r - 1,\ k = 1, ..., t. \tag{9}$$

Let $\mathcal{Q}(k)$ be the set of probability distributions produced by all constraints (9) by a fixed value of k. Denote

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

[1] We assume for simplicity 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 $i \neq k$.

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

Here the index j corresponds to the j-th criterion of the second level selected 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 the fixed distributions q and $q(k)$, $k=1,...,t$,

$$\mathrm{Bel}_{q,q(k)}(B_l)=\sum_{i:B_i\subseteq B_l} m_{q,q(k)}(B_i)=\sum_{k=1}^{t} q_k\sum_{j=1}^{r} q_j(k)\cdot a_{jl}(k), \tag{10}$$

$$\mathrm{Pl}_{q,q(k)}(B_l)=\sum_{i:B_i\cap B_l\neq\varnothing} m_{q,q(k)}(B_i)=\sum_{k=1}^{t} q_k\sum_{j=1}^{r} q_j(k)\cdot b_{jl}(k). \tag{11}$$

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

$$\mathrm{Bel}(B_l)=\min_{q,q(k)}\mathrm{Bel}_{q,q(k)}(B_l), \tag{12}$$

$$\mathrm{Pl}(B_l)=\max_{q,q(k)}\mathrm{Pl}_{q,q(k)}(B_l), \tag{13}$$

subject to (8) and (9).

At first sight, (12) and(13) are typical quadratic programming problems having linear constraints and nonlinear objective functions. However, we can show that every optimization problem can be solved by considering a set of $t+1$ linear programming problems.

Denote

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

Note that the multiplier $\mathbb{E}_{q(k)}a_l(k)$ in (10) 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 (9) or $q(k)\in\mathcal{Q}(k)$ and the linear programming problem

$$\mathrm{Bel}(B_l)=\min_{q}\sum_{k=1}^{t} q_k\cdot\underline{\mathbb{E}}a_l(k) \tag{14}$$

under constraints (8) or $q \in \mathcal{Q}$.

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

$$\mathrm{Pl}(B_l) = \max_q \sum_{k=1}^{t} q_k \cdot \overline{\mathbb{E}} b_l(k) \tag{15}$$

under constraints (8) or $q \in \mathcal{Q}$, where $\overline{\mathbb{E}} b_l(k)$, $k = 1, ..., t$, is obtained by solving t simple linear programming problems

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

under constraints (9) 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 the firms have different levels of the delivery reliability and the freight charge. The corresponding hierarchical decision tree with two levels of criteria is depicted in Fig. 2. 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 the 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 hold for* $l = 1$ $(B_1 = \{A_1\})$, $k = 1$,

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

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

Constraints are of the form (see (9)):

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

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. Now we can compute the lower unconditional belief function* $\mathrm{Bel}(B_l)$ *from (14) 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 (8):*

$$\begin{aligned}0.4 \le q_1 \le 1, \\ 0 \le q_2 \le 0.6, \\ 1 = q_1 + q_2.\end{aligned}$$

Table 6: Expert preferences related to criteria on the second level with respect to the first firm

	$D_1(1)$	$D_2(1)$	$D_3(1)$
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 with respect to the second firm

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

In the same way, we can find the upper unconditional plausibility function $\mathrm{Pl}(B_l)$ *from (15) 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".*

7 Arbitrary number of levels of criteria

Suppose that there are v levels of the criteria hierarchy. In order to define a criterion on the i-th level of the hierarchy, we have to write a way to this criterion by starting from the first level. This way is determined by vector $\mathbf{k}_i = (k_1, ..., k_{i-1})$ of integers, which means that the way starts from the criterion with the number k_1 on the first level, passes through the criterion having the number k_2 on the second level with the "parent" criterion with the number k_1, etc. The first level contains t_1 criteria, the second level contains $t_1 t_2$ criteria, etc. The i-th level contains a set $\mathbb{C}(\mathbf{k}_i) = (C_1(\mathbf{k}_i), ..., C_{t_i}(\mathbf{k}_i))$ of t_i criteria such that the "parent" criterion is determined by the way $\mathbf{k}_i$. Experts select some subsets $D_j(\mathbf{k}_i) \subseteq \mathbb{C}(\mathbf{k}_i)$ of criteria as favorable groups of criteria from the set $\mathbb{C}(\mathbf{k}_i)$ on the i-th level of the hierarchy.

