Belief function combination and conflict management

Eric Lefevre, Olivier Colot, Patrick Vannoorenberghe

To cite this version:

Eric Lefevre, Olivier Colot, Patrick Vannoorenberghe. Belief function combination and conflict management. Information Fusion, 2002, 3 (2), pp.149-162. 10.1016/S1566-2535(02)00053-2 . hal-03354198

HAL Id: hal-03354198

https://hal.science/hal-03354198

Submitted on 24 Sep 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Belief Functions Combination and Conflict Management

E. Lefevre[1], O. Colot[1] and P. Vannoorenberghe[2]

[1] *Laboratoire **P**erception **S**ystèmes **I**nformation (PSI), UPRES EA 2120*
Université & INSA de Rouen
Place Emile Blondel, 76131 Mont-Saint-Aignan Cedex
Tél: +33.(0)2.35.52.84.05, Fax:+33.(0)2.35.52.84.83

[2] *Laboratoire HEUDIASYC, UMR 6599 CNRS*
Université de Technologie de Compiègne
BP 20529, 60205 Compiègne Cedex
Tél: +33.(0)3.44.23.49.53, Fax:+33.(0)3.44.23.44.77
{Eric.Lefevre, Olivier.Colot}@insa-rouen.fr; Patrick.Vannoorenberghe@hds.utc.fr

Abstract

Within the framework of evidence theory, data fusion consists in obtaining a single belief function by the combination of several belief functions resulting from distinct information sources. The most popular rule of combination, called Dempster's rule of combination (or the orthogonal sum), has several interesting mathematical properties such as commutativity or associativity. However, combining belief functions with this operator implies normalizing the results by scaling them proportionally to the conflicting mass in order to keep some basic properties. Although this normalization seems logical, several authors have criticized it and some have proposed other solutions. In particular, Dempster's combination operator is a poor solution for the management of the conflict between the various information sources at the normalization step. Conflict management is a major problem especially during the fusion of many information sources. Indeed, the conflict increases with the number of information sources. That is why a strategy for re-assigning the conflicting mass is essential. In this paper, we define a formalism to describe a family of combination operators. So, we propose to develop a generic framework in order to unify several classical rules of combination. We also propose other combination rules allowing an arbitrary or adapted assignment of the conflicting mass to subsets.

Key words: Data Fusion, Theory of Evidence, Rules of Combination, Conflict.

1 Introduction

Information fusion has been the object of much research over the last few years [1–11]. Generally, it is based on the confidence measure theory (possibility theory, evidence theory, probability theory and fuzzy set theory) and has the advantage of:

- using redundant information,
- using the complementarity of the available information,
- achieving more reliable information,
- improving the decision making.

Data fusion is used in many application fields, such as multi sensor fusion [12,13], image processing and analysis [4–7,11,14,15], classification [16–18] or target tracking [19]. It takes into account heterogeneous information (numerical or symbolic) which is often imperfect (imprecise, uncertain and incomplete) and modeled by means of sources which have to be combined or aggregated. In the framework of evidence theory, information fusion relies on the use of a combination rule allowing the belief functions for the different propositions to be combined. The basic rule of combination is Dempster's rule of combination (orthogonal sum). It needs a normalization step in order to preserve the basic properties of the belief functions. In [20], Zadeh has underlined that this normalization involves counter-intuitive behaviours. In order to solve the problem of conflict management, R. Yager [21], D. Dubois [22] and Ph. Smets [23] and more recently C. Murphy [24] have proposed other combination rules. However, these rules have more or less satisfactory behaviours. In particular, Dubois' rule or Yager's rule of combination hold that the conflicting mass must be distributed over all all subsets. Smets proposes that the conflicting mass results from the non-exhaustivity of the frame of discernment. We propose another approach, in which we define a generalized framework for the fusion of information sources by means of a generic axiomatic. This framework enables a large family of combination rules to be obtained.

This paper is organized as follows. The basic concepts of evidence theory are first briefly introduced (Section 2) including the problem of conflict in Dempster's rule of combination. In section 3, we define the generic framework allowing classical combination operators to be unified and we propose a family of new combination rules. Finally, some methods to determine weighting factors for the conflicting mass distribution process for each proposition implied in the conflict are proposed (section 3.3.2). Tests are given in section 4.

2 Background

Evidence theory is initially based on Dempster's work [25] concerning lower and upper probability distribution families. From these mathematical foundations, Shafer [26] has shown the ability of the belief functions to modelize uncertain knowledge. The usefulness of belief functions, as an alternative to subjective probabilities, was later demonstrated axiomatically by Smets [27,28] with the *Transferable Belief Model* (TBM) giving a clear and coherent interpretation of the underlying concept of the theory.

2.1 Knowledge model

Evidence theory first supposes the definition of a set of hypotheses Θ called the frame of discernment, defined as follows:

$$\Theta = \{H_1, \ldots, H_n, \ldots, H_N\}. \tag{1}$$

It is composed of N exhaustive and exclusive hypotheses. From the frame of discernment Θ, let us denote 2^Θ, the power set composed with the 2^N propositions A of Θ:

$$2^\Theta = \{\emptyset, \{H_1\}, \{H_2\}, \ldots, \{H_N\}, \{H_1 \cup H_2\}, \{H_1 \cup H_3\}, \ldots, \Theta\}. \tag{2}$$

A key point of evidence theory is the *basic belief assignment* (bba). The mass of belief in an element of Θ is quite similar to a probability distribution, but differs by the fact that the unit mass is distributed among the elements of 2^Θ, that is to say not only on the singletons H_n in Θ but on composite hypotheses too. The belief m_j assigned to an information source S_j is thus defined by:

$$m_j : 2^\Theta \rightarrow [0, 1]. \tag{3}$$

This function verifies the following properties:

$$m_j(\emptyset) = 0, \tag{4}$$

$$\sum_{A \subseteq \Theta} m_j(A) = 1. \tag{5}$$

The mass $m_j(A)$ represents how strongly the evidence supports A which, in the case of a disjunction of hypotheses, has not been assigned to a subset of A because of insufficient information. This mass can be re-assigned more precisely to the subsets of A if additional information is available. Each subset $A \subseteq \Theta$ such as $m_j(A) > 0$ is called a focal element of m. Let us denote $\mathcal{F}_j$ the set of the focal elements associated to a belief function m_j. From this bba, a belief function Bel_j and a plausibility function Pl_j are defined, respectively, as:

$$Bel_j(A) = \sum_{B \subseteq A} m_j(B) \tag{6}$$

and

$$Pl_j(A) = \sum_{A \cap B \neq \emptyset} m_j(B). \tag{7}$$

The quantity $Bel_j(A)$ can be interpreted as a measure of one's belief that hypothesis A is true. The plausibility $Pl_j(A)$ can be viewed as the total amount of belief that could be potentially placed in A. Note that functions m_j, Bel_j and Pl_j are in one-to-one correspondence [26], and can be seen as three facets of the same piece of information.

In evidence theory, one of the main difficulties lies in modelling the knowledge of the problem by initializing the belief functions m_j as well as possible. Generally, the model depends on the application. In [29], A. Appriou proposes two models in order to manage the uncertain learning in the framework of evidence theory. These models are consistent with the Bayesian approach when the belief mass is only allocated to singletons. Other models, also based on likelihood functions, have been proposed [30–32]. Another method based on the use of a a neighbourhood information was introduced by T. Denœux [17,18,33,34].

2.2 Dempster's rule of combination

In the case of imperfect data (uncertain, imprecise and incomplete), fusion is an interesting solution to obtain more relevant information. Evidence theory offers appropriate aggregation tools. From the basic belief assignment denoted m_j obtained for each information source S_j, it is possible to use a combination rule in order to provide combined masses synthesizing the knowledge of the different sources. These belief masses can then be used by a decision process with the benefit of the whole knowledge contained in the belief functions given by each source.

