

## Combining partially independent belief functions

Mouna Chebbah, Arnaud Martin, Boutheina Ben Yaghlane

► **To cite this version:**

Mouna Chebbah, Arnaud Martin, Boutheina Ben Yaghlane. Combining partially independent belief functions. Decision Support Systems, Elsevier, 2015, pp.37-46. <10.1016/j.dss.2015.02.017>. <hal-01132564>

**HAL Id: hal-01132564**

**<https://hal.archives-ouvertes.fr/hal-01132564>**

Submitted on 17 Mar 2015

**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.

# Combining partially independent belief functions

Mouna Chebbah<sup>1,a,b</sup>, Arnaud Martin<sup>b</sup>, Boutheina Ben Yaghlane<sup>c</sup>

<sup>a</sup>LARODEC Laboratory, University of Tunis, ISG Tunis, Tunisia

<sup>b</sup>IRISA, University of Rennes1, Lannion, France

<sup>c</sup>LARODEC Laboratory, University of Carthage, IHEC Carthage, Tunisia

---

## Abstract

The theory of belief functions manages uncertainty and also proposes a set of combination rules to aggregate opinions of several sources. Some combination rules mix evidential information where sources are independent; other rules are suited to combine evidential information held by dependent sources. In this paper we have two main contributions: First we suggest a method to quantify sources' degree of independence that may guide the choice of the more appropriate set of combination rules. Second, we propose a new combination rule that takes consideration of sources' degree of independence. The proposed method is illustrated on generated mass functions.

*Keywords:* Theory of belief functions, Combination rules, Clustering, Independence, Sources independence, Combination rule choice

---

## 1. Introduction

Uncertainty theories like the *theory of probabilities*, the *theory of fuzzy sets* [1], the *theory of possibilities* [2] and the *theory of belief functions* [3, 4] model and manage uncertain data. The theory of belief functions can deal with imprecise

---

*Email addresses:* Mouna.Chebbah@univ-rennes1.fr (Mouna Chebbah),  
Arnaud.Martin@univ-rennes1.fr (Arnaud Martin),  
boutheina.yaghlane@ihec.rnu.tn (Boutheina Ben Yaghlane)

and/or uncertain data provided by several belief holders and also combine them.

Combining several evidential information held by distinct belief holders aggregates their points of view by stressing common points. In the theory of belief functions, many combination rules are proposed, some of them like [2, 5, 6, 7, 8, 9] are fitted to the aggregation of evidential information provided by cognitively independent sources whereas the *cautious*, *bold* [10] and *mean* combination rules can be applied when sources are *cognitively dependent*. The choice of combination rules depends on sources independence.

Some researches are focused on doxastic independence of variables such as [11, 12]; others [4, 13] tackled cognitive and evidential independence of variables. This paper is focused on measuring the independence of sources and not that of variables. We suggest a statistical approach to estimate the independence of sources on the bases of all evidential information that they provide. The aim of estimating the independence of sources is to guide the choice of the combination rule to be used when combining their evidential information.

We propose also a new combination rule to aggregate evidential information and take into account the independence degree of their sources. The proposed combination rule is weighted with that degree of independence leading to the conjunctive rule [14] when sources are fully independent and to the cautious rule [10] when they are fully dependent.

In the sequel, we introduce in Section 2 preliminaries of the theory of belief functions. In the Section 3, an evidential clustering algorithm is detailed. This clustering algorithm will be used in the first step of the independence measure process. Independence measure is then detailed in Section 4. It is estimated in four steps: In the first step the clustering algorithm is applied. Second a mapping between clusters is performed; then independence of clusters and sources are deduced in the last two steps. Independence is learned for only two sources and then generalized for

a greater number of sources. A new combination rule is proposed in the Section 5 taking into account the independence degree of sources. The proposed method is tested on random mass functions in Section 6. Finally, conclusions are drawn.

## 2. Theory of belief functions

The theory of belief functions was introduced by Dempster [3] and formalized by Shafer [4] to model imperfect data. The *frame of discernment* also called *universe of discourse*,  $\Omega = \{\omega_1, \omega_2, \dots, \omega_N\}$ , is an exhaustive set of  $N$  mutually exclusive hypotheses  $\omega_i$ . The *power set*  $2^\Omega$  is a set of all subsets of  $\Omega$ ; it is made of hypotheses and unions of hypotheses from  $\Omega$ . The *basic belief assignment* (BBA) commonly called *mass function* is a function defined on the power set  $2^\Omega$  and spans the interval  $[0, 1]$  such that:

$$\sum_{A \subseteq \Omega} m(A) = 1 \quad (1)$$

A *basic belief mass* (BBM) also called *mass*,  $m(A)$ , is a degree of faith on the truth of  $A$ . The BBM,  $m(A)$ , is a degree of belief on  $A$  which can be committed to its subsets if further information justifies it [7].

Subsets  $A$  having a strictly positive mass are called *focal elements*. Union of all focal elements is called *core*. Shafer [4] assumed a normality condition such that  $m(\emptyset) = 0$ , thereafter Smets [14] relaxed this condition in order to tolerate  $m(\emptyset) > 0$ .

The frame of discernment can also be a focal element; its BBM,  $m(\Omega)$ , is interpreted as a degree of *ignorance*. In the case of *total ignorance*,  $m(\Omega) = 1$ .

A *simple support function* is a mass function with two focal elements including the frame of discernment. A simple support function  $m$  is defined as follows:

$$m(A) = \begin{cases} 1 - w & \text{if } A = B \text{ for some } B \subset \Omega \\ w & \text{if } A = \Omega \\ 0 & \text{otherwise} \end{cases} \quad (2)$$

Where  $A$  is a *focus* of that simple support function and  $w \in [0, 1]$  is its *weight*. A simple support function is simply noted  $A^w$ . A nondogmatic mass function can be obtained by the combination of several simple support functions. Therefore, any nondogmatic mass function can be decomposed into several support functions using the *canonical decomposition* proposed by Smets [15].

The *belief function* ( $bel$ ) is computed from a BBA  $m$ . The amount  $bel(A)$  is the minimal belief on  $A$  justified by available information on  $B$  ( $B \subseteq A$ ):

$$bel(A) = \sum_{B \subseteq A, B \neq \emptyset} m(B) \quad (3)$$

The *plausibility function* ( $pl$ ) is also derived from a BBA  $m$ . The amount  $pl(A)$  is the maximal belief on  $A$  justified by information on  $B$  which are not contradictory with  $A$  ( $A \cap B \neq \emptyset$ ):

$$pl(A) = \sum_{A \cap B \neq \emptyset} m(B) \quad (4)$$

Pignistic transformation computes pignistic probabilities from mass functions in the purpose of making a decision. The pignistic probability of a single hypothesis  $A$  is given by:

$$BetP(A) = \sum_{B \subseteq \Omega, B \neq \emptyset} \frac{|B \cap A|}{|B|} \frac{m(B)}{1 - m(\emptyset)}. \quad (5)$$

Decision is made according to the maximum pignistic probability. The single point having the greatest  $BetP$  is the most likely hypothesis.

### 2.1. Discounting

Sources of information are not always reliable, they can be unreliable or even a little bit reliable. Taking into account reliability of sources, we adjust their beliefs proportionally to degrees of reliability. Discounting mass functions is a way of taking consideration of sources' reliabilities into their mass functions. If reliability

rate  $\alpha$  of a source is known or can be quantified; discounting its mass function  $m$  is defined as follows:

$$\begin{cases} m^\alpha(A) &= \alpha \times m(A) \\ m^\alpha(\Omega) &= 1 - \alpha \times (1 - m(\Omega)) \end{cases}, \forall A \subset \Omega \quad (6)$$

This discounting operator can be used not only to take consideration of source's reliability, but also to consider any information which can be integrated into the mass function,  $(1 - \alpha)$  is called *discounting rate*.

## 2.2. Combination rules

In the theory of belief functions, a great number of combination rules are used to summarize a set of mass functions into only one. Let  $s_1$  and  $s_2$  be two distinct and cognitively independent sources providing two different mass functions  $m_1$  and  $m_2$  defined on the same frame of discernment  $\Omega$ . Combining these mass functions induces a third one  $m_{12}$  defined on the same frame of discernment  $\Omega$ .

