A new measure of dissimilarity between two basic belief assignments

Zhun-Ga Liu, Jean Dezert, Quan Pan

A new measure of dissimilarity between two basic belief assignments

Zhun-ga Liu[1], Jean Dezert[2], Quan Pan[1]
1. School of Automation, Northwestern Polytechnical University, Xi'an, China, Email: liuzhunga@gmail.com
2. French Aerospace Lab, 29 Avenue de la Division Leclerc, 92320 Châtillon, France, Email: jean.dezert@onera.fr

***Abstract*-The measures of dissimilarity between basic belief assignments (bba's) in the framework of the theories of evidence have been regularly studied in the literature until recently. Nevertheless, the question is still open and it is very difficult to represent efficiently the dissimilarity between bba's through a single scalar measure. In this paper, we analyze the limitation of conflicting beliefs and the classical distance when they are used directly as dissimilarity measures. The main problem for evaluating the dissimilarity between two bba's lies in the relationship among their focal elements. Probabilistic transformations allow to approximate any bba into a subjective probability measure based on an underlying frame of discernment whose atomic elements are exhaustive and exclusive. In this paper, we propose to use both a distance based on probabilistic transformations (mainly the pignistic transformation $BetP$ and the Dezert-Smarandache transformation $DSmP$) and also a conflict coefficient in order to charracterize and measure the dissimilarity between two bba's. Our approach takes into account both the difference between bba's (through the probabilistic distance) and the degree of divergence (through the conflict coefficient) of hypothesis that two belief functions strongly support. These two aspects of disimilarity are complementary in the evaluation of our dissimilarity measure. The method proposed in this paper is applied for evaluating the reliability of sources of evidence and selecting a rule of combination rule of bba's. Simple numerical examples are given to illustrate the interest of the proposed approach.**

Keywords: belief functions, dissimilarity measure, evidential distance, conflict coefficient, pignistic transformation, DST, DSmT.

I. Introduction

The theories of evidence [15], [16], also called theories of belief functions, are widely used in information fusion as soon as the information to deal with are uncertain and possibly conflicting and represented by basic belief assignments (bba's). The serach for an efficient measure of dissimilarity between two bba's is still an open and challenging question. In last years, many works on the dissimilarity have emerged, and a many proposals for the definition of a measure dissimilarity have been proposed [7], [8], [10], [14]. The dissimilarity measures are usually applied for evaluating fusion algorithms or for the optimization of fusion systems [8]. For instance, this appears in [1] for belief functions approximation algorithms, in [2], [6] for defining the agreement between sources of evidence as a basis for discounting factors, or also in [11] in a criterion for selecting an adapted rule of combination. The dissimilarity between two bba's is actually difficult to quantify because several aspects of dissimilarity need to be involved when establishing a real efficient and precise measure of dissimilarity. In previous published works, a particular attention has been paid to search a scalar measure to represent the dissimilarity measure, but these proposed measures did capture only one aspect of the dissimilarity between bba's mainly associated with a distance metric. From authors opinion, the dissimilarity between two bba's is not only represented by some well chosen distance between bba's, but also by another aspect which reflects the level of conflict between the bba's. So both aspects must count when defining a measure of dissimilarity between bba's. So, the basic idea presented in this paper is to define the dissimilarity measure between two bba's from both a distance[1] between the bba's, and also from their intrinsic level of conflict. The proposed distance measures the mathematical difference between the belief assignments whereas the conflict measures the degree of divergences of hypothesis the sources of evidence strongly support. Therefore, the distance and the conflict captures and measure the two different aspect of the dissimilarity between two bba's, and they are mutually compensable in a certain sense.

The degree of conflict is generally used to evaluate the interaction between conflicting beliefs [18], but it is not appropriate in some cases (especially for the two equal belief functions when they can be considered as cognitively independent). Moreover, the conflicting beliefs can't precisely reflect the divergence of the hypothesis that the belief functions strongly support. An evidential distance proposed by Jousselme et al. in [7] is commonly considered as an interesting and valuable distance measure since it takes both into account the value of the mass of belief and the relative specificity of focal elements of each bba. This distance measure however is not good enough to capture the different aspects of the dissimilarity between bba's as it will be clearly shown in our examples 1 and 2 in the sequel. Moreover its computation burden can become a bit expensive in time and in memory requirements. When working in the probabilistic framework, the atomic

[1]For simplicity, we suggest to use the L_1 distance between the approximate subjective probability measures of the bba's obtained with $BetP$ or $DSmP$ transfromations.

elements are exclusive and independent, and the degree of the conflict and distance become easier to measure regardless the intrinsic relationship between bba's. The two most common probabilistic transformations are $BetP$ (proposed by Smets in [19]) and $DSmP$ (proposed by Smarandache and Dezert in [4]). These transformations allow to approximate any bba into a subjective probability measure. $BetP$ approximates the bba in a prudent way, and that's why we call it a pessimistic probabilistic transformation. The approximation done with $DSmP$ transformation is more specific and it allows to reach the highest probabilistic information content (PIC) [20] and so $DSmP$ is more satisfactory from the theoretical point of view. $DSmP$ can thus be considered as an optimistic probabilistic approximation of a bba. In this paper, we define the dissimilarity measure from both the level of conflict and from the distance between bba's based on $BetP$ and $DSmP$ transformations. The interest of the new measurements are illustrates by some numerical examples, and the dissimilarity mesure is shown for the selection of an adapted rule of combination among Dempster's rule and its main alternatives, and for the determination of the reliability factor for the sources of evidence to be discounted.

II. Preliminaries

A. Basics of Dempster-Shafer theory (DST)

DST [15] is based on a given set $\Theta = \{\theta_1, \theta_2, \dots, \theta_n\}$ of n mutually exclusive and exhaustive elements θ_i, $i = 1, 2, \dots, n$. Θ is called the frame of discernment of the fusion problem. The set of all subsets of Θ is called the power set of Θ, and it is denoted 2^Θ. For instance, if $\Theta = \{\theta_1, \theta_2, \theta_3\}, 2^\Theta = \{\emptyset, \{\theta_1\}, \{\theta_2\}, \{\theta_3\}, \{\theta_1, \theta_2\}, \{\theta_1, \theta_3\}, \{\theta_2, \theta_3\}, \Theta\}$. A basic belief assignment (bba), also called mass of belief, is a mapping $m(.) : 2^\Theta \to [0, 1]$ associated to a given source of evidence such that $m(\emptyset) = 0$ and $\sum_{A \in 2^\Theta} m(A) = 1$. The credibility function $Bel(.)$, commonality function $q(.)$ and the plausibility function $Pl(.)$ are also defined by Shafer and do not need to be reported here, for details see [15]. The functions $m(.)$, $Bel(.)$, $q(.)$ and $Pl(.)$ are in one-to-one correspondence.

Let $m_1(.)$ and $m_2(.)$ be two bba's provided by two independent sources of evidence over the frame of discernment Θ. The combination of $m_1(.)$ with $m_2(.)$, denoted $m(.) = [\mathbf{m}_1 \oplus \mathbf{m}_2](.)$ is obtained in DST framework by Dempster's rule of combination as follows:

$$\begin{cases} m(\emptyset) = 0 \\ m(A) = \frac{\sum\limits_{X_1 \cap X_2 = A} m_1(X_1) m_2(X_2)}{\sum\limits_{X_1 \cap X_2 \neq \emptyset} m_1(X_1) m_2(X_2)} \quad \forall A \neq \emptyset, A \in 2^\Theta \end{cases} \tag{1}$$

The total degree of conflict between the two sources of evidence is defined by

$$m_{12}(\emptyset) \triangleq \sum_{\substack{X_1, X_2 \in 2^\Theta \\ X_1 \cap X_2 = \emptyset}} m_1(X_1) m_2(X_2) \tag{2}$$

Dempster's rule can be directly extended to the combination of n independent and equally reliable sources. It is a commutative and associative rule of combination.

For decision-making, probabilistic transformations including BetP [3], [19], and DSmP [4] are commonly used to approximate any bba resulting of a fusion rule to a subjective probability measure which can then be used classically with the utility theory.