Dempster's rule of combination [26] is the first one defined within the framework of evidence theory. Using the rule implies that the independence condition for the sources to be combined must be respected. Dempster's rule of combination, the so called orthogonal sum, is commutative and associative. Let us denote $m_\oplus$, the belief function resulting from the combination of J information sources S_j defined as:

$$m_\oplus = m_1 \oplus \ldots \oplus m_j \ldots \oplus m_J \tag{8}$$

where $\oplus$ represents the operator of combination. With two information sources

S_1 and S_2, the combination rule is defined as:

$$m_{\oplus}(A) = \frac{m_{\cap}(A)}{1 - m(\emptyset)} \qquad \forall A \subseteq \Theta \tag{9}$$

where $m_{\cap}$ corresponds to the conjunctive rule of combination defined by:

$$m_{\cap}(A) = \sum_{B \cap C = A} m_1(B).m_2(C) \qquad \forall A \subseteq \Theta \tag{10}$$

and where the mass $m(\emptyset)$ assigned to the empty set is defined by:

$$m(\emptyset) = \sum_{B \cap C = \emptyset} m_1(B).m_2(C). \tag{11}$$

In equations (9) and (11), the coefficient $m(\emptyset)$ reflects the conflict between the two sources S_1 and S_2. Assuming the normality of the bba's ($m(\emptyset) = 0$), the use of this rule is possible only if m_1 and m_2 are not totally conflicting, i.e., if there exist two focal elements B and C of m_1 and m_2 satisfying $B \cap C \neq \emptyset$. This rule verifies some interesting properties and its use has been justified theoretically by several authors [35–37] according to specific axioms. However, in some situations, this operator cannot be used. It is the case when:

- the independence constraint [38–41] of the information sources is not respected. Indeed, the combination is not idempotent and its use would reinforce abusively the propositions supported by the bba,
- the sources are not perfectly reliable and when the mass function model is also imprecise, a conflict $m(\emptyset)$ appears. The normalization coefficient depends on this conflict and so induces a combination rule sensitivity to small imprecisions of the mass functions as Zadeh proved [20]. We give some illustrations of such a behaviour in the subsection *4.1.*

2.3 Conflict origins and solutions

Conflict management in belief functions has been already studied in the past. But why does the evidence conflict? In *2.3.1*, we present the main origins of the conflict. Some classical solutions are given in *2.3.2.*

2.3.1 Origins: why evidence conflicts.

There are three main reasons why a conflict appears when combining evidence. The first one corresponds to an aberrant measurement given by a sensor. In fact, an abnormal measurement (denoted by outliers in pattern recognition applications) can generate a conflicting mass $m(\emptyset)$ during the combination step. This is often due to:

- a sensor defect during the acquisition step,
- a poor calibration of the sensor during the learning phase. If the sensor has a correct behaviour, this situation can correspond to a non-exhaustive frame of discernment (an unknown class for example).

A second reason relies on the belief function model. Thus, imprecise model of the belief functions may provide a conflict. In fact, most of the models for determining basic belief assignments are derived from neighbourhood information according to a distance [33] or to likelihood functions [26,29]. An inappropriate choice of the metric in the distance-based approaches or a poor estimation of the likelihood functions for the likelihood-based methods can induce variations in the belief functions. Consequently, these variations provide a conflicting mass during the combination.

Finally, when the information sources to be aggregated are numerous, a conflicting mass can be induced even if these sources agree. For example, let us consider a set of J information sources with the following basic belief assignments:

$$m_j(\{H_1\}) = 0.80 \qquad m_j(\{H_2\}) = 0.15 \qquad m_j(\Theta) = 0.05.$$

According to these belief assignments, we can note that the majority of the belief supports the hypothesis H_1. Figure Fig.1 shows the conflicting mass evolution according to the number of information sources to be aggregated. This figure shows that the conflicting mass is approximately 25% when 2 sources ($J = 2$) are combined and this mass is close to 80% for 10 aggregated sources ($J = 10$)! These three main reasons plead for an adaptive distribution or assignment of the conflicting mass provided by the combining process.

2.3.2 Solutions

Several rules of combination have been introduced in order to manage the conflict problem. These solutions can be divided into two main categories corresponding to two strategies for the conflict distribution. The first one includes rules of combination which require reliable information sources (see 2.3.2.1), conjunctive operators are used [25,30]. The second family states that one information source tells the truth but without knowing exactly which of them it is (see 2.3.2.2). For this second category, the operators have conjunctive and disjunctive behaviours [21,22].

2.3.2.1 Combination of reliable sources - As Dempster postulates, Smets supposes that all the information sources are reliable. The idea is that the conflict can only come from a bad definition of the frame of discernment (ill-conditioned frame of discernment). In this case, Smets keeps the conflicting

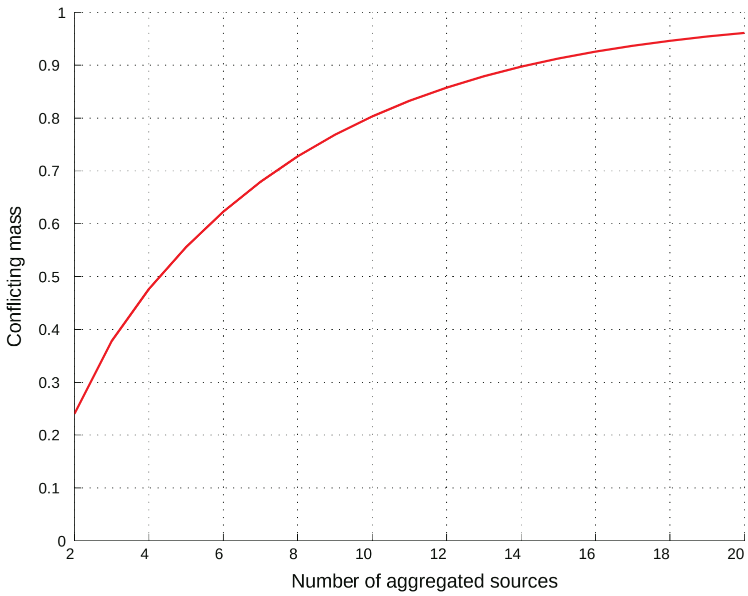

Fig. 1. Conflict vs. number of sources to combine.

mass $m(\emptyset)$ and does not use it for normalization. Thus, $\emptyset$ can be interpreted as one or several hypotheses which are not taken into account in the initial frame of discernment. The rule of combination proposed by Smets is thus defined by:

$$\begin{cases} m_S(A) = m_\cap(A) & \forall A \subseteq \Theta \\ m_S(\emptyset) = m(\emptyset). \end{cases} \tag{12}$$

Note that a similar approach is proposed by Yager in [21] which rests on the introduction of a new hypothesis in the frame of discernment. The conflicting mass is then given to this new hypothesis. These operators have conjunctive behaviours.

2.3.2.2 Combination of non-reliable sources - The conflicting mass can be provided by non-reliable information sources. This point of view has been introduced by Yager [21] and by Dubois and Prade [22]. Yager postulates that the frame of discernment is exhaustive (closed-world assumption). Yager's idea consists in assigning the conflicting mass $m(\emptyset)$ to the whole set Θ. The resulting mass, denoted m_Y, for the combination of two information sources S_1 and S_2 is obtained with the following equations:

$$\begin{cases} m_Y(A) = m_\cap(A) & \forall A \subset \Theta \\ m_Y(\Theta) = m_\cap(\Theta) + m(\emptyset). \end{cases} \tag{13}$$

The combination operator proposed by Dubois and Prade [22] can be explained as follows. Assume the source S_1 supports the subset B with a mass of belief

$m_1(B)$ and the source S_2 supports C with a mass of belief $m_2(C)$. When the intersection of subsets B and C is empty, the minimum specificity principle can be applied. According to this principle, the resulting mass $m_1(B).m_2(C)$ is then assigned to the subset $B \cup C$. The rule of combination proposed by Dubois and Prade is then defined by:

$$m_D(A) = m_\cap(A) + \sum_{\substack{B \cup C = A \\ B \cap C = \emptyset}} m_1(B).m_2(C) \qquad \forall A \subseteq \Theta. \tag{14}$$

This rule of combination is better adapted and more specific than Yager's rule of combination concerning the assignment of the conflicting mass.

Another way to solve the conflict in the case of non reliable sources is to use discounting coefficients in the model. So, let m_j be a belief mass given by the source S_j and let α_j be a coefficient which represents the confidence degree one has in source S_j. Let us denote $m_{\alpha_j,j}$ the belief mass m_j discounted by a coefficient $(1 - \alpha_j)$ and defined as:

$$\begin{cases} m_{\alpha_j,j}(A) = \alpha_j m_j(A) \qquad \forall A \subset \Theta \\ m_{\alpha_j,j}(\Theta) = 1 - \alpha_j + \alpha_j m_j(\Theta). \end{cases} \tag{15}$$

What does it mean? The value assigned to α_j leads to different interpretations:

- $\alpha_j = 0$ means a complete calling in question of the reliability of the source S_j,
- $\alpha_j = 1$ means a total confidence in S_j.

