Journal of Uncertain Systems
Vol.2, No.4, pp.255-266, 2008
Online at: www.jus.org.uk

Multi-criteria Decision Making by Incomplete Preferences

Lev V. Utkin* Natalia V. Simanova
Department of Computer Science, St.Petersburg State Forest Technical Academy
Institutski per. 5, 194021, St.Petersburg, Russia

Received 26 October 2007; Accepted 1 December 2007

Abstract

A method for solving a multi-criteria decision problem in the frameworks of analytic hierarchy process and Dempster-Shafer theory under condition that the groups of experts and decision makers supply comparisons of arbitrary groups of decision alternatives and criteria is proposed in the paper. An algebra of comparative preferences with the corresponding set-theoretical operations is developed. A rule for combining the preferences for alternatives and for criteria by using sets of probability distributions and the total probability theorem is proposed. The cautious decisions with using the imprecise Dirichlet model are used for some cases of initial data. Numerical examples explain and illustrate the proposed method.

Keywords: multi-criteria decision problem, analytic hierarchy process, Dempster-Shafer theory, imprecise Dirichlet model, preferences, pairwise comparison

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 [9]. 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 DA.

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 DA's, and it is also difficult to obtain a convincing consistency index with an increasing number of attributes or DA's. 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 some difficulties and to extend the AHP on a more real elicitation procedures, Beynon *et al.* [3, 4] proposed a method using *Dempster-Shafer theory* (DST) and is 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. An extension of the method was proposed by Tervonen *et al.* [12]. It should be noted that the main excellent idea underlying the DS/AHP method is not applying Dempster-Shafer theory to the AHP. Beynon *et al.* [3, 4] proposed to compare groups of DA's by means of their separate comparisons with the set of all DA's and assignments different rates to the comparisons. The such type of comparisons is equivalent to the preferences stated by the DM. The DS/AHP method has many advantages. However, it does not allow us to take into account possible comparisons of groups of DA's each other. The second shortcoming is that weights of criteria in the DS/AHP method are obtained by using the standard comparison procedure used in the AHP. The third disadvantage concerns the procedure of computing the basic probability assignments. The fourth problem is that it is difficult to assign a numerical value of the favorable opinion for a particular group of DA's.

Therefore, we propose a method for solving a multi-criteria decision problem in the framework of AHP under condition that the groups of experts and DM's supply comparisons of arbitrary groups of DA's and

*Corresponding author. Email: lev.utkin@mail.ru (L.V. Utkin).

criteria. Moreover, we use this approach for obtaining preferences sets on all levels of the hierarchy in AHP, i.e., for criteria and DA's. We develop an algebra of comparative preferences with the corresponding set-theoretical operations. At that, the comparisons consisting of single DA's produce a universal set, the incomplete preferences are represented by set of simplest preferences, and, hence, the comparisons consisting of groups of DA's can be processed in the framework of DST.

The paper is organized as follows. Some definitions and elements of DST are given in Section 2. A very short description of the AHP can be found in Section 3. The main idea of the method for processing incomplete preferences is discussed in Section 4. In this section, the set-theoretic operations with preferences are introduced and it is shown how DST can be applied for processing the comparative expert judgments. A rule for combining the preferences for DA's and for criteria by using sets of probability distributions and the total probability theorem is considered in Section 5. Some extreme cases of initial data concerning the criteria are analyzed in this section. The cautious decisions with using the imprecise Dirichlet model are considered in Section 6. A decision problem is numerically solved under different initial data for illustrative purposes in Section 7.

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, 11]:

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

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

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

As pointed out in [6], 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., $\mathrm{Bel}(A) \leq \Pr(A) \leq \mathrm{Pl}(A)$.

Let us explain the belief and plausibility functions in terms of a *multivalued sampling process*. Consider a probability measure $P(\omega)$ defined on a universal set Ω (which can be thought of as the set of our observations) related to U (the set of the values of our measurements) through a multivalued mapping $G : \Omega \to \mathcal{P}o(U)$. Then the BPA is [5]:

$$m(A_i) = P(\omega_i) = c_i/N,\ \omega_i \in \Omega.$$

Let A be a subset of U. If we define X_* as the subset of Ω whose elements must lead to A:

$$X_* = \{\omega \in \Omega : G(\omega) \subseteq A\},$$

then the lower probability of A, according to Dempster's principle of inductive reasoning, is defined by

$$\underline{P}(A) = \mathrm{Bel}(A) = P(X_*).$$

If we define X^* as the subset of Ω whose elements may lead to A:

$$X^* = \{\omega \in \Omega : G(\omega) \cap A \neq \emptyset\},$$

then the upper probability of A is given by

$$\overline{P}(A) = \mathrm{Pl}(A) = P(X^*).$$

3 The AHP Method

Briefly, the AHP decomposes a decision problem into elements, according to their common characteristics, and levels, which correspond to the common characteristic of the elements. The topmost level or "focus" of the problem is the main goal; the intermediate levels correspond to criteria, while the lowest level contains the DA's. For simplicity, we assume that there are only one level of DA's and one level of criteria (see fig. 1 for example). The AHP is based on paired comparisons and the use of ratio scales in preference judgements. The scales used in the AHP allow to convert the qualitative judgments into numerical values. The DA's are compared pairwise with respect to a specific criteria. The criteria are also compared pairwise. At that, the DM is asked to give the ratio of alternatives' weight $a_{ij}^{(k)}$ which represents the pairwise comparison rating between the element i and element j of a level with respect to a criterion, say the k-th criterion. The results of paired comparisons are presented in a comparison matrix $A_k = [a_{ij}^{(k)}]$. The entries $a_{ij}^{(k)}$ are governed by the following rules:

$$a_{ij}^{(k)} > 0; a_{ij}^{(k)} = 1/a_{ij}^{(k)}; a_{ii}^{(k)} = 1$$

for all i and j.