B. Probabilistic transformations

BetP transformation: Let m be a given bba related with Θ satisfying Shafer's model[2]. Its pignistic probability, denoted $BetP$, is given by [19]:

$$BetP(Y) = \sum_{X \subset 2^\Theta, Y \subseteq X} \frac{|X \cap Y|}{|X|} m(X) \tag{3}$$

where $|X|$ is the cardinality of subset X. The formula (4) can be rewritten for any singleton $Y \in \Theta$ as

$$BetP(Y) = \sum_{X \subset 2^\Theta, Y \subseteq X} \frac{1}{|X|} m(X) \tag{4}$$

DSmP transformation: $DSmP$ [4], [16] is a generalized probabilistic transformation which can work both when the frame Θ satisfies Shafer's model or when it satisfies any other hybrid DSm model[3]. $DSmP$ is more efficient than $BetP$ in the sense that it allows to reach the highest PIC value of the approximate subjective probability in a consistent way as shown in details in [4], however it is a bit more complicated to implement than $BetP$ transformation. Let's consider a discrete frame Θ with a given model, the $DSmP$ transformation is defined by $\forall X \in G^\Theta \setminus \{\emptyset\}$

$$DSmP(X) = \sum_{Y \in G^\Theta} \frac{\sum\limits_{\substack{Z \subseteq X \cap Y \\ |Z| = 1}} m(Z) + \epsilon \cdot |X \cap Y|}{\sum\limits_{\substack{Z \subseteq Y \\ |Z| = 1}} m(Z) + \epsilon \cdot |Y|} m(Y) \tag{5}$$

where $\epsilon \geq 0$ is a tuning parameter and G^Θ corresponds to the hyper-power set including eventually all the integrity constraints (if any) of the model. Since we consider here for simplicity only Shafer's model for the frame Θ, then $G^\Theta = 2^\Theta$ in this case.

The formula (6) can be rewritten for any singleton $X \in \Theta$ in Shafer's model as

$$DSmP(X) = \sum_{Y \subset 2^\Theta} \frac{m(X) + \epsilon}{\sum\limits_{\substack{z \subseteq Y \\ |z| = 1}} m(z) + \epsilon \cdot |Y|} m(Y) \tag{6}$$

It is worth to note that $BetP$ transfers the mass of belief of an ignorance (partial or total) to the singletons involved in that ignorance but proportionally only with respect to the singleton cardinals to reach the maximal entropy. $BetP$ doesn't provide

[2]i.e. all elements of Θ are truly exclusive and exhaustive.

[3]For simplicity, we consider in this work Shafer's model only since most of readers are acquainted with DST, for more details with examples see [16].

the highest PIC in general as pointed out by Sudano [20] and it transfers the belief in a prudent way and that's why it can be regarded as a pessimistic transformation. $DSmP$ redistributes the ignorance mass with respect to both the singleton masses and the singletons cardinals. $DSmP$ is justified by the maximization of the PIC criterion. Stricto sensu, DSmP doesn't maximize strictly the PIC value but it provides the largest PIC value with the numerical robustness of the result. It has been proved recently in [9] that the solution obtained from the strict maximization of the PIC doesn't provide robust numerical results and thus cannot be used as a valid and useful approach in practice. The parameter ϵ allows to reach the maximum value PIC of the probabilistic approximation of $m(.)$ in very specific degenerate cases, see [4] and [16], Vol. 3 for details with many examples. The smaller ϵ is, the bigger PIC value is. However, in some particular degenerate cases the $DSmP_{\epsilon=0}(.)$ values cannot be derived, while the $DSmP_{\epsilon>0}(.)$ values can always be derived by choosing ϵ as a very small positive number. We define $\epsilon = 1/1000$ by example in this paper in order to be as close as we want to the maximum of the PIC. $DSmP$ does a more specific transfer of masses committed to ignorances than $BetP$ and can be considered as an optimistic probabilistic transformation. As soon as the bba $m(.)$ is Bayesian (i.e. its focal elements are only singletons) $DSmP(.)$ coincides with $BetP(.)$.

C. Discounting source of evidence

When the sources of evidences are not considered equally reliable, it is reasonable to discount each unreliable source s_i, $i = 1, 2, \ldots, n$ by a reliability factor $\alpha_i \in [0, 1]$. Following the classical discounting method [15], a new discounted bba $m'(.)$ is obtained from the initial bba $m(.)$ provided by the unreliable source s_i as follows [15]

$$\begin{cases} m'(A) = \alpha_i \cdot m(A), \quad A \neq \Theta \\ m'(\Theta) = 1 - \sum_{\substack{A \in 2^\Theta \\ A \neq \Theta}} m'(A) \end{cases} \tag{7}$$

$\alpha_i = 1$ means the total confidence in the source s_i, and the original bba doesn't need to be discounted. $\alpha_i = 0$ means that the source is s_i is totally unreliable and its bba is revised as a vacuous bba $m'(\Theta) = 1$, which must have a neutral impact in the fusion process if (as we expect) the fusion rule statisfies the neutrality of the vacuous belief assigment. In practice, the discounting method can be used efficiently if one has a good estimation of the reliability factor of each source.

III. Distances between basic belief assignments

Usually a distance between two bba's is defined to represent the dissimilarity measure between two sources of evidence. The choice for a well-adapted distance is no easy and many proposals for distances have been proposed in the literature which will not be reported here since this has been recently pusblished in detail in [8]. In this paper, we present only the most commonly used distance proposed by Jousselme et al. in [7] and the distance based on probabilistic transformations of bba's suggested in [10] since it is involved in the approach proposed in this paper.

A. Jousselme's distance

Jousselme's distance [8], denoted d_J is since recent years commonly used because it takes judiciously into account both the mass and the cardinality of focal elements of each bba's. d_J between two bba's $\mathbf{m}_1$ and $\mathbf{m}_2$ defined on the same power set 2^Θ is defined by:

$$d_J(\mathbf{m}_1, \mathbf{m}_2) = \sqrt{\frac{1}{2}(\mathbf{m}_1 - \mathbf{m}_2)^T \mathbf{D} (\mathbf{m}_1 - \mathbf{m}_2)} \tag{8}$$

where $\mathbf{m}_1$ and $\mathbf{m}_2$ denote the vectors of bba's $m_1(.)$ and $m_2(.)$ and where $\mathbf{D}$ is a $2^{|\Theta|} \times 2^{|\Theta|}$ positive matrix whose elements are defined by Jaccard's indexes $D_{ij} \triangleq \frac{|A_i \cap B_j|}{|A_i \cup B_j|}$ where A_i and B_j are elements of the power set 2^Θ. $d_J(\mathbf{m}_1, \mathbf{m}_2) \in [0, 1]$ is a distance[4] which measures the similarity between $\mathbf{m}_1$ and $\mathbf{m}_2$ taking into account both the values and the relative specificity of focal elements of each bba.

However and as shown in the following example, this distance doesn't work well to in some cases and cannot reveal the difference between belief of a single element and of non specific element in some application. Moreover, its computational complexity is a bit large.

Example 1. Let's consider the frame $\Theta = \{\theta_1, \theta_2, \cdots, \theta_n\}$ with Shafer's model and the following three independent bba's

$$\begin{aligned} \mathbf{m}_1 : &\quad m_1(\theta_1) = m_1(\theta_2) = \cdots = m_1(\theta_n) = \frac{1}{n} \\ \mathbf{m}_2 : &\quad m_2(\Theta) = 1 \\ \mathbf{m}_3 : &\quad m_3(\theta_l) = 1, \quad \text{for some } l \in \{1, 2, \ldots n\} \end{aligned}$$

According to (8), one gets:

$$d_J(\mathbf{m}_1, \mathbf{m}_2) = d_J(\mathbf{m}_1, \mathbf{m}_3) = \sqrt{\frac{1}{2}(1 - \frac{1}{n})}. \tag{9}$$

One sees that $\mathbf{m}_3$ is absolutely confident in θ_l and it is very different from $\mathbf{m}_1$ and from $\mathbf{m}_2$. Moreover, $\mathbf{m}_1$ is rather different from $\mathbf{m}_2$ even if they can be considered both as uncertain sources. The source $\mathbf{m}_2$ is truly fully ignorant since it corresponds to the vacuous belief assignment, whereas the source $\mathbf{m}_1$ is much more specific than $\mathbf{m}_2$ since it is a Bayesian belief assignment. It turns out that $\mathbf{m}_1$ corresponds actually to nothing but a "probabilistic" fully ignorant source having uniform probability mass function (pmf). As one sees from (9), the dissimilarity measure based on Jousselme's distance doesn't discriminate (as we consider) well the difference between these two very different cases for dealing efficiently with the specificity of the information.

[4]More precisely, one can only conjecture that $d_J(.)$ is a true distance measure since no proof that $\mathbf{D}$ is a positive definite matrix has been given so far in the literature.