When we are full confident in the reliability of S_j, the information provided by this source is not supposed to generate any conflict when combined with the information given by the other sources. The coefficient α_j is then equal to 1 and the belief function is thus not modified. Conversely, if one supposes that one source S_j is not reliable, it may provide conflicting information when it is combined with other sources. By introducing a coefficient $\alpha_j = 0$, the belief function m_j associated to the source S_j becomes a belief function of total ignorance ($m_{\alpha_j,j}(\Theta) = 1$) and so a neutral element for Dempster's rule of combination. So, discounting is useful for managing the source influences according to their reliability before aggregation. Several methods have been developed in order to define the discounting coefficients [42,43]. The main problem with this kind of solution in conflict management is finding an appropriate technique for tuning the discounting coefficients properly.

2.3.3 What to do with the conflict?

There are several strategies for solving or managing the conflict. In practice, the main question is: "What to do with the conflict?". There is no single answer to this question. Different solutions are at our disposal:

- the information sources to be combined are perfectly reliable, so we can use either Dempster's rule of combination or Smets' rule of combination if we are not sure that the frame of discernment is exhaustive,
- the information sources to be combined are not reliable, so we must apply discounting if possible, or use one of the disjunctive rules of combination.

Are these combination rules the only ones possible? In the next section, we propose a generic formalism for the combination allowing the rules of combination cited previously to be retrieved and allowing others to be derived.

3 Generic framework

We propose a generic framework in order to unify the classical combination operators. Furthermore, this framework allows others rules of combination for the assignment of the conflicting mass to be defined. The idea defended here, is to assign the conflicting mass with weighting factors on the non-concordant hypotheses or possibly on the composite hypotheses (disjunctions). These weighting factors can be defined by means of expert knowledge or by means of cost functions. We focus on the problem of looking for a relevant weighted assignment of the conflicting mass on subsets A.

3.1 Presentation

The aim of these combination rules is to distribute the conflicting mass $m(\emptyset)$ on a set of propositions denoted $\mathcal{P}$. Part of the mass $m(\emptyset)$ is assigned to each subset $A \subseteq \mathcal{P}$ according to a weighting factor called $w(A, \mathbf{m})$ with $\mathbf{m} = \{m_1, \ldots, m_j, \ldots, m_J\}$. This weighting factor can be a function of the considered subset A and subsets which have caused the conflict. So, the total mass after aggregation for a subset A is the sum of two masses and is expressed as follows:

$$m(A) = m_{\cap}(A) + m^c(A) \qquad \forall A \subseteq \Theta. \tag{16}$$

In equation (16), the first term, $m_{\cap}(A)$, corresponds to the conjunctive rule of combination. The second one, denoted $m^c(A)$, is the part of the conflicting mass assigned to the subset A. It can be written as follows:

$$\begin{cases} m^c(A) = w(A, \mathbf{m}).m(\emptyset) & \forall\ A \subseteq \mathcal{P} \\ m^c(A) = 0 \quad \text{otherwise} \end{cases} \tag{17}$$

with the following constraint:

$$\sum_{A \subseteq \mathcal{P}} w(A, \mathbf{m}) = 1 \tag{18}$$

so as to respect the property that the sum of mass functions must be equal to 1 (see equation (5)).

This generic framework allows Dempster's rule of combination and other proposed by Smets [23], Yager [21] and Dubois and Prade [22] to be rewritten. For each operator, we need only to define the set $\mathcal{P}$ on which the conflicting mass must be distributed and the weighting factors $w(A, \mathbf{m})$ associated to each subset $A \subseteq \mathcal{P}$. Three ways to obtain the weighting factors are possible:

- fixing the values of the weighting factors,
- computing the weighting factors,
- learning the weighting factors.

The first way allows Smets' rule of combination and Yager's rule of combination to be retrieved (see 3.2). The second one allows Dempster's rule of combination and Dubois and Prade's rule of combination to be retrieved (see 3.3). Finally, the third approach is new and allows an adapted conflicting mass assignment to be achieved (see 3.4).

3.2 Weighting factors with fixed values

3.2.1 Classical rules of combination

According to the equations presented in 2.3.2.1, Smets' rule can be defined as follows. The set on which the conflicting mass is distributed is the empty set and so we obtain:

$$\mathcal{P} = \{\emptyset\} \tag{19}$$

and the weighting factor associated to the empty set is equal to 1:

$$w(\emptyset, \mathbf{m}) = 1. \tag{20}$$

The empty set can be viewed as a reject class. A similar approach based on the introduction of a new hypothesis in the frame of discernment was proposed in [44]. The aggregation operator proposed by Smets verifies the properties of commutativity and associativity. Finally, let us emphasize that in [45], Smets defines the α-junctions as a unified framework for purely conjunctive combination operators and for disjunctive combination operators.

The method proposed by Yager [21] can be defined as follows. Indeed, considering that at least one of the sources concerned by the fusion process is reliable but without knowing which one, Yager proposes to assign the mass of conflict to the set Θ. According to the generic framework previously presented, we obtain a set $\mathcal{P}$ given by:

$$\mathcal{P} = \{\Theta\}. \tag{21}$$

The weighting factor $w(\Theta, \mathbf{m})$ associated to this set is equal to 1. The conflicting mass is thus placed on Θ. This method involves the separation of the whole conflicting mass and furthermore, implies that it participates in the decision process for distinguishing the hypotheses. This rule of combination is commutative but not associative. It is therefore necessary to define an order for the fusion process.

3.2.2 Conflict distribution based on an expert valuation

The proposed formalism can be useful in the case of additional knowledge given by an expert specialized in the application. In medical fields, target tracking or obstacle detection, non-detection can have important consequences in decision making. In these kinds of application, the conflicting mass can be assigned to the most cautious hypothesis. As an example, let us consider an obstacle detection system equipped with two distance sensors placed at the front of a car. Suppose a measurement is taken and that the information from the 2 sensors is in conflict (the first sensor says that an obstacle is 1 metre from the car and the second one says it is 10 metres away). It is advisable in this case to single out the information which gives the smaller distance in order to avoid putting the driver's life at risk. Thus, the expert can decide to allocate, by fixing the weighting factors, most of the conflicting mass to one of the subsets. When no additional information can be provided by an expert, one can adopt a cautious strategy consisting in distributing the conflict uniformly or by learning the weighting factors as in the subsection 3.4.

3.3 Computed weighting factors

3.3.1 Classical rules of combination

Within the proposed generic framework, Dempster's rule of combination is defined as follows. The set on which the conflicting mass is distributed is Θ, so:

$$\mathcal{P} = \Theta \tag{22}$$

and the associated weighting factors are defined as follows:

$$w(A, \mathbf{m}) = \frac{m_{\cap}(A)}{1 - m(\emptyset)} \qquad \forall A \subseteq \Theta. \tag{23}$$

So, the rule of combination is then expressed as follows:

$$m_{\oplus}(A) = m_{\cap}(A) + \frac{m_{\cap}(A).m(\emptyset)}{1 - m(\emptyset)} \qquad \forall A \subseteq \Theta. \tag{24}$$

It can easily be seen that it is similar to equation (9).

The distribution of the conflicting mass proposed by Dubois and Prade can be considered in the proposed formalism. In order to describe this rule of combination, we introduce the notion of *partial conflicting mass*. Each information source S_j, with $j \in \{1, \dots, J\}$, gives a degree of belief to each focal element belonging to $\mathcal{F}_j$. When the focal elements are compatible, that is to say when the intersections between these subsets in $\mathcal{F}_j$ are not empty, the mass product assigned to these sets is assigned to the intersection. If the propositions are incompatible, that is to say when their intersection is equal to the empty set, a partial conflict denoted m^* appears. It is expressed as follows:

$$m^* = m_1(A_1) \times m_2(A_2) \times \dots m_J(A_J) \quad \text{with} \quad A_1 \cap A_2 \cap \dots \cap A_J = \emptyset. \tag{25}$$