The priorities or weights of the elements can be estimated by finding the principal eigenvector W_k of the matrix A_k, that is $A_k W_k = \lambda_{\max} W_k$. Here $\lambda_{\max}$ is the largest eigenvalue of the matrix A_k. When the vector W_k is normalized, it becomes the vector of priorities of DA's with respect to the k-th criterion. In the same way, criteria are compared and the normalized principal eigenvector V as the vector of weights of criteria is computed.

Once the local priorities of DA's with respect to all criteria are available and the weights of criteria are obtained, in order to obtain final priorities of the DA's, the priorities are aggregated as follows:

$$w_i = \sum_k v_k w_i^{(k)}.$$

Here v_k is the weight the k-th criterion (the k-th element of the vector V); $w_i^{(k)}$ is the priority of the i-th DA with respect to the k-th criterion.

The main virtues and shortcomings of the AHP have been pointed out in the introductory section. Therefore, we do not consider them here.

4 Algebra of Preferences

We suppose that there is a set of DA's $A = \{A_1, ..., A_n\}$ consisting of n elements and a set of criteria $C = \{C_1, ..., C_r\}$ consisting of r elements. Denote the set of all subsets of A (the power set) by $\mathcal{P}o(A)$ and the set of all subsets of C by $\mathcal{P}o(C)$. Let B_k be the short notation of a subset of A, i.e., $B_k \subseteq A$ or $B_k \in \mathcal{P}o(A)$. Let D_k be the short notation of a subset of C, i.e., $D_k \subseteq C$ or $D_k \in \mathcal{P}o(C)$. Here the index k is an order number of the corresponding subset of A or C. For example, the possible correspondences between subsets of DA's and B_k for $n = 3$ are given in Table 1.

An expert chooses some subset $B_k \subseteq A$ of DA's from the set A and compares this subset with another subset $B_i \subseteq A$ of DA's with respect to a certain criterion. In other words, experts set up the preferences $B_k \succeq B_i$. If B_k or B_i in the supplied preference $B_k \succeq B_i$ are not single elements of A, then we will say that there is an *incomplete preference*. In the same way, the DM chooses some subset $D_k \subseteq C$ of criteria from the set C and compares this subset with another subset $D_i \subseteq C$ of criteria. In other words, DM's set up the preferences $D_k \succeq D_i$. Every expert choice adds "1" to the corresponding preference, i.e. the preference rate is 0 or 1.

For example, if $A = \{A_1, A_2, A_3\}$, $B_k = \{A_3\}$, and $B_i = \{A_1, A_3\}$, then the preference $B_k \succeq B_i$ means that an expert chooses the DA A_3 from DA's A_1 and A_3. This is equivalent to the preference $\{A_3\} \succeq \{A_1\}$ which means that the DA A_3 is more preferable than A_1. If $B_k = \{A_2\}$ and $B_i = \{A_1, A_3\}$, then the preference

Table 1: The extended matrix of pairwise comparisons of DA's

	$\{A_1\}$	$\{A_2\}$	$\{A_3\}$	$\{A_1, A_2\}$	$\{A_1, A_3\}$	$\{A_2, A_3\}$	$\{A_1, A_2, A_3\}$
	B_1	B_2	B_3	B_4	B_5	B_6	B_7
B_1	-	a_{12}	a_{13}	a_{14}	a_{15}	a_{16}	a_{17}
B_2	a_{21}	-	a_{23}	a_{24}	a_{25}	a_{26}	a_{27}
B_3	a_{31}	a_{32}	-	a_{34}	a_{35}	a_{36}	a_{37}
B_4	a_{41}	a_{42}	a_{43}	-	a_{45}	a_{46}	a_{47}
B_5	a_{51}	a_{52}	a_{53}	a_{54}	-	a_{56}	a_{57}
B_6	a_{61}	a_{62}	a_{63}	a_{64}	a_{65}	-	a_{67}
B_7	a_{71}	a_{72}	a_{73}	a_{74}	a_{75}	a_{76}	-

Table 2: The extended matrix of pairwise comparisons of criteria

	$\{C_1\}$	$\{C_2\}$	$\{C_1, C_2\}$
	D_1	D_2	D_3
D_1	-	c_{12}	c_{13}
D_2	c_{21}	-	c_{23}
D_3	c_{31}	c_{32}	-

$B_k \succeq B_i$ means that the DA A_2 is more preferable than A_1 or A_3. In such the way, experts compare groups of DA's from the set A and DM's compare groups of criteria from the set C.