There is a great number of combination rules [2, 5, 6, 7, 8, 9], but we enumerate in this section only Dempster, conjunctive, disjunctive, Yager, Dubois and Prade, mean, cautious and bold combination rules. The first combination rule was proposed by Dempster in [3] to combine two distinct mass functions  $m_1$  and  $m_2$  as follows:

$$m_{1 \oplus 2}(A) = (m_1 \oplus m_2)(A) = \begin{cases} \frac{\sum_{B \cap C = A} m_1(B) \times m_2(C)}{1 - \sum_{B \cap C = \emptyset} m_1(B) \times m_2(C)} & \forall A \subseteq \Omega, A \neq \emptyset \\ 0 & \text{if } A = \emptyset \end{cases} \quad (7)$$

The BBM of the empty set is null ( $m(\emptyset) = 0$ ). This rule verifies the normality condition and works under a *closed world* where  $\Omega$  is exhaustive.

In order to solve the problem highlighted by Zadeh's counter example [16]

where Dempster's rule of combination produced unsatisfactory results, many combination rules appeared. Smets [14] proposed an *open world* where a positive mass can be allocated to the empty set. Hence the conjunctive rule of combination for two mass functions  $m_1$  and  $m_2$  is defined as follows:

$$m_1 \odot_2(A) = (m_1 \odot m_2)(A) = \sum_{B \cap C = A} m_1(B) \times m_2(C) \quad (8)$$

Even if Smets [17] interpreted the BBM,  $m_1 \odot_2(\emptyset)$ , as an amount of conflict between evidences that induced  $m_1$  and  $m_2$ ; that amount is not really a conflict because it includes a certain degree of auto-conflict due to the non-idempotence of the conjunctive combination [18].

The conjunctive rule is used only when both sources are reliable. Smets [14] proposed also to use a disjunctive combination when an unknown source is unreliable. The disjunctive rule of combination is defined for two BBAs  $m_1$  and  $m_2$  as follows:

$$m_1 \cup_2(A) = (m_1 \cup m_2)(A) = \sum_{B \cup C = A} m_1(B) \times m_2(C) \quad (9)$$

Yager in [8] interpreted  $m(\emptyset)$  as an amount of ignorance; consequently it is allocated to  $\Omega$ . Yager's rule of combination is also defined to combine two mass functions  $m_1$  and  $m_2$  as follows:

$$\begin{cases} m_Y(X) = m_1 \odot_2(X) & \forall X \subset \Omega, X \neq \emptyset \\ m_Y(\Omega) = m_1 \odot_2(\Omega) + m_1 \odot_2(\emptyset) \\ m_Y(\emptyset) = 0 \end{cases} \quad (10)$$

Dubois and Prade's solution [2] was to affect the mass resulting from the combination of conflicting focal elements to the union of these subsets:

$$\begin{cases} m_{DP}(B) = m_1 \odot_2(B) + \sum_{A \cap X = \emptyset, A \cup X = B} m_1(X) m_2(A) & \forall A \subseteq \Omega, A \neq \emptyset \\ m_{DP}(\emptyset) = 0 \end{cases} \quad (11)$$

Conjunctive, disjunctive and Dempster's rules are associative and commutative, but Yager and Dubois and Prade's rules are not associative, even if they are commutative. Unfortunately, all combination rules described above are not idempotent because  $m \oplus m \neq m$  and  $m \odot m \neq m$ .

Mean combination rule detailed in [6],  $m_{Mean}$ , of two mass functions  $m_1$  and  $m_2$  is the average of these ones. Therefore, for each focal element  $A$  of  $M$  mass functions, the combined one is defined as follows:

$$m_{Mean}(A) = \frac{1}{M} \sum_{i=1}^M m_i(A) \quad (12)$$

Besides idempotence, this combination rule verifies normality condition ( $m(\emptyset) = 0$ ) if combined mass functions are normalized ( $\forall i \in M, m_i(\emptyset) = 0$ ). We note also that this combination rule is commutative but not associative.

All combination rules described above work under a strong assumption of cognitive independence since they are used to combine mass functions induced by two distinct sources. This strong assumption is always assumed but never verified. Denoeux [10], proposed a family of conjunctive and disjunctive rules based on triangular norms and conorms. Cautious and bold rules are members of that family and combine mass functions for which independence assumption is not verified. Cautious combination of two mass functions  $m_1$  and  $m_2$  issued from probably dependent sources is defined as follows:

$$m_1 \otimes m_2 = \oplus_{A \subset \Omega} A^{w_1(A) \wedge w_2(A)} \quad (13)$$

Where  $A^{w_1(A)}$  and  $A^{w_2(A)}$  are simple support functions focused on  $A$  with weights  $w_1$  and  $w_2$  issued from the canonical decomposition [15] of  $m_1$  and  $m_2$  respectively, note also that  $\wedge$  is a *min operator* of simple support functions weights. The bold and cautious combination rules are commutative, associative and idempotent.

To summarize, the choice of the combination rule is based on the dependence



of sources. Combination rules like [2, 5, 6, 7, 8] combine mass functions which sources are independent, whereas cautious, bold and mean rules are the most fitted to combine mass functions issued from dependent sources.

In this paper, we propose a method to quantify sources' degrees of independence that may be used in a new mixed combination rule. In fact, we propose a statistical approach to learn sources' degrees of independence from all provided evidential information. Indeed, two sets of evidential information assessed by two different sources are classified into two sets of clusters. Clusters of both sources are matched and the independence of each couple of matched clusters is quantified in order to estimate sources' degrees of independence. Therefore, a clustering technique is used to gather similar objects into the same cluster in order to study the source's overall behavior. Before introducing our learning method, we detail in the next section the evidential clustering algorithm that will be used in the learning of sources' degrees of independence.

### **3. Evidential clustering**

In this paper, we propose a new clustering technique to classify objects; their attributes values are evidential and classes are unknown. Proposed clustering algorithm uses a distance on belief functions given by Jousselme et al. [19] such as proposed by Ben Hariz et al. [20].

Ben Hariz et al. [20] detailed a belief  $K$ -modes classifier in which Jousselme distance [19] is adapted to quantify distances between objects and clusters modes. These are sets of mass functions; each one is the combination of an attribute's values of all objects classified into that cluster. An object is attributed to the cluster having the minimum distance to its mode.

Temporal complexity of clustering algorithm proposed by Ben Hariz et al. [20]

is quite high as clusters modes and distances are computed in each iteration. The combination by the mean rule to compute modes values leads to mass functions with a high number of focal elements. Hence, the bigger the cluster is, the least significant is the distance.

We propose a clustering technique to classify objects that attributes values are uncertain. However uncertainty is modeled with the theory of belief functions detailed in Section 2. In the proposed algorithm, we do not use any cluster mode to avoid the growth of focal elements number in clusters modes. Temporal complexity is also significantly reduced because all distances are computed only once.

In this section,  $K$  is the number of clusters  $Cl_k$  ( $1 \leq k \leq K$ );  $n$  is the number of objects to be classified;  $n_k$  is the number of objects classified into cluster  $Cl_k$ ;  $o_i$  are objects to classify  $o_i : 1 \leq i \leq n$ ;  $c$  is the number of evidential attributes  $a_j : 1 \leq j \leq c$  which domains are  $\Omega_{a_j}$  and finally  $m_{ij}$  is a mass function value of attribute “ $j$ ” for object “ $i$ ”. Mass functions  $m_{ij}$  can be certain, probabilistic, possibilistic, evidential and even missing.