Example 2. Let's consider the frame $\Theta = \{\theta_1, \theta_2, \theta_3\}$ with Shafer's model and the following three independent bba's

$$\begin{aligned}
\mathbf{m}_1: \quad & m_1(\theta_1) = m_1(\theta_2) = m_1(\theta_3) = 1/3 \\
\mathbf{m}_2: \quad & m_2(\theta_1) = m_2(\theta_2) = m_2(\theta_3) = 0.1, m_2(\Theta) = 0.7 \\
\mathbf{m}_3: \quad & m_3(\theta_1) = m_3(\theta_2) = 0.1, m_3(\theta_3) = 0.8
\end{aligned}$$

In this example, one sees that it is impossible to take a rational decision from $\mathbf{m}_1$ because all masses of singletons are equal. Same problem occurs with $\mathbf{m}_2$ because this second source has a very high mass of belief of its total ignorance and the masses of singletons are also the same. The sources 1 and 2 correspond to two very different situations in term of the specificity of their informational content but they yield the same problem from decision-making point of view. $\mathbf{m}_3$ assigns its largest belief to θ_3. Intuitively, it seems natural to consider $\mathbf{m}_1$ and $\mathbf{m}_2$ more closer than $\mathbf{m}_1$ and $\mathbf{m}_3$ since $\mathbf{m}_1$ and $\mathbf{m}_2$ yields the same (impossible) choice in decision-making because of the ambiguity in choice among the singletons of the frame. Using Jousselme's distance as a measure of dissimilarity, one obtains the same dissimilarity, i.e. $d_J(\mathbf{m}_1, \mathbf{m}_2) = d_J(\mathbf{m}_1, \mathbf{m}_3) = 0.4041$ which we think is not very satisfactory for such case because it means that the dissimilarity between $\mathbf{m}_1$ and $\mathbf{m}_2$ is the same as between $\mathbf{m}_1$ and $\mathbf{m}_3$ which is obviously not acceptable, nor convincing.

Such very simple examples show that the most commonly used Jousselme's distance is not sufficient to fully measure the dissimilarity betweenbba's in general and that's why some other/better approaches need to be developed. The main reason of such unsatisfactory results comes from the fact that such dissimilarity measure doesn't consider all the aspects of the dissimilarity.

B. Probabilistic-based distances

Let $\mathbf{m}_1$ and $\mathbf{m}_2$ be two bba's defined with respect to a given frame Θ with Shafer's model. We propose to define probabilistic distance between $\mathbf{m}_1$ and $\mathbf{m}_2$ through their approximate subjective probability measures. Since many transformations exist to approximate a bba into a subjective probability, we concentrate only on the two most well known and used transformations, i.e. $BetP$ and $DSmP$ decribed in the section II-B. For notation convenience, we denote the first probabilistic transformation ($BetP$) and the second probabilistic transformation ($DSmP$) of any bba $m_i(.)$ by

$$\begin{cases} P^{(1)}_{m_i}(.) \triangleq BetP_{m_i}(.) \\ P^{(2)}_{m_i}(.) \triangleq DSmP_{m_i}(.) \end{cases}$$

where $BetP_{m_i}(.)$ is the pignistic transformation of $m_i(.)$ obtained with formula (3) in replacing $m(.)$ by $m_i(.)$ and where $DSmP_{m_i}(.)$ is the DSmP transformation of $m_i(.)$ obtained with formula (6) in replacing $m(.)$ by $m_i(.)$.

- **The MaxDiff distance**

In 2006, W. Liu has proposed in [10], the *MaxDiff* distance between two bba's as

$$MaxDiff(BetP_{\mathbf{m}_1}, BetP_{\mathbf{m}_2}) = \max_{A \in \Theta} |BetP_{\mathbf{m}_1}(A) - BetP_{\mathbf{m}_2}(A)| \quad (10)$$

Such kind of distance can be defined also using DSmP transformation (or any other probabilistic transformations) as well. The MaxDiff distance reflects the variation only by the maximal distance between the (pignistic) probabilities of a pair of the individual element. However, it is not adapted for measuring precisely the total amount of difference between two bba's as shown in the next example.

Example 3. Let's consider the frame of discernment $\Theta = \{\theta_1, \theta_2, \cdots \theta_n, \theta_{n+1}, \cdots, \theta_{n+t}\}$ with Shafer's model and the following two pairs of bba's from different sources:

$$\begin{cases} m^1_1(\theta_1 \cup \theta_2 \cup \ldots \cup \theta_n) = 1 \\ m^1_2(\theta_{n+1} \cup \theta_{n+2} \cup \ldots \cup \theta_{n+t}) = 1 \end{cases}$$
$$\begin{cases} m^2_1(\theta_1 \cup \theta_2 \cup \ldots \cup \theta_l) = 1 \\ m^2_2(\theta_1 \cup \theta_2 \cup \cdots \cup \theta_f) = 1 \end{cases}$$

As we can see, $\mathbf{m}^1_1$ and $\mathbf{m}^1_2$ totally contradict with each other, but $MaxDiff(BetP_{\mathbf{m}^1_2}, BetP_{\mathbf{m}^1_1}) = \max\{\frac{1}{n}, \frac{1}{t}\} \to 0$ when $n, t \to \infty$. $\mathbf{m}^2_1$ and $\mathbf{m}^2_2$ are much more similar and consistent. Let's now take $\frac{f}{2} < l < f$, then

$$MaxDiff(BetP_{\mathbf{m}^2_1}, BetP_{\mathbf{m}^2_2}) = \frac{1}{f}.$$

For instance, if one takes $t = n = 10$, and $f = 3$, $l = 2$, then

$$\begin{aligned} MaxDiff(BetP_{\mathbf{m}^1_2}, BetP_{\mathbf{m}^1_1}) &= 0.1 \\ &< MaxDiff(BetP_{\mathbf{m}^2_1}, BetP_{\mathbf{m}^2_2}) = 0.333 \end{aligned}$$

According to the value of $MaxDiff$, the difference between $\mathbf{m}^1_1$ and $\mathbf{m}^1_2$ is larger than the difference between $\mathbf{m}^2_1$ and $\mathbf{m}^2_2$. This result shows that the MaxDiff distance doesn't work well in such very simple case like this one.

- **Minkowski distances**

In this paper, we propose to use Minkowski's distances denoted $DistP^{(k)}_t$ for $k = 1, 2$ (the index k specifies the type of the probabilistic transformation under consideration; $k = 1$ means BetP, whereas $k = 2$ means DSmP) and defined by

$$DistP^{(k)}_t(m_1, m_2) = (\frac{1}{2} \sum_{\substack{\theta_i \in \Theta \\ |\theta_i| = 1}} |P^{(k)}_{m_1}(\theta_i) - P^{(k)}_{m_2}(\theta_i)|^t)^{\frac{1}{t}} \quad (11)$$

for $t \geq 1$, and $k = 1, 2$.
The coefficient $\frac{1}{2}$ in (11) is to satisfy $DistP^{(k)}_t(m_1, m_2) \in [0, 1]$. This can be proven as follows: since $P^{(k)}_{m_i}(w_i) \in [0, 1]$

and $t \geq 1$ then the following inequalities hold

$$\begin{aligned} 0 \leq & \sum_{\substack{w_i \in \Theta \\ |w_i|=1}} |P^{(k)}_{m_1}(w_i) - P^{(k)}_{m_2}(w_i)|^t \\ \leq & \sum_{\substack{w_i \in \Theta \\ |w_i|=1}} |P^{(k)}_{m_1}(w_i) - P^{(k)}_{m_2}(w_i)| \\ \leq & \sum_{\substack{w_i \in \Theta \\ |w_i|=1}} |P^{(k)}_{m_1}(w_i)| + \sum_{\substack{w_i \in \Theta \\ |w_i|=1}} |P^{(k)}_{m_2}(w_i)| = 2 \end{aligned}$$

Therefore

$$\begin{aligned} & 0 \leq \frac{1}{2} \sum_{\substack{w_i \in \Theta \\ |w_i|=1}} |P^{(k)}_{m_1}(w_i) - P^{(k)}_{m_2}(w_i)|^t \leq 1 \Rightarrow \\ & 0 \leq (\frac{1}{2} \sum_{\substack{w_i \in \Theta \\ |w_i|=1}} |P^{(k)}_{m_1}(w_i) - P^{(k)}_{m_2}(w_i)|^t)^{\frac{1}{t}} \leq 1 \end{aligned}$$

which completes the proof.

Note that when $t = 1$, this Minkovski's distance corresponds to the well-known city-block (a.k.a. Manhattan distance) and when $t = 2$ it corresponds to the classical Euclidean distance. For some cases, it can happen that $DistP^{(1)}_t = DistP^{(2)}_t$ whenever $DSmP(.) = BetP(.)$.

The type of distance characterized by the choice of the parameter t can be tuned by the user. The larger t leads to the larger complexity burden. When two sources of evidence are in total conflict, such distances do not work well if $t > 1$.