Suppose that the k_i-th criterion on the i-th level defined by the way $\mathbf{k}_i$ is selected by experts with some unknown probability $q_k(\mathbf{k}_i)$ such that the condition $\sum_{k=1}^{t_i} q_k(\mathbf{k}_i) = 1$ is valid. Then the probabilities of the criteria satisfy the following system of inequalities

$$\mathrm{Bel}(D_k(\mathbf{k}_i)) \leq \sum_{j: C_j(\mathbf{k}_i) \in D_k(\mathbf{k}_i)} q_j(\mathbf{k}_i) \leq \mathrm{Pl}(D_k(\mathbf{k}_i)),$$
$$k = 1, ..., 2^{t_i} - 1, \ i = 1, ..., v. \tag{16}$$

Let $\mathcal{Q}(\mathbf{k}_i)$ be the set of probability distributions $q(\mathbf{k}_i) = (q_1(\mathbf{k}_i), ..., q_{t_i}(\mathbf{k}_i))$ produced by all the above constraints for a fixed $\mathbf{k}_i$.

Denote

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

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
$\mathrm{Bel}(B_l)$	0.245	0.089	0.156	0.557	0.533	0.289	1
$\mathrm{Pl}(B_l)$	0.71	0.467	0.443	0.844	0.91	0.755	1

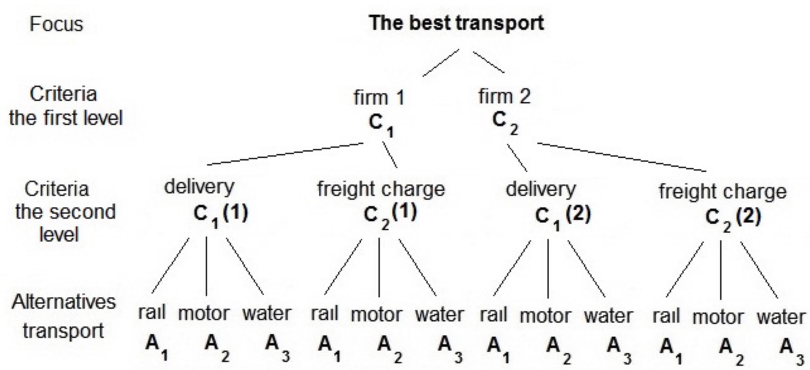

Figure 2: A hierarchical decision tree with two levels of criteria

If we fix the probability distributions $q(\mathbf{k}_i)$ for all possible $\mathbf{k}_i$ and $i = 1, ..., v$, then we can write

$$\mathrm{Bel}_q(B_l) = \sum_{k=1}^{t_1} q_k(\mathbf{k}_1) \cdot \sum_{k=1}^{t_2} q_k(\mathbf{k}_2) \cdots \sum_{k=1}^{t_v} q_k(\mathbf{k}_v) \cdot a_{kl}(\mathbf{k}_v),$$

$$\mathrm{Pl}_q(B_l) = \sum_{k=1}^{t_1} q_k(\mathbf{k}_1) \cdot \sum_{k=1}^{t_2} q_k(\mathbf{k}_2) \cdots \sum_{k=1}^{t_v} q_k(\mathbf{k}_v) \cdot b_{kl}(\mathbf{k}_i).$$

By using the approach proposed for the case $v = 2$, we get the set of

$$1 + t_1 + t_1 t_2 + ... + t_1 \cdots t_v = 1 + \sum_{i=0}^{v-1} \prod_{k=1}^{v-i} t_k$$

linear programming problems for computing the lower belief function and the same number of problems for computing the upper plausibility function. At first step, we solve $t_1 \cdots t_v$ linear problems for computing the lower bounds for

$$\mathbb{E}_{q_k(\mathbf{k}_v)} a_l(\mathbf{k}_v) = \sum_{k=1}^{t_v} q_k(\mathbf{k}_v) \cdot a_{kl}(\mathbf{k}_v).$$