To classify objects  $o_i$  into  $K$  clusters, we use a clustering algorithm with a distance on belief functions given by [19]. The number of clusters  $K$  is assumed to be known. Proposed clustering technique is based on a distance which quantifies how much is far an object  $o_i$  from a cluster  $Cl_k$ . This distance is the mean of distances between  $o_i$ , and all objects  $o_q$  that are classified into cluster  $Cl_k$  as follows:

$$D(o_i, Cl_k) = \frac{1}{n_k} \sum_{q=1}^{n_k} dist(o_i, o_q) \quad (14)$$

and

$$dist(o_i, o_q) = \frac{1}{c} \sum_{j=1}^c d(m_{ij}, m_{qj}) \quad (15)$$

with :

$$d(m_{ij}, m_{qj}) = \sqrt{\frac{1}{2} (m_{ij} - m_{qj})^t \underline{\underline{D}} (m_{ij} - m_{qj})} \quad (16)$$

such that :

$$\underline{D}(A, B) = \begin{cases} 1 & \text{if } A = B = \emptyset \\ \frac{|A \cap B|}{|A \cup B|} & \forall A, B \in 2^{\Omega_{a_j}} \end{cases} \quad (17)$$

Each object is affected to the most similar cluster in an iterative way till reaching an unchanged cluster partition. It is obvious that clusters number  $K$  must be known. Temporal complexity of the proposed algorithm is significantly optimized as pairwise distances are computed once a time from the beginning. We do not use any cluster mode. Consequently, there will be no problem of increasing number of focal elements because attributes values are not combined. Indeed, the evidential clustering algorithm provides a cluster partition that minimizes distances between objects into the same cluster and maximizes the distance between objects classified into different clusters. The main asset of the evidential clustering algorithm according to the belief  $K$ -modes proposed by Ben Hariz et al. [20] is the optimization of the temporal complexity. In fact, run-time of the evidential clustering algorithm is improved. The optimization of run-time depends on the size of the frame of discernment  $|\Omega_{a_j}|$ , the number of clusters  $K$  and number of objects  $n$ . For example, figure 1 shows a big gain in the run-time of evidential clustering according to the belief  $K$ -modes when the number of mass functions varies,  $n \in [10, 1000]$ . Temporal complexity of the evidential clustering algorithm is optimized and that optimization is especially noticed when the number of mass functions to classify is high and also when the frame of discernment contains many hypotheses. Thanks to the improve of the temporal complexity, this clustering algorithm is used in the following sections.

#### 4. Learning sources independence degree

In this section we extend paper [21] for many sources, and propose a combination rule emphasizing sources independence degree. In the theory of prob-

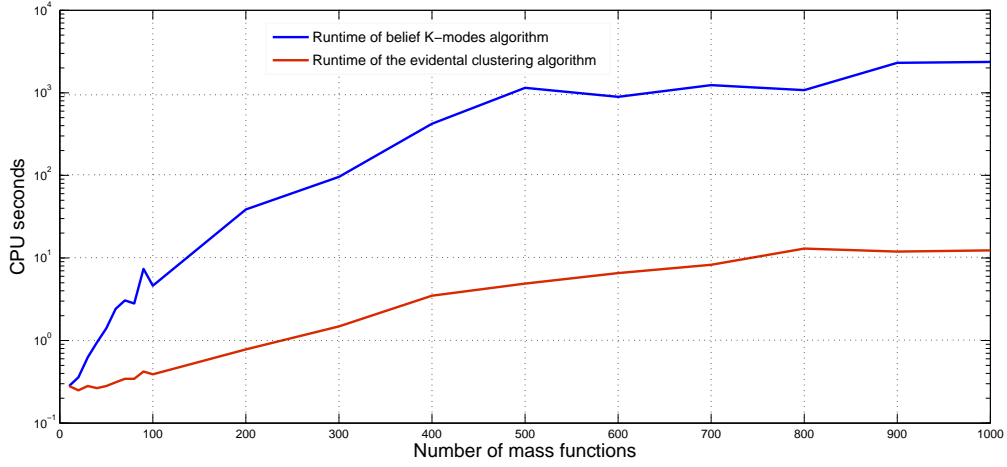


Figure 1: Run-time optimization of the evidential clustering and the belief  $K$ -modes [20] according to  $n \in [10, 1000]$ ,  $|\Omega_{a_j}| = 5$  and  $K = 5$

abilities, two hypotheses  $X$  and  $Y$  are assumed to be statistically independent if  $P(X \cap Y) = P(X) \times P(Y)$  or  $P(X|Y) = P(X)$ . In the context of the theory of belief functions, Shafer [4] defined cognitive and evidential independence.

**Definition 1.** “Two frames of discernment may be called cognitively independent with respect to the evidence if new evidence that bears on only one of them will not change the degree of support for propositions discerned by the other”<sup>1</sup>.

The cognitive independence is a *weak independence*; two variables are independent with respect to a mass function if new evidence that bears on only one of the two variables does not change propositions discerned by the other one. For two variables  $X$  and  $Y$  such that  $\Omega_X$  and  $\Omega_Y$  their domains (frames of discernment) and  $\Omega_X \times \Omega_Y$  the product space of domains  $\Omega_X$  and  $\Omega_Y$ . Variables  $X$  and  $Y$  are cognitively independent with respect to  $m^{\Omega_X \times \Omega_Y}$  if:

$$pl^{\Omega_X \times \Omega_Y}(x, y) = pl^{\Omega_X \times \Omega_Y \downarrow \Omega_X}(x) \times pl^{\Omega_X \times \Omega_Y \downarrow \Omega_Y}(y) \quad (18)$$

---

<sup>1</sup>[4], page 149

Note that  $\Omega_X \times \Omega_Y \downarrow \Omega_X$  is the marginalization of  $\Omega_X \times \Omega_Y$  in  $\Omega_X$  [7, 22].

Shafer [4] defined also a *strong independence* called evidential independence as follows:

**Definition 2.** “Two frames of discernment are evidentially independent with respect to a support function if that support function could be obtained by combining evidence that bears on only one of them with evidence that bears on only the other”.

Two variables are evidentially independent if their joint mass function can be obtained by combining marginal mass functions that bears on each one of them. Variables  $X$  and  $Y$  are evidentially independent with respect to  $m^{\Omega_X \times \Omega_Y}$  if:

$$\begin{cases} pl^{\Omega_X \times \Omega_Y}(x, y) = pl^{\Omega_X \times \Omega_Y \downarrow \Omega_X}(x) \times pl^{\Omega_X \times \Omega_Y \downarrow \Omega_Y}(y) \\ bel^{\Omega_X \times \Omega_Y}(x, y) = bel^{\Omega_X \times \Omega_Y \downarrow \Omega_X}(x) \times bel^{\Omega_X \times \Omega_Y \downarrow \Omega_Y}(y) \end{cases} \quad (19)$$

Independence can also be defined in terms of *irrelevance*. The knowledge of the value of one variable does not change the belief on the other one. In the theory of belief functions, irrelevance is based on the conditioning. Variables  $X$  and  $Y$  are irrelevant with respect to  $m$ ,  $IR_m(X, Y)$  if the marginal mass function on  $X$  is obtained by conditioning the joint mass function on values  $y$  of  $Y$  and marginalizing this conditioned joint mass function on  $X$ :

$$m_{[y]}^{\Omega_X \times \Omega_Y \downarrow \Omega_X}(x) \propto m^{\Omega_X \times \Omega_Y \downarrow \Omega_X}(x) \quad (20)$$

Note that proportionality  $\propto$  is replaced by equality when  $m_{[y]}^{\Omega_X \times \Omega_Y \downarrow \Omega_X}$  and  $m^{\Omega_X \times \Omega_Y \downarrow \Omega_X}$  are normalized.

Doxastic independence is especially proposed in the theory of belief functions by [11, 12] and it is defined as follows:

**Definition 3.** “Two variables are considered as doxastically independent only when they are irrelevant and this irrelevance is preserved under Dempster’s rules of combination”.

In other words, two variables  $X$  and  $Y$  are doxastically independent if they are irrelevant with respect to  $m \oplus m_0$  when they are irrelevant with respect to  $m$  and  $m_0$ . Indeed, if  $X$  and  $Y$  are irrelevant according to any mass function  $m$  and if they are also irrelevant with respect to another mass function  $m_0$ ; they are assumed to be doxastically independent if they are irrelevant with respect to the orthogonal sum of  $m$  and  $m_0$ . Thus, if  $IR_m(X, Y)$ ,  $IR_{m_0}(X, Y)$  and  $IR_{m \oplus m_0}(X, Y)$  is verified then  $X$  and  $Y$  are doxastically independent.