Example 4. Let's consider the frame $\Theta = \{\theta_1, \theta_2, \ldots, \theta_{2n}\}$ with Shafer's model and the following two independent bba's

$$\begin{aligned} \mathbf{m}_1 : & \quad m_1(\theta_1) = m_1(\theta_2) = \cdots = m_1(\theta_n) = 1/n \\ \mathbf{m}_2 : & \quad m_2(\theta_{n+1}) = m_2(\theta_{n+2}) = \cdots = m_2(\theta_{2n}) = 1/n \end{aligned}$$

In this example $\mathbf{m}_1$ and $\mathbf{m}_2$ totally contradict with each other. The distance measures between bba's $\mathbf{m}_1$ and $\mathbf{m}_2$ is shown in the Fig.1.

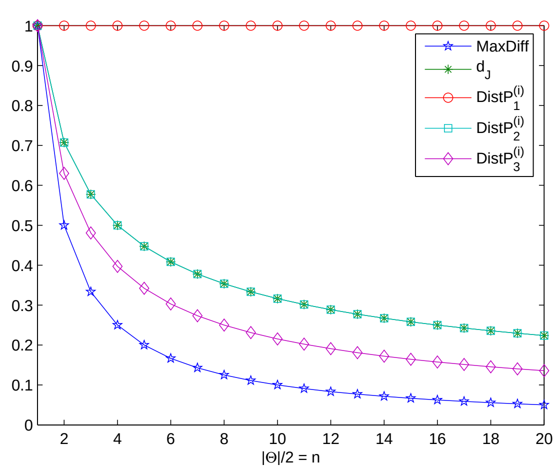

Fig.1: Distances measures between $\mathbf{m}_1$ and $\mathbf{m}_2$.

$DistP^{(1)}_t = DistP^{(2)}_t$ in Fig 1 and $DistP^{(i)}_2 = d_J$, since $\mathbf{m}_1$ and $\mathbf{m}_2$ are Bayesian bba's, and $BetP(.) = DSmP(.)$. Therefore, the plots for $DistP^{(i)}_2$ and d_J coincide on the figure. The values of $DistP^{(i)}_2$ and $DistP^{(i)}_3$ tends towards 0, meaning that $\mathbf{m}_1$ and $\mathbf{m}_2$ are closer and closer with the increase of n, which is obviously abnormal. The larger value of the tuning parameter t makes the dissimilarity decrease faster. $MaxDiff$ converges towards 0 with the fastest rate when $|\Theta|/2 = n$ increases. Only $DistP^{(i)}_1 = 1$ is invariable, and it shows that $\mathbf{m}_1$ and $\mathbf{m}_2$ are completely different, which correctly reflects their dissimilarity. Moreover, the computation burden is low when using $t = 1$. So we did choose $t = 1$ in this paper.

Lemma 1: Let $\mathbf{m}_1$, $\mathbf{m}_2$ be two bba's defined on 2^Θ. The probabilistic-based distance $DistP(\mathbf{m}_1, \mathbf{m}_2) \in [0, 1]$. If $\mathbf{m}_1 = \mathbf{m}_2$, then $DistP(\mathbf{m}_1, \mathbf{m}_2) = 0$, but its reciprocal is not true. If $DistP(\mathbf{m}_1, \mathbf{m}_2) = 1$, then $\mathbf{m}_1$ and $\mathbf{m}_2$ totally contradict and therefore there is none compatible elements supported by the both bba's, and its reciprocal is true.

In the example 1, one has $DistP(\mathbf{m}_1, \mathbf{m}_2) = 0$ and $DistP(\mathbf{m}_1, \mathbf{m}_3) = DistP(\mathbf{m}_2, \mathbf{m}_3) = (n-1)/n$. In the example 2, one has $DistP(\mathbf{m}_1, \mathbf{m}_2) = 0$ and $DistP(\mathbf{m}_2, \mathbf{m}_3) = 0.4666$. In the example 3, let assume $f \geq l$, $DistP(\mathbf{m}^1_1, \mathbf{m}^1_2) = 1 > DistP(\mathbf{m}^2_1, \mathbf{m}^2_2) = 1 - \frac{l}{f}$.

For the bba's taken in the examples 1, 2 and 3, one has $DistP^{(1)} = DistP^{(2)}$. The distances $DistP^{(k)}$, $k = 1, 2$ say that $\mathbf{m}_1$, $\mathbf{m}_2$ is closer than $\mathbf{m}_2$ and $\mathbf{m}_3$ in both Example 1 and 2, and $\mathbf{m}^1_1$ and $\mathbf{m}^1_2$ is much more similar than that of $\mathbf{m}^2_1$ and $\mathbf{m}^2_2$ in Example 3. The results are consistent with what is intuitively expected as an acceptable behavior. Moreover, the computation is much simpler than that the use of d_J measure.

$DistP$ characterizes the dissimilarity between bba's by the absolute distance between their associate subjective probabilities. In this dissimilarity measure, the degree of the divergence of hypothesis that different sources of evidence strongly support plays an important role. Unfortunately, $DistP$ is unable to reveal the divergence at all.

Lemma 2: Even if the distance/disimilarity measures are high, the two bba's can however possibly and strongly support the same hypothesis and even if the distance measures are low, the bba's can still assign the most belief to different incompatible elements of the frame.

This lemma is illustrated/proved through the following simple examples.

Example 5. Let's consider the frame $\Theta = \{\theta_1, \theta_2, \dots, \theta_{2n-1}\}$ with Shafer's model and the following three independent bba's

$$\mathbf{m}_1: \begin{cases} m_1(\theta_1) = \cdots = m_1(\theta_{n-1}) = \frac{1}{n} - \frac{\epsilon}{n-1} \\ m_1(\theta_n) = \frac{1}{n} + \epsilon, \end{cases}$$

$$\mathbf{m}_2: \begin{cases} m_2(\theta_1) = \frac{1}{n} + \epsilon, \\ m_2(\theta_2) = \cdots = m_2(\theta_n) = \frac{1}{n} - \frac{\epsilon}{n-1} \end{cases}$$

$$\mathbf{m}_3: \begin{cases} m_3(\theta_n) = \frac{1}{n} + \epsilon, \\ m_3(\theta_{n+1}) = \cdots = m_2(\theta_{2n-1}) = \frac{1}{n} - \frac{\epsilon}{n-1} \end{cases}$$

In this example, one has $DistP^{(1)} = DistP^{(2)}$ and the various distances between these bba's are ($\epsilon \to 0$) :

$$\begin{cases} d_J(\mathbf{m}_1, \mathbf{m}_2) = \epsilon + \frac{\epsilon}{n-1}, \\ d_J(\mathbf{m}_1, \mathbf{m}_3) = (\frac{1}{n} - \frac{\epsilon}{n-1})\sqrt{n-1} \end{cases}$$

$$\begin{cases} MaxDiff(BetP_{\mathbf{m}_1}, BetP_{\mathbf{m}_2}) = \epsilon + \frac{\epsilon}{n-1}, \\ MaxDiff(BetP_{\mathbf{m}_1}, BetP_{\mathbf{m}_3}) = \frac{1}{n} - \frac{\epsilon}{n-1} \end{cases}$$

$$\begin{cases} DistP(\mathbf{m}_1, \mathbf{m}_2) = \epsilon + \frac{\epsilon}{n-1}, \\ DistP(\mathbf{m}_1, \mathbf{m}_3) = 1 - \frac{1}{n} - \epsilon \end{cases}$$

Although $\mathbf{m}_1$ and $\mathbf{m}_3$ strongly support the same hypothesis θ_n, and $\mathbf{m}_2$ is different from $\mathbf{m}_1$ and $\mathbf{m}_3$ supporting θ_1, the dissimilarity between $\mathbf{m}_1$ and $\mathbf{m}_3$ is larger than the dissimilarity between $\mathbf{m}_1$ and $\mathbf{m}_2$ according to distance measures.

The divergence of the hypothesis that the two bba's strongly support cannot be taken into account efficiently from $DistP$ measure of dissimilarity. This remark implies that $DistP$ is not sufficient enough to measure the dissimilarity, and another measure as the complement of probabilistic-based distance to reflect the degree of divergence/conflict among the belief functions is necessary.

IV. Intrinsic conflict among belief functions

As in [10], a qualitative definition of conflict between two beliefs in the context of DST is given.

Definition 2: A conflict between two beliefs can be interpreted qualitatively as one source strongly supports one hypothesis and the other strongly supports another hypothesis, and the two hypotheses are not compatible (their intersection is empty).