Then we solve $t_1 \cdots t_{v-1}$ problems for computing the lower bounds for

$$\mathbb{E}_{q_k(\mathbf{k}_{v-1})} a_l(\mathbf{k}_{v-1}) = \sum_{k=1}^{t_{v-1}} q_k(\mathbf{k}_{v-1}) \cdot \underline{\mathbb{E}}_{q_k(\mathbf{k}_v)} a_l(\mathbf{k}_v).$$

The above procedures are repeated v times and, finally, we solve one problem

$$\mathrm{Bel}(B_l) = \sum_{k=1}^{t_1} q_k \cdot \underline{\mathbb{E}}_{q_k(\mathbf{k}_1)} a_l(\mathbf{k}_1).$$

In the similar way, we get the plausibility function $\mathrm{Pl}(B_l)$.

8 Conclusion

An extension of the DS/AHP method has been proposed in the paper. The extension uses groups of experts or DM's, takes into account the possible selection of groups of criteria, does not require to assign favorability values for groups of DA's and criteria. It also assumed that there are several levels of criteria and estimates of experts on every level can be incomplete and imprecise. The main advantage of the proposed approach for decision making is the rather simple procedure for computing the belief and plausibility functions of the DA's and their groups. This procedure is based on solving a set of linear programming problems, which can be carried out by means of standard tools. The procedure has been explained and illustrated by various numerical examples.

It should be noted that the proposed approaches can simply be extended on the case when experts are asked to supply the favorability values for groups of DA's and criteria. If the j-th expert provides the preference rate $x_{ki}^{(j)} \in \{0, 1, ..., m\}$ (the value 0 is used if the corresponding preference is not selected by experts) for DA's, then the BPA of the subset B_k is computed as

$$m(B_k) = \sum_{j=1}^{a_k} x_k^{(j)} \ / \ N,$$

where N is the total sum of the preference rates of all subsets with respect to a criterion.

The same procedure can be applied to criteria.

At the same time, the procedure of decision making could be improved by considering all possible judgments, including comparisons of different groups of DA's. This is a direction for further work, which, in our opinion, could be solved in the framework of DST or imprecise probability theory[29]. Another direction for further work is to study different types of independence conditions. In the considered extension, we have investigated only the so-called strong independence. However, by having imprecise estimates or judgments, we might use several types of independence conditions, including the random set independence, the unknown interaction, etc.

Acknowledgement

The authors would like to express their appreciation to the anonymous referees whose very valuable comments have improved the paper.

References

[1] T. Augustin, Expected utility within a generalized concept of probability - a comprehensive framework for decision making under ambiguity, *Stat. Pap.*, **43**(2002) 5–22.

[2] M. Beynon, DS/AHP method: A mathematical analysis, including an understanding of uncertainty, *Eur. J. Oper. Res.*, **140**(2002) 148–164.

[3] M. Beynon, A method of aggregation in DS/AHP for group decision-making with the non-equivalent importance of individuals in the group, *Comput. Oper. Res.*, **32**(2005) 1881–1896.

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

[5] M. Danielson, L. Ekenberg, and A. Larsson, Distribution of expected utility in decision trees, *Int. J. Approx. Reason.*, **46**(2007) 387–407.

[6] A.P. Dempster, Upper and lower probabilities induced by a multi-valued mapping, *Ann. Math. Stat.*, **38**(1967) 325–339.

[7] M. Deng, W. Xu, J.-B. Yang, Estimating the attribute weights through evidential reasoning and mathematical programming, *Int. J. Inf. Tech. Decis.*, **3**(2004) 419–428.

[8] D.K. Despotis and D. Derpanis, A minmax goal programming approach to priority derivation in AHP with interval judgements, *Int. J. Inf. Tech. Decis.*, **7**(2008) 175–182.