This paper is not focused on variables independence [11, 12, 4] but on *sources independence*. Sources independence is computed according to a set of different belief functions provided by each source separately. Sources are dependent when all their beliefs are correlated, there is a link between all mass functions they provide. This problem is not tackled till now, we noticed a lack of references treating this problem. To study sources independence, a great number of mass functions provided by both sources is needed. This set of mass functions must be defined on the same frame of discernment according to the same problems. For example, two distinct doctors provide  $n$  diagnoses in the examination of the same  $n$  patients. In that case, the frame of discernment contains all diseases and is already the same for both doctors. We define sources independence as follows:

**Definition 4.** Two sources are cognitively independent if they do not communicate and if their evidential corpora are different.

**Definition 5.** Evidential corpus is the set of all pieces of evidence held by a source.

Not only communicating sources are considered dependent but also sources having the same background of knowledge since their beliefs are correlated. The aim of estimating sources independence is either to guide the choice of combination rules when aggregating their beliefs, or to integrate this degree of independence in a new combination rule.

In this paper, mass functions provided by two sources are studied in order to reveal any dependence between them. In the following, we define an independence measure  $I_d$ , ( $I_d(s_1, s_2)$ ), as the independence of  $s_1$  on  $s_2$  verifying the following axioms:

1. Non-negativity: The independence of a source  $s_1$  on another source  $s_2$ ,  $I_d(s_1, s_2)$  cannot be negative, it is either positive or null.
2. Normalization: The degree of independence  $I_d$  is a degree over  $[0, 1]$ , it is null when the first source is dependent on the second one, equal to 1 when it is completely independent and a degree from  $[0, 1]$  otherwise.
3. Non-symmetry: In the case where  $s_1$  is independent on  $s_2$ ,  $s_2$  is not necessarily independent on  $s_1$ . Even if  $s_1$  and  $s_2$  are mutually independent, degrees of independence are not necessarily equal.
4. Identity: Any source is completely dependent on itself and  $I_d(s_1, s_1) = 0$ .

If  $s_1$  and  $s_2$  are independent, there will be no correlation between their mass functions. The main idea of this paper is: First, classify mass functions provided by each source separately. Then, study similarities between cluster partitions to reveal any dependence between sources. By using clustering algorithm, sources overall behavior is studied. The proposed method is in three steps: First, mass functions of each source are classified. Then, similar clusters are matched. Finally, weights of linked clusters and sources independence are quantified.

#### 4.1. Clustering

Clustering algorithm detailed in Section 3 is used to classify two sets of  $n$  mass functions respectively provided by sources  $s_1$  and  $s_2$ . Clustering algorithm is performed on all mass functions of  $s_1$  independently of the clustering performed on those of  $s_2$ . We remind that all mass functions of both sources are defined on the

same frame of discernment and so considered as values of only one attribute when classifying their corresponding objects. For the same example of doctors, patients are objects to classify according to an attribute *disease*. Values of this attribute are mass functions defined on the frame of discernment enumerating all possible diseases. Distance (14) can be simplified as follows because we have only one attribute:

$$D(o_i, Cl_k) = \frac{1}{n_k} \sum_{q=1}^{n_k} d(m_i, m_q) \quad (21)$$

In this paper, we fix the number of clusters to the number of hypotheses in the frame of discernment. In a classification point of view, number of hypotheses is the number of possible classes. For example, the frame of discernment of the attribute *disease* enumerates all possible diseases. Hence, when a doctor examines a patient, he gives a mass function as a classification of the patient in some possible diseases.

#### 4.2. Cluster matching

After clustering technique, both mass functions provided by  $s_1$  and  $s_2$  are distributed separately on  $K$  clusters. In this section, we try to find a mapping between clusters in order to link those containing the same objects. If clusters are perfectly linked, meaning all objects are classified similarly for both sources, we can conclude that sources are dependent as they are choosing similar focal elements (not contradictory at least) when providing mass functions for same objects. If clusters are weakly linked, sources choose similar focal elements for different objects and so they are independent. Clusters independence degree is proportional to the number of objects similarly classified. More clusters contain the same objects, more they are dependent as they are correlated.

We note  $Cl_{k_1}^1$  where  $1 \leq k_1 \leq K$  for clusters of  $s_1$  and  $Cl_{k_2}^2$  where  $1 \leq k_2 \leq K$  for those of  $s_2$ . The similarity between two clusters  $Cl_{k_1}^1$  and  $Cl_{k_2}^2$  is the proportion



of objects simultaneously classified into  $Cl_{k_1}^1$  and  $Cl_{k_2}^2$ :

$$\beta_{k_i k_j}^i = \beta^i(Cl_{k_i}^i, Cl_{k_j}^j) = \frac{|Cl_{k_i}^i \cap Cl_{k_j}^j|}{|Cl_{k_i}^i|} \quad (22)$$

with  $i, j \in \{1, 2\}$  and  $i \neq j$ .  $\beta_{k_1 k_2}^1$  quantifies a proportion of objects classified simultaneously in clusters  $Cl_{k_1}^1$  and  $Cl_{k_2}^2$  with regard to objects in  $Cl_{k_1}^1$ , analogically  $\beta_{k_2 k_1}^2$  is a proportion of objects simultaneously in  $Cl_{k_1}^1$  and  $Cl_{k_2}^2$  with regard to those in  $Cl_{k_2}^2$ . Note that  $\beta_{k_1 k_2}^1 \neq \beta_{k_2 k_1}^2$  since the number of objects classified into  $Cl_{k_1}^1$  and  $Cl_{k_2}^2$  are different ( $|Cl_{k_1}^1| \neq |Cl_{k_2}^2|$ ).

We remind that  $\beta^1$  are similarities towards  $s_1$  and  $\beta^2$  are those towards  $s_2$ . It is obvious that  $\beta^i(Cl_{k_i}^i, Cl_{k_j}^j) = 0$  when  $Cl_{k_i}^i$  and  $Cl_{k_j}^j$  do not contain any common object; however they are completely different.  $\beta^i(Cl_{k_i}^i, Cl_{k_j}^j) = 1$  when these clusters are strongly similar so they contain the same objects. A similarity matrix  $M_1$  containing similarities of clusters of  $s_1$  according to those of  $s_2$  ( $\beta^1$ ), and  $M_2$  the similarity matrix between clusters of  $s_2$  and those of  $s_1$  ( $\beta^2$ ) are defined as follows:

$$M_1 = \begin{pmatrix} \beta_{11}^1 & \beta_{12}^1 & \dots & \beta_{1K}^1 \\ \dots & \dots & \dots & \dots \\ \beta_{k1}^1 & \beta_{k2}^1 & \dots & \beta_{kK}^1 \\ \dots & \dots & \dots & \dots \\ \beta_{K1}^1 & \beta_{K2}^1 & \dots & \beta_{KK}^1 \end{pmatrix} \quad \text{and} \quad M_2 = \begin{pmatrix} \beta_{11}^2 & \beta_{12}^2 & \dots & \beta_{1K}^2 \\ \dots & \dots & \dots & \dots \\ \beta_{k1}^2 & \beta_{k2}^2 & \dots & \beta_{kK}^2 \\ \dots & \dots & \dots & \dots \\ \beta_{K1}^2 & \beta_{K2}^2 & \dots & \beta_{KK}^2 \end{pmatrix} \quad (23)$$

We note that  $M_1$  and  $M_2$  are different since  $\beta_{k_1 k_2}^1 \neq \beta_{k_2 k_1}^2$ . Clusters of  $s_1$  are matched to those of  $s_2$  according to maximum of  $\beta^1$  such that each cluster  $Cl_{k_1}^1$  is linked to only one cluster  $Cl_{k_2}^2$  and each cluster  $Cl_{k_2}^2$  has only one cluster  $Cl_{k_1}^1$  linked to it. The idea is to link iteratively clusters having the maximal  $\beta^1$  in  $M_1$  then eliminate these clusters and the corresponding line and column from the matrix until having a bijective cluster matching. Algorithm 1 details cluster matching process. We note that different matchings are obtained for  $s_1$  and  $s_2$  because  $M_1$  and  $M_2$

are different. This algorithm is iterative and the number of iteration is equal to the

---

**Algorithm 1** Cluster matching

---

**Require:** Similarity matrix  $M$ .

- 1: **while**  $M$  is not empty **do**
  - 2: Find  $\max(M)$  and indexes  $c$  and  $l$  of clusters having this maximal similarity.
  - 3: Map clusters  $l$  and  $c$ .
  - 4: Delete line  $l$  and column  $c$  from  $M$ .
  - 5: **end while**
  - 6: **return** Cluster matching.
- 

number of clusters  $K$ . Even if this algorithm is quite simple, it provides a matching of clusters in order to compare evidential information provided by both sources. The assignment algorithm proposed in [23] for square matrices and that for rectangular matrices [24] can also be used to minimize the dissimilarity between matched clusters. Other methods for cluster matching [25] and [26] can also be used.