The extended matrix of pairwise comparisons of DA's in this case has $2^n - 1$ columns and rows (the empty element of $\mathcal{P}o(A)$ is not considered here). An example of such the matrix by $n = 3$ is shown in Table 1. Similarly, the extended matrix of comparisons of criteria has $2^r - 1$ columns and $2^r - 1$ rows. An example of such the matrix by $r = 2$ is shown in Table 2. It is supposed that experts and DM's compare only subsets of DA's and criteria, but they do not provide preference values or weights. At that, if an expert supplies the comparison assessment $B_k \succeq B_i$, then the value 1 is added to the corresponding cell in the comparison matrix (k-th row and i-th column). In this case, the preference value a_{ki} can be regarded as the number of experts chosen the comparison assessment $B_k \succeq B_i$. The same can be said about criteria where c_{ki} is the number of expert judgments concerning the preference $D_k \succeq D_i$. It should be noted that experts or DM's do not need to fill all cells of the extended matrices. An extreme case is when one of the matrices is empty, i.e., all a_{ki} or c_{ki} are 0.

First of all, we consider how to process the extended matrix of pairwise comparisons for DA's. We define the set $\mathcal{L}$ of *basic preferences of DA's*

$$\mathcal{L} = \{\{A_i\} \succeq \{A_k\},\ \forall i, k \in \{1, 2, ..., n\},\ i \neq k\}.$$

One can see that $\mathcal{L}$ consists of preferences between single DA's. We also define the set $\mathcal{M}$ of *basic preferences of criteria*

$$\mathcal{M} = \{\{C_i\} \succeq \{C_k\},\ \forall i, k \in \{1, 2, ..., r\},\ i \neq k\}.$$

Note that the preferences $\{A_i\} \succeq \{A_k\}$ and $\{A_j\} \succeq \{A_k\}$ follow from the preference $\{A_i, A_j\} \succeq \{A_k\}$ if $i \neq k$ and $j \neq k$. By generalizing the above and assuming that $B_k = \{A_v, ..., A_w\}$ consists of n_k DA's and $B_i = \{A_t, ..., A_l\}$ consists of n_i DA's such that $B_k \cap B_i = \emptyset$, we can say that the preference $B_k \succeq B_i$ implies $n_k \cdot n_i$ basic preferences of the form:

$$A_v \succeq A_t, A_{v+1} \succeq A_t, ..., A_v \succeq A_l, ..., A_w \succeq A_t, ..., A_w \succeq A_l.$$

At the same time, the preference $\{A_j\} \succeq \{A_k\}$ follows from the preference $\{A_k, A_j\} \succeq \{A_k\}$ if $j \neq k$. It can be seen from the above that the comparison of common parts of subsets B_k and B_i makes no sense. Experts have to compare different subsets of DA's, i.e., $B_k \cap B_i = \emptyset$ for all k and i. Nevertheless, we would not like to restrict experts to supply only "permitted" judgment. However, by processing the preference $B_k \succeq B_i$ with $B_k \cap B_i = \widetilde{B}_{ki} \neq \emptyset$, we will replace them by the following two preference $B_k \succeq B_i \backslash \widetilde{B}_{ki}$ and $B_k \backslash \widetilde{B}_{ki} \succeq B_i$. For instance, the preference $\{A_1, A_2\} \succeq \{A_1, A_2, A_3\}$ can be represented as the preference $\{A_1, A_2\} \succeq \{A_3\}$ ($B_k \succeq B_i \backslash \widetilde{B}_{ki}$, $\widetilde{B}_{ki} = \{A_1 A_2\}$), which can be represented as the subset of the set $\mathcal{L}$ consisting of the basic preferences $\{A_1\} \succeq \{A_3\}$, $\{A_2\} \succeq \{A_3\}$.

In sum, we can represent every preference by the set of simplest preferences of the from $\{A_i\} \succeq \{A_k\}$ from the set $\mathcal{L}$ with the corresponding operations.

The set $\mathcal{L}$ and all intersections of its elements can be regarded as the universal set. At the same time, we can not precisely assign probabilities to elements of the universal set because it is necessary to take into account the so-called incomplete preferences or estimates concerning the groups of DA's. We do not have complete probabilistic information about basic preferences from the set $\mathcal{L}$ in this case. For instance, by having the probability p for the incomplete preference $\{A_2\} \succeq \{A_1 A_3\}$ and by representing this preference as a subset of two basic preferences $\{A_2\} \succeq \{A_1\}$ and $\{A_2\} \succeq \{A_3\}$, we do not know how the probability p is distributed among the preferences $\{A_2\} \succeq \{A_1\}$ and $\{A_2\} \succeq \{A_3\}$. In this case, we can apply the framework of DST to the considered preferences.