The total conflict $m(\emptyset)$ is the sum of the partial conflicts and is expressed as follows:

$$m(\emptyset) = \sum\nolimits^* m^* \tag{26}$$

where $\sum^*$ is a countable sum which depends on the focal elements of $\mathcal{F}_j$. So, with this formalism, we are able to write the combination principle for two sources as it follows.
Let S_1 be a source which supports subset $A_1 \subseteq \Theta$ with the belief mass $m_1(A_1)$ and let S_2 be a source which supports the subset $A_2 \subseteq \Theta$ with a belief mass $m_2(A_2)$. If the proposition A_1 is in contradiction with A_2, that is to say if $A_1 \cap A_2 = \emptyset$, although it is impossible to decide between the sources, then one of the two propositions must be true. The partial conflicting mass m^* defined by:

$$m^* = m_1(A_1).m_2(A_2) \tag{27}$$

is then assigned to the proposition $A_1 \cup A_2$. In the general case of this kind of combination, we have a proposition A on which the partial conflicting masses are assigned. The set of all the subsets on which the conflicting mass is distributed is defined by:

$$\mathcal{P} = \{A \subseteq \Theta \backslash \exists A_1 \in \mathcal{F}_1, \exists A_2 \in \mathcal{F}_2, A = A_1 \cup A_2 \quad \text{and} \quad A_1 \cap A_2 = \emptyset\}. \tag{28}$$

Part of the conflicting mass is assigned to the subset $A \subseteq \mathcal{P}$ by means of a weighting factor $w(A, \mathbf{m})$ with $\mathbf{m} = \{m_1, m_2\}$. This weighting factor, in the case of the operator of combination considered, is expressed as follows:

$$\forall A \subseteq \mathcal{P} \qquad w(A, \mathbf{m}) = \frac{\sum\limits_{\substack{A_1, A_2 \backslash A_1 \cup A_2 = A \\ A_1 \cap A_2 = \emptyset}} m^*}{m(\emptyset)}. \tag{29}$$

We can observe that the computation of the weighting factors does not depend exclusively on propositions with which they are associated, but depends on belief mass functions which have cause the partial conflicts. The belief mass functions leading to the conflict allow us to compute that part of the conflicting mass which must be assigned to the subsets in $\mathcal{P}$. We can note that this rule of combination uses a conjunctive approach when the sources agree and a disjunctive approach when evidence conflicts. Like Yager's rule of combination, Dubois and Prade's rule of combination is commutative but is not associative.

3.3.2 Other solutions for weighting factor computation

We have seen that the proposed formalism allowed some of the classical operators of the literature to be retrieved. On the basis of $\mathcal{P}$ and the associated weighting factors $w(A, \mathbf{m})$ for $A \subseteq \mathcal{P}$, we can derive other operators. In [46], two particular operators have been presented. We have seen above (cf. 2.3.2) that we can manage the conflict by means of discounting. We can also obtain relationships between the weighting factors $w(., \mathbf{m})$ and the discounting factors (see the appendix). This demonstrates that conflict management by means of discounting is just a particular case of the redistribution of the conflicting mass by means of computed weighting factors.

3.4 Automatic learning of the weighting factors

We propose here another way to manage conflicting mass distribution. It is based on an automatic learning of the weighting factors involved in the assignment of the conflicting mass. Before dealing with this approach, let us underline a function defined by Smets for decision making in the framework of evidence theory.

Smets [28,47] defines a particular probability distribution function, called the pignistic probability. It is obtained by distributing the belief mass $m(A)$ equally between the different elements of A. So, we have:

$$\forall H_n \in \Theta \qquad BetP(H_n) = \sum_{A \subseteq \Theta} \frac{|H_n \cap A|}{|A|}.m(A) \tag{30}$$

where $|.|$ represents the cardinal of the considered set.

We propose a learning of the weighting factors based on the use of training set and the minimization of an error criterion. This error criterion is defined by the mean square error between the pignistic probability $BetP$ computed according to both equation (30) and the membership indicator. The mean square error E_{MS} of the training set vectors is defined as follows:

$$E_{MS} = \sum_{i=1}^{I} \sum_{n=1}^{N} [BetP^{(i)}(H_n) - u_n^i]^2 \tag{31}$$

where I is the number of elements, $BetP^{(i)}$ represents the pignistic probability of a vector X_i in the learning set and $u_n^i \in \{0, 1\}$ is the membership indicator of the vector X_i to the hypothesis H_n. For example, $u_s^i = 1$ if the vector X_i belongs to the class H_s, and $u_n^i = 0$ for all $n \neq s$. We then determine each weighting factor $w(A, \mathbf{m})$ for $A \subseteq \mathcal{P}$ by minimizing the criterion given in equation (31).

In other words, one has I cases in the training set for which one knows the class and the basic belief assignments provided by J sources. For each case, one can combine the J basic belief assignments with re-assignment of the conflicting mass $m(\emptyset)$ by means of weighting factors $w(., \mathbf{m})$ to obtain a new mass m. Having obtained the belief assignment after combination, one computes the pignistic probability $BetP$. The final step consists in comparing $BetP$ to the truth. The weighting factors $w(., \mathbf{m})$ can be obtained by using a technique such as the gradient descent.

4 Results

We present here several results which describe the behaviour of the weighted combination strategies for the conflicting mass distribution relative to the classical rules of combination. First, a comparison between these rules and our strategy in terms of the resulting belief mass interpretation is proposed (subsection 4.1). Finally, results in the field of pattern recognition with weighting factor learning are presented in subsection 4.2.

4.1 Combination rules and conflicting mass assignment

Let us consider the following well-known example. Suppose a murder case with three suspects H_1, H_2 and H_3 and such that $\Theta = \{H_1, H_2, H_3\}$ is the frame of discernment. Let S_1 and S_2 be two witnesses who are two information sources each providing two basic belief assignments m_1 and m_2 defined respectively as:

$$\begin{aligned} m_1(\{H_1\}) &= \epsilon & m_2(\{H_1\}) &= 1 - k - \epsilon \\ m_1(\{H_2\}) &= k & m_2(\{H_2\}) &= k \\ m_1(\{H_3\}) &= 1 - k - \epsilon & m_2(\{H_3\}) &= \epsilon \end{aligned} \tag{32}$$

with $0 \leq k \leq 1$. These masses represent the degrees of belief of each witness about who might be the murderer. We present below some results given by the rules of combination according to the value of ϵ (see Table 1 up to Table 3). In this test, k is equal to 0.1. The weighting factors associated to each rule of combination are also specified.

Consider, first, that the information sources are totally reliable. Table 1 gives the results obtained with Dempster's rule of combination and Smets' rule of combination for different values for ϵ. For Dempster's rule of combination, the belief mass assigned to H_2 increases when ϵ decreases. If the conflicting mass $m(\emptyset)$ is not analyzed, H_2 is chosen by the decision making process. With Smets' rule of combination, the belief masses are weak and do not allow reliable decision making.

Consider now that the information sources are not reliable. The results obtained for the different strategies are presented in Table 2 and Table 3.
The first strategy consists in discounting the two sources S_1 and S_2 according to the following discounting factor $\alpha_1 = 0.2$ and $\alpha_2 = 0.8$, and then using Dempster's rule of combination. In this case, the source S_2 is supposed to be telling the truth and part of the unit mass is assigned to Θ which represents the amount of uncertainty. The two other strategies are Yager's rule of combination and Dubois and Prade's rule of combination.
For the first strategy, the belief mass is a little modified according to ϵ because the reliability coefficient of the source S_1 is weak, so the decision relies mainly on the source S_2. The weighting factors corresponding to this strategy are little modified with respect to the evolution of ϵ and show that most of the conflicting mass is assigned to H_1.
For Yager's rule of combination, the belief mass assigned to Θ increases when ϵ decreases. This is illustrated by the fact that the credibilities are weak and the plausibilities are high. The belief function is weakly specific leading to unreliable decision making.
For Dubois and Prade's rule of combination, the plausibility for H_1 and the plausibility for H_3 increase when ϵ decreases such that one cannot decide between H_1 or H_3. The weighting factors corresponding to this strategy are quite stable with respect to ϵ and show that most of the conflicting mass is assigned to $H_1 \cup H_3$.

To summarize, it appears that the classical rules of combination are designed for conflicts with different origins. All these classical rules of combination are perfectly retrieved with the proposed generic framework according to the values of w presented in the third row of Table 1, Table 2 and Table 3.

Table 1
Credibility and plausibility functions results for Dempster and Smets rules of combination according to ϵ.