#### 4.3. Cluster independence

Once cluster matching is obtained, a degree of independence/dependence of matched clusters is quantified in this step. A set of matched clusters is obtained for both sources and a mass function can be used to quantify each couple of clusters independence. Assume that cluster  $Cl_{k_1}^1$  is matched to  $Cl_{k_2}^2$ , a mass function  $m^{\Omega_I^2}$  defined on the frame of discernment  $\Omega_I = \{Dependent\ Dep, Independent\ Ind\}$  describes how much this couple of clusters is independent or dependent as follows:

$$\begin{cases} m_{k_i k_j}^{\Omega_I, i}(Dep) = \alpha_{k_i}^i \beta_{k_i k_j}^1 \\ m_{k_i k_j}^{\Omega_I, i}(Ind) = \alpha_{k_i}^i (1 - \beta_{k_i k_j}^1) \\ m_{k_i k_j}^{\Omega_I, i}(Dep \cup Ind) = 1 - \alpha_{k_i}^i \end{cases} \quad (24)$$

---

<sup>2</sup>We note the frame of discernment in the mass functions to avoid confusion.

A mass function quantifies the degree of independence of each couple of clusters according to each source;  $m_{k_i k_j}^{\Omega, i}$  is a mass function for the independence of each linked clusters  $Cl_{k_i}^i$  and  $Cl_{k_j}^j$  according to  $s_i$  with  $i, j \in \{1, 2\}$  and  $i \neq j$ . Coefficient  $\alpha_{k_i}^i$  is used to take into account of number of mass functions in each cluster  $Cl_{k_i}^i$  of the source  $i$ . Reliability factor  $\alpha_{k_i}^i$  is not the reliability of any source but it can be seen as the reliability of the clusters independence estimation. Consequently, independence estimation is more reliable when clusters contain enough mass functions. For example, assume two clusters; one containing only one mass function and the second one containing 100 mass functions. It is obvious that the independence estimation of the second cluster is more precise and significant than the independence estimation of the first one.

Reliability factors  $\alpha_{k_i}^i$  are proportional to the number of hypotheses in the frame of discernment  $|\Omega|$ , and the number of objects classified in  $Cl_{k_i}^i$  as follows:

$$\alpha_{k_i}^i = f(|\Omega|, |Cl_{k_i}^i|) \quad (25)$$

The bigger  $|\Omega|$  is, the more mass functions are needed to have a reliable cluster independence estimation. For example, if  $|\Omega| = 5$  then there are  $2^5$  possible focal elements, also independence estimation of a cluster containing 20 objects cannot be precise. No existing method to define such function  $f$ . Hence, we use simple heuristics as follows:

$$\alpha_{k_i}^i = 1 - \frac{1}{|Cl_{k_i}^i|^{\frac{1}{|\Omega|}}} \quad (26)$$

As shown in figure 2, if  $|\Omega|$  and number of mass functions in a cluster are big enough, cluster independence mass function is almost not discounted. Reliability factor is an increasing function of  $|\Omega|$  and  $|Cl_{k_i}^i|$  which favors big clusters<sup>3</sup>.

---

<sup>3</sup>Big clusters are those containing enough mass functions according to  $|\Omega|$ .

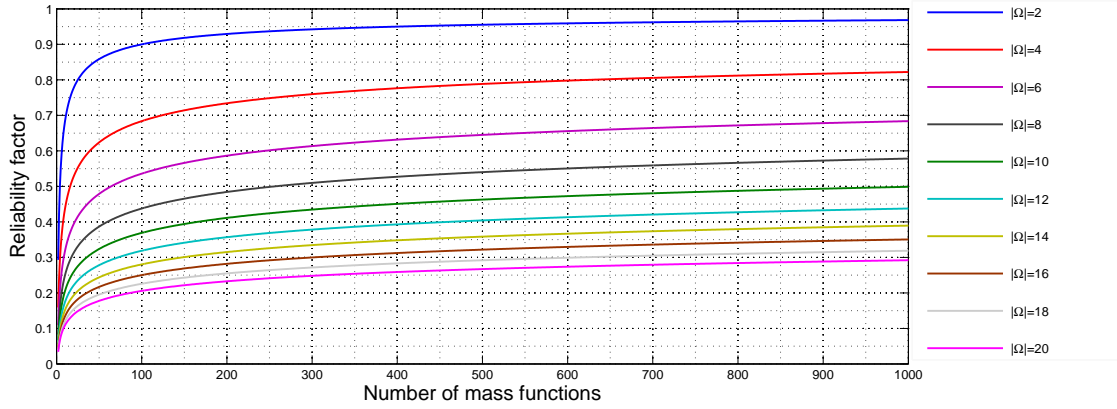


Figure 2: Reliability factors  $\alpha_{k_i}^i$

#### 4.4. Sources independence

Obtained mass functions quantify each matched clusters independence according to each source. Therefore,  $K$  mass functions are obtained for each source such that each mass function quantifies the independence of each couple of matched clusters. The combination of  $K$  mass functions for each source using the mean, defined by equation (12), is a mass function  $m^{\Omega_I}$  defining the whole independence of one source on another one:

$$m^{\Omega_I, s_i}(A) = \frac{1}{K} \sum_{k_i=1}^K m_{k_i k_j}^{\Omega_I, i}(A) \quad \forall A \subseteq 2^\Omega \quad (27)$$

With  $k_j$  is the cluster matched to  $k_i$  according to  $s_j$ . Two different mass functions  $m^{\Omega_I, s_1}$  and  $m^{\Omega_I, s_2}$  are obtained for  $s_1$  and  $s_2$  respectively. We note that  $m^{\Omega_I, s_1}$  is the combination of  $K$  mass functions representing the independence of matched clusters according to  $s_1$  defined using equation (24). Mass functions  $m^{\Omega_I, s_1}$  and  $m^{\Omega_I, s_2}$  are different since cluster matchings are different which verifies the axiom of non-symmetry.  $\beta_{k_1 k_2}^1, \beta_{k_2 k_1}^2 \in [0, 1]$  verify the non-negativity and the normalization axioms. Finally, pignistic probabilities are computed from these mass functions in

order to decide about sources independence  $I_d$  such that:

$$\begin{cases} I_d(s_1, s_2) = \text{BetP}(Ind) \\ \bar{I}_d(s_1, s_2) = \text{BetP}(Dep) \end{cases} \quad (28)$$

If  $I_d(s_1, s_2) > \bar{I}_d(s_1, s_2)$  we claim that sources  $s_1$  and  $s_2$  are independent otherwise they are dependent.

#### 4.5. General case

The method detailed above estimates the independence of one source on another one. Independence measure is non-symmetric because if a source  $s_1$  is independent on a source  $s_2$  then  $s_2$  is not necessarily independent on  $s_1$  and even if it is the case, degrees of independence are not necessarily the same.

It is wise to choose the minimum independence from  $I_d(s_1, s_2)$  and  $I_d(s_2, s_1)$  as the overall independence. Consequently, if at least one of two sources is dependent on the other, then sources are considered dependent. In other words, two sources are independent only if they are mutually independent. Hence, overall independence that is denoted  $I(s_1, s_2)$  is given by:

$$I(s_1, s_2) = \min(I_d(s_1, s_2), I_d(s_2, s_1)) \quad (29)$$

We note that  $I(s_1, s_2)$  is non-negative, normalized, symmetric and identical.