For brevity, we will denote the preferences $B_k \succeq B_i$ and $D_k \succeq D_i$ by $\mathcal{B}_{ki}$ and $\mathcal{D}_{ki}$, respectively. For every pairwise comparison in the extended comparison matrix, we define its BPA as follows:

$$m(B_k \succeq B_i) = m(\mathcal{B}_{ki}) = \frac{a_{ki}}{N}, \ N = \sum_{k,i \in \{1,2,\dots,n\}, k \neq i} a_{ki}$$

or

$$m(D_k \succeq D_i) = m(\mathcal{D}_{ki}) = \frac{c_{ki}}{M}, \ M = \sum_{k,i \in \{1,2,\dots,r\}, k \neq i} c_{ki}.$$

Since every preference $B_k \succeq B_i$ ($D_k \succeq D_i$) is represented by a set of basic preferences, we denote this set of basic preferences $\mathcal{L}_{ki} \subseteq \mathcal{L}$ ($\mathcal{M}_{ki} \subseteq \mathcal{M}$). Now we can define some rules of set-theoretic operations.

We will say that the preference $B_j \succeq B_l$ is a subset of the preference $B_k \succeq B_i$ or $B_k \succeq B_i$ includes $B_j \succeq B_l$ if there holds $\mathcal{L}_{jl} \subseteq \mathcal{L}_{ki}$. This means that the set of basic preferences $\mathcal{L}_{jl}$ produced by $B_j \succeq B_l$ is a subset of basic preferences $\mathcal{L}_{ki}$ produced by $B_k \succeq B_i$.

We will also say that the preference $B_j \succeq B_l$ intersects the preference $B_k \succeq B_i$ if there holds $\mathcal{L}_{ki} \cap \mathcal{L}_{jl} \neq \emptyset$. This means that the sets $\mathcal{L}_{jl}$ and $\mathcal{L}_{ki}$ produced by $B_j \succeq B_l$ and $B_k \succeq B_i$ have common basic preferences.

Then the belief and plausibility functions for the preference $B_k \succeq B_i$ can be defined as follows:

$$\text{Bel}(\mathcal{B}_{ki}) = \sum_{j,l:\mathcal{L}_{jl} \subseteq \mathcal{L}_{ki}} m(\mathcal{B}_{jl}), \ \text{Pl}(\mathcal{B}_{ki}) = \sum_{j,l:\mathcal{L}_{jl} \cap \mathcal{L}_{ki} \neq \emptyset} m(\mathcal{B}_{jl}). \quad (3)$$

The belief and plausibility functions for the preference $D_k \succeq D_i$ can be defined in the same way

$$\text{Bel}(\mathcal{D}_{ki}) = \sum_{j,l:\mathcal{M}_{jl} \subseteq \mathcal{M}_{ki}} m(\mathcal{D}_{jl}), \ \text{Pl}(\mathcal{D}_{ki}) = \sum_{j,l:\mathcal{M}_{jl} \cap \mathcal{M}_{ki} \neq \emptyset} m(\mathcal{D}_{jl}). \quad (4)$$

Note that the plausibility functions $\text{Pl}(\mathcal{B}_{ki})$ and $\text{Pl}(\mathcal{D}_{ki})$ can be expressed through the belief functions of complementary preferences $\mathcal{B}_{ki}^c$ and $\mathcal{D}_{ki}^c$, respectively, i.e.

$$\text{Pl}(\mathcal{B}_{ki}) = 1 - \text{Bel}(\mathcal{B}_{ki}^c)$$

and

$$\text{Pl}(\mathcal{D}_{ki}) = 1 - \text{Bel}(\mathcal{D}_{ki}^c).$$

5 Combination of DA's and Criteria

By introducing the special algebra of sets of preferences, we have to define a rule for combining the preferences for DA's and for criteria. First of all, let us write the initial information we have after elicitation procedures.

Experts provide comparison judgments concerning the DA's with respect to the j-th criterion C_j, $j = 1, ..., r$. As a result, we can compute the conditional BPA's $m(\mathcal{B}_{ki} \mid C_j) = a_{ki}^{(j)}/N^{(j)}$, $j = 1, ..., r$, where $a_{ki}^{(j)}$ is the number of experts supplied the preference $\mathcal{B}_{ki}$ with respect to the j-th criterion, $N^{(j)}$ is the total number of experts supplied the preferences with respect to the j-th criterion.

Then DM's provide comparison judgments concerning the criteria. As a result, we can compute the BPA's $m(\mathcal{D}_{ki}) = c_{ki}/M$, $j = 1, ..., r$, where c_{ki} is the number of DM's supplied the preference $\mathcal{D}_{ki}$, M is the total number of DM's supplied the preferences.

The first difficulty here is that we have preferences concerning criteria and groups of criteria instead of single ones. Therefore, our first task is to find a way for replacing the preferences by separate criteria.

One of the possible solutions of this task is to use a nice idea of Beynon *et al.* [3, 4] in the DS/AHP method. According to this idea, the comparison of groups of objects with a whole set of objects, say C, is equivalent to the identification of the most favorable objects from the set C. This idea applied by Beynon *et al.* to DA's can be used for processing the criteria. For instance, if $C = \{C_1, C_2\}$, then the preference $\{C_1\} \succeq \{C_1, C_2\}$ means that the DM chooses C_1 from all criteria $\{C_1, C_2\}$. So, by computing the belief and plausibility functions of the preference $D_k \succeq C$, we determine the lower and upper probabilities of the group D_k of criteria or one criterion C_k if D_k consists of one element C_k. So, we have to compute the belief and plausibility functions of all possible preferences $D_k \succeq C$, $k = 1, ..., 2^r - 1$, by using (4).