Value of ϵ	Rule of Combination	Associated weighting factors	$Bel(\{H_1\})$	$Bel(\{H_2\})$	$Bel(\{H_3\})$	$Bel(\Theta)$	$Pl(\{H_1\})$	$Pl(\{H_2\})$	$Pl(\{H_3\})$
$\epsilon = 0.1$	*Dempster*	$w(\{H_1\}, m_1, m_2) = 0.47$ $w(\{H_2\}, m_1, m_2) = 0.06$ $w(\{H_3\}, m_1, m_2) = 0.47$	0.47	0.06	0.47	0	0.47	0.06	0.47
	Smets	$w(\emptyset, m_1, m_2) = 1$	0.08	0.01	0.08	0.17	0.08	0.01	0.08
$\epsilon = 0.01$	*Dempster*	$w(\{H_1\}, m_1, m_2) = 0.32$ $w(\{H_2\}, m_1, m_2) = 0.36$ $w(\{H_3\}, m_1, m_2) = 0.32$	0.32	0.36	0.32	0	0.32	0.36	0.32
	Smets	$w(\emptyset, m_1, m_2) = 1$	0.009	0.01	0.009	0.028	0.009	0.01	0.009
$\epsilon = 0.001$	*Dempster*	$w(\{H_1\}, m_1, m_2) = 0.076$ $w(\{H_2\}, m_1, m_2) = 0.848$ $w(\{H_3\}, m_1, m_2) = 0.076$	0.076	0.848	0.076	0	0.076	0.848	0.076
	Smets	$w(\emptyset, m_1, m_2) = 1$	0.001	0.01	0.001	0.012	0.001	0.01	0.001
$\epsilon = 0.0001$	*Dempster*	$w(\{H_1\}, m_1, m_2) = 0.009$ $w(\{H_2\}, m_1, m_2) = 0.982$ $w(\{H_3\}, m_1, m_2) = 0.009$	0.009	0.982	0.009	0	0.009	0.982	0.009
	Smets	$w(\emptyset, m_1, m_2) = 1$	0.001	0.01	0.001	0.01	0.001	0.01	0.001

Table 2
Credibility and plausibility functions results for different rules of combination according to ϵ.

Value of ϵ	Rule of combination	Associated weighting factors	$Bel(\{H_1\})$	$Bel(\{H_2\})$	$Bel(\{H_3\})$	$Pl(\{H_1\})$	$Pl(\{H_2\})$	$Pl(\{H_3\})$
$\epsilon = 0.1$	*Discounting & Dempster*	$w(\{H_1\}, m_1, m_2) = 0.61$ $w(\{H_2\}, m_1, m_2) = 0.08$ $w(\{H_3\}, m_1, m_2) = 0.125$ $w(\Theta, m_1, m_2) = 0.184$	0.61	0.08	0.125	0.794	0.265	0.31
	Yager	$w(\Theta, m_1, m_2) = 1$	0.08	0.01	0.08	0.91	0.84	0.91
	Dubois et al.	$w(H_1 \cup H_2, m_1, m_2) = 0.108$ $w(H_1 \cup H_3, m_1, m_2) = 0.784$ $w(H_2 \cup H_3, m_1, m_2) = 0.108$	0.080	0.010	0.080	0.82	0.19	0.82
$\epsilon = 0.01$	*Discounting & Dempster*	$w(\{H_1\}, m_1, m_2) = 0.677$ $w(\{H_2\}, m_1, m_2) = 0.082$ $w(\{H_3\}, m_1, m_2) = 0.051$ $w(\Theta, m_1, m_2) = 0.189$	0.677	0.082	0.051	0.866	0.272	0.241
	Yager	$w(\Theta, m_1, m_2) = 1$	0.009	0.01	0.009	0.981	0.982	0.981
	Dubois et al.	$w(H_1 \cup H_2, m_1, m_2) = 0.093$ $w(H_1 \cup H_3, m_1, m_2) = 0.814$ $w(H_2 \cup H_3, m_1, m_2) = 0.093$	0.009	0.010	0.009	0.891	0.19	0.891

Table 3
Credibility and plausibility functions results for different rules of combination according to ϵ.

Value of ϵ	Rule of combination	Associated weighting factors	$Bel(\{H_1\})$	$Bel(\{H_2\})$	$Bel(\{H_3\})$	$Pl(\{H_1\})$	$Pl(\{H_2\})$	$Pl(\{H_3\})$
$\epsilon = 0.001$	*Discounting & Dempster*	$w(\{H_1\}, m_1, m_2) = 0.683$ $w(\{H_2\}, m_1, m_2) = 0.08$ $w(\{H_3\}, m_1, m_2) = 0.044$ $w(\Theta, m_1, m_2) = 0.19$	0.683	0.083	0.044	0.874	0.273	0.234
	Yager	$w(\Theta, m_1, m_2) = 1$	0.001	0.01	0.001	0.989	0.998	0.989
	Dubois et al.	$w(H_1 \cup H_2, m_1, m_2) = 0.091$ $w(H_1 \cup H_3, m_1, m_2) = 0.818$ $w(H_2 \cup H_3, m_1, m_2) = 0.091$	0.001	0.01	0.001	0.899	0.190	0.899
$\epsilon = 0.0001$	*Discounting & Dempster*	$w(\{H_1\}, m_1, m_2) = 0.684$ $w(\{H_2\}, m_1, m_2) = 0.083$ $w(\{H_3\}, m_1, m_2) = 0.043$ $w(\Theta, m_1, m_2) = 0.19$	0.684	0.083	0.043	0.874	0.273	0.233
	Yager	$w(\Theta, m_1, m_2) = 1$	0.001	0.01	0.001	0.989	0.998	0.989
	Dubois et al.	$w(H_1 \cup H_2, m_1, m_2) = 0.09$ $w(H_1 \cup H_3, m_1, m_2) = 0.82$ $w(H_2 \cup H_3, m_1, m_2) = 0.09$	0.001	0.01	0.001	0.899	0.190	0.899

4.2 Pattern classification

We consider hereafter a pattern recognition problem. Basically, this is a problem of decision making under uncertainty. Different classical rules of combination are compared with the strategy based on weighting factors learning in order to illustrate how the classical rules behave in pattern recognition problems.

4.2.1 Pattern recognition problem

A pattern recognition problem consists in assigning an input pattern $\mathbf{x}$ to a class H_n, given a learning set $\mathcal{L}$ composed of n patterns $\mathbf{x}^i$ with known classification. Each pattern in $\mathcal{L}$ is represented by a p-dimensional feature vector $\mathbf{x}^i$ and its corresponding class label H^i. In the last ten years, several solutions to this problem have been proposed, based on the Dempster-Shafer theory of evidence. Let us assume that the belief functions are derived from the evidential k-NN classifier proposed by Denœux [48]. In this method, a bba is constructed directly, using as a source of information the training patterns $\mathbf{x}^i$ situated in the neighborhood of the pattern $\mathbf{x}$ to be classified. If the k nearest neighbors (according to a distance measure) are considered, we thus obtain k bba's that are combined using Dempster's rule of combination. The initial method was later refined to allow parameter optimization [17]. Each neighbor can be viewed as a piece of evidence that influences the belief concerning the membership class of $\mathbf{x}$. A belief function m^i associated to each neighbor i is then defined for all $n \in \{1, \cdots, N\}$ as:

$$m^i(\{H_n\}) = \alpha\phi(d^i) \tag{33}$$

$$m^i(\Theta) = 1 - \alpha\phi(d^i) \tag{34}$$

$$m^i(A) = 0 \;\; \forall\, A \in 2^\Theta \setminus \{\{H_n\}, \Theta\} \tag{35}$$

where d^i is the Euclidean distance to the i-th neighbor, α is a discounting parameter and $\phi(.)$ is a decreasing function defined as $\phi(d^i) = \exp[-\gamma(d^i)^2]$. In this expression, γ is a positive parameter. The focal elements of each belief function m^i are singletons of Θ and Θ itself. The belief functions m^i for each neighbor are then aggregated using Dempster's combination rule.

4.2.2 Decision rules

The decision to classify $\mathbf{x}$ as class H_n depends on a decision rule generally based on the plausibility function as defined in equation (7) or the pignistic probability as defined in equation (30).