We define an independence measure, noted  $I$ , generalizing the independence for more than two sources verifying the following axioms:

1. Non-negativity: Many sources independence  $\{s_1, s_2, s_3, \dots, s_{ns}\}$ , noted  $I_d(s_1, s_2, \dots, s_{ns})$  cannot be negative, it is either positive or null.
2. Normalization: Sources independence  $I$  is a degree in  $[0, 1]$ . The minimum 0 is reached when sources are completely dependent and the maximum 1 is reached when they are completely independent.

3. Symmetry:  $I(s_1, s_2, s_3, \dots, s_{ns})$  is the sources' overall independence and  $I(s_1, s_2, s_3, \dots, s_{ns}) = I(s_2, s_1, s_3, \dots, s_{ns}) = I(s_3, s_1, s_2, \dots, s_{ns})$ .
4. Identity:  $I(s_1, s_1, s_1) = 0$ . It is obvious that any source is completely dependent on itself.
5. Increasing with inclusion:  $I(s_1, s_2) \leq I(s_1, s_2, s_3)$ , more there are sources, more they are likely to be independent.

To compute the overall independence of  $ns$  sources  $\{s_1, s_2, \dots, s_{ns}\}$ , independencies of pairs of sources are computed and the maximum<sup>4</sup> independence is the sources overall independence:

$$I(s_1, s_2, \dots, s_{ns}) = \max(I(s_i, s_j)), \quad \forall i \in [1, ns], j \in [i, ns] \quad (30)$$

or equivalently:

$$I(s_1, s_2, \dots, s_{ns}) = \max(\min(I_d(s_i, s_j), I_d(s_j, s_i))), \quad \forall i, j \in [1, ns] \quad i \neq j \quad (31)$$

Independence degree of sources is then integrated in the combination step using the following mixed combination rule.

## 5. Combination rule

Combination rules using conjunctive and/or disjunctive rules such as [2, 5, 6, 7, 8] are used when sources are completely independent but cautious and bold rules [10] tolerate redundant information and consequently can be used to combine mass functions which sources are dependent. In the combination step, sources dependence or independence hypothesis is intuitively made without any possibility of check. Sources independence degree is neither 0 nor 1 but a level over  $[0, 1]$ .

---

<sup>4</sup>The maximum is used to insure the property of increasing with inclusion.

The main question is “which combination rule to use when combining partially independent\dependent mass functions?”

In this paper, we propose a new mixed combination rule using conjunctive and cautious rules detailed in equations (8) and (13). In the case of totally dependent sources (where independence is 0), the cautious and proposed mixed combination rules are similar; whereas in the case of totally independent sources (independence is 1), the conjunctive and proposed combination rules are similar. In the case of an independence degree in  $]0, 1[$ , combined mass function is the average of conjunctive and cautious combinations weighted by sources’ independence degree.

Assume that two sources  $s_1$  and  $s_2$  are independent with a degree  $\gamma$  such that  $\gamma = I(s_1, s_2)$ ;  $m_1$  and  $m_2$  are mass functions provided by  $s_1$  and  $s_2$ . The proposed mixed combination rule is defined as follows:

$$m_{Mixed}(A) = \gamma * m_{\odot}(A) + (1 - \gamma) * m_{\wedge}(A), \quad \forall A \subseteq \Omega \quad (32)$$

The degree of independence of a set of sources is given by equation (30), and the mixed combination of a set of mass functions  $\{m_1, m_2, \dots, m_{ns}\}$  provided by sources  $\{s_1, s_2, \dots, s_{ns}\}$  is also a weighted average such that:

$$\gamma = I(s_1, s_2, \dots, s_{ns}) \quad (33)$$

**Properties of the proposed mixed combination rule:**

- **Commutativity:** Conjunctive and cautious rules are commutative. Independence measure is symmetric because sources’ degree of independence is the same for a set of sources. Then the proposed rule is commutative.
- **Associativity:** Conjunctive and cautious rule are associative but the proposed rule is not because independence degree of  $n$  sources and  $n + 1$  ones is not necessarily the same.

- Idempotent: Degree of independence of one source to itself is 0, in that case the proposed rule is equivalent to the cautious rule. As the cautious rule is idempotent, it is the case of the proposed mixed rule.
- Neutral element: Mixed combination rule does not have any neutral element.
- Absorbing element: No absorbing element also.

*Example.* Assume a frame of discernment  $\Omega = \{a, b, c\}$  and two sources  $s_1$  and  $s_2$  providing two mass functions  $m_1$  and  $m_2$ . Table 1 illustrates conjunctive and cautious combinations as well as mixed combination in the cases where  $\gamma = 0$ ,  $\gamma = 0.3$ ,  $\gamma = 0.6$  and  $\gamma = 1$ . When  $\gamma = 0$ , mixed and cautious combinations are equivalent; when  $\gamma = 1$ , mixed and conjunctive combinations are equivalent, otherwise it is a weighted average by  $\gamma \in ]0, 1[$ .

Finally, to illustrate the proposed mixed combination rule and compare it to other combination rules, three mass functions are generated randomly using algorithm 2. These mass functions are combined with conjunctive, Dempster, Yager, disjunctive, cautious and mean combination rules. They are also combined with the mixed combination rule with different independence levels.

Figure 3 illustrates distances<sup>5</sup> between the mixed combination with several degrees of independence and combined mass functions using conjunctive, Dempster, Yager, disjunctive, cautious and mean combination rules. Distances between mixed combination with several independence degrees; and Yager, disjunctive, mean and Dempster's rules are linear and decreasing proportionally to  $\gamma$ .

---

<sup>5</sup>Jousselme distance detailed in equation (16).



Table 1: Combination of two mass functions

$2^\Omega$	$m_1$	$m_2$	$m_{\wedge}$	$m_{\ominus}$	$m_{Mixed}$ $\gamma = 0$	$m_{Mixed}$ $\gamma = 0.3$	$m_{Mixed}$ $\gamma = 0.6$	$m_{Mixed}$ $\gamma = 1$
$\emptyset$	0	0	0.1071	0.06	0.1071	0.093	0.0789	0.06
$a$	0.3	0.3	0.2679	0.45	0.2679	0.3225	0.3771	0.45
$b$	0	0	0	0	0	0	0	0
$a \cup b$	0	0	0	0	0	0	0	0
$c$	0.2	0	0.1786	0.14	0.1786	0.167	0.1554	0.14
$a \cup c$	0.2	0.4	0.2551	0.26	0.2551	0.2566	0.2580	0.26
$b \cup c$	0	0	0	0	0	0	0	0
$a \cup b \cup c$	0.3	0.3	0.1913	0.09	0.1913	0.1609	0.1305	0.09

## 6. Experiments

Because of the lack of real evidential data, we use generated mass functions to test the method detailed above. Moreover, it is difficult to simulate all situations with all possible combinations of focal elements for several degrees of independence between sources. First, we generate two sets of mass functions for two sources  $s_1$  and  $s_2$ ; then we illustrate for three sources.

### 6.1. Generated data depiction

Generating sets of  $n$  mass functions for several sources depends on sources independence. We discern cases of independent and dependent sources.

#### 6.1.1. Independent sources

In general, to generate mass functions some information are needed: the number of hypotheses in the frame of discernment,  $|\Omega|$  and the number of mass func-

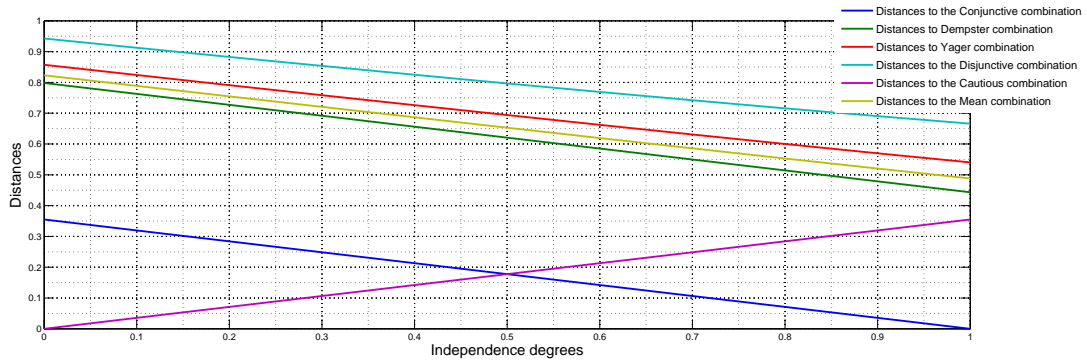


Figure 3: Distances between combined mass functions

tions. We note that number of focal elements, and masses are chosen randomly.