Suppose that $C \backslash D_k = \{C_{i_1}, ..., C_{i_v}\}$. It is very important to point out here that the preference $D_k \succeq C$ means $D_k \succeq C_{i_1}$ *and* $D_k \succeq C_{i_2}$ *and, ..., and* $D_k \succeq C_{i_v}$. In other words, by dealing with preferences $D_k \succeq C$ below, we should consider the intersection of the above basic preferences, but not their union.

The second task is how to use the belief $\mathrm{Bel}(D_k \succeq C)$ and plausibility $\mathrm{Pl}(D_k \succeq C)$ functions of preferences $D_k \succeq C$ for combining the DA's and criteria. Suppose that the j-th criterion is chosen with the probability p_j such that $\sum_{j=1}^{r} p_j = 1$. Then the probabilities of criteria satisfy the following system of inequalities

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

By viewing the belief and plausibility functions as lower and upper probabilities, respectively, we can say that the set of the above inequalities produces a set $\mathcal{P}$ of possible distributions $p = (p_1, ..., p_r)$ satisfying all the inequalities (5). Let us fix a distribution p from $\mathcal{P}$. Then, by applying the total probability theorem, we can write the combined BPA's of preferences $\mathcal{B}_{ki}$ as follows:

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

It should be noted that the obtained BPA depends on the probability distribution $p \in \mathcal{P}$ and can not be considered as a final result. We return to the question what to do with these BPA's later.

After the BPA's of all possible preferences depending on $p \in \mathcal{P}$ are computed, we have to choose the "best" DA. Therefore, the third task is to choose the "best" DA depending on $p \in \mathcal{P}$. This task can be solved by using the same idea of Beynon *et al.* [3, 4] applied to DA's. In other words, we have to find the belief and plausibility functions of preferences $B_k \succeq A$ by using (3)

$$Bel_p(B_k \succeq A) = \sum_{i,l:\mathcal{L}_{il} \subseteq \mathcal{L}_{kL}} m_p(\mathcal{B}_{il}) = \sum_{j=1}^{r} p_j \cdot \left(\sum_{i,l:\mathcal{L}_{il} \subseteq \mathcal{L}_{kL}} m(\mathcal{B}_{il} \mid C_j) \right),$$

$$Pl_p(B_k \succeq A) = \sum_{i,l:\mathcal{L}_{il} \cap \mathcal{L}_{kL} \neq \emptyset} m_p(\mathcal{B}_{il}) = \sum_{j=1}^{r} p_j \cdot \left(\sum_{i,l:\mathcal{L}_{il} \cap \mathcal{L}_{kL} \neq \emptyset} m(\mathcal{B}_{il} \mid C_j) \right),$$

$$L = 2^n - 1.$$

Note that the obtained belief and plausibility functions linearly depend on p. Therefore, we can compute the lower belief and upper plausibility functions by solving the linear programming problems

$$\mathrm{Bel}(B_k \succeq A) = \inf_{p \in \mathcal{P}} \sum_{j=1}^{r} p_j \cdot \left(\sum_{i,l:\mathcal{L}_{il} \subseteq \mathcal{L}_{kL}} m(\mathcal{B}_{il} \mid C_j) \right),$$

$$\mathrm{Pl}(B_k \succeq A) = \sup_{p \in \mathcal{P}} \sum_{j=1}^{r} p_j \cdot \left(\sum_{i,l:\mathcal{L}_{il} \cap \mathcal{L}_{kL} \neq \emptyset} m(\mathcal{B}_{il} \mid C_j) \right)$$

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

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 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 can be computed as follows:

$$\mathrm{Bel}(B_k \succeq A) = \min_{j=1,\dots,r} \sum_{i,l:\mathcal{L}_{il} \subseteq \mathcal{L}_{kL}} m(\mathcal{B}_{il} \mid C_j), \tag{6}$$

$$\mathrm{Pl}(B_k \succeq A) = \max_{j=1,\dots,r} \sum_{i,l:\mathcal{L}_{il} \cap \mathcal{L}_{kL} \neq \emptyset} m(\mathcal{B}_{il} \mid C_j). \tag{7}$$

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 minimal "combined" belief function of every DA over all criteria in accordance with (6). 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 (7).

By having the belief and plausibility functions of all preferences $B_k \succeq A$, $k = 1,...,2^n - 1$, 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 of intervals. We propose to use the most justified method based on the so-called caution parameter [10, 18] or the parameter of pessimism $\eta \in [0,1]$ which has the same meaning as the optimism parameter in Hurwicz criterion [7]. 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 [1, 8]. If $\eta = 0$, then we analyze only plausibility functions and make optimistic decision.

6 Cautious Decision Making with the Imprecise Dirichlet Model

One of the main difficulty of the proposed method is the possible small number of experts. Expression (1) can be used when the number of expert judgments is rather large. In order to overcome this difficulty, the *imprecise Dirichlet model* (IDM) [17] can be applied to extend belief and plausibility functions such that a lack of sufficient statistical data can be taken into account [13, 14].