[9] L. Ekenberg, J. Thorbioernson, Second-Order Decision Analysis, *Int. J. Uncertain. Fuzz.*, **9**(2001) 13–38.

[10] J.Y. Halpern and R. Fagin, Two views of belief: Belief as generalized probability and belief as evidence, *Artif. Intell.*, **54**(1992) 275–317.

[11] M.-S. Kuo, G.-S. Liang, and W.-C. Huang, Extensions of the multicriteria analysis with pairwise comparison under a fuzzy environment, *Int. J. Approx. Reason.*, **43**(2006) 268–285.

[12] C.-T. Lin, C. Lee, and C.-S. Wu, Fuzzy group decision making in pursuit of a competitive marketing strategy, *Int. J. Inf. Tech. Decis.*, **9**(2010) 281–300.

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

[14] L. Mikhailov, H. Didehkhani, and S. Sadi-Nezhad, Weighted prioritization models in the fuzzy analytic hierarchy process, *Int. J. Inf. Tech. Decis.*, **10**(2011) 681–694.

[15] J. Mustajoki, R.P. Hamalainen, and M.R.K. Lindstedt, Using intervals for global sensitivity and worst-case analyses in multiattribute value trees, *Eur. J. Oper. Res.*, **174**(2006) 278–292.

[16] J. Mustajoki, R.P. Hamalainen, and A. Salo, Decision support by interval SMART/SWING – Incorporating imprecision in the SMART and SWING methods, *Decision Sci.*, **36**(2005) 317–339.

[17] A. Nieto-Morote and F. Ruz-Vila, A fuzzy AHP multi-criteria decision-making approach applied to combined cooling, heating, and power production systems, *Int. J. Inf. Tech. Decis.*, **10**(2011) 497–517.

[18] K.-M. Osei-Bryson, Supporting knowledge elicitation and consensus building for Dempster-Shafer decision models, *Int. J. Intell. Syst.*, **18**(2003) 129–148.

[19] K.S. Park and S.H. Kim, Tools for interactive multi-attribute decision making with incompletely identified information, *Eur. J. Oper. Res.*, **98**(1997) 111–123.

[20] Y. Peng, G. Kou, G. Wang, W. Wu and Y. Shi, Ensemble of software defect predictors: An AHP-based evaluation method, *Int. J. Inf. Tech. Decis.*, **10**(2011) 187-206.

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

[22] A. Salo and R.P. Hamalainen, Preference ratios in multiattribute evaluation (PRIME) – Elicitation and decision procedures under incomplete information, *IEEE T. Syst. Man. Cyb.*, **31**(2001), 533–545.

[23] A. Salo and A. Punkka, Rank inclusion in criteria hierarchies, *Eur. J. Oper. Res.*, **163**(2005) 338–356.

[24] J. Schubert, On ρ in a decision-theoretic apparatus of Dempster-Shafer theory, *Int. J. Approx. Reason.*, **13**(1995) 185–200.

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

[26] 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.

[27] L.V. Utkin and Th. Augustin, Efficient algorithms for decision making under partial prior information and general ambiguity attitudes, in *Proc. of the 4th Int. Symposium on Imprecise Probabilities and Their Applications, ISIPTA'05*, eds. T. Seidenfeld, F.G. Cozman, R. Nau (Carnegie Mellon University, SIPTA, Pittsburgh, USA, 2005) pp. 349–358.

[28] L.V. Utkin and Th. Augustin, Decision making under incomplete data using the imprecise Dirichlet model, *Int. J. Approx. Reason.*, **44**(2007) 322–338.

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

[30] Y.-M. Wang, J.-B. Yang, and D.-L. Xu, Interval weight generation approaches based on consistency test and interval comparison matrices, *Appl. Math. Comput.*, **167**(2005) 252–273.

[31] M. Weber, Decision making with incomplete information, *Eur. J. Oper. Res.*, **28**(1987) 44–57.

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