In the case of independent sources, masses can be anywhere and focal elements of both sources are chosen independently. Mass functions of  $s_1$  and  $s_2$  are generated following algorithm 2. We note that focal elements, their number and BBMs are chosen randomly according to the universal law.

---

**Algorithm 2** Independent mass functions generating

---

**Require:**  $|\Omega|, n$  : number of mass functions

- 1: **for**  $i = 1$  to  $n$  **do**
  - 2:   Choose randomly  $|F|$ , the number of focal elements on  $[1, |2^\Omega|]$ .
  - 3:   Choose randomly  $|F|$  focal elements noted  $F$ .
  - 4:   Divvy the interval  $[0, 1]$  into  $|F|$  continuous sub-intervals.
  - 5:   Focal elements BBMs are intervals sizes.
  - 6: **end for**
  - 7: **return**  $n$  mass functions
- 

### 6.1.2. Dependent sources

The case of dependent sources is a bit difficult to simulate as several scenarios can occur. In this section, we will try to illustrate the most common situations.

Generated mass functions for dependent sources are supposed to be consistent and do not enclose any internal conflict [27]. Consistent mass functions contain at least one focal element common to all focal sets. Figure 4 illustrates a consistent mass function where all focal elements  $\{A, B, C, D\}$  intersect.

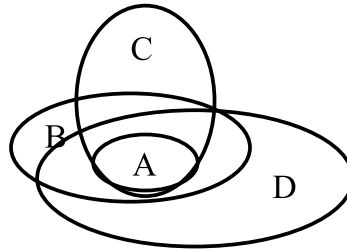


Figure 4: Consistent belief function

Algorithm 3 generates a set of  $n$  consistent mass functions<sup>6</sup> defined on a frame of discernment of size  $|\Omega|$ . In the case of dependent sources, they are almost consistent and at least one of them is dependent on the other. To simulate the case where one source is dependent on another one, consistent mass functions of the first one are generated following algorithm 3, then those of the second source are generated knowing decisions of the first one. Algorithm 4 generates a set of mass functions that are dependent on another set of mass functions. Dependence is due to the knowledge of other source's decisions.

## 6.2. Results of tests

Algorithms detailed in the previous section are used to test some cases of sources' dependence and independence. We note that in extreme cases where mass functions are certain or even when focal elements do not intersect; maximal values of independence are obtained. In the case of perfect dependence; mass functions

---

<sup>6</sup>Conflict within such mass functions is null.

---

**Algorithm 3** Consistent mass functions generating

---

**Require:**  $|\Omega|, n$  : number of mass functions

- 1: **for**  $i = 1$  to  $n$  **do**
  - 2:   Choose randomly a focal set  $\omega_i$  (it can be a single point) from  $\Omega$ .
  - 3:   Find the set  $S$  of all focal sets including  $\omega_i$ .
  - 4:   Choose randomly  $|F|$ , the number of focal elements on  $[1, |S|]$ .
  - 5:   Choose randomly  $|F|$  focal elements from  $S$  noted  $F$ .
  - 6:   Divvy the interval  $[0, 1]$  into  $|F|$  continuous sub-intervals.
  - 7:   BBMs of focal elements are intervals sizes.
  - 8: **end for**
  - 9: **return**  $n$  consistent mass functions
- 

have the same focal elements; however, clusters contain mass functions with consistent focal elements. Clustering is performed according to focal elements and clusters are perfectly linked.

### 6.2.1. Independent sources

In this paragraph, mass functions are independent. Focal elements and BBMs are randomly chosen ensuing algorithm 2. For tests, we choose  $|\Omega| = 5$  which is considered as medium-sized frame of discernment and  $n = 100$ . Table 2 illustrates the mean of 100 tests in the case of independent sources. The mean of 100 tests for two dependent sources yields to a degree of independence  $\gamma = 0.68$ , thus sources are independent. Assume that  $m_1$  and  $m_2$ , given in table 1, are provided by two sources  $s_1$  and  $s_2$  which independence degree is given in table 2. Combination of  $m_1$  and  $m_2$  is given in table 3.

To illustrate the case of three independent sources, three sets of 100 independent mass functions are generated following algorithm 2 with  $|\Omega| = 5$ . The mean of 100 tests are illustrated in table 4.

---

**Algorithm 4** Dependent mass functions generating

---

**Require:**  $|\Omega|, n$  : number of mass functions,  $d$  decision of another source

- 1: **for**  $i = 1$  to  $n$  **do**
  - 2: Find the set  $S$  of all focal sets including  $d$ .
  - 3: Choose randomly  $|F|$ , the number of focal elements on  $[1, |S|]$ .
  - 4: Choose randomly  $|F|$  focal elements from  $S$  noted  $F$ .
  - 5: Divvy the interval  $[0, 1]$  into  $|F|$  continuous sub-intervals.
  - 6: Focal elements BBMs are intervals sizes.
  - 7: **end for**
  - 8: **return**  $n$  consistent mass functions
- 

Table 2: Mean of 100 tests on 100 generated mass functions for two sources

Dependence type	Degree of independence	Overall independence
Independence	$I_d(s_1, s_2) = 0.68, \bar{I}_d(s_1, s_2) = 0.32$	$\gamma = 0.68$
	$I_d(s_2, s_1) = 0.68, \bar{I}_d(s_2, s_1) = 0.32$	
Dependence	$I_d(s_1, s_2) = 0.34, \bar{I}_d(s_1, s_2) = 0.66$	$\gamma = 0.34$
	$I_d(s_2, s_1) = 0.35, \bar{I}_d(s_2, s_1) = 0.65$	

### 6.2.2. Dependent sources

In the case of dependent sources, mass functions are generated ensuing algorithms 3 and 4. For tests, we choose  $|\Omega| = 5$  and  $n = 100$ . We generate 100 mass functions of both  $s_1$  and  $s_2$  for 100 times and then compute the average of  $I_d(s_1, s_2)$ ,  $I_d(s_2, s_1)$  and  $I(s_1, s_2)$ . Table 2 illustrates the mean of 100 independence degrees of two dependent sources providing each one 100 randomly generated mass functions. These sources are dependent with a degree  $1 - \gamma = 0.66$ . In table 3,  $m_1$  and  $m_2$  are combined using the mixed rule when  $\gamma = 0.34$ .

To illustrate the case of three dependent sources, three sets of 100 dependent mass

Table 3: Mixed combination of  $m_1$  and  $m_2$ 

$2^\Omega$	$m_1$	$m_2$	$m_{Mixed}$ $\gamma = 0.68$	$m_{Mixed}$ $\gamma = 0.34$
$\emptyset$	0	0	0.092	0.076
$a$	0.3	0.3	0.3262	0.3881
$b$	0	0	0	0
$a \cup b$	0	0	0	0
$c$	0.2	0	0.1662	0.1531
$a \cup c$	0.2	0.4	0.2567	0.2583
$b \cup c$	0	0	0	0
$a \cup b \cup c$	0.3	0.3	0.1589	0.1244

functions are generated following algorithms 3 and 4 when  $|\Omega| = 5$ . The mean of 100 degrees of independence are illustrated in table 5.

Finally, assume that  $m_1$ ,  $m_2$  and  $m_3$  of table 6 are three mass functions defined on a frame of discernment  $\Omega = \{a, b, c\}$  and provided by three dependent sources. The mixed combined mass function when their degree of independence is  $\gamma = 0.35$  is also given in table 6.