Let $\mathcal{A}$ be a finite set of actions defined as $\mathcal{A} = \{a_1, \cdots, a_L\}$. If we choose action a_i whereas the pattern is of class H_j, we incur a loss $\lambda(a_i \mid H_j)$. For a basic belief assignment m, we obtain the following expressions for the risk associated with each possible action $a \in \mathcal{A}$:

$$R_\star(a \mid \mathbf{x}) = \sum_{A \subseteq \Theta} m(A) \min_{H \in A} \lambda(a \mid H). \tag{36}$$

Moreover, the risk with respect to the pignistic probability $BetP$ derived from m is equal to:

$$R_{BetP}(a \mid \mathbf{x}) = \sum_{A \subseteq \Theta} m(A) \frac{1}{|A|} \sum_{H \in A} \lambda(a \mid H). \tag{37}$$

The above considerations lead to different decision rules relying on the principle of the minimization of the expected loss. Thus, according to the above equations, we obtain the following two decision rules:

$$D_\star(\mathbf{x}) = a_\star \quad \text{with} \quad R_\star(a_\star) = \min_{a \in \mathcal{A}} R_\star(a \mid \mathbf{x}) \tag{38}$$

and:

$$D_{BetP}(\mathbf{x}) = a_{Bet} \quad \text{with} \quad R_{BetP}(a_{Bet}) = \min_{a \in \mathcal{A}} R_{BetP}(a \mid \mathbf{x}). \tag{39}$$

Details are available in [16]. Assume that the learning set is such that it contains patterns from all classes H_n with $n \in \{1, \cdots, N\}$. In the case of a decision rule with rejection, the typical actions are the assignment a_n to each class H_n and rejection a_0. By considering only the plausibility function and the pignistic probability which are the most useful, the conditions for rejection are expressed as follows [16]:

$$D_\star = a_0 \Leftrightarrow \max_{n=1,\cdots,N} Pl(\{H_n\}) < 1 - \lambda_0, \tag{40}$$

$$D_{Bet} = a_0 \Leftrightarrow \max_{n=1,\cdots,N} BetP(\{H_n\}) < 1 - \lambda_0 \tag{41}$$

where $\lambda_0 \geq 0$ is the rejection cost. The classification rate depends on the value λ_0 and on either the plausibility function or the pignistic probability. Because the plausibility and the pignistic probability depend on the belief mass m, the classification rate clearly depends on the chosen combination as demonstrated in the following example.

4.2.3 Application

For the following simulations, a learning set $\mathcal{L}$ was generated using 3 classes, each containing 100 bidimensional vectors. Each vector from class H_n was generated using Gaussian distributions. The means of the 3 distributions were taken as: $\mu_1 = (0,0)'$, $\mu_2 = (2,2)'$, $\mu_3 = (10,10)'$ and the variance matrices were of the form:

$$\Sigma_1 = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}, \qquad \Sigma_2 = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}, \qquad \Sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}. \tag{42}$$

Let $\mathcal{T}$ be the test database made of 250 vectors for each hypothesis H_n generated as for the learning set. In order to compare the different rules of combination, we show the results obtained on the mean of 10 trials. Figure Fig.2 shows the error according to the rejection rate in the case of a decision based on the plausibility function. Figure Fig.3 shows the error rate according to the rejection rate in the case of a decision based on the pignistic probability. The basic belief assignment is carried out according to the equation (35) with $k = 15$, $\alpha = 0.99$ and where γ_n is the intra-class distance.

Let us suppose that three rules of combination are considered respectively:

- Dempster's combination rule,
- Yager's combination rule,
- and a combination rule based on the learning of the weighting factors [1].

Thus, the results show that the weighted combination is more accurate than Dempster's combination rule whatever the decision rule. In the case of a decision rule based on the pignistic probability, Yager's rule of combination and the weighted combination have similar behaviours. This is coherent because the decision rule distributes the mass assigned to Θ among the singletons.

5 Conclusion

In this paper, we have presented a generic framework for the fusion of information sources modeled by means of belief mass functions. From this framework, we retrieve the classical combination operators used in evidence theory. Furthermore, this generic framework allows a family of combination operators to be defined, so it is possible to derive different operators based on:

- the definition of a set $\mathcal{P}$ collecting the subsets A where the conflicting mass will be distributed,
- weighting factors denoted $w(A, \mathbf{m})$ assigned to each subset $A \subseteq \mathcal{P}$.

[1] A validation database is built in the same way as the test database. It enables the values of the weighting factors $w(A \subseteq \mathcal{P}, \mathbf{m})$ with $\mathbf{m} = \{m_1, \ldots, m_{15}\}$ to be learnt. The weighting factor $w(A \subseteq \mathcal{P}, \mathbf{m})$ is obtained according to the methodology described in subsection 3.4.

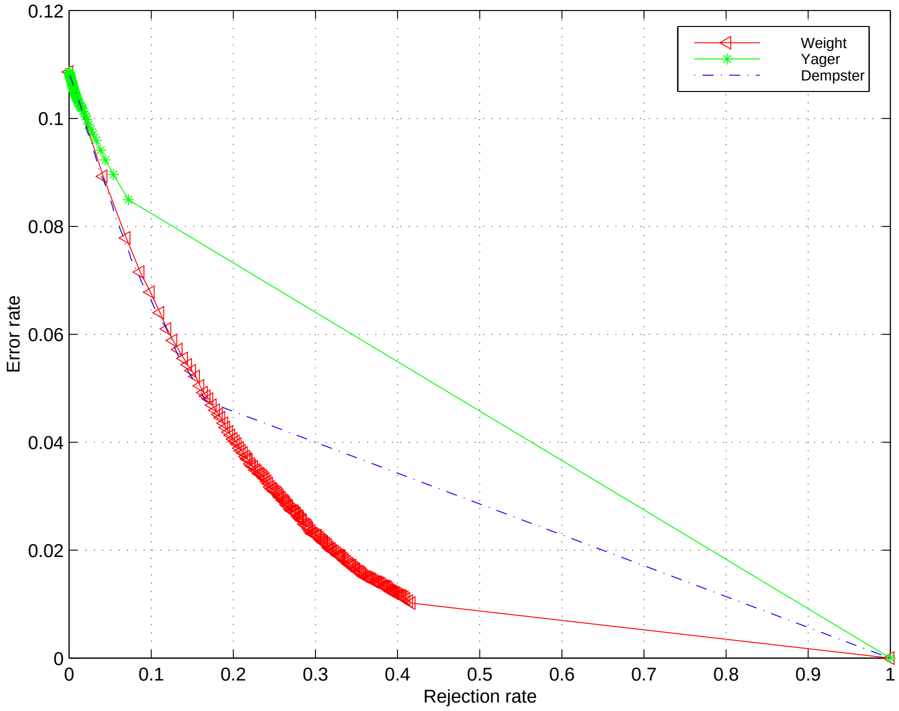

Fig. 2. Error rate vs. rejection rate for decision making based on the plausibility.

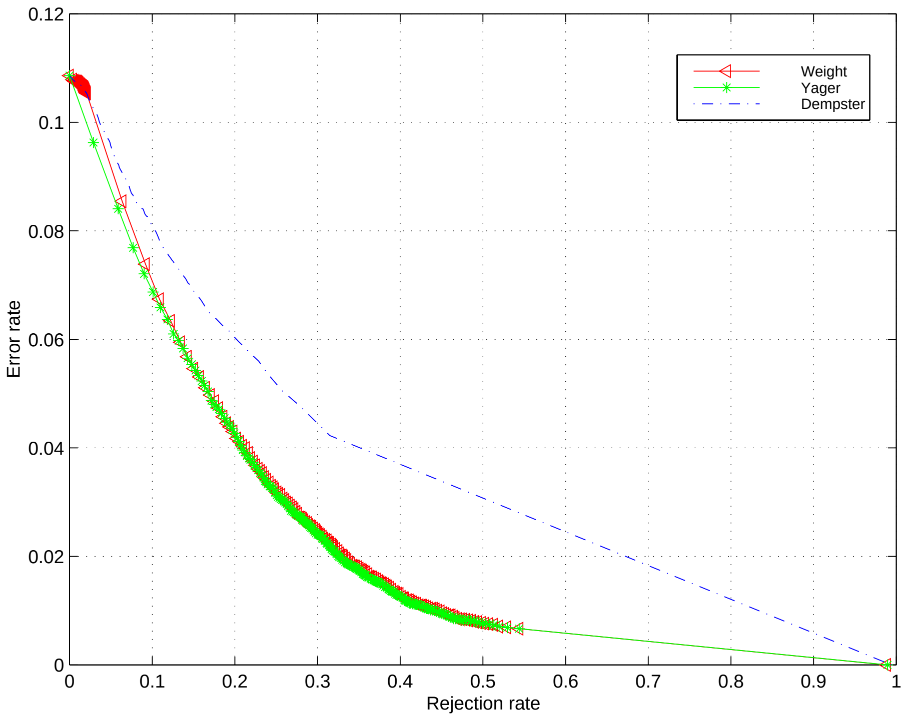

Fig. 3. Error rate vs. rejection rate for decision making based on the pignistic probability.