For brevity, we will not consider in detail what this model is and how to obtain it. The interested reader should refer to [2, 17] and [13, 14, 15, 16]. We point out only that using the imprecise Dirichlet model leads to the extended belief and plausibility functions of the form:

$$\mathrm{Bel}_s(A) = \frac{N \cdot \mathrm{Bel}(A)}{N+s},\ \mathrm{Pl}_s(A) = \frac{N \cdot \mathrm{Pl}(A) + s}{N+s}. \tag{8}$$

Here the *hyperparameter* $s > 0$ determines how quickly upper and lower probabilities of events converge as statistical data accumulate; N is the number of expert judgments. Smaller values of s produce faster convergence and stronger conclusions, whereas large values of s produce more cautious inferences. Walley [17] and Bernard [2] argue that the parameter s should be taken to be 1 or 2.

It should be noted that the simple modification of the belief and plausibility functions with using the IDM has a number of nice properties [14]. For example, if we have N identical estimates, then the belief and plausibility functions are the same $Bel(A) = \mathrm{Pl}(A) = 1$. This implies that the belief and plausibility functions do not depend on the value N while $Bel_s(A)$ and $\mathrm{Pl}_s(A)$ are $N/(N+s)$ and 1, respectively. However, the main advantage of the IDM is that it produces the cautious inference. In particular, if $N = 0$, then $Bel_s(A) = 0$ and $\mathrm{Pl}_s(A) = 1$. In the case $N \to \infty$, it can be stated for any s: $Bel_s(A) = Bel(A)$, $\mathrm{Pl}_s(A) = \mathrm{Pl}(A)$.

The extended belief and plausibility functions are obtained from the BPA's $m^*(A) = c/(N+s)$ for every A and the additional BPA $m^*(A \succeq A) = s/(N+s)$, i.e., $Bel_s(A)$ and $\mathrm{Pl}_s(A)$ can be obtained as standard belief and plausibility functions under condition that there are s additional observations $A = A \succeq A$. This fact allows us to change the BPA's of B_i and to make the cautious decision.

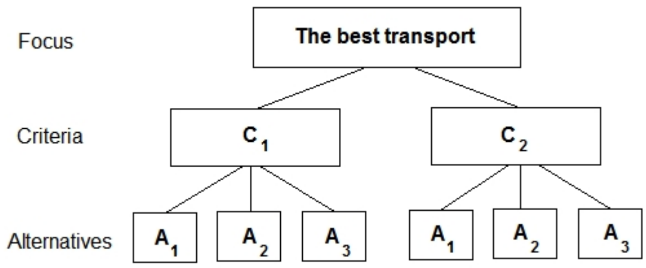

Figure 1: A hierarchical decision tree

7 Numerical Examples

Let us study a decision problem where the DM has to choose which one of three types of transport to use. Three alternatives (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). Here $A = \{A_1, A_2, A_3\}$ and $C = \{C_1, C_2\}$. A hierarchical decision tree of the problem is depicted in Fig.1.

Five experts provide the following preferences with respect to the first criterion:
two experts ($a_{16} = 2$): $\{A_1\} \succeq \{A_2A_3\} = \mathcal{B}_{16}$,
three experts ($a_{47} = 3$): $\{A_1A_2\} \succeq \{A_1A_2A_3\} = \mathcal{B}_{47}$.

The same experts provide the following preferences with respect to the second criterion:
one expert ($a_{52} = 1$): $\{A_1, A_3\} \succeq \{A_2\} = \mathcal{B}_{52}$,
three experts ($a_{37} = 3$): $\{A_3\} \succeq \{A_1, A_2, A_3\} = \mathcal{B}_{37}$,
one experts ($a_{13} = 1$): $\{A_1\} \succeq \{A_3\} = \mathcal{B}_{13}$.

Note that the condition

$$\{A_1A_2\} \succeq \{A_1A_2A_3\}$$

is equivalent to

$$\{A_1A_2\} \succeq \{A_3\},$$

and the preference

$$\{A_3\} \succeq \{A_1A_2A_3\}$$

is equivalent to

$$\{A_3\} \succeq \{A_1A_2\},$$

i.e.,

$$\mathcal{B}_{47} = \mathcal{B}_{43}$$

and

$$\mathcal{B}_{37} = \mathcal{B}_{34}.$$

The BPA's of all preferences are:

$$m(\mathcal{B}_{16} \mid C_1) = 0.4,\ m(\mathcal{B}_{43} \mid C_1) = 0.6,$$

$$m(\mathcal{B}_{52} \mid C_2) = 0.2,\ m(\mathcal{B}_{34} \mid C_2) = 0.6,\ m(\mathcal{B}_{13} \mid C_2) = 0.2.$$

Two DM's provide their preferences related to criteria:
The first DM ($c_{21} = 1$): $\{C_2\} \succeq \{C_1\} = \mathcal{D}_{21}$. The BPA of $\mathcal{D}_{21}$ is $m(\mathcal{D}_{21}) = 0.5$.

The second DM ($c_{33} = 1$) could not compare the criteria, i.e. the second DM provides preference $\mathcal{D}_{33} = C \succeq C$ and $m(\mathcal{D}_{33}) = 0.5$.