## 7. Conclusion

In this paper, we proposed a method to learn sources cognitive independence in order to use the appropriate combination rule either when sources are cognitively dependent or independent. Sources are cognitively independent if they are different; not communicating and they have distinct evidential corpora. The proposed statistical approach is based on a clustering algorithm applied to mass functions

Table 4: Mean of 100 tests on 100 generated mass functions for three independent sources

Sources	Degree of independence	Pairwise independence	Overall independence
$s_1-s_2$	$I_d(s_1, s_2) = 0.67, \bar{I}_d(s_1, s_2) = 0.33$	$I(s_1, s_2) = 0.67$	$\gamma = 0.68$
	$I_d(s_2, s_1) = 0.67, \bar{I}_d(s_2, s_1) = 0.33$		
$s_1-s_3$	$I_d(s_1, s_3) = 0.68, \bar{I}_d(s_1, s_3) = 0.32$	$I(s_1, s_3) = 0.68$	
	$I_d(s_3, s_1) = 0.68, \bar{I}_d(s_3, s_1) = 0.32$		
$s_2-s_3$	$I_d(s_2, s_3) = 0.68, \bar{I}_d(s_2, s_3) = 0.32$	$I(s_2, s_3) = 0.68$	
	$I_d(s_3, s_2) = 0.68, \bar{I}_d(s_3, s_2) = 0.32$		

provided by several sources. A pair of sources independence is deduced from weights of linked clusters after a matching of their clusters. Independence degree of sources can either guide the choice of the combination rule if it is either 1 or 0; when it is a degree over  $]0, 1[$ , we propose a new combination rule that weights the conjunctive and cautious combinations with sources' independence degree.

## References

- [1] L. A. Zadeh, Fuzzy sets, *Information and Control* 8 (3) (1965) 338–353.
- [2] D. Dubois, H. Prade, Representation and combination of uncertainty with belief functions and possibility measures, *Computational Intelligence* 4 (3) (1988) 244–264.
- [3] A. P. Dempster, Upper and lower probabilities induced by a multivalued mapping, *The Annals of Mathematical Statistics* 38 (2) (1967) 325–339.

Table 5: Mean of 100 tests on 100 generated mass functions for three dependent sources

Sources	Degree of independence	Pairwise independence	Overall independence
$s_1-s_2$	$I_d(s_1, s_2) = 0.35, \bar{I}_d(s_1, s_2) = 0.65$	$I(s_1, s_2) = 0.34$	$\gamma = 0.35$
	$I_d(s_2, s_1) = 0.34, \bar{I}_d(s_2, s_1) = 0.66$		
$s_1-s_3$	$I_d(s_1, s_3) = 0.32, \bar{I}_d(s_1, s_3) = 0.68$	$I(s_1, s_3) = 0.31$	
	$I_d(s_3, s_1) = 0.31, \bar{I}_d(s_3, s_1) = 0.69$		
$s_2-s_3$	$I_d(s_2, s_3) = 0.36, \bar{I}_d(s_2, s_3) = 0.64$	$I(s_2, s_3) = 0.35$	
	$I_d(s_3, s_2) = 0.35, \bar{I}_d(s_3, s_2) = 0.65$		

- [4] G. Shafer, A mathematical theory of evidence, Princeton University Press, 1976.
- [5] A. Martin, C. Osswald, Toward a combination rule to deal with partial conflict and specificity in belief functions theory, in: International Conference on Information Fusion, Québec, Canada, 2007, pp. 1–8.
- [6] C. K. Murphy, Combining belief functions when evidence conflicts, Decision Support Systems 29 (1) (2000) 1–9.
- [7] P. Smets, R. Kennes, The transferable belief model, Artificial Intelligence 66 (2) (1994) 191–234.
- [8] R. R. Yager, On the Dempster-Shafer framework and new combination rules, Information Sciences 41 (2) (1987) 93–137.
- [9] E. Lefèvre, Z. Elouedi, How to preserve the conflict as an alarm in the combination of belief functions?, Decision Support Systems 56 (2013) 326–333.



Table 6: Mixed combination of  $m_1$ ,  $m_2$  and  $m_3$

$2^\Omega$	$m_1$	$m_2$	$m_3$	$m_{Mixed}$ $\gamma = 0.35$
$\emptyset$	0	0	0	0
$a$	0	0	0	0
$b$	0	0	0	0
$a \cup b$	0	0	0	0
$c$	0.03	0.05	0	0.32
$a \cup c$	0.39	0.07	0.04	0.24
$b \cup c$	0.3	0.47	0.22	0.29
$a \cup b \cup c$	0.28	0.41	0.74	0.15

- [10] T. Denœux, Conjunctive and disjunctive combination of belief functions induced by nondistinct bodies of evidence, *Artificial Intelligence* 172 (2-3) (2008) 234–264.
- [11] B. Ben Yaghlane, P. Smets, K. Mellouli, Belief function independence: I. The marginal case, *International Journal of Approximate Reasoning* 29 (1) (2002) 47–70.
- [12] B. Ben Yaghlane, P. Smets, K. Mellouli, Belief function independence: II. The conditional case, *International Journal of Approximate Reasoning* 31 (1-2) (2002) 31–75.
- [13] P. Smets, Belief functions: The disjunctive rule of combination and the generalized Bayesian theorem, *International Journal of Approximate Reasoning* 9 (1) (1993) 1–35.

- [14] P. Smets, The combination of evidence in the transferable belief model, *IEEE Transactions on Pattern Analysis and Machine Intelligence* 12 (5) (1990) 447–458.
- [15] P. Smets, The canonical decomposition of a weighted belief, in: *International Joint Conference on Artificial Intelligence*, Vol. 2, Morgan Kaufman, Montréal, Québec, Canada, 1995, pp. 1896–1901.
- [16] L. A. Zadeh, A mathematical theory of evidence (book review), *AI magazine* 5 (3) (1984) 81–83.
- [17] P. Smets, The nature of the unnormalized beliefs encountered in the transferable belief model, in: D. Dubois, M. P. Wellman (Eds.), *International conference on Uncertainty in Artificial Intelligence*, Morgan Kaufmann, Stanford, California, USA, 1992, pp. 292–297.
- [18] A. Martin, A.-L. Jousselme, C. Osswald, Conflict measure for the discounting operation on belief functions, in: *International Conference on Information Fusion*, Cologne, Germany, 2008, pp. 1–8.
- [19] A.-L. Jousselme, D. Grenier, E. Bossé, A new distance between two bodies of evidence, *Information Fusion* 2 (2) (2001) 91–101.
- [20] S. Ben Hariz, Z. Elouedi, K. Mellouli, Clustering approach using belief function theory, in: J. Euzenat, J. Domingue (Eds.), *7th Conference of the European Society for Fuzzy Logic and Technology*, Vol. 4183 of *Lecture Notes in Computer Science*, Atlantis Press, Varna, Bulgaria, 2006, pp. 162–171.
- [21] M. Chebbah, A. Martin, B. Ben Yaghlane, About sources dependence in the theory of belief functions, in: T. Dencœux, M.-H. Masson (Eds.), *International Conference on Belief Functions*, Vol. 164 of *Advances in Intelligent and Soft*

- Computing, Springer Berlin Heidelberg, Compiègne, France, 2012, pp. 239–246.
- [22] P. Smets, R. Kruse, *Uncertainty Management in Information Systems: From Needs to Solutions*, Springer US, Boston, 1997, Ch. The Transferable Belief Model for Belief Representation, pp. 343–368.
- [23] J. Munkres, Algorithms for the Assignment and Transportation Problems, *Journal of the Society for Industrial and Applied Mathematics* 5 (1) (1957) 32–38.
- [24] F. Bourgeois, J.-C. Lassalle, An Extension of the Munkers Algorithm for the Assignment Problem to Rectangular Matrices, *Communication of the ACM* 12 (14) (1971) 802–804.
- [25] C. Wemmert, P. Gañarski, A multi-view voting method to combine unsupervised classifications, in: *IASTED International Conference on Artificial Intelligence and Applications*, Málaga, Spain, 2002, pp. 447–453.
- [26] P. Gañarski, C. Wemmert, Collaborative multi-strategy classification: Application to per-pixel analysis of images, in: *International Workshop on Multimedia Data Mining: Mining Integrated Media and Complex Data*, Chicago, Illinois, USA, 2005, pp. 15–22.
- [27] M. Daniel, Conflicts within and between belief functions, in: *IPMU, 2010*, pp. 696–705.