This definition is intuitively consistent, and it will be adopted here. According this definition, the conflict mainly comes from pairs of incompatible hypothesis which are separately strongly supported by two different sources of evidence. So the extent of conflict should mainly be reflected by the conflicting beliefs of the pair of incompatible hypothesis.

The total degree of conflict, denoted $m_\oplus(\emptyset) \equiv m_{12}(\emptyset)$, is generally used to evaluate the level of conflict [18] between the two sources of evidence. From (2), one sees that $m_\oplus(\emptyset)$ is the sum of all the masses of belief committed to the empty through the conjunctive rule of combination. It is not so appropriate to measure the conflict between bba's, particularly in case of two equal bba's, and it cannot show the divergence of hypothesis in which two sources of evidence commits the most possibility as shown on the next example.

Example 6. Let's consider the frame $\Theta = \{\theta_1, \theta_2, \dots, \theta_n\}$ with Shafer's model and the following three independent bba's

$$\mathbf{m}_1: \begin{cases} m_1(\theta_1) = \frac{1}{n} + \epsilon, \\ m_1(\theta_2) = \cdots = m_1(\theta_n) = \frac{1}{n} - \frac{\epsilon}{n-1} \end{cases}$$

$$\mathbf{m}_2: \begin{cases} m_2(\theta_1) = \frac{1}{n} + \epsilon, \\ m_2(\theta_2) = \cdots = m_2(\theta_n) = \frac{1}{n} - \frac{\epsilon}{n-1} \end{cases}$$

$$\mathbf{m}_3: \quad m_3(\theta_l) = 1, \text{ for } l = 2, \cdots n$$

It is clear that $\mathbf{m}_1 = \mathbf{m}_2 \neq \mathbf{m}_3$, and $\mathbf{m}_1$, $\mathbf{m}_2$ are very ambiguous when $\epsilon \ll 1$, while $\mathbf{m}_3$ absolutely supports θ_l. Therefore, $\mathbf{m}_3$ is highly conflicting with $\mathbf{m}_1$ and $\mathbf{m}_2$. From (2), one sees that the total degrees of conflict are given by

$$m_\oplus^{12}(\emptyset) = 1 - \frac{1}{n} - \epsilon^2 - \frac{\epsilon^2}{n-1}$$
$$m_\oplus^{13}(\emptyset) = 1 - \frac{1}{n} + \frac{\epsilon}{n-1}$$

and therefore, when ϵ tends towards zero, one gets in the limit case $m_\oplus^{12}(\emptyset) \overset{\epsilon=0}{=} m_\oplus^{13}(\emptyset)$.

According to the value of $m_\oplus(\emptyset)$, if $\epsilon \ll 1$, $\mathbf{m}_1$ and $\mathbf{m}_2$ tends towards the total conflict when n increases. The large number of pairs of incompatible hypotheses leads to a large value of $m_\oplus(\emptyset)$. Moreover, $m_\oplus(\emptyset)$ cannot distinguish the level of conflict between $\mathbf{m}_1$ and $\mathbf{m}_2$, and between $\mathbf{m}_1$ and $\mathbf{m}_3$ at all as soon as $\epsilon = 0$.

Example 7. Let's consider the frame $\Theta = \{\theta_1, \theta_2, \theta_3\}$ with Shafer's model and the following three independent bba's

$$\mathbf{m}_1: \quad m_1(\theta_1) = 0.5, m_1(\theta_2 \cup \theta_3) = 0.4, m_1(\Theta) = 0.1$$
$$\mathbf{m}_2: \quad m_2(\theta_1) = 0.5, m_2(\theta_2) = 0.3, m_2(\theta_3) = m_2(\Theta) = 0.1$$
$$\mathbf{m}_3: \quad m_3(\theta_2) = 0.6, m_3(\Theta) = 0.4$$

From (2), one gets $m_\oplus^{12}(\emptyset) = 0.4 > m_\oplus^{23}(\emptyset) = 0.36$. The amount of conflict between $\mathbf{m}_1$ and $\mathbf{m}_2$ is higher than that of $\mathbf{m}_2$ and $\mathbf{m}_3$. Nevertheless, both $\mathbf{m}_1$ and $\mathbf{m}_2$ strongly support the same hypothesis θ_1, whereas, $\mathbf{m}_3$ distributes the most of its mass of belief to θ_2, which is obviously different from $\mathbf{m}_1$ and $\mathbf{m}_2$. Such examples shows that $m_\oplus(\emptyset)$ doesn't measure efficiently the degree of divergence of a hypothesis that two sources strongly support.

Lemma 3: If two bba's commit the most of their masses of belief onto compatible or same elements, the level of conflict between them still may be very large and even can tend towards 1 according to conflicting beliefs.

The example 6 can be used to illustrate the lemma 3. Indeed, $m_\oplus^{12}(\emptyset) = 1 - \frac{1}{n} - \epsilon^2 - \frac{\epsilon^2}{n-1} \to 1$ when $n \to \infty, \epsilon \to 0$. This implies that $m_{\oplus}(\emptyset)$ cannot correctly represent the degree of conflict in some cases.

A. A new measure of level of conflict

We want to pay more attention to the hypothesis which gets the most credibility in the bba's. If two sources of evidence commit the most possibility to compatible or same elements, we argue that they are consistent in the element they strongly support, and they do not contradict with each other. Otherwise, they are considered in conflict. In order to overcome the limitation of $m_{\oplus}(\emptyset)$, a new measure of level of conflict, called *conflict coefficient* is proposed in this paper in using probabilistic-based transformations and based on the definition 2.

Definition 3 (conflict coefficient): Let $\mathbf{m}_1$ and $\mathbf{m}_2$ be two bba's on 2^{Θ}. Their associated subjective probabilities are $P^{(k)}_{m_i}(.), i = 1,2; k = 1,2$ as defined in Section II-B. The Conflict coefficient, denoted $ConfP \triangleq ConfP^{(k)}$, $k = 1,2$ is defined by

$$ConfP^{(k)}(m_1, m_2) = \Psi(m_1, m_2)P^{(k)}_{m_1}(x_1)P^{(k)}_{m_2}(x_2) \quad (12)$$

where

$$\Psi(m_1, m_2) = \begin{cases} 0, & \text{if} \quad X_1 \cap X_2 \neq \emptyset, x_i \in X_i \\ 1, & \text{if} \quad X_1 \cap X_2 = \emptyset, x_i \in X_i \end{cases}$$

and $X_i = \{x_i | P^{(k)}_{m_i}(x_i) = \max(P^{(k)}_{m_i}(.))\}$.

Naturally, it can happen in some particular cases that $ConfP^{(1)} = ConfP^{(2)}$ whenever $DSmP(.) = BetP(.)$. The conflict coefficient is defined in using the maximal approximate subjective probability of the bba's. If two sources of evidence distribute their most mass of belief to compatible elements, there is no conflict between the two sources in such conditions. Otherwise, the amount of conflict will be represented by the product of the pair of maximal subjective probability from different sources (as illustrated again by the examples 6 and 7).

Indeed, in the example 6 one has $ConfP^{(1)}(.) = ConfP^{(2)}(.)$ and

$$\begin{cases} ConfP(\mathbf{m}_1, \mathbf{m}_2) = 0, \\ ConfP(\mathbf{m}_2, \mathbf{m}_3) = \frac{1}{n} + \epsilon \end{cases}$$

In the example 7, one has

$$\begin{cases} ConfP^{(1)}(\mathbf{m}_1, \mathbf{m}_2) = 0, \\ ConfP^{(1)}(\mathbf{m}_2, \mathbf{m}_3) = 0.3911 \end{cases}$$

and

$$\begin{cases} ConfP^{(2)}(\mathbf{m}_1, \mathbf{m}_2) = 0, \\ ConfP^{(2)}(\mathbf{m}_2, \mathbf{m}_3) = 0.55 \end{cases}$$

Our new conflict coefficient indicates that $\mathbf{m}_1$ and $\mathbf{m}_2$ strongly support a compatible hypothesis, but $\mathbf{m}_3$ is in conflict with $\mathbf{m}_2$ for these examples 6 and 7.

Lemma 4: Let m_1 and m_2 are independent bba's on 2^{Θ}. $m^{12}_{\oplus}(\emptyset) \in (0,1)$, even if $ConfP(m_1, m_2) = 0$. Also, $m^{12}_{\oplus}(\emptyset) = 1$, when $ConfP(m_1, m_2) = 1$.