The belief and plausibility functions of preferences $D_k \succeq C$ are

$$Bel(D_1 \succeq C) = 0,$$

$$Pl(D_1 \succeq C) = m(\mathcal{D}_{33}) = 0.5,$$

$$Bel(D_2 \succeq C) = m(\mathcal{D}_{21}) = 0.5,$$
$$Pl(D_2 \succeq C) = m(\mathcal{D}_{21}) + m(\mathcal{D}_{33}) = 1,$$
$$Bel(D_3 \succeq C) = 1,$$
$$Pl(D_3 \succeq C) = 1.$$

Now we define the sets $\mathcal{L}_{ki}$ for every criterion.

Criterion C_1:

$$\mathcal{L}_{16} = \{\{A_1\} \succeq \{A_2\}, \{A_1\} \succeq \{A_3\}\},$$
$$\mathcal{L}_{43} = \{\{A_1\} \succeq \{A_3\}, \{A_2\} \succeq \{A_3\}\}.$$

Criterion C_2:

$$\mathcal{L}_{52} = \{\{A_1\} \succeq \{A_2\}, \{A_3\} \succeq \{A_2\}\},$$
$$\mathcal{L}_{34} = \{\{A_3\} \succeq \{A_1\}, \{A_3\} \succeq \{A_2\}\},$$
$$\mathcal{L}_{13} = \{\{A_1\} \succeq \{A_3\}\}.$$

By having the above information, we can write the combined belief and plausibility functions of preferences $A_k \succeq A$:

$$\mathrm{Bel}_p(B_1 \succeq A) = m(\mathcal{B}_{16} \mid C_1)p_1 + 0 \cdot p_2 = 0.4p_1 + 0 \cdot p_2,$$
$$\mathrm{Pl}_p(B_1 \succeq A) = 1p_1 + (1 - m(\mathcal{B}_{34} \mid C_2))\, p_2 = p_1 + 0.4p_2,$$
$$\mathrm{Bel}_p(B_2 \succeq A) = 0p_1 + 0p_2 = 0,$$
$$\mathrm{Pl}_p(B_2 \succeq A) = m(\mathcal{B}_{43} \mid C_1)p_1 + (1 - m(\mathcal{B}_{13} \mid C_2))\, p_2 = 0.6p_1 + 0.2p_2,$$
$$\mathrm{Bel}_p(B_3 \succeq A) = 0p_1 + m(\mathcal{B}_{34} \mid C_2)p_2 = 0.6p_2,$$
$$\mathrm{Pl}_p(B_3 \succeq A) = 0p_1 + (1 - m(\mathcal{B}_{13} \mid C_2) - m(\mathcal{B}_{52} \mid C_2))\, p_2 = 0.8p_2.$$

Every function above is the objective function for a linear programming problem with the following constraints:

$$0 \le p_1 \le 0.5,$$
$$0.5 \le p_2 \le 1,$$
$$p_1 + p_2 = 1.$$

After solving the corresponding linear programming problems, we get

$$Bel(B_1 \succeq A) = 0,\ \mathrm{Pl}(B_1 \succeq A) = 0.5 + 0.4 \cdot 0.5 = 0.7,$$
$$Bel(B_2 \succeq A) = 0,\ \mathrm{Pl}(B_2 \succeq A) = 0.6 \cdot 0.5 + 0.2 \cdot 0.5 = 0.4,$$
$$Bel(B_3 \succeq A) = 0.6 \cdot 0.5 = 0.3,\ \mathrm{Pl}(B_3 \succeq A) = 0.8 \cdot 1 = 0.8.$$

In the same way, we can compute the combined belief and plausibility functions of all preferences $B_k \succeq A$, but it is not necessary because our main aim is to choose one of the DA's. Finally, we can conclude that the third DA is the "best" one.

It should be noted that the number of experts and DM's is rather small. Therefore, we apply the IDM for making the cautious decision. The modified BPA's by $s = 1$ are

$$m^*(\mathcal{B}_{16} \mid C_1) = 0.333,\ m^*(\mathcal{B}_{43} \mid C_1) = 0.5,\ m^*(\mathcal{B}_{77} \mid C_1) = 0.167,$$
$$m^*(\mathcal{B}_{52} \mid C_2) = 0.167,\ m^*(\mathcal{B}_{34} \mid C_2) = 0.5,$$
$$m^*(\mathcal{B}_{13} \mid C_2) = 0.167,\ m^*(\mathcal{B}_{77} \mid C_2) = 0.166.$$