The relationship between the discounting and the weighted combination has been emphasized (see the appendix). Several methods are possible in order to obtain the weighting factors. One of these methods determines the weighting factor by minimizing a mean square error and this is possible for pattern recognition problems. The method has been checked and compared with Dempster's classical combination rule. The tests also show that our approach, as well as being particularly well-suited to pattern recognition, is also valid for the interpretation of the resulting mass. With this formalism the most suitable solution will always be applied in each different conflict management strategy. The application of the proposed formalism to the case of partially known labeling is under study.

References

[1] M. Abidi, R. Gonzales (Eds.), Data Fusion in Robotics and Machine Intelligence, Academic Press, New York, 1992.

[2] D. Dubois, H. Prade, Combination of Fuzzy Information in the framework of Possibility Theory, Data Fusion in Robotics and Machine Intelligence (1992) 481–505.

[3] S. Deveughèle, B. Dubuisson, Adaptative Aggregation: Decomposing Before Combining, in: Proc. of the 4th IEEE International Conference on Fuzzy Systems (FUZZ-IEEE'95), Yokohama, 1995, pp. 1589–1596.

[4] H. Li, S. Munjanath, S. Mitra, Multisensor Image Fusion Using the Wavelet Transform, Graphical Models and Image Processing 57 (3) (1995) 235–245.

[5] J. Desachy, L. Roux, E. Zahzah, Numeric and Symbolic Data Fusion: A Soft Computing Approach to Remote Sensing Images Analysis, Pattern Recognition Letters 17 (1996) 1361–1378.

[6] S. L. Hégarat-Mascle, I. Bloch, D. Vidal-Madjar, Application of Dempster-Shafer evidence theory to unsupervised classification in multisource remote sensing, IEEE Transactions on Geoscience and remote Sensing 35 (4) (1997) 1018–1032.

[7] S. L. Hégarat-Mascle, I. Bloch, D. Vidal-Madjar, Introduction of Neighborhood Information in Evidence Theory and Application to Data Fusion of Radar and Optical Images with Partial Cloud Cover, Pattern Recognition 34 (1998) 1811–1823.

[8] L. Cholvy, About Merged Information, in: D. Dubois, H. Prade (Eds.), Handbook of Defeasible Reasoning and Uncertainty Management Systems, Vol. 3, Kluwer Acad. Publ., Dordrecht, 1998, pp. 233–263.

[9] J. Gebhardt, R. Kruse, Parallel Combination of Information Sources, in: D. Dubois, H. Prade (Eds.), Handbook of Defeasible Reasoning and Uncertainty Management Systems, Vol. 3, Kluwer Acad. Publ., Dordrecht, 1998, pp. 393–439.

[10] J. Bezdek, J. Keller, R. Krishnapuram, N. Pal, Fuzzy Models and Algorithms for Pattern Recognition and Image Processing, in: H. Zimmermann (Ed.), The Handbook of Fuzzy Sets Series, Kluwer Acad. Publ., Dordrecht, 1999.

[11] L. Fouque, A. Appriou, An evidential Markovian model for data fusion and unsupervised image classification, in: Proc. of the Third International Conference on Information Fusion (FUSION 2000), Paris, France, 2000, pp. TuB4 25–32.

[12] A. Appriou, Multisensor Signal Processing in the Framework of the Theory of Evidence, in: Application of Mathematical Signal Processing Techniques to Mission Systems, Research and Technology Organization (Lecture Series 216), November 1999, pp. (5–1)(5–31).

[13] C. Pohl, J. van Genderen, Multisensor Image Fusion in Remote Sensing: Concepts, Methods and Applications, International Journal of Remote Sensing 19 (5) (1998) 823–854.

[14] I. Bloch, Some Aspects of Dempster-Shafer Evidence Theory for Classification of Multi-Modality Medical Images Taking Partial Volume Effect Into Account, Pattern recognition Letters 17 (1996) 905–919.

[15] E. Lefevre, P. Vannoorenberghe, O. Colot, About the Use of Dempter-Shafer Theory for Color Image Segmentation, in: First International Conference on Color in Graphics and Image Processing (CGIP'2000), October 2000, pp. 164–169.

[16] T. Denoeux, Analysis of Evidence-Theory Decision Rules for Pattern Classification, Pattern Recognition 30 (7) (1997) 1095–1107.

[17] L. M. Zouhal, T. Denoeux, An Evidence-Theoretic K-NN Rule with Parameter Optimization, IEEE Transactions on Systems, Man and Cybernetics - Part C 28 (2) (1998) 263–271.

[18] T. Denoeux, A Neural Network Classifier Based on Demspter-Shafer Theory, IEEE Transactions on Systems, Man and Cybernetics Part A 30 (2) (2000) 131–150.

[19] Y. Bar-Shalom, X. Li, Multitarget-Multisensor Tracking: Principles and Techniques, Storrs, CT: YBS Publishing, 1995.

[20] L. Zadeh, A simple view of the Dempster-Shafer theory of evidence and its implication for the rule of combination, AI magazine 7 (1986) 85–90.

[21] R. Yager, On the Dempster-Shafer Framework and New Combination Rules, Information Sciences 41 (1987) 93–138.

[22] D. Dubois, H. Prade, Representation and Combination of Uncertainty with Belief Functions and Possibility Measures, Comput. Intell. 4 (1988) 244–264.

[23] P. Smets, The Combination of Evidence in the Transferable Belief Model, IEEE Trans. on Pattern Analysis and Machine Intelligence 12 (5) (1990) 447–458.

[24] C. K. Murphy, Combining Belief Functions When Evidence Conflicts, Decision Support Systems 29 (2000) 1–9.

[25] A. Dempster, Upper and Lower Probabilities Induced by Multivalued Mapping, Annals of Mathematical Statistics AMS-38 (1967) 325–339.

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

[27] P. Smets, What is Dempster-Shafer's Model?, in: R. Yager, M. Fedrizzi, J. Kacprzyk (Eds.), Advances in the Dempster-Shafer Theory of Evidence, Wiley, 1994, pp. 5–34.

[28] P. Smets, R. Kennes, The Transferable Belief Model, Artificial Intelligence 66 (2) (1994) 191–234.

[29] A. Appriou, Probabilités et incertitude en fusion de données multi-senseurs, Revue Scientifique et Technique de la Défense 11 (1991) 27–40.

[30] P. Smets, Belief Functions: The Disjonctive Rule of Combination and the Generalized Bayesian Theorem, Int. J. of Approximate Reasoning 9 (1993) 1–35.

[31] E. Lefevre, P. Vannoorenberghe, O. Colot, Using Information Criteria in Dempster-Shafer's Basic Belief Assignment, in: Proceeding of Fuzz'ieee 99, 1999, pp. 173–178.

[32] P. Walley, S. Moral, Upper Probabilities Based Only on the Likelihood Function, Journal of Royal Statistical Society, Serie B 61 (Part 4) (1999) 831–847.

[33] T. Denoeux, A K-Nearest Neighbour Classification Rule Based on Dempster-Shafer Theory, IEEE Transactions on Systems, Man and Cybernetics 25 (5) (1995) 804–813.

[34] T. Denoeux, L. M. Zouhal, Handling Possibilistic Labels in Pattern Classification Using Evidential Reasoning, Fuzzy Sets and Systems 122 (3) (2001) 47–62.

[35] F. Voorbraak, On the Justification of Dempster's Rule of Combinations, Artificial Intelligence 48 (1991) 171–197.

[36] F. Klawonn, E. Schwecke, On the axiomatic justification of Dempster's rule combination, International Journal of Intelligent Systems 7 (1992) 469–478.

[37] D. Dubois, H. Prade, On the unicity of Dempster rule of combination, International Journal of Intelligent System 1 (1986) 133–142.

[38] P. Quinio, T. Matsuyama, Random Closed Sets : A Unified Approach to the Representation of Imprecision and Uncertainty, in: P. Siegel, R. Kruse (Eds.), Symbolic and Quantitative Approaches to Uncertainty (ECSQARU), Springer Verlag, Marseille, France, 1991, pp. 282–286.

[39] P. Smets, Resolving Misunderstandings About Belief Functions, International Journal of Approximate Reasoning 6 (1992) 321–344.

[40] B. Yaghlane, P. Smets, K. Mellouli, Independence and Non-Interactivity in the Transferable Belief Model, in: F. Matus, M. Studeny (Eds.), Workshop on Conditional Independence Structures and Graphical Models, Toronto, Canada, 1999, pp. 4–5.

[41] B. Yaghlane, P. Smets, K. Mellouli, Independence Concepts for Belief Functions, in: 8th International Conference Information Processing and Management of Uncertainty in Knowledge-based Systems (IPMU), Vol. 1, 2000, pp. 357–364.