The former part of the Lemma 4 is consistent with the Lemma 3. The later part can be easily proven. This lemma implies $m_{\oplus}(\emptyset)$ is not quite efficient when the bba's are not in conflict, and $m_{\oplus}(\emptyset)$ is similar with $ConfP$ in case of highly conflicting situations. The conflict coefficient reflects well the divergence of incompatible hypothesis that two sources of evidence commit most belief on. However, it ignores the other elements of bba's, so that it cannot capture the total difference among the belief of the compatible elements in the belief functions.

Example 8. Let's consider the frame $\Theta = \{\theta_1, \theta_2\}$ with Shafer's model and the following three independent bba's

$$\begin{aligned} \mathbf{m}_1: &\quad m_1(\theta_1) = 1 \\ \mathbf{m}_2: &\quad m_2(\Theta) = 1 \\ \mathbf{m}_3: &\quad m_3(\theta_1) = 0.9, m_3(\Theta) = 0.1 \end{aligned}$$

$\mathbf{m}_1$ and $\mathbf{m}_3$ are much closer than $\mathbf{m}_1$ and $\mathbf{m}_2$, since $\mathbf{m}_1$ and $\mathbf{m}_3$ distribute most of their mass of belief onto the same hypothesis θ_1, whereas $\mathbf{m}_2$ is fully ignorant (i.e. $\mathbf{m}_2$ is the vacuous belief assignment). Nevertheless, from the formula (2) and (12), one gets

$$ConfP(\mathbf{m}_1, \mathbf{m}_2) = ConfP(\mathbf{m}_1, \mathbf{m}_3) = 0$$
$$m^{12}_{\oplus}(\emptyset) = m^{13}_{\oplus}(\emptyset) = 0.$$

So in such case, we cannot make a distinction between $\mathbf{m}_1$ and $\mathbf{m}_2$, and between $\mathbf{m}_1$ and $\mathbf{m}_3$ at all only from these conflict measures.

If the proposed probabilistic-based distance is used in this example, one gets

$$\begin{cases} DistP(\mathbf{m}_1, \mathbf{m}_2) = 0.5, \\ DistP(\mathbf{m}_1, \mathbf{m}_3) = 0.05 \end{cases}$$

Naturally, the dissimilarity between $\mathbf{m}_1$ and $\mathbf{m}_2$ is quite larger than between $\mathbf{m}_1$ and $\mathbf{m}_3$ according to the probabilistic-based distance measure. Actually, probabilistic-based distance and the conflict coefficient are complementary and they separately capture different aspects involved in the dissimilarity. Altogether they can help to define a better measure the dissimilarity between bba's rather than taking only one measure separately (say the probabilistic-based distance, or the conflict coefficient).

Definition 4 (A new dissimilarity mesure): Let $\mathbf{m}_1$, $\mathbf{m}_2$ be bba's on 2^{Θ}. The new measurement of dissimilarity between $\mathbf{m}_1$ and $\mathbf{m}_2$ is defined by a 2D vector as follows:

$$DismP(\mathbf{m}_1, \mathbf{m}_2) = < DistP(\mathbf{m}_1, \mathbf{m}_2), ConfP(\mathbf{m}_1, \mathbf{m}_2) > \quad (13)$$

$k = 1, 2$. The first component of $DismP$ is the probabilistic-based distancee and the second component is the conflict coefficient. Sometimes, it is much more convenient to consider only a scalar measure of the dissimilarity. This can be obtained easily using the weighted arithmetic mean as follows:

$$\begin{aligned} dismP(\mathbf{m}_1, \mathbf{m}_2) = \gamma_D DistP(\mathbf{m}_1, \mathbf{m}_2) \\ + \gamma_C ConfP(\mathbf{m}_1, \mathbf{m}_2) \end{aligned} \quad (14)$$

where the weighting factors γ_C and γ_D are in $[0, 1]$ and such that $\gamma_C + \gamma_D = 1$ and must be tuned according to the application. Since $DistP(\mathbf{m}_1, \mathbf{m}_2)$ and $ConfP(\mathbf{m}_1, \mathbf{m}_2)$ are also in [0,1], then $dismP(\mathbf{m}_1, \mathbf{m}_2) \in [0, 1]$. If $DistP$ is considered more important than $ConfP$ in the $dismP$, we will take $\gamma_D > \gamma_C$, as for example in the evaluation of reliability of sources of evidence. Otherwise, one can take $\gamma_C > \gamma_D$, for example as in the selection of adapted combination rules.

In this new dissimilarity measure, $DSmP$ provides the most specific/optimistic transfer of masses of uncertainties to singletons of the frame and by tuning the parameter ϵ one can reach the maximum of the PIC while preserving the numerical robustness of the result. So it is appropriate to use $DSmP$ instead of $BetP$ if specific results are needed in some applications. However $DSmP$ requires more computation resources than $BetP$. $BetP$ transfers the belief committed to ignorances in a very prudent/pessimistic way onto the singletons of the frame. If we want to keep the nonspecific results in case of very uncertain information, $BetP$ can be chosen instead of DSmP.

V. Application and comparisons of our approach

A. Determination of reliability factors for sources of evidence

Usually the evidences arising from different independent sources are considered equally reliable in the combination process, when the prior knowledge about the reliability of each source is unknown. However, all the sources to combine can have different reliability in real applications[5]. If the sources of evidence are considered as equi-reliable, the unreliable ones may bring a very bad influence in the combination result, and even leads to inconsistent results and wrong decisions. Thus, the reliability of each source must be taken into account in the fusion process as best as possible to provide a useful and unbiased result.

In this work, we propose to evaluate on the fly the reliability of the sources to combine based on the dissimilarity measure. From this reliability measure, one can then discount the unreliable sources accordingly before applying a rule of combination of bba's. The reliability of the sources is related with the difference among the belief functions under the underlying principle that the "Truth lies in the majority opinion". That's why the distance measure plays a more important role in the dissimilarity than the conflict coefficient so that the weighting factor in $dismP$ are chosen such that $1 \geq \gamma_D > \gamma_C \geq 0$ with $\gamma_D + \gamma_C = 1$.

If there is a batch of n sources of evidence to combine, the dissimilarity between each pair of sources can be obtained following from (14). The mutual support degree among these sources is then given by:

$$\sup(m_i, m_j) = (1 - dismP(m_i, m_j)^{\lambda})^{\frac{1}{\lambda}} \quad (15)$$

For simplicity, one suggest to take $\lambda = 1$. The mutually support degree $n \times n$ matrix is then defined by

$$\mathbf{S} = \begin{bmatrix} 1 & \sup_{12} & \dots & \sup_{1n} \\ \sup_{21} & 1 & \dots & \sup_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ \sup_{n1} & \sup_{n2} & \dots & 1 \end{bmatrix} \quad (16)$$

where $\sup_{ij} \triangleq \sup(m_i, m_j)$.

The Perron-Frobenius vector (the eigen vector associated to the maximal positive eigen value) of $\mathbf{S}$ is used as the credibility factor denoted $\beta = [\beta_1, \beta_2, \dots, \beta_n]'$, that is $\lambda_{\max} \cdot \beta = \mathbf{S} \cdot \beta$.

The source with the largest reliability factor is considered as totally credible, and there is no need to revise this source. The other sources are discounted classically as follows:

$$\begin{cases} m'_i(w_j) = \alpha_i \cdot m'_i(w_j), & \text{for } w_j \neq \Theta \\ m'_i(\Theta) = 1 - \sum m'_i(w_j) \end{cases} \quad (17)$$

where $\alpha_i = \beta_i / \max(\beta)$ which can be called the *relative reliability factor* of the source no. i.

Example 9. Let's consider the frame $\Theta = \{\theta_1, \theta_2, \theta_3\}$ with Shafer's model and the following five independent bba's

$$\begin{aligned}
\mathbf{m}_1 : &\quad m_1(\theta_1) = 0.8, m_1(\theta_2) = 0.1, m_1(\Theta) = 0.1 \\
\mathbf{m}_2 : &\quad m_2(\theta_1) = 0.4, m_2(\theta_2) = 0.25, m_2(\theta_3) = 0.2, \\
&\quad m_2(\theta_2, \theta_3) = 0.15 \\
\mathbf{m}_3 : &\quad m_3(\theta_2) = 0.9, m_3(\theta_3) = 0.1 \\
\mathbf{m}_4 : &\quad m_4(\theta_1) = 0.35, m_4(\theta_2) = 0.1, m_4(\theta_3) = 0.35, \\
&\quad m_4(\Theta) = 0.2 \\
\mathbf{m}_5 : &\quad m_5(\theta_1) = 0.3, m_5(\theta_1, \theta_2) = 0.25, m_5(\theta_3) = 0.1, \\
&\quad m_5(\Theta) = 0.35
\end{aligned}$$

We assume no prior knowledge about reliability of these five sources of evidence. The weighting factors needed in $dismP$ formula are chosen here as $\gamma_D = 1/1.5 = 2/3$ and $\gamma_C = 0.5/1.5 = 1/3$. The $BetP$ and $DSmP$ probabilistic transformations of the bba's are given in the following tables.