Here

$$\mathcal{B}_{77}^{(1)} = \mathcal{B}_{77}^{(2)} = A \succeq A.$$

The belief and plausibility functions of preferences $D_k \succeq C$ are

$$Bel_1(D_1 \succeq C) = 0,$$

$$Pl_1(D_1 \succeq C) = m(\mathcal{D}_{33}) = 0.666,$$

$$Bel_1(D_2 \succeq C) = m(\mathcal{D}_{21}) = 0.333,$$

$$Pl_1(D_2 \succeq C) = m(\mathcal{D}_{21}) + m(\mathcal{D}_{33}) = 1,$$

$$Bel_1(D_3 \succeq C) = 0.666,$$

$$Pl_1(D_3 \succeq C) = 1.$$

The combined belief and plausibility functions of preferences $A_k \succeq A$ are:

$$Bel_{p,1}(B_1 \succeq A) = 0.333p_1 + 0p_2,$$

$$Pl_{p,1}(B_1 \succeq A) = p_1 + 0.5p_2,$$

$$Bel_{p,1}(B_2 \succeq A) = 0p_1 + 0p_2 = 0,$$

$$Pl_{p,1}(B_2 \succeq A) = 0.667p_1 + 0.333p_2,$$

$$Bel_{p,1}(B_3 \succeq A) = 0.5p_2,$$

$$Pl_{p,1}(B_3 \succeq A) = 0p_1 + 0.833p_2.$$

Note that the above belief and plausibility functions can be obtained without using the modified BPA's, but by means of (8). Every function above is the objective function for a linear programming problem with the following constraints:

$$0 \leq p_1 \leq 0.666,$$

$$0.333 \leq p_2 \leq 1,$$

$$p_1 + p_2 = 1.$$

Hence

$$Bel_1(B_1 \succeq A) = 0, Pl_1(B_1 \succeq A) = 0.833,$$

$$Bel_1(B_2 \succeq A) = 0, Pl_1(B_2 \succeq A) = 0.555,$$

$$Bel_1(B_3 \succeq A) = 0.167, Pl_1(B_3 \succeq A) = 0.833.$$

It can be seen from the results that the fact that the third DA is optimal is not so clear here. This implies that additional judgments are required for reducing the risk of decision making.

Let us consider an extreme case when the DM can not to choose a preferable criterion or a group of criteria. In this case, we have $m(\mathcal{D}_{ki}) = 0$ for all k and i except for $(k, i) = (3, 3)$, $m(\mathcal{D}_{33}) = 1$. Hence, the set of constraints for the distribution p is reduced to one constraint $p_1 + p_2 = 1$. This leads to the pessimistic decision. In particular, by taking $s = 0$, we get the following belief and plausibility functions:

$$Bel(B_1 \succeq A) = 0,\ Pl(B_1 \succeq A) = 1,$$

$$Bel(B_2 \succeq A) = 0,\ Pl(B_2 \succeq A) = 0.6,$$

$$Bel(B_3 \succeq A) = 0,\ Pl(B_3 \succeq A) = 0.8.$$

It can be seen from the above results that the first DA is optimal. If $s = 1$, then there hold

$$Bel_1(B_1 \succeq A) = 0,\ Pl_1(B_1 \succeq A) = 1,$$

$$Bel_1(B_2 \succeq A) = 0,\ Pl_1(B_2 \succeq A) = 0.667,$$

$$Bel_1(B_3 \succeq A) = 0,\ Pl_1(B_3 \succeq A) = 0.833.$$

We again have the same optimal DA.

8 Conclusion

The method for solving a multi-criteria decision problem in the frameworks of the AHP and DST has been proposed in the paper. The method can be regarded as a generalization of the DS/AHP method. It significantly extends the procedure of the expert elicitation in the AHP. The proposed method does not require hard computations. Moreover, the optimization problem for computing the lower belief and upper plausibility functions is linear and can simply be solved by means of well-known methods, for instance, by means of the simplex method.

It should be noted that the method is not changed if experts provide not only preferences, but also some rates of preferences by using a m-point scale $(1-m)$ for the pairwise comparisons. 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 was not chosen by experts) for DA's, then the BPA of the preference $\mathcal{B}_{ki}$ is computed as

$$m(\mathcal{B}_{ki}) = \sum_{j=1}^{a_{ki}} x_{ki}^{(j)} \; / \; N,$$

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

The same procedure can be applied to criteria.

At the same time, we have to point out that there are many combination rules used in DST in addition to the linear rule used in the paper and based on the total probability theorem. However, these rules can lead to the non-linear optimization problems whose solution might be complicated. Therefore, algorithms for using other combination rules and their comparative analysis are questions for further research.

References

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

[2] Bernard, J.M., An introduction to the imprecise Dirichlet model for multinomial data, *International Journal of Approximate Reasoning*, vol.39, nos.2-5, pp.123–150, 2005.

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

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

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

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

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

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

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

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

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

[12] Tervonen, T., 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, 2004.

[13] Utkin, L.V., Probabilities of judgments provided by unknown experts by using the imprecise Dirichlet model, *Risk, Decision and Policy*, vol.9, no.4, pp.391–400, 2004.

[14] Utkin, L.V., Extensions of belief functions and possibility distributions by using the imprecise Dirichlet model, *Fuzzy Sets and Systems*, vol.154, no.3, pp.413–431, 2005.

[15] Utkin, L.V., Ranking procedures by pairwise comparison using random sets and the imprecise Dirichlet model, *Applied Mathematics and Computations*, vol.183, no.1, pp.394–408, 2006.

[16] Utkin, L.V., and Th. Augustin, Decision making under incomplete data using the imprecise Dirichlet model, *International Journal of Approximate Reasoning*, vol.44, no.3, pp.322–338, 2007.

[17] Walley, P., Inferences from multinomial data: Learning about a bag of marbles, *Journal of the Royal Statistical Society, Series B*, vol.58, pp.3–57, 1996.

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