[42] E. Lefevre, O. Colot, P. Vannoorenberghe, D. de Brucq, Contribution des mesures d'information à la modélisation crédibiliste de connaissances, Revue Traitement du Signal 17 (2) (2000) 87–97.

[43] S. Fabre, A. Appriou, X. Briottet, Presentation and description of two classification methods using data fusion based on sensor management, Information Fusion 2 (2001) 49–71.

[44] C. Royère, D. Gruyer, V. Cherfaoui, Data Association with Belief Theory, in: Thrid International Conference on Information Fusion (FUSION'2000), July 2000, pp. TuD2 3–9.

[45] P. Smets, The Alpha-Junctions : The Commutative Combination Operators Applicable to Belief Functions, D. Gabbay, R. Kruse, A. Nonengart and H.J. Ohlbach Edition, Springer Verlag, Berlin, 1997, Ch. Qualtitative and quantitative practical reasoning, pp. 131–153.

[46] E. Lefevre, O. Colot, P. Vannoorenberghe, D. D. Brucq, A Generic Framework for Resolving the Conflict in the Combination of Belief Structures, in: Thrid International Conference on Information Fusion (FUSION'2000), July 2000, pp. MOD4 11–18.

[47] P. Smets, Constructing the Pignistic Probability Function in a Context of Uncertainty, in: M. Henrion, R. D. Schachter, L. Kanal, J. Lemmer (Eds.), Uncertainty in Artificial Intelligence 5, North-Holland, Amsterdam, 1990, pp. 29–40.

[48] T. Denoeux, An Evidence-Theoric Neural Network Classifier, in: IEEE International Conference on Systems, Man and Cybernetics, Vol. 3, Vancouver, 1995, pp. 712–717.

Appendix: discounting vs. combination

This appendix is devoted to the detailed calculus allowing to emphasize the link between the conflict management based on a discounting (case of not reliable sources) and a management based on an adapted assignment of the conflicting mass by means of weighting factors associated to each subset.

Let m_j a belief mass function issued from an information source S_j. The commonality function q_j associated to m_j is defined as:

$$q_j(A) = \sum_{A \subseteq B} m_j(B) \qquad \forall A \subseteq \Theta. \tag{43}$$

Furthermore, the inverse Möbius transform allows to retrieve the mass functions from the commonality function q_j by means of the following equation:

$$m_j(A) = \sum_{A \subseteq B} (-1)^{|B-A|} q_j(B) \qquad \forall A \subseteq \Theta. \tag{44}$$

Results of the combination of discounted belief mass functions

Let $\{m_1, \ldots, m_j, \ldots, m_J\}$ be a set of belief functions. We denote $m_{\alpha_j,j}$ the belief function m_j discounted by a coefficient α_j. Then, the function $m_{\alpha_j,j}$ can be written as it follows:

$$\begin{cases} m_{\alpha_j,j}(A) = \alpha_j m_j(A) \qquad \forall A \subset \Theta \\ m_{\alpha_j,j}(\Theta) = 1 - \alpha_j + \alpha_j m_j(\Theta). \end{cases} \tag{45}$$

The commonality function $q_{\alpha_j,j}$, associated to $m_{\alpha_j,j}$, can be written as:

$$\begin{aligned} \forall A \subseteq \Theta \qquad q_{\alpha_j,j}(A) &= \textstyle\sum_{A \subseteq B} m_{\alpha_j,j}(B) \\ &= \textstyle\sum_{\substack{A \subseteq B \\ B \neq \Theta}} (\alpha_j m_j(B)) + 1 - \alpha_j + \alpha_j m_j(\Theta) \\ &= \alpha_j \textstyle\sum_{A \subseteq B} m_j(B) + 1 - \alpha_j \\ &= \alpha_j q_j(A) + 1 - \alpha_j \\ q_{\alpha_j,j}(A) &= \alpha_j(q_j(A) - 1) + 1 \end{aligned} \tag{46}$$

Belief functions resulting from the combination

One can express the combination of the J information sources by means of commonality functions. The result of this fusion is denoted q_α and can be written as follows:

$$\begin{aligned} q_\alpha(A) &= K_\alpha \times q_{\alpha_1,1}(A) \times \ldots \times q_{\alpha_j,j}(A) \times \ldots q_{\alpha_J,J}(A) \qquad \forall A \subseteq \Theta \\ &= K_\alpha \times \prod_{j=1}^{J} q_{\alpha_j,j}(A) \qquad \forall A \subseteq \Theta \end{aligned} \tag{47}$$

where K_α is the normalization coefficient of the combination. This coefficient is as follows:

$$K_\alpha = \frac{1}{-\sum\limits_{\substack{B\subseteq\Theta \\ B\neq\emptyset}} (-1)^{|B|} q_\alpha(B)}. \tag{48}$$

The mass resulting from Dempster's rule of combination (normalized combination) can be written as follows:

$$\begin{aligned}
\forall A \subseteq \Theta \qquad m_\alpha(A) &= \frac{1}{-\sum\limits_{\substack{B\subseteq\Theta \\ B\neq\emptyset}} (-1)^{|B|} q_\alpha(B)} \sum_{A\subseteq B} (-1)^{|B-A|} q_\alpha(B) \\
&= \frac{1}{-\sum\limits_{\substack{B\subseteq\Theta \\ B\neq\emptyset}} (-1)^{|B|} \prod\limits_{j=1}^{J} q_{\alpha_j,j}(B)} \sum_{A\subseteq B} (-1)^{|B-A|} \prod_{j=1}^{J} q_{\alpha_j,j}(B) \\
m_\alpha(A) &= \frac{\sum\limits_{A\subseteq B} (-1)^{|B-A|} \prod\limits_{j=1}^{J} [\alpha_j(q_j(B)-1)+1]}{-\sum\limits_{\substack{B\subseteq\Theta \\ B\neq\emptyset}} (-1)^{|B|} \prod\limits_{j=1}^{J} [\alpha_j(q_j(B)-1)+1]}.
\end{aligned} \tag{49}$$

Belief mass function resulting from the proposed rule of combination

Let m_c be the belief function resulting from the proposed combination of the J belief functions m_j. This one can be written as follows:

$$m_c(A) = m_\cap(A) + w(A, \mathbf{m}) m(\emptyset) \qquad \forall A \subseteq \Theta \tag{50}$$

where $m_\cap(.)$ is the mass resulting of the conjunctive combination and where $w(A, \mathbf{m})$, with $\mathbf{m} = \{m_1, \ldots, m_J\}$, is the weighting factor associated to the assignment of the conflicting mass $m(\emptyset)$ to the subset A. This equation (50) can be written by means of the commonality function. Indeed, the result of the conjunctive combination can be written:

$$m_\cap(A) = \sum_{A\subseteq B} (-1)^{|B-A|} \prod_{j=1}^{J} q_j(B) \qquad \forall A \subseteq \Theta. \tag{51}$$

The conflicting mass generated by this conjunctive combination can be written as:

$$m(\emptyset) = 1 + \sum_{\substack{B\subseteq\Theta \\ B\neq\emptyset}} (-1)^{|B|} \prod_{j=1}^{J} q_j(B). \tag{52}$$

The resulting belief function of the proposed combination is defined $\forall A \subseteq \Theta$:

$$m_c(A) = \sum_{A\subseteq B} (-1)^{|B-A|} \prod_{j=1}^{J} q_j(B) + w(A, \mathbf{m}).[1 + \sum_{\substack{B\subseteq\Theta \\ B\neq\emptyset}} (-1)^{|B|} \prod_{j=1}^{J} q_j(B)]. \tag{53}$$

Taking into account the equations (49) and (53), we obtain the weighting factor values for the assignment of the conflicting mass according to the discounting coefficients α_j and the belief masses m_j:

$$w(A,\mathbf{m}) = \zeta \left[\frac{\sum_{A\subseteq B} (-1)^{|B-A|} \prod_{j=1}^{J} [\alpha_j(q_j(B)-1)+1]}{-\sum_{\substack{B\subseteq\Theta\\ B\neq\emptyset}} (-1)^{|B|} \prod_{j=1}^{J} [\alpha_j(q_j(B)-1)+1]} - \sum_{A\subseteq B} (-1)^{|B-A|} \prod_{j=1}^{J} q_j(B) \right] \tag{54}$$

with:

$$\zeta = \frac{1}{1 + \sum_{\substack{B\subseteq\Theta\\ B\neq\emptyset}} (-1)^{|B|} \prod_{j=1}^{J} q_j(B)}. \tag{55}$$