[5]They can also have different importances as well in some applications and a new method for dealing with the different importances of sources has been developped recently in [17]. For sake of simplicity, we consider only the reliability of the sources in this paper.

Table 1. Probabilistic transformation by using $BetP$.

	θ_1	θ_2	θ_3	$PIC(.)$
$BetP_{m_1}(.)$	0.8333	0.1333	0.0333	0.5140
$BetP_{m_2}(.)$	0.5000	0.4250	0.0750	0.1767
$BetP_{m_3}(.)$	0	0.9000	0.1000	0.7041
$BetP_{m_4}(.)$	0.4167	0.1667	0.4167	0.0641
$BetP_{m_5}(.)$	0.5417	0.2417	0.2167	0.0837

Table 2. Probabilistic transformation by using $DSmP$.

	θ_1	θ_2	θ_3	$PIC(.)$
$DSmP_{m_1}(.)$	0.8887	0.1112	0.0001	0.6813
$DSmP_{m_2}(.)$	0.5230	0.4764	0.0006	0.3659
$DSmP_{m_3}(.)$	0	0.9000	0.1000	0.7041
$DSmP_{m_4}(.)$	0.4374	0.1252	0.4374	0.1048
$DSmP_{m_5}(.)$	0.8106	0.0017	0.1877	0.5494

From the tables 1 and 2, one sees that the PIC value based on $DSmP$ is larger than the one based on $BetP$, and the probabilities from $DSmP$ are all more specific than that from $BetP$. So we argue that the results of $DSmP$ is optimistic, whereas the results of $BetP$ is more pessimistic. The computation burden of $DSmP$ is a bit larger than of $BetP$.

The dissimilarity measures among these bba's can be obtained according to the formula (13) and (14). The degree of mutually support is calculated by the formula (15) and the support degree matrices are then given by

$$\mathbf{S}^{(1)} = \begin{bmatrix} 1.0000 & 0.7778 & 0.1944 & 0.7222 & 0.8056 \\ 0.7778 & 1.0000 & 0.5167 & 0.7722 & 0.8778 \\ 0.1944 & 0.5167 & 1.0000 & 0.3861 & 0.3986 \\ 0.7222 & 0.7722 & 0.3861 & 1.0000 & 0.8667 \\ 0.8056 & 0.8778 & 0.3986 & 0.8667 & 1.0000 \end{bmatrix}$$

$$\mathbf{S}^{(2)} = \begin{bmatrix} 1.0000 & 0.7562 & 0.1409 & 0.6991 & 0.8749 \\ 0.7562 & 1.0000 & 0.4944 & 0.7088 & 0.6835 \\ 0.1409 & 0.4944 & 1.0000 & 0.3522 & 0.1580 \\ 0.6991 & 0.7088 & 0.3522 & 1.0000 & 0.7512 \\ 0.8749 & 0.6835 & 0.1580 & 0.7512 & 1.0000 \end{bmatrix}$$

The two matrices are very similar, but there are still some little differences. The mutual support degree among bba's by $DSmP$ in matrix $\mathbf{S}^{(2)}$ is smaller than that by $BetP$ in matrix $\mathbf{S}^{(1)}$ in most cases. It indicates the dissimilarity by $DSmP$ is larger than that by $BetP$, especially for the dissimilarity between $\mathbf{m_3}$ with respect to the others, since the probabilistic transformation of $m_3(.)$ is invariant since $m_3(.)$ is already a Bayesian bba, but the probabilities of the other bba's become more specific with the $DSmP$ than with the $BetP$.

The eigenvectors of $\mathbf{S}^{(1)}$ and $\mathbf{S}^{(2)}$ associated with the maximum positive eigenvalue are respectively

$$\beta^{(1)} = [0.4505\ 0.4942\ 0.2761\ 0.4746\ 0.5013]'$$
$$\beta^{(2)} = [0.4898\ 0.4879\ 0.2341\ 0.4788\ 0.4878]'$$

The relative reliability factors of the sources are then given by

$$\alpha^{(1)} = [0.8987\ 0.9859\ 0.5508\ 0.9468\ 1.0000]'$$
$$\alpha^{(2)} = [1.0000\ 0.9961\ 0.4778\ 0.9774\ 0.9959]'$$

As we can see $\mathbf{m}_1$, $\mathbf{m}_2$, $\mathbf{m}_4$ and $\mathbf{m}_5$ all assign their most belief on θ_1, but $\mathbf{m}_3$ oppositely commits its largest mass of belief in θ_2. $\mathbf{m}_3$ is then considered as less reliable based on the aforementioned underlying principle, and this corresponding source may be considered as a noisy source. The result of the relative reliability factors agrees pretty well with our intuition. The reliability factors of $\mathbf{m}_3$ in both $\alpha^{(1)}(.)$ and $\alpha^{(2)}(.)$ are much smaller that the others, which leads an important discounting of $\mathbf{m}_3$. Moreover, the factor of $\mathbf{m}_3$ in $\alpha^{(2)}(.)$ is even smaller than in $\alpha^{(1)}(.)$ which indicates that $\mathbf{m}_3$ will be discounted much more when using $DSmP$.

B. Selection of a combination rule

Many rules, like Dempster's rule [15] and its alternatives can be used to combine sources of evidences expressed by bba's, but they all have their drawbacks and advantages. Dempster's rule, is usually considered well adapted for combining the evidences in low conflict situations and provide a good comprise of complexity and specificity, but it involves counter-intuitive behaviors when the sources of evidences become highly conflicting. To palliate this drawback, several interesting alternatives have been proposed when Dempster's rule doesn't work well, mainly: Yager's rule [21], DP rule [5], and PCR5 [16]. The choice among PCR5, DP, and Yager's rule should depend on the actual application. PCR5 is very appropriate to use in general for decision-making because it provides the most specific fusion results, but it requires more computational resources than other rules. If we want to keep uncertain results and don't necessarily need a very specific decision in case of high conflict between sources, Yager's rule or DP rule can be selected instead.

The criteria of selection mainly concentrates on the amount of conflict between belief functions. Therefore, the conflict coefficient should be more effectual in the dissimilarity measures when applied in selection of rules, and the weighting factors must be selected such that $0 \leq \gamma_D < \gamma_C \leq 1$ with $\gamma_D + \gamma_C = 1$.

According to the properties of the combination rules, several simple general principles are present in the selection process by using the dissimilarity measures. Let $\mathbf{m}_1$ and $\mathbf{m}_2$ be two independent bba's, then the selection of the rule can be done according the following algorithm:

1) If $ConfP^{(k)}(\mathbf{m}_1, \mathbf{m}_2) = 0$, Dempster's rule is used.
2) If $ConfP^{(k)}(\mathbf{m}_1, \mathbf{m}_2) \in (0, \eta^{(k)})$ and if $dismP(\mathbf{m}_1, \mathbf{m}_2) \in (0, \mu^{(k)})$, Dempster's rule could be used only with caution, and its alternatives should be used instead.
3) If $ConfP^{(k)}(\mathbf{m}_1, \mathbf{m}_2) \geq \eta^{(k)}$, Dempster's alternative rules must be used.

The tuning of thresholds $\eta^{(k)}$, $\mu^{(k)}$ is not easy in general. If the thresholds are too large, one takes the risk to get counterintuitive results, whereas if they are set to too low values the non specificity of the result will increase and even will lead to take a decision under a big uncertainty. Therefore, they need to be determined by users' experience depending on the actual application.

Example 10. Let's consider the frame $\Theta = \{\theta_1, \theta_2, \theta_3\}$ with Shafer's model and the following four pairs of independent bba's

$$\begin{cases} m_1^1(\theta_1) = 0.6, m_1^1(\theta_2 \cup \theta_3) = 0.4 \\ m_2^1(\theta_2) = 0.9, m_2^1(\theta_3) = 0.1 \end{cases}$$
$$\begin{cases} m_1^2(\theta_2 \cup \theta_3) = 0.7, m_1^2(\Theta) = 0.3 \\ m_2^2(\theta_1) = 0.4, m_2^2(\Theta) = 0.6 \end{cases}$$
$$\begin{cases} m_1^3(\theta_1) = 0.5, m_1^3(\theta_2) = 0.3, m_1^3(\theta_3) = 0.2 \\ m_2^3(\theta_1) = 0.5, m_2^3(\theta_2) = 0.3, m_2^3(\theta_3) = 0.2 \end{cases}$$
$$\begin{cases} m_1^4(\theta_1) = 0.8, m_1^4(\theta_2) = 0.1, m_1^4(\Theta) = 0.1 \\ m_2^4(\theta_1) = 0.7, m_2^4(\theta_2 \cup \theta_3) = 0.3 \end{cases}$$

In this example, the weighting factors have been chosen as $\gamma_C = 1/1.5 = 2/3$, $\gamma_D = 0.5/1.5 = 1/3$ reflecting $\gamma_C > \gamma_D$, and the thresholds for selecting the combination rule have been taken as $\eta^{(2)} = \eta^{(1)} = 0.5$ and $\mu^{(1)} = \mu^{(2)} = 0.5$.

Table 3. Selection of rules based on $BetP$.

bba's	$m_\oplus(\emptyset)$	$ConfP$	$DistP$	$dismP$	$selection$
$\mathbf{m}_1^1, \mathbf{m}_2^1$	0.6	0.54	0.7	0.5933	Alternative
$\mathbf{m}_1^2, \mathbf{m}_2^2$	0.28	0.27	0.5	0.3467	Caution
$\mathbf{m}_1^3, \mathbf{m}_2^3$	0.62	0	0	0	Dempster's
$\mathbf{m}_1^4, \mathbf{m}_2^4$	0.31	0	0.1333	0.0444	Dempster's

Table 4. Selection of rules based on $DSmP$.

bba's	$ConfP$	$DistP$	$dismP$	$selection$
$\mathbf{m}_1^1, \mathbf{m}_2^1$	0.54	0.7	0.5933	Alternative
$\mathbf{m}_1^2, \mathbf{m}_2^2$	0.4487	0.8970	0.5981	Alternative
$\mathbf{m}_1^3, \mathbf{m}_2^3$	0	0	0	Dempster's
$\mathbf{m}_1^4, \mathbf{m}_2^4$	0	0.1877	0.0629	Dempster's

$\mathbf{m}_1^1$ and $\mathbf{m}_2^1$ are in high conflict, and Dempster's rule will be involved counterintuitive results. So its alternatives should be selected. The degree of conflict between $\mathbf{m}_1^2$ and $\mathbf{m}_2^2$ lies in the caution zone by using $BetP$, but if $DSmP$ is applied, the dissimilarity will becomes larger, and even over the threshold. Therfore, an alternative rule will be safer to use in such conditions. Although the conflicting beliefs between $\mathbf{m}_1^3$ and $\mathbf{m}_2^3$ is large, the two belief functions are actually the same, and thus Dempster's rule can be used according to the choice of our setting parameters. For the fourth pair of bba's, the dissimilarity between $\mathbf{m}_1^4$ and $\mathbf{m}_2^4$ is small, and therefore Dempster's rule could be used to combine these two bba's.

It is worth to note that the results based on dissimilarity mesure with $BetP$ or $DSmP$ are very similar, but small differences still exist. We have to select the probabilistic transformation, $BetP$ or $DSmP$, according the application under consideration, and maybe they could be used together in a more sophisticated process in future.

VI. Conclusions

In this paper, a new measurement of dissimilarity between basic belief assignments (bba's) has been proposed. The notion of dissimilarity is rather difficult to represent by an efficient scalar measure. After analyzing the limitation of the classical dissimilarity measures based either on the degree of conflict or on the distance measures, we have proposed a new dissimilarity measure which mixes both the probabilistic-based distances and a conflict coefficient and which uses $BetP$ and $DSmP$ transformations. The distance measures mainly captures the difference between two belief functions, whereas the conflict coefficient measures the degree of divergence of hypothesis that two belief functions strongly support. The distance and conflict measures can be seen as complementary for characterizing the different aspects of the dissimilarity between two bba's. The selection between $BetP$ and $DSmP$ in the dissimilarity measure depends on the application. $DSmP$ is appropriate to use if specific results are needed since it provides robust numerical results with highest PIC, but it requires more computation resources. If one prefers to get nonspecific results in case of very uncertain information, $BetP$ can be used instead. We have shown how to use the dissimilarity to estimate the reliability factors of sources when no prior knowledge is given on their reliabilities. Also several simple principles for the selection of a combination rule have been defined based on our new dissimilarity measure. The numerical examples presented in this paper illustrate clearly the potential interest of this new approach for applications dealing with evidential reasoning.

Acknowledgement

This work was supported by the grant from the National Natural Science Foundation of China(NO.60634030) and PhD Thesis Innovation Fund from Northwestern Polytechnical University(No.cx201015).

References

[1] F. Cuzzolin, *A Geometric Approach to the Theory of Evidence* IEEE Transactions on Systems, Man, and Cybernetics - Part C : Applications and Reviews, vol. 38, no. 4, pp.522-534, 2008.

[2] Y. Deng, W.K. Shi, Z.F. Zhu, Q. Liu, *Combining belief functions based on DistPance of evidence*, Decision Support Systems, Vol. 38, No. 3, pp.489-493, 2004.

[3] J. Dezert, F. Smarandache, M. Daniel, *The Generalized Pignistic Transformation*, Proceedings of 7th International Conference on Information Fusion, Fusion 2004, Stockholm, Sweden, June, 2004.

[4] J. Dezert, F. Smarandache, *A new probabilistic transformation of belief mass assignment*, in Proceedings of Fusion 2008, Cologne, Germany, July 2008.

[5] D. Dubois, H. Prade, *On the unicity of Dempster's rule of combination*, International Journal of Intelligent Systems, Vol. 1, No. 2, pp.133-142, 1986.

[6] M. C. Florea, E. Bosse, and A.-L. Jousselme, *Metrics, DistPances and Dissimilarity Measures within Dempster-Shafer Theory to Characterize Sources' Reliability*in Cognitive Systems with Interactive Sensors (COGIS '09), 2009.

[7] A.L. Jousselme, D. Grenier, E. Bossé, *A new DistPance between two bodies of evidence*, Information Fusion, Vol. 2, No. 1, pp. 91-101, 2001.

[8] A.L. Jousselme, P.Maupin, *On some properties of DistPances in evidence theory*, Proceeding of workshop on the theory of belief functions, Apr.1-2, 2010 Brest, France

[9] D. Han, J. Dezert, C. Han, Y. Yang, *Is Entropy Enough to Evaluate the Probability Transformation Approach of Belief Function?*, to appear in Proc. of Fusion 2010 Conference, Edinburgh, UK, July 26-29th, 2010.

[10] W. Liu, *Analyzing the degree of conflict among belief functions*, Artificial Intelligence, Vol. 170, No. 11, pp. 909–924, 2006.

[11] Z. Liu, Q.Pan, Y.Cheng, J.Dezert, *Sequential adaptive combination of unreliable sources of evidence*, Proceeding of workshop on the theory of belief functions, Apr. 1-2, 2010 Brest, France

[12] A. Martin, A.L. Jousselme, C. Osswald, *Conflict measure for the discounting operation on belief functions*, in Proc. of Fusion 2008 Conference, Cologne, Germany, July 2008.

[13] C.K. Murphy, *Combining belief functions when evidence conflicts*, Decision Support Systems, Vol. 29, No. 1, pp. 1–9, 2000.

[14] B. Ristic and P. Smets, *The TBM global DistPance measure for the association of uncertain combat ID declarations*, Information Fusion, vol. 7, pp. 276-284, 2006.

[15] G. Shafer, *A Mathematical Theory of Evidence*, Princeton Univ. Press, 1976.

[16] F. Smarandache, J. Dezert (Editors), *Advances and Applications of DSmT for Information Fusion*, American Research Press, Rehoboth, Vol.1-3, 2004-2009.

[17] F. Smarandache, J. Dezert, J.-M. Tacnet, *Fusion of sources of evidence with different importances and reliabilities*, to appear in Proc. of Fusion 2010 Conference, Edinburgh, UK, July 26-29th, 2010.

[18] P. Smets, *Analyzing the combination of conflicting belief functions*, Information Fusion, vol. 8, pp. 387-412, 2007.

[19] Ph. Smets, *Decision making in the TBM: the necessity of the pignistic transformation* Int. Jour. Approx. Reasoning, vol. 38, 2005.

[20] J. Sudano, *The system probability information content (PIC) relationship to contributing components, combining independent multi-source beliefs, hybrid and pedigree pignistic probabilities*, Proc. of Fusion 2002, Annapolis, July 2002.

[21] R.R. Yager, *Hedging in the combination of evidence*, Journal of Information & Optimization Sciences, Vol. 4, No. 1, pp.73-81, 1983.